Journal of Memory and Language 146 (2026) 104716

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Exploring effects of foveal load and preview restrictions for single and multiple parafoveal words in Chinese reading

Manman Zhang [a,b,c], Zhichao Zhang [a,b,c], Fang Li [b], Xuejun Bai [a,b,c], Chuanli Zang [d,*], Simon P. Liversedge [e]

[a] Key Research Base of Humanities and Social Sciences of the Ministry of Education, Academy of Psychology and Behavior, Tianjin Normal University, Tianjin 300387, China
[b] Faculty of Psychology, Tianjin Normal University, Tianjin 300387, China
[c] Tianjin Key Laboratory of Student Mental Health and Intelligence Assessment, Tianjin 300387, China
[d] School of Psychology, Liverpool Hope University, Liverpool L16 9JD, United Kingdom
[e] School of Psychology, Northumbria University, Newcastle-upon-Tyne NE1 8ST, United Kingdom

ARTICLE INFO

Keywords:
Chinese reading
Eye movements
Foveal load
Parafoveal processing

ABSTRACT

Two experiments are reported that used the boundary paradigm to investigate how foveal lexical processing load (high/low frequency) of a pre-target word influences parafoveal processing of upcoming target word(s) with either zero-, one-, two- or three-character, or full preview in Chinese reading. In Experiment 1, the three characters comprised a single word as the target while in Experiment 2 they formed multiple words (two or three words). Pre-target word analyses showed an effective foveal load manipulation with low frequency pre-targets being fixated for longer than high frequency pre-targets in both experiments. Both experiments showed robust preview extent effects at the target words, such that fixation times increased, and landing positions shortened dramatically with reduced preview extent. Modulatory influences of foveal load effects were obtained on both fixation times and landing positions at the target region. These effects themselves were consistent, but reduced, for parafoveal character strings comprised of multiple words relative to a single word, consistent with the MCU hypothesis (Zang, 2019). Our findings demonstrate that increased foveal load reduces the disruptive influence of restrictive parafoveal windows and reduces preview extent in relation to saccadic targeting. The current findings align at a very basic level with the Foveal Load Hypothesis (Henderson & Ferreira, 1990), though the results indicate that a more nuanced theoretical account is necessary to capture all aspects of the results in respect of Chinese reading.

Introduction

As we read, when the eye pauses temporarily during a fixation, readers extract visual information from the fixated, foveal word, as well as from upcoming parafoveal words. In the foveal area, with high visual acuity, detailed visual information is obtained allowing readers to readily recognize the fixated word. Although visual acuity declines in parafoveal and peripheral areas, some amount of preprocessing of upcoming words occurs prior to their fixation, that is, the upcoming text is previewed. One critical issue that currently remains in dispute concerns how foveal processing load, or difficulty, influences preprocessing of parafoveal information, both in terms of the spatial extent of parafoveal processing (i.e., preview extent) and the degree to which a parafoveal stimulus is processed prior to inspection (this may be considered as the preview depth).[1]

The original theoretical ideas underpinning how foveal load might constrain the spatial extent of parafoveal processing in reading were put forward by Rayner (1986). In his experiment, Rayner adopted the moving window technique (text within a pre-specified width window is shown correctly, whereas text outside the window is masked by invalid

* Corresponding author.
E-mail address: zangc@hope.ac.uk (C. Zang).

[1] It is important to be clear, here, that when we refer to the depth to which a parafoveal stimulus may be processed, we are not referring to differences in the “linguistic” depth of processing (e.g., a stimulus being processed to an orthographic level vs. a semantic level). Indeed, in the current experiments, all the preview stimuli were pseudocharacters for which it was not possible to undertake linguistic processing. Instead, we use depth of processing here to refer to how a reader might be able to pre-process a parafoveal stimulus and evaluate to a greater or lesser degree whether it is a pseudocharacter, or a real character.

https://doi.org/10.1016/j.jml.2025.104716
Received 20 November 2024; Received in revised form 14 November 2025; Accepted 16 November 2025
Available online 21 November 2025

letters or symbols as the eyes move) to examine the perceptual span of second-, fourth-, sixth-grade children and adult readers, as well as the perceptual span of fourth-grade children using easy and difficult sentences. The results showed that the right extent of the perceptual span reduced when text was difficult to read, or when reading skill was reduced (perceptual extent to the left of fixation did not change). Rayner assumed that this occurred because greater processing difficulty meant readers allocated increased attention to foveal word processing (at the cost of parafoveal processing). Note, though, that presenting texts with different levels of difficulty using the moving window paradigm in Rayner's study, potentially, led to variations in processing difficulty in foveal and parafoveal regions simultaneously as the eyes moved. Therefore, the reduced preview extent may have been caused by processing difficulty deriving from the fovea, the parafovea, or both (see Henderson & Ferreira, 1990 for discussion). In order to eliminate possible concurrent influences of parafoveal processing difficulty, and to directly examine the effects of foveal processing difficulty alone on parafoveal preview, the majority of studies examining foveal load effects have adopted the boundary paradigm (Rayner, 1975), whereby foveal and parafoveal load can be manipulated independently. Henderson and Ferreira manipulated foveal processing load associated with the pre-target word (via word frequency in Experiment 1 and via syntactic difficulty in Experiment 2) and preview validity (identical, visually similar or dissimilar) of the target word by using boundary paradigm. Their results showed that when foveal load was low, identical or visually similar previews significantly reduced fixation time on the target words compared to dissimilar previews; by contrast, when foveal load was high, valid previews did not produce shorter fixation times at the target words, that is, readers obtained a greater amount of preview benefit when foveal load was low compared to when it was high. Based on these findings, Henderson and Ferreira concluded that the perceptual span is variable in its extent and is a function of the foveal processing difficulty (c.f., Rayner, 1986), and they termed this the Foveal Load Hypothesis. Note that in both experiments, the visual characteristics of the parafoveal preview were manipulated in relation to the word's identity and were imposed on only the parafoveal word adjacent to the boundary. Despite their claims, Henderson and Ferreira did not directly manipulate the preview extent. Thus, a more rigorous examination of claims that foveal load affects the extent of parafoveal processing might directly assess limitations associated with information further into the parafovea (i.e., in relation to words beyond the adjacent word). The present study examined this question.

Based on Henderson and Ferreira's (1990) findings and the Foveal Load Hypothesis, current models of eye movement control during reading such as the E-Z Reader model (Reichle, 2011; Reichle et al., 1998; see also the Uber Reader model, Reichle, 2020), the SWIFT model (Engbert et al., 2005; Engbert & Kliegl, 2011; Schad & Engbert, 2012) and the Chinese Reading Model (CRM; Li & Pollatsek, 2020) include respective mechanisms stipulating that foveal load influences parafoveal processing. Each model, to different extents, emphasizes that foveal load affects parafoveal processing depth and/or spatial extent. However, to our knowledge, there has been no experimental work manipulating foveal load and using the boundary paradigm to manipulate preview extent and parafoveal processing. Part of the rationale for the present study was to examine this issue.

Several previous studies have used the boundary paradigm to examine how foveal load affects parafoveal preview effects and these have reported mixed results. For example, Drieghe et al. (2005) manipulated foveal load (high- or low-frequency) and preview validity (correct or incorrect) for three-letter parafoveal words. Drieghe et al. found preview benefit effects of greater magnitude than Henderson and Ferreira (1990), however, they failed to find interactive effects between foveal load and parafoveal processing. By contrast, White et al. (2005) obtained an interaction between foveal load and preview benefit using four-letter words as targets. However, the interaction only occurred for participants who were unaware of the display changes. It did not occur for those who were sensitive to the display changes. Other studies that used longer target words (e.g., 4.6 letters in Experiment 2 of Veldre & Andrews, 2018; 5.9 letters in Experiment 2a of Vasilev et al., 2018) like Henderson and Ferreira, observed interactive effects between foveal load and preview validity. Based on a *meta*-analysis reported in Veldre and Andrews, only 6 of 16 experiments demonstrated robust interactive effects of foveal load and parafoveal preview validity. Although the existing findings are mixed, it does appear that foveal load effects occur more reliably for long than short words. Given this, it is possible that foveal load effects might actually be caused by restrictions in the extent of parafoveal processing; under low foveal load conditions, readers might obtain parafoveal information about the whole of a longer parafoveal word, whereas under high foveal load conditions, readers might only obtain information about the initial part of a longer parafoveal word. And if so, this may be one reason for inconsistencies in previous findings. Other evidence that is supportive of this conclusion comes from a study by Kennison and Clifton (1995), in which foveal load effects occurred only for trials in which participants initiated saccades to target words from near compared to distant launch sites. This finding implies that the magnitude of foveal load effects may be related to the spatial extent of parafoveal preview, at least to some degree.

There have also been some studies that attempted to explore foveal load effects on parafoveal processing using the moving window paradigm. For example, Luke (2018) adopted a corpus-based approach to examine how the frequency of the fixated word n modulates parafoveal processing of word $n + 1$ during paragraph reading. Texts were presented with either a no preview window or a one-word preview window, while two words leftward were always visible across conditions. He found word frequency interacted with preview window such that word n frequency affected processing of word $n + 1$ in the one-word preview window condition, but not the no preview window condition. Similarly, Meixner et al. (2022) used the moving window paradigm to manipulate the size of a moving window (3, 4, 7, 10, 14 letters visible to both sides of the fixated character, or full line) to examine the momentary size of the perceptual span as a function of foveal word frequency during sentence reading among children. Their findings revealed two critical patterns in relation to foveal load effects on parafoveal preview. First, fixation times on word $n + 1$ were shorter when word n was high compared to low frequency when preview of upcoming words was (at least partially) available. Second, the magnitude of such effects was greater with full relative to partial previews. We note that, while the two findings suggest differential fixation patterns under different preview windows, Meixner et al. did not quantify the differential preview spans associated with levels of foveal load – a fundamental premise of both Rayner's (1986) claims and the Foveal Load Hypothesis (Henderson & Ferreira, 1990). Additionally, their results were based on children's data. Thus, the nature of foveal load effects on preview extent in adult readers remains unexplored.

In alphabetic languages like English, there is considerable word length variability. However, relative to such alphabetic languages, word length variability in Chinese is much reduced (the majority of words are one or two characters long, Li et al., 2015; Zang et al., 2011). Thus, to minimize potential influences of parafoveal word length, Zhang et al. (2019) examined foveal load effects associated with depth of parafoveal processing during Chinese reading. They used the boundary paradigm with target words formed from a single character. Even though Zhang et al.'s target words were very short and their manipulations resulted in readers launching more than 80 % of their saccades to the target word from the pre-target word, thereby maximizing the possibility that readers effectively processed the preview prior to fixation, they failed to find interactive effects of preview and foveal load. Zhang et al.'s results suggested that regardless of foveal processing difficulty, readers pre-processed the upcoming single character parafoveal word similarly.

Although Zhang et al.'s work suggests that foveal load did not modulate the depth to which a parafoveal word was processed, research has indicated that foveal load might influence the range of characters, that is, the extent over which parafoveal processes operate. For example, it has been shown that Chinese readers extract information from word $n + 2$ in the parafovea (i.e., from the word two words to the right of the fixated word), and that the extent to which the reader does this is modulated by processing difficulty associated with word $n + 1$ (e.g., M. Yan et al., 2010; see also both Yang et al., 2009 and Yang, Rayner et al., 2012 for complementary evidence). The study by M. Yan et al. explored whether parafoveal processing of word $n + 2$ (with either an identical, orthographically related, semantically related, or an unrelated preview), was affected by processing load manipulated through the frequency of the upcoming single character word $n + 1$, whilst the eyes remained fixating word *n*. M. Yan et al. found that preview benefit for word $n + 2$ with an identity preview benefit (identical vs. unrelated) was greater when word $n + 1$ was frequent relative to when it was infrequent, suggesting that processing load (associated with an upcoming word, rather than the word under fixation) modulates preview extent. This said, however, whether foveal processing load, that is, load associated with the fixated word, directly constrains preview extent during Chinese reading remains empirically unexplored.

The present study was conducted in Chinese to investigate whether, and how, foveal lexical processing load influences parafoveal preview extent both in relation to temporal (processing time) and spatial (saccadic targeting) eye movement measures. And, as we noted earlier, given that the moving window paradigm is, arguably, not the ideal method for assessing foveal load effects as the characteristics of foveal and parafoveal words vary concurrently from one fixation to the next, we chose to adopt the boundary paradigm in our experiments. The boundary paradigm allowed us to manipulate foveal processing load independent of parafoveal processing load in relation to preview extent. Thus, in our experiments we manipulated whether the pre-target word was high or low frequency (foveal load) and used the boundary paradigm (with the boundary immediately following the pre-target word) to manipulate preview extent of tightly controlled target words following the boundary. Since the rightward extent of the perceptual span in Chinese reading is approximately 3 characters from fixation (Inhoff & Liu, 1998; G. Yan, Zhang, Zhang et al., 2013), we manipulated rightward preview extent such that it was 0, 1, 2, or 3 characters, or a full preview of the sentence to the right of the boundary immediately followed the pre-target word. To align with the experimental conditions adopted by Rayner (1986), other than in our full preview condition, we replaced all the characters beyond the preview window with pseudocharacters prior to the boundary change. Note, however, that in all our experimental conditions, once the eyes crossed the boundary, the sentence was always presented in full. In this way, we considered that our experimental situation best approximated the situation in Rayner's original study, whilst allowing us to independently manipulate foveal load simultaneously with the spatial extent to which parafoveal characters were available. Thus, the current study allowed us to directly examine the claim of Rayner (1986) and predictions of the Foveal Load Hypothesis (Henderson & Ferreira, 1990) that increased foveal processing difficulty reduces the spatial extent of parafoveal preview.

In the present study, two experiments were conducted with a further important difference between these. In Experiment 1, the three characters in the parafovea comprised a single word, while in Experiment 2 they comprised two or more than two words (for 77 % of stimuli there were three single-character words in the parafovea and for 90 % of stimuli the first character was a single character word in Experiment 2).[2] A number of studies have demonstrated that single words (or Multi-Constituent Units, MCUs, that are processed as single lexicalized elements) are processed more effectively than multiple words within the same perceptual extent during Chinese reading (e.g., He et al., 2021; Zang et al., 2021; Zang, Fu et al., 2024; Zang, Wang et al., 2024). Therefore, it is possible that the lexical status of the three parafoveal target characters may determine how they are processed, and this in turn may influence the parafoveal preview effects in relation to foveal load. Thus, a priori, we planned to undertake *meta*-analyses across the data sets for the two experiments to investigate whether we might obtain interactive effects of foveal load, preview extent and experiment. Specifically, first, in line with the basic claims associated with the foveal load hypothesis, we predicted that preview extent would be wider when foveal load was low compared to when it was high, that is, an interaction between foveal load and preview extent at the target word. The current manipulations also allowed us to examine whether there were differences in the depth to which parafoveal characters that lay further into the parafovea were processed under high and low foveal load conditions. That is to say, for the one-, two-, and three- character preview conditions separately, we might predict increased preview effects when foveal load was low relative to when it was high. Finally, on the assumption that it takes longer to lexically identify two or three separate words than it takes to lexically identify a single word (see He et al., 2021), one might anticipate that any foveal load effects in relation to parafoveal processing of the upcoming character string would be increased in magnitude when those characters formed a single word than when they formed multiple words. Critically, however, we predicted that the nature of the effects, qualitatively, should remain consistent across experiments. We examined fixation times (first fixation, single fixation and gaze duration) as well as landing positions on the target word to assess whether foveal load and preview extent affected processing time and saccadic targeting in a similar manner.

Experiment 1

Method

Participants

Two hundred and fifty-six students from Tianjin Normal University participated in this experiment. Six participants were removed from analysis due to tracking loss. Therefore, there were 250 valid participants[3] (33 males; Mean Age: 22 years, $SD = 2$ years). All participants were native Chinese speakers with normal or corrected-to-normal vision. They were naïve with respect to the purpose of the experiment. Each participant was paid for their participation after testing.

Apparatus

Participants' eye movements were recorded using an SR Research Eyelink 1000 system at a sampling rate of 1000 Hz. Due to experimental testing constraints, the experiment was run during two different sessions using two different eye trackers (an Eyelink 1000 tower eye tracker and an Eyelink 1000 Plus eye tracker). All sentences were displayed in black in Song font with a white background on a 19-inch CRT monitor with a 1024×768 pixel resolution and a refresh rate of 150 Hz during the first testing session, and on a 24-inch ASUS VG248QE monitor with a 1024×768 pixel resolution and a refresh rate of 144 Hz in the second session. The viewing distance was 65 cm, and one Chinese character subtended 1.1° of visual angle. Experiment Builder software (SR Research) was used to control the presentation of the experimental stimuli and to

[2] To be explicit about the number of words that comprised the three character parafoveal region, 77% of stimuli had a "1+1+1" word structure, 15% had a "1+2" or a "2+1" word structure and 8% had a "1+1+2" word structure (wherein the third character was the first character of a two character word that extended beyond the three character parafoveal window).

[3] In our original power analyses, we computed the power by using the effect size for the interactive effect of foveal load and parafoveal preview (0.52) in the original study from Henderson and Ferreira (1990). The results indicated that at least 80 participants in both experiments are required to achieve a power of 0.8. However, given the quite mixed findings in previous studies, we wished to maximize the chances of finding robust effects if at all possible, and therefore, we tested a larger number of participants (250 valid) in each experiment.

undertake the calibration procedures.

Materials and design

Foveal load was manipulated via word frequency as in previous studies (e.g., Drieghe et al., 2005; Henderson & Ferreira, 1990; Veldre & Andrews, 2018; Zhang et al., 2019). One hundred pairs of high and low frequency two-character pre-target words were selected from Cai and Brysbaert (2010). High frequency words were significantly more frequent than low frequency words, F (1, 99) = 61.68, $p < .001$, $\eta_p^2 = 0.43$. To minimize differences in visual processing, each pair of pre-target words shared the same second character and first characters were matched in respect of stroke number, F (1, 99) = 1.11, $p = .295$. The basic descriptive characteristics of the pre-target and target words are shown in Table 1. One hundred three-character target words were also selected, each of which fitted well with both versions of the pre-target word within the sentence frame that we developed for each item (see Fig. 1).

Each of the pair of pre-target words and the target word were embedded into a sentence frame for which the content up to the pre-target word was identical. The pre-target and target words were positioned in approximately the middle of the sentence. All the sentences were between 16 and 23 characters in length ($M = 21$, $SD = 2$). Three separate groups of participants from Tianjin Normal University who did not take part in the eye tracking experiments were asked to rate the naturalness of all sentences, the predictabilities of the pre-target words, and the predictabilities of the target words respectively. For these rating, 42 participants (21 in each of the counterbalanced low- and high-load conditions) rated the naturalness of sentences on a 5-point scale in which "1" meant "very unnatural", and "5" meant "very natural". The mean naturalness of all sentences was 3.89 ($SD = 0.31$), with no difference between low- and high-load conditions, F (1, 99) = 1.06, $p = .305$. A second group of 19 participants assessed the predictability of the pre-target word by providing completions to sentence fragments up to and including the pre-target word in a cloze task. A third group of 38 participants rated the predictability of the target word, with 19 participants in each of the counterbalanced low- and high-load conditions performing the cloze task for sentence fragments up to and including the target word. Both predictabilities of the pre-target and the target words were very low[4] (see Table 1) with no significant differences between low and high load conditions, Fs < 0.24, ps > .05.

The boundary paradigm (Rayner, 1975) was used to manipulate parafoveal preview extent of the target word (i.e., the number of characters available for preview). There were 5 preview conditions (see Fig. 1), that is, zero (0CP), one (1CP), two (2CP) or three characters (3CP) of the parafoveal target word (the remainder of the sentence was masked by pseudocharacters that were matched for stroke number with the character they replaced), or the full sentence preview (FP). Note that the sentence was presented in full in all conditions as soon as the eyes crossed the boundary. The current experiment, thus, was a 2 (Foveal Load: Low, High) × 5 (Parafoveal Preview Extent: 0CP, 1CP, 2CP, 3CP, FP) within subject design. Ten files were constructed. Conditions were rotated across files following a Latin square design. Each file consisted of 100 experimental sentences (10 for each condition) and 43 filler sentences. Thirty-five percent of the experimental and filler sentences were followed by a comprehension question that participants answer by pressing a "Yes" or "No" key. Sentences were presented in a random order. Also, 8 practice sentences, 4 of which were followed by a comprehension question, were presented at the beginning of the testing session.

Procedure

Each participant was tested individually and received stimuli from one file only. They were instructed to read sentences for comprehension, and to press a keyboard button to indicate that they had finished reading each sentence. Participants were told to answer the yes/no comprehension questions when they appeared after a proportion of the sentences. Before recording participants' eye movement data, a three-point horizontal calibration procedure occurred. The average calibration error was less than 0.30 degrees. During testing, each trial began with a drift check positioned at the beginning of the sentence. The participant was recalibrated if the value of the drift correct was greater than 0.35 degrees. After testing, participants were asked to complete a questionnaire to assess their awareness of the boundary display change (White et al., 2005). The experiment lasted approximately 50 min.

Results and discussion

The average comprehension accuracy for all participants was 91 % ($SD = 5$ %), indicating that all participants understood the sentences well. Before analyzing our data, we merged fixations below 80 ms with adjacent fixations that were less than 0.5 degrees of visual angle away, or fixations below 40 ms with fixations within 1.25 degrees of visual angle. Then, we removed fixations below 80 ms or above 1200 ms (4.8 % of total fixations). Trials were eliminated in which (a) a track loss occurred or there were fewer than four fixations in total (0.2 % of the data); (b) for the target word analyses, the display change occurred early (i.e., the eyes initially fixated very briefly slightly to the right of the boundary and then quickly moved back to the left of the boundary, 2.7 %), or was delayed (i.e., the display change occurred more than 10 ms after fixation onset on the target word; 3.7 %), or hooking occurred (i.e., when a saccade crossed the boundary from the left and triggered a display change but the fixation landed slightly to the left rather than to the right of the boundary, Degno et al., 2021; 5.5 %), or other incorrect triggering of display changes by blinks or saccades occurred (2.6 %); (c) for each measure and each participant, any observations more than three standard deviations from that participant's mean (pre-target word analyses: 1.1 %; target word analyses: 0.7 %) were also removed.

We examined first-pass reading times both on the pre-target word and the target word: first fixation duration (FFD; the duration of the first fixation made on a word), single fixation duration (SFD; the duration of a fixation when it was the only first pass fixation made on a word), gaze duration (GD; the sum of all fixations made on a word before the eyes moved to another word). Also, we analyzed the landing position (LP) on the target word that launched from the pre-target word with launch site as a covariate.[5] Means and standard deviations for each condition on these measures at the pre-target and target words are presented in Table 2.

To analyze the data, Linear Mixed Models (LMMs) were constructed by using the lme4 package (Bates et al., 2015) in R 4.4.2 (R Core Team, 2024) and R studio (2024). Fixation times and landing positions were log transformed before running the models. Both foveal load and parafoveal preview extent were treated as fixed factors. To reduce statistical tests, we treated preview extent as a numerical factor in the models (we

[4] Since our target words were three-character words in Experiment 1, and because three-character words do not occur very frequently in Chinese (less than 30% of words are 3 or more characters in length, see Li et al., 2015 for a review), they had relatively low predictability scores. To avoid confounds, we ensured that the predictability of the words in Experiment 2 was comparable to the predictability of the words in Experiment 1.

[5] For fixation times, the results did not change with, or without, launch site as a covariate. But the results for landing position did change slightly (interactive instead of additive effects for Experiment 2 and the *meta*-analysis of the two experiments). Given this, we reported analyses without launch site as a covariate for the fixation time results, and analyses with launch site as a covariate for landing position. For transparency, the landing position analyses without launch site as a covariate, and fixation times with launch site as a covariate are provided as Supplemental Analyses on OSF.

Table 1
The Mean Statistical Characteristics of the Pre-target and Target Word and the Experimental Sentence under Low and High Foveal Load Conditions (SDs in Parentheses, Frequency in counts per million).

	Pre-target word			Target word			Sentence
Foveal load	Frequency	Stroke No.	Predictability (%)	Frequency	Stroke No.	Predictability (%)	Naturalness
Low	133 (168)	18 (4)	2 (5)	5 (6)	23 (5)	1 (4)	3.9 (0.3)
High	0.4 (1.2)	18 (3)	2 (4)			1 (4)	3.9 (0.3)

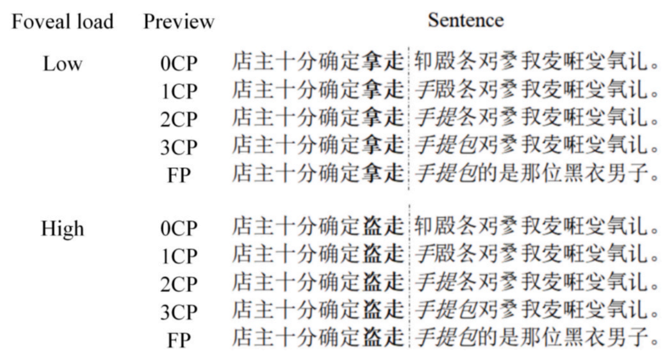

Fig. 1. An example set of the experimental stimuli under each condition. Pre-target words are presented in bold while target words are in italics (for illustration purposes only). The vertical black dotted line represents the position of the invisible boundary. The translation for the sentence is "The shopkeeper was pretty sure that it was the man with the black coat who **took/stole** the *handbag*".

also reported comparisons of critical pairs of conditions in the supplementary file on OSF). Additionally, participants and items were specified as random factors, with both random intercepts and random slopes (Barr et al., 2013). We always began running models specified in full including the maximum random effects structure. The slopes were removed when the model failed to converge indicating over-parameterization of the model. P-values were calculated based on Satterthwaite's approximations using the lmerTest package (Kuznetsova et al., 2017). Fixed effect estimations (i.e., beta values, confidence intervals, standard errors, t-values and p-values) for the eye movement measures at the pre-target word are shown in Table 3. Eye movement measures at the target word are shown in Table 4.

Given that our paradigm produced a very substantive visual change across a quite large portion of the sentence when the eyes crossed the boundary (see also Angele et al., 2016; Degno et al., 2019; Vasilev et al., 2018; Slattery et al., 2011; White et al., 2005), it is not at all surprising that 91 % of participants reported that they noticed some changes or flashes (26 % of trials, SD = 27 %) that occurred during their reading (see the supplementary file on OSF for analyses of effects for those who did, and those who did not perceive changes or flashes).

Pre-target word analyses

For all three fixation time measures, there were robust word frequency effects, with longer times on infrequent pre-target words than that on frequent pre-target words. Frequency effects are very well established in the literature (see Rayner, 2009; Rayner & Liversedge, 2011 for reviews). The pre-target word was less likely to be skipped in the low frequency condition (19 %) than the high frequency condition (23 %), $b = -0.30$, $SE = 0.06$, $t = -4.73$, $p < .001$, 95 %CI = [−0.42, −0.18]. The high fixation probability (79 %) on the two-character pre-target word indicates that for most trials, readers processed the parafoveal information in the target region (and beyond) from a relatively close position. Clearly, our manipulation of foveal lexical processing load, as indicated by the frequency effects, was highly effective.

We also obtained a small main effect of parafoveal preview (i.e., parafoveal-on-foveal effects of the pseudocharacter preview) for fixation time measures, specifically, all three fixation times on the pre-target word were generally and slightly increased (no more than 6 ms), as the imposed preview window narrowed. This was likely due to the reader's sensitivity to the extensive visually unusual pseudocharacter sequence in the parafovea. We note, that the disruption at the pre-target word was only significant for the zero-character preview condition (see the supplementary file on OSF) in which the visual disruption to the preview was maximal, which suggests that visually mediated PoF effects of this kind occur robustly only when disruption to preview is substantial. Such a finding is in line with results from numerous other studies (see Schotter et al., 2012 for a review). In sum, the evidence for PoF effects at this point in the sentence was limited, and for most preview conditions, was statistically unreliable.

Target word analyses

When considering the saccade launched from the pre-target word across the boundary, on 93 % of trials, the eyes landed on the target word region. For landing positions (as shown in Fig. 2, Panel B), we obtained a significant foveal load effect, such that readers landed nearer the beginning of the target word after reading an infrequent than a frequent pre-target word, which is entirely consistent with findings reported by White and Liversedge (2006). Landing positions also showed a reliable effect of parafoveal preview extent, such that landing positions

Table 2
M ± SD of Eye Movement Measures for the Pre-target Word and the Target Word across Conditions (Number of Observations in Parentheses) in Experiment 1.

		Pre-target word FFD (ms)	SFD (ms)	GD (ms)	Target word LP (char.)	FFD (ms)	SFD (ms)	GD (ms)
n		(19,462)	(15,766)	(19,357)	(15,948)	(19,398)	(11,989)	(19,337)
Low load	0CP	242 ± 41 (1908)	242 ± 42 (1610)	277 ± 66 (1905)	1.32 ± 0.67 (1572)	283 ± 53 (2013)	288 ± 67 (1065)	402 ± 98 (2011)
	1CP	237 ± 45 (1918)	237 ± 46 (1671)	265 ± 60 (1924)	1.46 ± 0.67 (1603)	280 ± 55 (1982)	291 ± 76 (1053)	396 ± 101 (1978)
	2CP	239 ± 41 (1904)	238 ± 43 (1639)	269 ± 64 (1899)	1.62 ± 0.77 (1541)	260 ± 48 (1914)	266 ± 68 (1136)	363 ± 99 (1903)
	3CP	237 ± 41 (1900)	237 ± 42 (1649)	267 ± 61 (1903)	1.66 ± 0.80 (1551)	247 ± 42 (1927)	247 ± 50 (1273)	328 ± 88 (1912)
	FP	236 ± 45 (1925)	235 ± 47 (1652)	267 ± 65 (1929)	1.63 ± 0.71 (1566)	236 ± 37 (1909)	235 ± 46 (1332)	312 ± 83 (1906)
High load	0CP	253 ± 50 (1968)	253 ± 55 (1490)	309 ± 90 (1941)	1.26 ± 0.69 (1635)	280 ± 51 (1982)	295 ± 72 (1191)	379 ± 86 (1987)
	1CP	252 ± 48 (1944)	252 ± 53 (1477)	310 ± 94 (1916)	1.37 ± 0.75 (1601)	274 ± 50 (1965)	278 ± 62 (1154)	380 ± 93 (1955)
	2CP	252 ± 47 (2003)	252 ± 51 (1517)	313 ± 94 (1991)	1.50 ± 0.80 (1656)	260 ± 45 (1916)	260 ± 62 (1167)	361 ± 96 (1906)
	3CP	251 ± 47 (2001)	251 ± 54 (1535)	309 ± 88 (1980)	1.59 ± 0.84 (1628)	249 ± 45 (1910)	252 ± 55 (1308)	327 ± 86 (1904)
	FP	251 ± 47 (1991)	250 ± 50 (1526)	307 ± 85 (1969)	1.60 ± 0.93 (1595)	239 ± 39 (1880)	235 ± 46 (1310)	313 ± 81 (1875)

Note. 0CP = zero-character preview, 1CP = one-character preview, 2CP = two-character preview, 3CP = three-character preview, FP = full preview; FFD = first fixation duration, SFD = single fixation duration, GD = gaze duration, LS = launch site, LP = landing position, char. = character.

Table 3
LMM Analyses and 95% Confidence Intervals (CI) for the Eye Movement Measures for the Pre-target Word in Experiment 1.

	Effect	*b*	*CI*	*SE*	*t*	*p*
FFD	(Intercept)	**5.45**	**[5.44, 5.47]**	**0.01**	**555.30**	**<.001**
	Foveal load (HL vs. LL)	**0.04**	**[0.03, 0.06]**	**0.01**	**5.61**	**<.001**
	Preview extent	**−0.00**	**[−0.01, −0.00]**	**0.00**	**−2.40**	**.017**
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	1.12	.263
SFD	(Intercept)	**5.46**	**[5.44, 5.48]**	**0.01**	**539.69**	**<.001**
	Foveal load (HL vs. LL)	**0.05**	**[0.03, 0.06]**	**0.01**	**5.24**	**<.001**
	Preview extent	**−0.00**	**[−0.01, −0.00]**	**0.00**	**−2.90**	**.004**
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	0.97	.331
GD	(Intercept)	**5.58**	**[5.55, 5.61]**	**0.01**	**397.34**	**<.001**
	Foveal load (HL vs. LL)	**0.11**	**[0.09, 0.13]**	**0.01**	**10.21**	**<.001**
	Preview extent	**−0.00**	**[−0.01, −0.00]**	**0.00**	**−2.24**	**.026**
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	0.97	.334

Note. Foveal load effects are word frequency effects. Significant terms featured in bold.

Table 4
LMM Analyses and 95% Confidence Intervals (CI) for the Eye Movement Measures for the Target Word in Experiment 1.

	Effect	*b*	*CI*	*SE*	*t*	*p*
FFD	(Intercept)	**5.59**	**[5.57, 5.61]**	**0.01**	**522.35**	**<.001**
	Foveal load (HL vs. LL)	−0.01	[−0.03, 0.00]	0.01	−1.92	.055
	Preview extent	**−0.04**	**[−0.05, −0.04]**	**0.00**	**−19.45**	**<.001**
	Foveal load × Preview extent	0.01	[−0.00, 0.01]	0.00	1.74	.082
SFD	(Intercept)	**5.62**	**[5.60, 5.64]**	**0.01**	**524.79**	**<.001**
	Foveal load (HL vs. LL)	−0.02	[−0.04, 0.00]	0.01	−1.70	.090
	Preview extent	**−0.05**	**[−0.06, −0.05]**	**0.00**	**−26.25**	**<.001**
	Foveal load × Preview extent	0.01	[−0.00, 0.01]	0.00	1.44	.149
GD	(Intercept)	**5.90**	**[5.87, 5.93]**	**0.02**	**394.17**	**<.001**
	Foveal load (HL vs. LL)	**−0.05**	**[−0.07, −0.03]**	**0.01**	**−5.08**	**<.001**
	Preview extent	**−0.07**	**[−0.07, −0.06]**	**0.00**	**−31.74**	**<.001**
	Foveal load × Preview extent	**0.02**	**[0.01, 0.02]**	**0.00**	**3.72**	**<.001**
LP	(Intercept)	**0.66**	**[0.60, 0.73]**	**0.03**	**20.01**	**<.001**
	Foveal load (HL vs. LL)	**−0.10**	**[−0.14, −0.06]**	**0.02**	**−5.01**	**<.001**
	Preview extent	**0.06**	**[0.05, 0.07]**	**0.00**	**15.20**	**<.001**
	Foveal load × Preview extent	−0.01	[−0.02, 0.01]	0.01	−1.09	.276
	Launch site	**−0.71**	**[−0.73, −0.69]**	**0.01**	**−65.77**	**<.001**

Note. Significant terms featured in bold.

were increasingly rightward across zero- to two-character previews but were similar for the two-, three- character previews and the full preview condition (see the supplementary file on OSF), which suggests the preview extent was two characters. However, foveal load did not interact with preview extent, which demonstrates that foveal load and parafoveal preview extent independently determine saccadic targeting decisions toward a single-lexical unit.

The main effect of foveal load at the target word was significant for gaze duration (−8 ms) but not for first fixation duration or single fixation duration. The low foveal load condition (i.e., high frequency) produced longer gaze durations on the target word compared to the high foveal load condition (i.e., low frequency), that is, a "reversed" spillover effect of word frequency. This result is inconsistent with other studies, most of which have either reported no spillover effect (e.g., Liu et al., 2015; White et al., 2005; Vasilev et al., 2018; Zhang et al., 2019), or a spillover effect whereby there are shorter fixation times on a target word when the pre-target word was high than when it was low frequency (e.g., Angele et al., 2016; Drieghe et al., 2005; Kennison & Clifton, 1995; Veldre & Andrews, 2018). In fact, our "reversed" spillover effect was likely not a spillover effect at all, but instead an effect caused by the parafoveal preview. This effect was mainly driven by zero- and one-character previews wherein high frequency pre-target words led to 23 ms and 15 ms longer gaze durations respectively relative to low frequency pre-target words (for the other preview conditions, the difference between high and low frequency was less than |3| ms). These results very likely arise due to participants perceiving pseudocharacters in the parafovea (and being particularly sensitive to these under minimal preview conditions, when the preview window was the smallest), and after a saccade to fixate those pseudocharacters, they experienced a visual change and then quickly regressed to re-read earlier portions of the sentence, doing this more often when that word was low (31 %) than high (21 %) frequency ($b = 0.70$, $SE = 0.06$, $z = 12.26$, $p < .001$, 95 %CI = [0.59, 0.81]).

There was a significant main effect of parafoveal preview extent on the three fixation time measures at the target. First pass reading times at the target region were reduced with increased preview extent, suggesting that Chinese readers obtain useful visual or orthographic information from an area of more than three characters to the right of the boundary, and clearly, more than one word, given that the three characters in the parafovea comprised a single word in this experiment. It appears that increased availability of parafoveal information (beyond two characters into the next word) did not affect fixation positions, though such preview information did affect processing of the target when it was actually fixated. Together, the landing position and first pass reading time results provide another example of the independence and differential nature of the systems responsible for controlling where and when the eyes move (see Rayner, 2009; Findlay & Walker, 1999).

Most importantly, there was a significant interactive effect between foveal load and parafoveal preview extent on gaze durations but not on first or single fixation durations. This effect was such that with increased preview extent, gaze durations were reduced, and the magnitude of these differences was greater under low than high foveal load conditions (see Fig. 2, Panel A, for an illustration). Analyses achieved statistical significance for comparisons between the full preview and the zero-character preview and between the full preview and the one-character preview (see the supplementary file) which reflect the more extreme differences between our conditions. The nature of the interactive pattern of effects for gaze durations at the target word is theoretically important in other respects. First, they were qualitatively different to the non-interactive patterns we observed for landing positions. If increased foveal load had reduced the extent of parafoveal processing in relation to fixation times, then we would have expected to observe different fixation times between zero- and other preview conditions compared with the full preview condition under low foveal load conditions, but more comparable fixation times for the different preview conditions under high load condition. Such a pattern would have indicated that readers

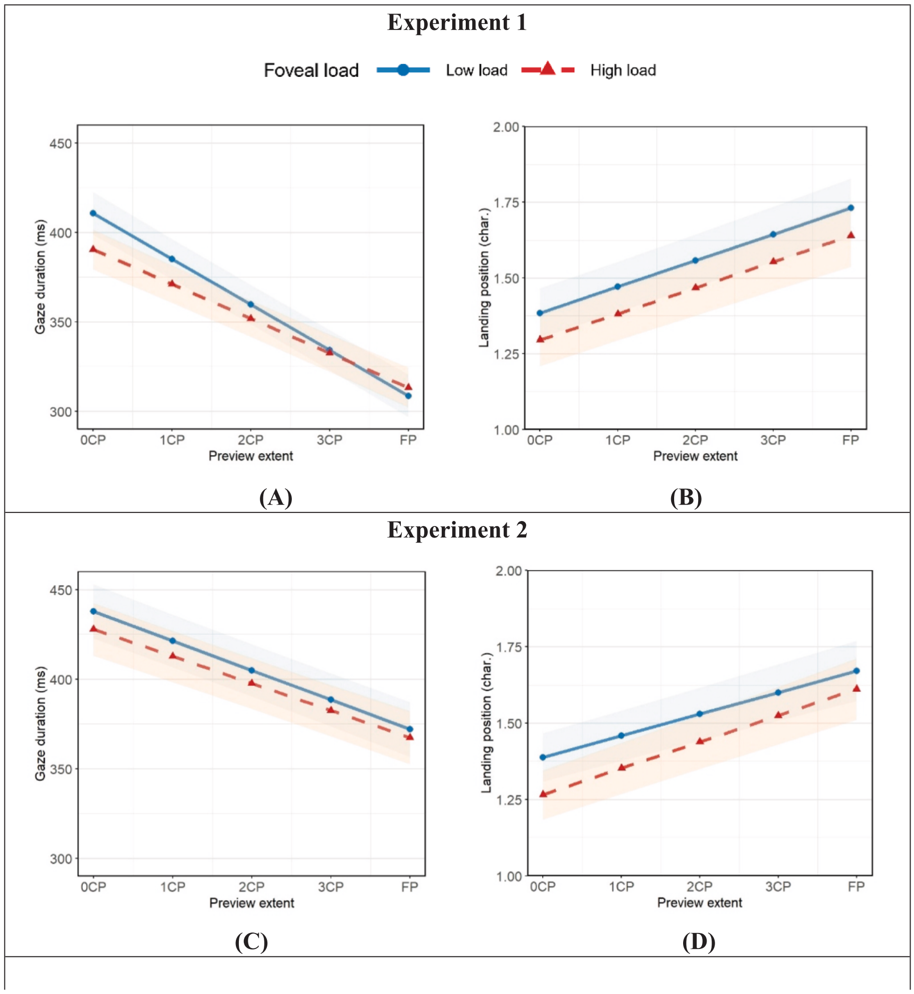

Fig. 2. Gaze durations and landing positions on the target word for the different preview conditions under low- and high- foveal load in Experiment 1 and Experiment 2 (the shaded areas represent 95% confidence intervals).

were able to extract useful parafoveal information post-boundary from a wider extent under low than high foveal load conditions. However, what we observed was that, for both load conditions, readers were able to extract meaningful information beyond three characters post-boundary. Nevertheless, increased foveal load still reduced the reader's sensitivity to restrictive preview windows that caused disruption to reading. Therefore, the pattern of the interactive effects in respect of gaze durations suggests that foveal load reduced the depth to which parafoveal information was processed, but not the spatial extent over which parafoveal processing took place.

Second, note that, the magnitude of the preview effects we report here (more than 50 ms between the full and zero- or one-character previews under conditions of low foveal load for single fixation and gaze durations) were on the larger side of those that are typically reported in the literature (e.g., 30 ~ 50 ms; see Rayner, 2009 for a review). It is likely that the increased magnitude of the effects here arose due to an inhibitory influence of the substantial display changes in our paradigm relative to other paradigms adopted in previous boundary studies. It is also likely that the effect reflected such a sensitivity rather than it being a spillover effect as might be predicted under basic Foveal Load Hypothesis. More specifically, as can be seen from Fig. 2 Panel A for gaze durations (but the pattern of effects is similar for the other fixation time measures), under the full preview, there was no difference in fixation times between high and low foveal load conditions, whilst for the zero- and one-character previews, foveal load did modulate preview effects such that the more characters in parafovea that were masked, then the greater the degree of disruption to reading. Most critically, the modulation by foveal load can be seen when cost of processing was greatest, that is, when preview extent was most restricted. Another reason why we obtained such large preview effects here might be because written Chinese is a densely-packed orthography, with a small number of characters conveying a relatively large amount of information. Thus, a relatively large amount of parafoveal linguistic information was available to Chinese readers under all our parafoveal window conditions (other than the zero character preview condition) and this might have contributed to the size of the effects that we observed (see M. Yan et al., 2009; Yang, Wang et al., 2012).

Overall, the results of Experiment 1 demonstrated a modulatory influence of foveal load on parafoveal preview that affected fixation times but not landing positions when the three characters that formed the target region comprised a single lexical unit. In Experiment 2, we examined whether such effects were comparable when the three parafoveal characters in the target region formed multiple lexical units. Research has shown that a single lexical unit is processed to a greater extent in the parafovea than a matched character string that is comprised of multiple lexical units (e.g., Zang et al., 2021; Zang, Fu et al., 2024; Zang, Wang et al., 2024). Thus, for Experiment 2, we predicted a pattern of preview effects that would be comparable to those we obtained in Experiment 1, but that the effects would be reduced in magnitude.

Experiment 2

Method

Participants

A different group of 256 students from Tianjin Normal University participated in Experiment 2, 6 of whom were removed from analyses due to tracking loss. The remaining valid 250 participants (37 were males) were native Chinese speakers with mean age of 21 ($SD = 2$) years and normal or corrected to normal vision and were naïve with respect to the purpose of experiment. Each participant was paid for their participation.

Apparatus

Participants' eye movements were recorded using the same EyeLink 1000 Plus eye tracker and experimental set up as in Experiment 1.

Materials and design

Foveal load and preview extent were manipulated in the same way as in Experiment 1, except that the target three characters comprised more than one word, most of which were three single-character words (about 80 %; see Footnote 2). More specifically, 180 pairs of high- and low-frequency two-character words were selected as pre-target words. Each of the pre-target words was followed by the same set of three characters which were treated as the target words (see Fig. 3). Therefore, there were 180 target words. Frequencies were reliably higher for high-frequency than low-frequency pre-targets [$F(1, 179) = 137.28$, $p < .001$, $\eta_p^2 = 0.43$], whereas, stroke numbers of characters across the two conditions did not differ [$F(1, 179) = 0.99$, $p = .322$]. The basic descriptive characteristics of the pre-target and target words are shown in Table 5.

Each pair of the pre-target words and the target word were embedded into the same sentence frame. The sentences were between 19 and 25 characters in length. Sentence naturalness and the predictabilities of the pre-target and target words were assessed by using the same approach as in Experiment 1. Sentence naturalness was rated by 24 (12 in each of the counterbalanced low- and high-load conditions) university students that did not participate in other ratings or testing. All sentences were natural with no significant difference between the low- and high-load conditions, $F(1,179) = 1.79$, $p = .182$. A second group of 12 participants assessed the predictabilities of pre-target words, and a third group of 22 participants (11 in each of the counterbalanced low- and high-load conditions) assessed the predictabilities of target words. Both predictabilities of the pre-target words and the target words were very low with no significant differences between the two foveal load conditions, $Fs < 1.00$, $ps > .05$.

The parafoveal preview extent of the target word was manipulated in exactly the same way as Experiment 1. Thus, five preview conditions were constructed, that is, zero, one, two or three characters of the parafoveal target region, or the full sentence preview. The content downstream from the preview window was also masked by pseudocharacters in the same way as in Experiment 1. Ten files were created, and the sentences were presented as in Experiment 1. Each file consisted of 180 experimental sentences (18 for each condition) and 70 fillers, 36 % of which followed by a comprehension question that participants were required to answer. Additionally, 6 practice sentences, with 4 followed by a comprehension question, were presented before the testing session.

Procedure

The procedure in Experiment 2 was identical to Experiment 1.

Results and discussion

The average comprehension accuracy for all participants was 93 % ($SD = 4$ %), indicating that all participants comprehended the sentences well. Data were excluded using the same criteria as in Experiment 1. Fixations below 80 ms or above 1200 ms were removed (6.0 % of total fixations). Trials were eliminated in which a track loss occurred (0.1 % of the data). For target region analyses, trials were eliminated in which the display change occurred early (0.8 %) or was delayed (6.6 %), or hooking (7.9 %), or other incorrect triggering of display changes by blinks or saccades occurred (1.1 %). Also, for each measure (except for skipping data) and each participant, any observations more than three standard deviations based on the participant's mean (pre-target word analyses: 1.1 %; target words analyses: 0.8 %) were excluded from analyses.

As in Experiment 1, we carried out analyses for the same measures on both the pre-target word and the target words. LMMs were again constructed to analyze the data using the lme4 package in R. All settings and procedures in the model were the same as in Experiment 1. Ninety-four percent of participants reported they noticed some changes or flashes (23 % of trials on average, $SD = 24$ %) occurring during their reading which was similar to Experiment 1 (again, see the supplemental file on OSF for analyses of effects for those who did, and did not perceive changes or flashes).

Pre-target word analyses

Means and standard deviations for all eye movement measures on the pre-target word in Experiment 2 are presented in Table 6 and the

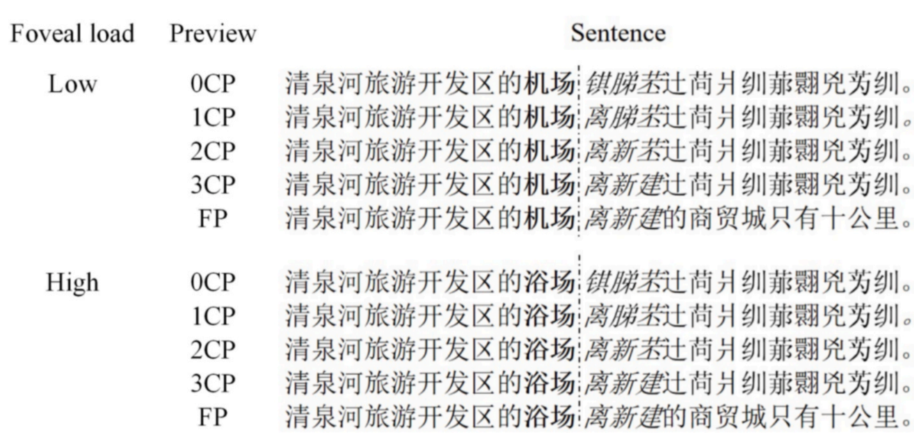

Fig. 3. An example of the experimental stimuli under different conditions in Experiment 2. Pre-target words are presented in bold, of which the following three characters in italics are target words. The vertical black dotted line represents the position of the invisible boundary. The translation of the sentence is "The **airport/bathing place** in the tourism development zone of Qingquanhe is only ten kilometers away *from the newly built* trade city.".

Table 5
The Means and Standard Deviations for Pre-target Words, Target Words and the Experimental Sentences' Characteristics under Low and High Foveal Load Conditions (SDs in Parentheses, Frequency in counts per million).

Foveal load	Pre-target word Frequency	Stroke No.	Predictability (%)	Target words Frequency 1st	2nd	3rd	Stroke No. 1st	2nd	3rd	Predictability (%)	Sentence Naturalness
Low	118 (134)	17 (4)	2 (5)	265 (569)	572 (1057)	619 (2377)	9 (3)	8 (3)	8 (3)	2 (4)	4.5 (0.3)
High	1 (1)	17 (3)	2 (4)							3 (5)	4.4 (0.3)

Table 6
M ± SD of Eye Movement Measures for the Pre-target Word and Target Words Across Conditions (Number of Observations in Parentheses) in Experiment 2.

		Pre-target word FFD (ms)	SFD (ms)	GD (ms)	Target words LP (char.)	FFD (ms)	SFD (ms)	GD (ms)
Observations		(34,576)	(28,802)	(34,407)	(27,608)	(34,321)	(19,352)	(34,218)
Low load	0CP	235 ± 35 (3403)	234 ± 37 (2935)	265 ± 55 (3390)	1.30 ± 0.60 (2779)	279 ± 50 (3565)	282 ± 58 (1830)	431 ± 117 (3554)
	1CP	234 ± 34 (3396)	234 ± 35 (3005)	259 ± 49 (3394)	1.52 ± 0.70 (2732)	273 ± 50 (3499)	279 ± 62 (1771)	425 ± 123 (3477)
	2CP	234 ± 34 (3374)	233 ± 35 (2958)	259 ± 48 (3375)	1.57 ± 0.75 (2720)	260 ± 41 (3444)	261 ± 52 (1866)	406 ± 125 (3425)
	3CP	234 ± 34 (3351)	233 ± 35 (2923)	261 ± 52 (3353)	1.58 ± 0.79 (2647)	255 ± 40 (3376)	253 ± 51 (1937)	394 ± 122 (3381)
	FP	235 ± 37 (3349)	234 ± 38 (2908)	262 ± 51 (3364)	1.62 ± 0.79 (2691)	246 ± 36 (3428)	241 ± 39 (2125)	368 ± 114 (3413)
High load	0CP	248 ± 38 (3541)	247 ± 40 (2734)	300 ± 73 (3481)	1.21 ± 0.60 (2845)	278 ± 48 (3482)	283 ± 60 (1909)	421 ± 116 (3479)
	1CP	251 ± 39 (3526)	250 ± 41 (2844)	295 ± 65 (3504)	1.39 ± 0.67 (2830)	270 ± 46 (3437)	275 ± 65 (1822)	422 ± 119 (3433)
	2CP	252 ± 39 (3504)	252 ± 41 (2801)	299 ± 65 (3487)	1.52 ± 0.77 (2797)	260 ± 40 (3399)	256 ± 50 (1983)	393 ± 119 (3391)
	3CP	248 ± 39 (3600)	247 ± 41 (2872)	294 ± 63 (3565)	1.53 ± 0.83 (2794)	257 ± 39 (3353)	250 ± 47 (2037)	383 ± 116 (3351)
	FP	250 ± 36 (3532)	249 ± 39 (2822)	294 ± 62 (3494)	1.56 ± 0.84 (2773)	248 ± 37 (3338)	242 ± 43 (2072)	369 ± 109 (3314)

Note. 0CP = zero-character preview, 1CP = one-character preview, 2CP = two-character preview, 3CP = three-character preview, FP = full preview; FFD = first fixation duration, SFD = single fixation duration, GD = gaze duration, LP = landing position, char. = character.

LMM results are shown in Table 7.

As with Experiment 1, word frequency effects on the pre-target words were robust on all fixation time measures. The word frequency effect on skipping probability was also reliable ($b = -0.25$, $SE = 0.04$, $z = -5.78$, $p < .001$, 95 %CI = [−0.33, −0.16]), it being 4 % lower under low (20 %) than high (24 %) frequency conditions. The pre-target words were fixated on most occasions (78 %), indicating that readers most often previewed parafoveal information in the target region from a close distance. Again, these findings suggest that our manipulation of word frequency, as an indicator of foveal lexical processing load, was highly effective. The numerical patterns of preview extent effects (no more than 5 ms) at the pre-target word in Experiment 2 were very similar to those obtained in Experiment 1, although the main effect of preview extent did not attain significance, nor was there a significant interaction at the pre-target word.

Table 7
LMM Analyses and 95% Confidence Intervals (CI) for the Eye Movement Measures for the Pre-target Word in Experiment 2.

	Effect	*b*	*CI*	*SE*	*t*	*p*
FFD	(Intercept)	**5.44**	**[5.42, 5.45]**	**0.01**	**644.84**	**<.001**
	Foveal load (HL vs. LL)	**0.06**	**[0.05, 0.07]**	**0.01**	**10.34**	**<.001**
	Preview extent	−0.00	[−0.00, 0.00]	0.00	−0.08	.934
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	0.76	.447
SFD	(Intercept)	**5.44**	**[5.42, 5.45]**	**0.01**	**623.32**	**<.001**
	Foveal load (HL vs. LL)	**0.06**	**[0.05, 0.07]**	**0.01**	**9.74**	**<.001**
	Preview extent	−0.00	[−0.00, 0.00]	0.00	−0.04	.972
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	0.51	.611
GD	(Intercept)	**5.55**	**[5.52, 5.57]**	**0.01**	**485.31**	**<.001**
	Foveal load (HL vs. LL)	**0.12**	**[0.10, 0.13]**	**0.01**	**15.77**	**<.001**
	Preview extent	−0.00	[−0.01, 0.00]	0.00	−1.54	.124
	Foveal load × Preview extent	−0.00	[−0.01, 0.00]	0.00	−1.07	.286

Note. Foveal load effects are word frequency effects. Significant terms featured in bold.

Target word analyses

Means and standard deviations for all eye movement measures on the target word in Experiment 2 are presented in Table 6 and the LMM results are shown in Table 8.

For landing positions, the word frequency effect was significant and showed a pattern consistent with Experiment 1. Also, the preview extent manipulation caused readers to land further into the target region when they received a preview with an increased number of characters (see Panel D of Fig. 2). Note that these differences were significant when comparing the full preview with zero- and one-character previews, but were not significant when comparing the full preview with two- and three- character previews. Inconsistent with Experiment 1, there was a significant interaction between foveal load and preview extent. Specifically, the difference between the full and two-character previews was reliable in the low load condition ($b = 0.03$, $SE = 0.01$, $t = 2.37$, $p = .018$), but not in the high load condition ($b = -0.01$, $SE = 0.01$, $t = -0.68$, $p = .495$) (see the supplementary file on OSF). These findings suggest that a wider preview extent was obtained under low load (three or more characters) than under high load conditions (approximately two characters), which suggests that foveal load interacted with parafoveal processing in the aspect of spatial extent as reflected in saccadic targeting toward upcoming characters comprised of multiple lexical units.

The foveal load effects on fixation times perfectly replicated the results in Experiment 1. Specifically, we obtained no frequency effect for first and single fixation durations, however, for gaze duration the frequency effect was reversed and very comparable in magnitude to the difference obtained in Experiment 1 (−7 ms vs. − 8 ms respectively). As per Experiment 1, it appears that the effect was mainly driven by the restricted previews (i.e., zero-, one-, two-, and three- character previews) wherein high frequency pretarget words led to longer gaze durations (4 ~ 13 ms) on the target than low frequency pretarget words. Note, again, that for the full preview condition, the difference between the high and low frequency condition was minimal (−1 ms). Also in line

Table 8
LMM Analyses and 95% Confidence Intervals (CI) for the Eye Movement Measures for the Target Word in Experiment 2.

	Effect	*b*	*CI*	*SE*	*t*	*p*
FFD	(Intercept)	**5.57**	**[5.55, 5.58]**	**0.01**	**624.94**	**<.001**
	Foveal load (HL vs. LL)	−0.01	[−0.02, 0.01]	0.01	−1.02	.308
	Preview extent	**−0.03**	**[−0.03, −0.03]**	**0.00**	**−23.76**	**<.001**
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	1.49	.136
SFD	(Intercept)	**5.58**	**[5.56, 5.60]**	**0.01**	**499.96**	**<.001**
	Foveal load (HL vs. LL)	−0.01	[−0.03, 0.01]	0.01	−1.20	.232
	Preview extent	**−0.04**	**[−0.04, −0.03]**	**0.00**	**−17.02**	**<.001**
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	0.71	.479
GD	(Intercept)	**5.94**	**[5.91, 5.97]**	**0.02**	**353.44**	**<.001**
	Foveal load (HL vs. LL)	**−0.03**	**[−0.05, −0.01]**	**0.01**	**−3.09**	**.002**
	Preview extent	**−0.05**	**[−0.05, −0.04]**	**0.00**	**−25.15**	**<.001**
	Foveal load × Preview extent	0.00	[−0.00, 0.01]	0.00	1.07	.283
LP	(Intercept)	**0.71**	**[0.65, 0.77]**	**0.03**	**22.56**	**<.001**
	Foveal load (HL vs. LL)	**−0.15**	**[−0.18, −0.12]**	**0.01**	**−10.01**	**<.001**
	Preview extent	**0.04**	**[0.04, 0.05]**	**0.00**	**10.63**	**<.001**
	Foveal load × Preview extent	**0.02**	**[0.00, 0.03]**	**0.01**	**2.75**	**.006**
	Launch site	**−0.68**	**[−0.69, −0.66]**	**0.01**	**−82.74**	**<.001**

Note. Significant terms featured in bold.

with Experiment 1, there were reliable preview extent effects on the first-pass reading times, such that smaller preview extent caused longer fixations on target words. More specifically, the restricted previews caused significantly longer fixations than full preview, indicating that readers obtained parafoveal visual or orthographic information from more than three characters to the right of the boundary.

Importantly, there were no statistically robust interactive effects of foveal load and preview extent, though we observed quite similar numerical patterns to those in Experiment 1. That is to say, the differences across preview extent conditions were numerically larger when foveal load was low compared with high (see Fig. 2, Panel D for an illustration for gaze duration). The patterns of effects we observed for the fixation time results appear to be quite different to those we observed for the landing position results, which, again, provides evidence for the differential mechanisms underlying the "where" and "when" oculomotor control systems (see Rayner, 2009).

Taking the results from Experiments 1 and 2 together, given the lack of formal statistical significance for the effects in Experiment 2, at this stage, it is difficult to have a very clear sense of the degree of consistency in the effects across experiments. Recall that previous studies (e.g., Cutter et al., 2014; Zang et al., 2021; Zang, Fu et al., 2024; Zang, Wang et al., 2024) have shown that a single lexical unit produced a larger preview effect compared to multiple lexical units (findings consistent with the MCU Hypothesis; Zang, 2019; see also He et al., 2021). Therefore, we might anticipate that any preview effects would be greater in magnitude in Experiment 1 where those effects would be observed across the full extent of the single lexical unit than the effects observed in Experiment 2 when the target region was comprised of multiple words that would be processed parafoveally in a more piecemeal manner. From our perspective, the critical statistical issue here concerns whether we might observe a two-way interaction between foveal load and preview extent demonstrating consistent, but reduced, patterns of effects between Experiments 1 and 2, or instead a three-way interaction between foveal load, preview extent and experiment indicating that we obtained patterns of effects that were qualitatively different between the two experiments. To examine these predictions, we conducted the supplementary analyses reported below in which we assessed the nature of the interactive effects across experiments.

Supplementary analyses

For our *meta*-analysis we combined the data sets from Experiments 1 and 2 for fixation times and landing positions, and added Experiment as a fixed factor into the LMMs.

There was main effect of Experiment at the target region on single fixation duration ($b = -0.04$, $SE = 0.01$, $t = -2.79$, $p = .005$, 95 %CI = [−0.07, −0.01]) and gaze duration ($b = 0.05$, $SE = 0.02$, $t = 2.31$, $p = .022$, 95 % CI = [0.01, 0.09]) but not on first fixation duration ($b = -0.02$, $SE = 0.01$, $t = -1.85$, $p = .065$, 95 %CI = [−0.05, 0.00]), such that single fixation duration was 3 ms longer and gaze duration was 45 ms shorter in Experiment 1 than in Experiment 2, in line with the findings of Zang et al. (2018) and He et al. (2021). There was a significant main effect of foveal load at the target region. As noted previously, this was likely due to the preview mask. It occurred for all fixation time measures ($|t|s > 1.98$, $ps < .05$). Also, a similar main effect of preview extent on all the fixation time measures was obtained such that small preview windows induced longer reading times than larger preview windows ($|t|s > 34.54$, $ps < .001$).

The interactive effect of foveal load and Experiment did not attain significance ($|t|s < 1.81$, $ps > .05$). The interaction between preview extent and Experiment was significant in all fixation time measures ($ts > 6.61$, $ps < .001$) such that the preview effects were larger in Experiment 1 than those observed in Experiment 2. This finding indicates that a restricted preview extent produced more disruption when the three upcoming characters formed a single word compared to when they formed multiple words, which is consistent with our prediction and supports the MCU hypothesis (e.g., Zang, 2019; Zang et al., 2021; Zang, Fu et al., 2024; Zang, Wang et al., 2024). The interactions between foveal load and parafoveal preview extent were significant for first fixation duration ($b = 0.00$, $SE = 0.00$, $t = 2.30$, $p = .021$, 95 % CI = [0.00, 0.01]) and gaze duration ($b = 0.01$, $SE = 0.00$, $t = 3.37$, $p = .001$, 95 % CI = [0.00, 0.02]), but not significant for single fixation duration ($b = 0.00$, $SE = 0.00$, $t = 1.50$, $p = .134$, 95 % CI = [−0.00, 0.01]). The interactive pattern was such that the magnitude of the preview effect (cost to processing at the target) was greater under low than high foveal load conditions. Crucially, whilst the three-way interaction was not significant on first and single fixation durations ($|t|s < 0.63$, $ps > .05$), it was significant on gaze duration ($b = -0.01$, $SE = 0.01$, $t = -2.02$, $p = .044$, 95 % CI = [−0.02, −0.00]). This result fits well with our previous findings in that for gaze duration we obtained a statistically reliable two-way interactive effect in Experiment 1 but not in Experiment 2. The *meta*-analysis supports our suggestion that foveal load effects on parafoveal processing operate in line with the claims of the MCU Hypothesis, namely, that any such effects will be more pronounced when parafoveal information is processed as a single lexical unit relative to when it is processed as multiple lexical units. We return to this issue in the General Discussion.

Unlike the results for reading times, there was no main effect of Experiment on landing positions ($b = -0.01$, $SE = 0.04$, $t = -0.19$, $p = .849$, 95 % CI = [−0.09, 0.07]). Unsurprisingly, there were main effects of foveal load ($b = -0.13$, $SE = 0.01$, $t = -11.10$, $p < .001$, 95 % CI = [−0.16, −0.11]) and preview extent ($b = 0.06$, $SE = 0.00$, $t = 20.53$, $p < .001$, 95 % CI = [0.05, 0.06]) on landing positions which showed the same pattern as the findings we reported earlier.

There was an interaction between foveal load and Experiment ($b = -0.05$, $SE = 0.02$, $t = -2.00$, $p = .046$, 95 % CI = [−0.09, −0.00]), with a smaller foveal load effect on landing positions in Experiment 2 than in

Experiment 1. We found no interactive effects of preview extent and Experiment ($b = -0.01$, $SE = 0.01$, $t = -1.03$, $p = .302$, 95 % CI = [−0.02, 0.01]), and foveal load and preview extent ($b = 0.00$, $SE = 0.00$, $t = 0.85$, $p = .398$, 95 % CI = [−0.01, 0.01]) for landing positions. However, we did find a three-way interaction ($b = 0.02$, $SE = 0.01$, $t = 2.56$, $p = .011$, 95 % CI = [0.01, 0.04]). The three-way interaction in the *meta*-analysis is complicated reflecting mean differences across all the twenty conditions of our two experiments. Furthermore, the patterns across the means, to us at least, do not point to a very straightforward explanation of the interactive effects. Nonetheless, an aspect of the results is that the largest effect of foveal load on landing positions in Experiment 1 was obtained with a two-character preview window, whereas the largest such effect in Experiment 2 was obtained with a one-character window. Assuming that readers target the preferred viewing location (PVL) of words (e.g., Rayner, 1979; see also Cutter et al., 2017, 2018), then a two-character preview is the smallest preview window that would provide accurate information pertaining to this location in Experiment 1, and a one-character preview is the smallest preview window providing information about this location in Experiment 2. Thus, it is possible that the three-way interaction for landing positions reflects differences in effects associated with PVL targeting under different preview conditions for each experiment. Clearly, this account is speculative and further research is required to better our understanding.

In order to further assess the three-way interactions for fixation times and landing positions, we conducted Bayesian analyses for linear mixed models by using BayesFactor package (Morey et al., 2018). Bayes factors both for the full model (i.e., BF_{Full}, the model containing the main effects of Experiment, foveal load, parafoveal preview extent, and their two- and three-way interactions) and the reduced model without the three-way interaction (i.e., BF_{Reduced}) were calculated. We used the default scale prior (.5) and 100,000 Monte Carlo iterations. A sensitivity analysis with different priors (i.e., .2, .3, .4, .5, .6, .7, and .8) was also conducted. The result of the Bayesian analysis on all three fixation time measures favored the null hypothesis, as well as the sensitivity analysis (all BFs < 0.001). This suggests that whilst there may have been differences in the magnitude to which foveal load modulated parafoveal preview between Experiments 1 and 2 in relation to fixation times, the nature of such effects between experiments was qualitatively similar. By contrast, for landing positions, both the results of Bayesian analysis ($BF = 3.94 \times 10^4$) and the sensitivity analysis ($BF > 1.50 \times 10^4$) favored the alternative hypothesis, providing strong evidence for the three-way interactive effect. The Bayesian analyses on fixation times and landing positions reinforces our earlier conclusion that foveal load and preview manipulations influence mechanisms controlling where and when to move the eyes differentially.

General discussion

Two experiments were conducted to investigate whether, and if so, how increased foveal load affects parafoveal processing of upcoming text in relation to both spatial and temporal oculomotor decision metrics during Chinese reading. As in previous studies (e.g., Drieghe et al., 2005; Henderson & Ferreira, 1990; White et al., 2005; Vasilev et al., 2018; Zhang et al., 2019), we manipulated word frequency of pre-target words to produce increased or decreased foveal processing load in the two experiments. Our results replicated typical word frequency effects consistent with numerous studies (see Rayner, 2009; Rayner & Liversedge, 2011, for reviews) and confirmed the effectiveness of the manipulation of foveal lexical processing load. Additionally, our adoption of two-character pre-target words ensured a fixation on the pre-target location during first-pass reading on 79 % of trials, thereby ensuring that readers obtained effective preview of upcoming stimuli in the target region. We also manipulated preview extent in relation to the three-character target region (i.e., zero-, one-, two-, three-character previews, as well as a full preview condition) in both experiments. The three target characters constituted a single word in Experiment 1 whereas they formed multiple words in Experiment 2. To reiterate, our manipulation of preview extent via the boundary paradigm allowed us to examine preview effects in relation to foveal load in the absence of possible changes in both foveal and parafoveal processing during reading (c.f., moving window studies; e.g., Luke, 2018; Meixner et al., 2022; Rayner,1986).

Both experiments showed robust preview extent effects at the target words, such that fixation times increased, and landing positions were substantially more leftward with reduced preview extent. Specifically, for fixation times, even when three characters were available in the parafovea, processing in the target word region was still disrupted relative to the full preview in both experiments, suggesting that readers extract and utilize information four or more characters to the right of fixation to aid identification of upcoming words irrespective of whether those upcoming characters comprise one, or more than one, lexical unit. For landing positions, based on the results of *meta*-analyses, the eyes appear to be targeted to land approximately one to two characters ahead even when additional meaningful information about characters beyond this area is available. Together the reading time and landing position results show that whilst there may be a sensitivity to characters three or more positions to the right of fixation, those characters differentially affect decisions regarding when and where to move the eyes. Such findings are consistent with previous studies that have shown the perceptual span in Chinese reading is at least 2–3 characters to the right of the fixation on average (Inhoff & Liu, 1998; G. Yan, Zhang, Zhang et al., 2013), and that the extent of the perceptual span depends on various factors, such as the ease with which words to the right of fixation may be processed, as well as the masking material used outside the preview window (e.g., McConkie & Rayner, 1975; Rayner, 1986, G. Yan, Zhang, & Bai, 2013; G. Yan, Zhang, Zhang et al., 2013; M. Yan et al., 2015). Some studies have demonstrated that the rightward preview span may extend beyond three characters to the right in Chinese reading (e.g., G. Yan, Zhang, & Bai, 2013; M. Yan et al., 2015; Zang, Wang et al., 2024). Given that the three characters in the present experiments were relatively frequent, and that the masks were pseudocharacters with stroke numbers similar to their counterpart identities, it is unsurprising that we found readers obtained useful parafoveal visual or orthographic information beyond three characters to the right of fixation (as shown on fixation times in the present study).

We also found that restricted preview extent caused more disruption to parafoveal processing when the parafoveal characters comprised one lexical unit (Experiment 1) than when they formed multiple words (Experiment 2). This finding is consistent with the MCU hypothesis (e.g., Zang, 2019; Zang et al., 2021; Zang, Fu et al., 2024; Zang, Wang et al., 2024). According to the MCU hypothesis, frequently used multiple word sequences can be lexicalized and processed as single representations, allowing the constituents within each sequence to be identified in parallel (Zang, 2019). Empirical evidence has shown that MCUs are processed as single representations both foveally (e.g., Yu et al., 2016) and parafoveally (e.g., Cutter et al., 2014; Zang et al., 2021; Zang, Fu et al., 2024; Zang, Wang et al., 2024). Therefore, in our Experiment 1, given that processing of later characters is licensed by the presence of earlier characters (when they form a lexical unit), when such processing is prevented due to preview window restrictions, more disruption occurred for processing the three characters that constituted a single word because the integrity of the target word was disrupted in the parafovea; by contrast, in Experiment 2 there was less disruption while processing the three characters that formed multiple distinct words. Such a claim was supported by the reliable interactive effect between Experiment and Preview for the specific comparison of the two-character and three-character preview conditions (for FFD, GD: ts > 3.26, ps $< .01$), such that larger preview effects (2CP vs. 3CP) occurred in Experiment 1 than in Experiment 2.

Importantly, we found the parafoveal preview effects on fixation times were modulated by foveal lexical processing load in Experiment 1 and the *meta*-analysis for the two experiments, such that, the more restricted the preview, the greater the cost to processing, and the magnitude of that cost was reduced when foveal load was high relative to when it was low. Veldre and Andrews (2018), Experiment 2, reported similar results in relation to processing costs and foveal load such that preview cost was reduced when foveal load was high compared with when it was low. Thus, our results align well with the findings of Veldre and Andrews. However, it is important to note that there are several studies that have failed to show such robust modulatory effects of foveal load on parafoveal processing. For example, Drieghe et al., (2005), Veldre and Andrews, Experiment 1, and Zhang et al., (2019) showed effects that were relatively diminished compared with those reported here. One possible reason that these studies showed such modest effects is because they adopted stimuli in which the target words were relatively short. The use of longer parafoveal words compared to shorter ones (e.g., Juhasz et al., 2008), might afford an increased opportunity to observe foveal load effects. Our findings that preview effects were larger in Experiment 1 than Experiment 2 and that interactive effects of foveal load and preview in Experiment 1 reached statistical significance while in Experiment 2 they did not, again, provide further support for such explanation. Moreover, we found that with a wider preview extent (i.e., two, or more than two, characters), the modulatory effects of foveal load on parafoveal processing were reduced, indicating that modulation of disruption caused by pseudocharacter previews was limited to an area that was fairly close to the foveal word. This finding is compatible with findings from Kennison and Clifton (1995), in which foveal load effects occurred for parafoveal stimuli at a close preview distance (i.e., near launch sites) but not for those that were more distant (i.e., far launch sites). Additionally, as already noted, both our experiments consistently indicated that Chinese readers extracted some parafoveal information beyond three characters regardless of foveal load.

It appears that the degree to which readers are more “effective” in their parafoveal processing (both in relation to processing of useful-information and disruption due to sensitivity to misinformation), then the more likely that modulatory influences of foveal load on such processing will be observed. We consider this effect occurs because activation based on an invalid preview will be overridden because the new visual input from the target is inconsistent with that perceived from the parafovea. Nonetheless, it is possible that there may be some degree of conflict between activation produced by the preview and that produced by the target after the boundary change (Li & Pollatsek, 2020). This is in line with *trans*-saccadic linguistic integration accounts (Cutter et al., 2015; Schotter et al., 2019), which posit that parafoveal preview will be integrated into foveal target processing once the eye fixates the target. The more dissimilar the preview and target information, the more conflict or integration difficulty there will be in respect of target processing; consequently, in such situations, fixations on the target will be longer. If this was the case, then presumably any such *trans*-saccadic integration difficulty would be greater under low than high foveal load conditions. And following on from this, the particular characteristics of the experimental paradigm that is adopted in an experiment will be a strong determinant of whether readers will engage in more, or less, “effective” parafoveal processing. For example, whether a moving window paradigm or a boundary paradigm is adopted; the particular extent of a preview manipulation; whether single or multiple words comprise the preview region; the precise nature of parafoveal masking stimuli, etc. will all influence parafoveal processing. To summarize, regarding our results pertaining to fixation times, increased foveal load reduced the depth to which parafoveal information was processed within the restricted parafoveal window.

By contrast, for landing positions, we only found significant interactive effects in Experiment 2 but not in Experiment 1, and this was supported by the three-way interaction in our *meta*-analyses suggesting that (in relation to landing positions) foveal load and preview extent effects are qualitatively different between parafoveal stimuli formed from a single lexical unit and those formed from multiple lexical units. This result strongly suggests that foveal load modulated sensitivity to differences in the lexical status of character sequences available in the parafovea and this directly affected the extent to which saccades were targeted into that region.

It is important to consider whether our results do, actually, support the Foveal Load Hypothesis. Let us start by reiterating the point that the present results do indicate that the degree of difficulty associated with processing the foveal stimulus influences parafoveal processing during reading. This was put forward in past papers (Henderson & Ferreira, 1990; Rayner, 1986), and to the extent that our results showed a relationship between foveal processing difficulty and parafoveal processing, then they are consistent with this basic idea. However, it is also the case that the present results (alongside other more recent relevant findings, e.g., Veldre & Andrews, 2018) do not fit neatly with the particular specifications of the Foveal Load Hypothesis, nor with how foveal load mechanisms have been specified within current eye movement control models such as the E-Z Reader model (Reichle et al., 2003; Reichle, 2011), the SWIFT model (Engbert et al., 2005) and the CRM (Li & Pollatsek, 2020). For example, according to E-Z Reader and SWIFT, we might have expected a spillover effect at the target region such that fixations were longer after a low than a high frequency pre-target word (due to less effective parafoveal processing of the target when foveal load was high). Furthermore, any such effect should have occurred most prominently under full preview conditions where availability of parafoveal information was maximal, and therefore, effects of load should be most visible (see also Veldre & Andrews, 2018, for discussion). In contrast to these predictions, first, we found no reliable difference between low and high foveal load conditions for full preview on fixation times, which provides no evidence for such a spillover effect. We did, however, obtain reliable differences between the two load conditions for full preview on landing positions, suggesting a dissociation in relation to decisions of when and where to move the eyes. Secondly, increased foveal lexical processing load reduced the disruptive influence of visually restrictive parafoveal windows rather than reducing benefit from a full preview. In other words, foveal load reduced Chinese readers' sensitivity to the parafoveal mask, and to reiterate, the effects we observed at the target word were almost certainly driven by disruption caused by the upcoming preview mask. A final interesting finding from our results is that foveal lexical processing load reduced preview extent across characters forming multiple lexical units but not when the same number of characters formed a single lexical unit. Thus, all three of these aspects of our results were not specified in the original Foveal Load Hypothesis as originally conceived (Henderson & Ferreira, 1990).

Let us next, also, briefly consider our results in relation to the CRM since this model seeks to directly explain eye movement control during reading in Chinese (the language in which we studied foveal load effects here). The CRM incorporates a flexible mechanism in relation to preview benefit and preview cost effects each respectively associated with a valid- or an invalid-preview (Li & Pollatsek, 2020), and this mechanism might allow for modulation of disruptive effects of invalid or restricted parafoveal information by foveal load. In the CRM, under conditions of invalid preview, some characters in the mental lexicon may be activated to a modest degree by the parafoveal stimulus. However, upon the eyes crossing the boundary, any such activation based on preview will be overridden because the new visual input from the target (presented after the boundary change) will be inconsistent with that processed in the parafovea. If this was the case, then presumably any such conflict would be greater under low than high foveal load conditions. Our findings provide evidence for such a possibility. Also, the CRM assumes that Chinese readers adopt a processing-based strategy for their saccade targeting such that saccade targeting is interactively determined by foveal processing and parafoveal preview. Our findings in relation to landing positions also provide evidence for such an assumption (though whether the CRM can explain differential effects across multiple words compared with MCUs remains an open question).

Finally, we must acknowledge that the experimental paradigm that we adopted in the present experiments was quite different to the paradigm originally used by Henderson and Ferreira (1990). To this extent, it might be unsurprising that the specific details of our results are somewhat difficult to align directly with the Foveal Load Hypothesis. Also, our results might have occurred because foveal load effects in Chinese operate in a somewhat different way to how they occur in alphabetic reading (perhaps due to the dense orthography, the unspaced nature of written Chinese and word boundary ambiguity, all of which affect parafoveal processing), or perhaps because frequency manipulations of foveal load may exert less of an influence over parafoveal processing during Chinese reading when that reading operates in an unconstrained manner (i.e., without constraint over the availability of parafoveal information). These issues clearly require future examination. One theoretical possibility that we do consider plausible, however, is that a reduction of preview extent by foveal load may represent a common mechanism across languages, and this itself may be linked to claims of a universal mechanism in relation to cognitive load on the functional visual field (e.g., Ikeda & Takeuchi, 1975).

CRediT authorship contribution statement

Manman Zhang: Writing – original draft, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Zhichao Zhang:** Investigation. **Fang Li:** Investigation. **Xuejun Bai:** Supervision. **Chuanli Zang:** Writing – review & editing, Supervision, Funding acquisition. **Simon P. Liversedge:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge support from National Natural Science Foundation of China Grants, China (31800920), and Economic and Social Research Council Grant, United Kingdom (ES/R003386/1).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2025.104716.

Data availability

All data sets and analysis scripts are publicly available at: https://osf.io/swcau/

References

Angele, B., Slattery, T. J., & Rayner, K. (2016). Two stages of parafoveal processing during reading: Evidence from a display change detection task. *Psychonomic Bulletin & Review, 23*, 1241–1249.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.
Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.
Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One1, 5*, Article e10729.
Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(6), 1778–1786.
Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across fixations in reading? In A. Pollatsek, & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 245–260). Oxford, UK: Oxford University Press.
Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2017). Is orthographic information from multiple parafoveal words processed in parallel: An eye-tracking study. *Journal of Experimental Psychology: Human Perception and Performance, 43*(8), 1550–1567.
Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2018). Reading sentences of uniform word length–II: Very rapid adaptation of the preferred saccade length. *Psychonomic Bulletin & Review, 25*(4), 1435–1440.
Degno, F., Loberg, O., & Liversedge, S. P. (2021). Co-registration of eye movements and fixation—Related potentials in natural reading: Practical issues of experimental design and data analysis. *Collabra. Psychology, 7*(1), 18032.
Degno, F., Loberg, O., Zang, C., Zhang, M., Donnelly, N., & Liversedge, S. P. (2019). Parafoveal previews and lexical frequency in natural reading: Evidence from eye movements and fixation-related potentials. *Journal of Experimental Psychology: General, 148*(3), 453–474.
Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 954–969.
Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777–813.
Engbert, R., & Kliegl, R. (2011). Parallel graded attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford Library of Psychology. The Oxford Handbook of Eye Movements* (pp. 787–800). New York, NY, US: Oxford University Press.
Findlay, J., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences, 22*(4), 661–721.
He, L., Song, Z., Chang, M., Zang, C., Yan, G., & Liversedge, S. P. (2021). Contrasting off-line word segmentation with on-line word segmentation during reading. *British Journal of Psychology, 112*(3), 662–689.
Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 417–429.
Ikeda, M., & Takeuchi, T. (1975). Influence of foveal load on the functional visual field. *Perception & Psychophysics, 18*, 255–260.
Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 20–34.
Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1560–1579.
Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 68–81.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear mixed Effects Models. *Journal of Statistical Software, 82*, 1–26.
Li, X., Zang, C., Liversedge, S. P., & Pollatsek, A. (2015). The role of words in Chinese reading. In A. Pollatsek, & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 232–244). New York, NY: Oxford University Press.
Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review, 127*(6), 1139–1162.
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.
Luke, S. G. (2018). Influences on and consequences of parafoveal preview in reading. *Attention, Perception, & Psychophysics, 80*, 1675–1682.
McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578–586.
Meixner, J. M., Nixon, J. S., & Laubrock, J. (2022). The perceptual span is dynamically adjusted in response to foveal load by beginning readers. *Journal of Experimental Psychology: General, 151*(6), 1219–1232.
Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2018). *BayesFactor: Computation of Bayes factors for common designs*. Retrieved from http://CRAN.R-project.org/packageBayesFactor.
R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna, Austria: R foundation for statistical computing. Retrieved from https://www.r-project.org/.
RStudio: Integrated development environment for R (Version 2023.12.0+369) [Computer software]. Boston, MA. Retrieved from http://www.rstudio.org/.
Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*(1), 65–81.
Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*(1), 21–30.
Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*, 211–236.
Rayner, K. (2009). The 35th 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., & Liversedge, S. P. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 751–766). New York, NY, US: Oxford University Press.
Reichle, E. D. (2011). Serial-attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford Library of Psychology. The Oxford Handbook of Eye Movements* (pp. 767–786). New York, NY, US: Oxford University Press.
Reichle, E. D. (2020). *Computational models of reading: A handbook*. Oxford University Press.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(4), 445–476.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*, 391–421.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics, 74*(1), 5–35.

Schotter, E. R., von der Malsburg, T., & Leinenger, M. (2019). Forced fixations, trans-saccadic integration, and word recognition: Evidence for a hybrid mechanism of saccade triggering in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(4), 677–688.

Slattery, T. J., Angele, B., & Rayner, K. (2011). Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1924–1938.

Vasilev, M. R., Slattery, T. J., Kirkby, J. A., & Angele, B. (2018). What are the costs of degraded parafoveal previews during silent reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 371–386.

Veldre, A., & Andrews, S. (2018). How does foveal processing difficulty affect parafoveal processing during reading? *Journal of Memory and Language, 103*, 74–90.

White, S. J., & Liversedge, S. P. (2006). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research, 46*(3), 426–437.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review, 12*, 891–896.

Yan, G. L., Zhang, Q. M., & Bai, X. J. (2013). Study on the influencing factors of perceptual span in Chinese reading [in Chinese]. *Psychological Development and Education, 29*(2), 121–130.

Yan, G. L., Zhang, Q. M., Zhang, L. L., & Bai, X. J. (2013). Effect of masking material on percetptual span in Chinese reading [in Chinese]. *Journal of Psychological Science, 36*(6), 1317–1322.

Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word N+ 1 modulates preprocessing effectiveness of word N+2 in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1669–1676.

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, M., Zhou, W., Shu, H., & Kliegl, R. (2015). Perceptual span depends on font size during the reading of Chinese sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 41*(1), 209–219.

Yang, J., Rayner, K., Li, N., & Wang, S. (2012). Is preview benefit from word n+ 2 a common effect in reading Chinese? Evidence from eye movements. *Reading and Writing, 25*, 1079–1091.

Yang, J., Wang, S., Tong, X., & Rayner, K. (2012). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading and Writing, 25*, 1031–1052.

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word n+ 2? evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 1192–1204.

Yu, L., Cutter, M. G., Yan, G., Bai, X., Fu, Y., Drieghe, D., & Liversedge, S. P. (2016). Word n + 2 preview effects in three-character Chinese idioms and phrases. *Language, Cognition and Neuroscience, 31*(9), 1130–1149.

Zang, C. (2019). New perspectives on serialism and parallelism in oculomotor control during reading: The multi-constituent unit hypothesis. *Vision, 3*(4). Article 50.

Zang, C., Fu, Y., Bai, X., Yan, G., & Liversedge, S. P. (2018). Investigating word length effects in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 44*(12), 1831–1841.

Zang, C., Fu, Y., Bai, X., Yan, G., & Liversedge, S. P. (2021). Foveal and parafoveal processing of Chinese three-character idioms in reading. *Journal of Memory and Language, 119*, Article 104243.

Zang, C., Fu, Y., Du, H., Bai, X., Yan, G., & Liversedge, S. P. (2024). Processing multiconstituent units: Preview effects during reading of Chinese words, idioms, and phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 50*(1), 169–188.

Zang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 961–978). New York, NY: Oxford University Press.

Zang, C., Wang, S., Bai, X., Yan, G., & Liversedge, S. P. (2024). Parafoveal processing of Chinese four-character idioms and phrases in reading: Evidence for multi-constituent unit hypothesis. *Journal of Memory and Language, 136*, Article 104508.

Zhang, M., Liversedge, S. P., Bai, X., Yan, G., & Zang, C. (2019). The influence of foveal lexical processing load on parafoveal preview and saccadic targeting during Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 45*(6), 812–825.