

Predictability effects and parafoveal processing of compound words in natural Chinese reading



Lei Cui¹ , Chuanli Zang^{2,3}, Xiaochen Xu¹, Wenxin Zhang¹, Yuhan Su¹ and Simon P Liversedge⁴

Quarterly Journal of Experimental Psychology
2022, Vol. 75(1) 18–29
© Experimental Psychology Society 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17470218211048193
qjep.sagepub.com



Abstract

We report a boundary paradigm eye movement experiment to investigate whether the predictability of the second character of a two-character compound word affects how it is processed prior to direct fixation during reading. The boundary was positioned immediately prior to the second character of the target word, which itself was either predictable or unpredictable. The preview was either a pseudocharacter (nonsense preview) or an identity preview. We obtained clear preview effects in all conditions, but more importantly, skipping probability for the second character of the target word and the whole target word from pretarget was greater when it was predictable than when it was not predictable from the preceding context. Interactive effects for later measures on the whole target word (gaze duration and go-past time) were also obtained. These results demonstrate that predictability information from preceding sentential context and information regarding the likely identity of upcoming characters are used concurrently to constrain the nature of lexical processing during natural Chinese reading.

Keywords

Predictability; preview effects; compound word; Chinese reading

Received: 8 September 2020; revised: 5 March 2021; accepted: 6 March 2021

The predictability of an upcoming word or constituent in a sentence is usually considered to be the likelihood that the constituent will be anticipated based on the context up to the point at which it appears. Word predictability is known to be influenced by prior sentential context. When sentential context provides a high constraint, then only a limited number of words may form a plausible continuation of that sentence. In contrast, when sentential context provides a low, or weak, predictability constraint, then many potential words might plausibly form a continuation. Considerable research has shown that the predictability of a word directly influences the time needed to process that word. Generally, the more predictable a word is given the context, the faster that word is to be processed. This robust effect has been demonstrated using eye movement recordings during reading, the methodology we used here. The basic eye movement effects are that less predictable words are fixated for longer (e.g., Ashby et al., 2005; Ehrlich & Rayner, 1981), elicit more regressions (e.g., Rayner & Well, 1996; Staub, 2011) and are skipped less often (e.g., Abbott et al., 2015; Drieghe et al., 2004; Ehrlich & Rayner, 1981) than more predictable words.

Although it is well established that contextual constraint influences eye movement behaviour during reading, detailed understanding of the time course of this influence is not yet clear. Several studies have shown that the manipulation of contextual constraint influences measures of eye movements reflecting later stages of linguistic processing during reading (second pass and total reading time; likelihood of a regression for rereading, for example, Calvo & Meseguer, 2002; Kretzschmar et al., 2015; Rayner & Well, 1996). Furthermore, evidence suggesting earlier influences is mixed, with some studies showing first fixation predictability effects (e.g., Altarriba et al., 1996; Hand et al., 2010; Lee et al., 1999; Rayner et al., 2004, 2011;

¹Shandong Normal University, Jinan, China

²University of Central Lancashire, Preston, UK

³Tianjin Normal University, Tianjin, China

⁴University of Central Lancashire, Preston, UK

Corresponding author:

Simon P Liversedge, School of Psychology and Computer Science, University of Central Lancashire, Preston PR1 2HE, UK.
Email: SPLiversedge@uclan.ac.uk

White et al., 2005), and others reporting null effects (e.g., Balota et al., 1985; Drieghe et al., 2004; Rayner & Well, 1996). For example, for initial skipping probability, Bélanger and Rayner (2013) found increased skipping probability for high-predictable words compared with low-predictable words (0.20 vs. 0.14). These results were confirmed by Kretzschmar et al. (2015). However, some researchers did not find similar differences between high- and low-predictable words (Choi et al., 2017; Gollan et al., 2011; Hand et al., 2010; Miell et al., 2007). More research is required to better understand the time course of predictability effects on eye movements during reading.

Beyond sentential context constraints, McDonald and Shillcock (2003a, 2003b) argued that constraint over a word's predictability can also derive from transitional probabilities between words, that is, the statistical likelihood that a particular word $n + 1$ will follow a particular word n . McDonald and Shillcock demonstrated that transitional probabilities exerted an immediate influence on reading, affecting first fixation durations (FFDs) on words. Relatedly, Zola (1984) conducted a study in which the strength of a local constraint was manipulated such that the predictability of an upcoming word was determined almost entirely by the preceding word. The adjective preceding a target word (e.g., *beautiful gardens* vs. *botanical gardens*) produced predictability effects on fixation durations on the target (but no effect for target skipping). More recently, Fitzsimmons and Drieghe (2013) undertook an experiment in which they investigated the influence of a word's global predictability (based on preceding sentential context) and its local predictability (the likelihood of the next word given the current word). Fitzsimmons and Drieghe suggested that the lack of a skipping effect in Zola's study may have been due to the target words being relatively long (7–8 letters) and therefore likely to attract fixations based on their length. For these reasons, Fitzsimmons and Drieghe ran an experiment using shorter target words and demonstrated that local predictability affected both fixation durations and skipping probability. Note, though, that global and local predictability in this experiment were manipulated separately by Fitzsimmons and Drieghe, and therefore, the joint influence of these two variables on eye movement behaviour was not assessed. However, together these studies do indicate that different sources of predictability constraint (local influences, as well as more global influences derived from preceding context) can reduce the likely lexical candidates in respect of upcoming word identification. The purpose of this study was to utilise stimuli in which predictability constraints were maximised (via manipulations of both sentential, global, predictability cues and local predictability cues) to determine how those constraints, along with constraints derived from parafoveal information, influenced processing during normal Chinese reading.

From the discussion above, it should be apparent that word length can affect the magnitude of a local predictability effect. For alphabetic languages, word length varies significantly (particularly in agglutinative languages such as Finnish or German) and, therefore, it has a substantial effect on reading behaviour. However, Chinese is a character-based, ideographic script. Chinese written sentences are formed from strings of characters that are equally spaced and there are no word spacing cues to demarcate word boundaries. Some characters can be words by themselves, or they can be combined with other characters to form different multicharacter words. In Chinese, approximately 6% of words comprised one character, 72% comprised two characters, and the remaining words are three or more characters in length (*Lexicon of common words in contemporary Chinese*, 2008). Thus, word length variability is much reduced in Chinese relative to alphabetic languages, and consequently, written Chinese provides an opportunity to probe predictability effects in conditions where word length influences are substantially reduced (cf. Fitzsimmons & Drieghe, 2013). In our view, therefore, it may be advantageous to investigate influences of global, sentential predictability and local predictability in Chinese relative to alphabetic languages.

As we have already noted, written Chinese is an unspaced language without word boundary demarcations. It is also the case that compound words abound in Chinese, wherein multiple character words quite often comprise multiple morphemes. This means that to effectively process a written Chinese sentence, readers must engage in word segmentation processes, that is, they must make decisions about where words begin and end in order that they might be identified and saccades targeted effectively during reading. Let us consider an example. In normally presented Chinese text, there are no spaces to indicate whether a two-character string such as “*马虎*” should be processed as a single compound word meaning (meaning *careless* in English), or instead as two separate words (the first meaning *horse*, and the second meaning *tiger*). It should, therefore, be apparent that the Chinese reader frequently faces significant word constituent ambiguity that must be resolved rapidly in order for reading to proceed effectively. Note that this level of ambiguity does not arise in alphabetic languages like English where word spacing cues very often disambiguate whether a string is a single two-constituent compound word (e.g., *fireworks* as in *I saw the fireworks in the sky*), or instead two separate constituents (e.g., *fire* and *works* in *Cleaning the metal with fire works very effectively*). Of course, there are a limited number of spaced compound words in English (e.g., *traffic light*) that are considered to be associated with a single lexical entry, and to this extent, English readers must, on some occasions, form decisions as to where word boundaries lie irrespective of word spacing cues (cf. Cutter et al., 2014; Zang, 2019). Nonetheless, the fact remains that in

the majority of occasions, word spacing cues in English override potential influence that transitional probability relations between the morphemes of a compound might exert in the determination of a word's constituency. However, due to the prevalence of word boundary ambiguity in Chinese, the issue of how sentential predictability and local predictability influence word segmentation decisions, and therefore, eye movement behaviour during processing of Chinese compound words is of particular significance. To be explicit, given that the first character of a two-character Chinese compound word has the potential to be either a word in and of itself, or the first constituent of a compound word, and given that there are no physical demarcations to indicate lexical constituency, then there is significant potential in Chinese reading for local predictability relations between characters (as well as contextual predictability constraints) to exert an influence on word segmentation commitments and processing more generally.

A final characteristic of Chinese that is important in relation to this experiment is that the written form is visually dense and less spatially extended than, for example, alphabetic languages (see Liversedge et al., 2016). Therefore, parafoveal information is closer to the fixation point in Chinese than in alphabetic languages, meaning that, potentially, stronger and earlier parafoveal preview effects might occur (e.g., Cui et al., 2013a, 2013b). All of the issues discussed above are important reasons to consider local and more global, contextual predictability influences in Chinese reading (as distinct from such effects in alphabetic languages).

Contextual constraint and local predictability influences on word identification are derived from text that (in languages read from left to right) lies at, or to the left of, the point of fixation. However, there is a further source of constraint over word identification that derives from text that lies to the right of the point of fixation. Parafoveal processing is central to efficient reading. Information about the visual and linguistic characteristics of the upcoming words is extracted and used to facilitate their subsequent lexical identification. A critical issue concerns the precise time course of the influence of such information on word identification, both intrinsically and in relation to predictability constraints (global and local) derived from preceding context. Thus, on any particular fixation, a reader is preprocessing the upcoming word for identification, and such preprocessing is influenced by (1) the constraint provided by sentential context, and the local predictability of the upcoming character being a constituent that forms a word with the currently fixated character, and (2) the influence of visual and linguistic information about the upcoming word derived from parafoveal vision. In this study, not only did we examine the combined influence of sentential and local predictability constraints, but we also considered how these influences interact with

information derived from parafoveal preview to jointly influence the processes of word segmentation and identification and eye movement behaviour during natural Chinese reading. To this extent, this study follows directly from earlier work (e.g., Balota et al., 1985) that showed readers more effectively parafoveally process high- compared with low-predictable words. This study also directly relates to experiments in which the boundary paradigm was adopted with the boundary positioned at a point within a word (rather than immediately before a word) to manipulate the availability of morphemic information (e.g., Hyönä et al., 2004). Hyönä et al. were the first to adopt the boundary paradigm to study processing of compound words in Finnish, and they were the first researchers to position the invisible boundary within a word between two morphemes. This approach allows for investigation of the nature and time course of processing of the constituent morphemes of compound words during natural reading, that is, whether each morpheme of a compound is processed sequentially and independently, or instead interactively and in parallel during the identification of the word as a whole.

In respect of Chinese reading, recently, Chang, Hao, et al. (2020) demonstrated an interplay between predictability and preview processes in early reading time measures, but that word-skipping was not strongly mediated by sentential predictability. They placed an invisible boundary before a high- or low-predictable two-character target compound word. When readers' eyes were located prior to the boundary, readers received either a valid identity preview of the entire word or an invalid preview of a very low-frequency two-character word (effectively a non-word) that was visually dissimilar to the target word. Substantial preview effects were obtained with this manipulation (102 ms for high predictability previews and 81 ms for low predictability previews) compared with previous studies in which preview manipulations have been made for single-character words (e.g., 31 ms in Cui et al., 2013a). It is very likely that the increased magnitude of the preview effects reported by Chang, Hao, et al. was due to differences in the magnitude of change (due to the manipulation operating over two characters rather than a single character). The results indicate that the sentential predictability of a word affects the degree to which it is processed parafoveally in natural Chinese reading (though see Li et al., 2018, who showed inhibitory preview effects for single-character target words).

Additional research has demonstrated that the local predictability between the constituent characters of a word also affects preview processing during Chinese normal reading. Cui et al. (2013a) investigated whether the linguistic category of a two-character Chinese string (monomorphemic word, compound word or phrase) affected how the second character of that string was processed prior to fixation during reading. They found clear preview effects in all conditions; however, a parafoveal-on-foveal effect

only occurred for monomorphemic words. This effect was likely caused by the increased local predictability of the second character based on the first character for monomorphemic words (88%) relative to compound words (21%) and phrases (14%). Also, recently Zang et al. (2016) manipulated the probability that a character would either appear as a single-character word or be the first character of a two-character word. They found that readers were sensitive to probabilistic information of this type such that larger preview effects occurred when a character was more likely to be the first character of a two-character word than when it was a single character.

A further study that more directly investigated the relationship between local character predictability and parafoveal processing in Chinese was reported by Cui et al. (2013b). In this experiment, the target word was a two-character compound that shared the same second constituent that was either highly predictable given the first constituent or was less predictable given the first constituent. They found a significant preview benefit for both kinds of compound; however, critically, a parafoveal semantic preview benefit and a significant parafoveal-on-foveal effect only occurred when the second constituent was highly predictable. From the discussion above, it seems reasonable to suggest that both sentential predictability and local predictability constrain the degree to which upcoming characters and words are processed in the parafovea.

In this study, we manipulated the predictability of the second character of a two-character compound target word based on preceding context (sentential predictability) and the second character's likelihood given the first character of the compound (local predictability). We used the boundary paradigm, whereby readers were provided with a preview of the second character of the target (identity condition), or alternatively a pseudocharacter. The invisible boundary was located immediately prior to the second character of the target, and when the reader's eye crossed the boundary, the preview was replaced by the target. In this way, we manipulated the availability of the second character of the target word prior to direct fixation. Thus, our manipulations allowed us to explore how the predictability of the second character of a compound word (based on constraints from preceding sentential context and local predictability) and the availability of preview of the second constituent of a compound jointly influenced word identification. We predicted that if both sources of influence constrain the identification of an upcoming character, then we would observe interactive effects of predictability and preview such that preview benefit would be larger for high-predictable compared with low-predictable characters. Any such effects that occurred prior to the boundary change will reflect parafoveal-on-foveal influences and any effects that occur after the boundary change will more likely reflect the ease with which the target word is integrated with context.

There is an additional aspect of processing in relation to Chinese reading that we should also consider. As Chinese is

unspaced and overt word segmentation cues are unavailable, then readers must also make segmentation commitments. Thus, in relation to positioning the boundary in the middle of a two-character target word, we were careful to design our stimuli so that the first character of the target could have been a single-character word, and the second character of the target could, in principle, have been the first character of a subsequent word, or even a single-character word in itself. It is also important to note that the status of the two characters forming the target (in relation to them being two separate single-character words, or a single word comprised of two characters) depends upon the particular segmentation commitment that the reader makes. To this extent, in the current experiment, the segmentation decisions the reader makes are a determinant of how the first character of the target region is processed (and specifically in relation to how it is processed with respect to the postboundary character of the target). This nontrivial point underlines the fact that the current experiment is as much about word segmentation commitments and concomitant implications for lexical processing as it is about preview effects. That is, the segmentation commitments that the reader forms themselves determine the extent to which a dependency relation exists between the two characters that comprise the target region, and thus, influences of the preview of the second character are dependent on the status of this character with respect to the preboundary character.

Method

Participants

Forty-four native Chinese-speaking Shandong Normal University undergraduates with normal or corrected to normal vision participated in the experiment.

Apparatus

Eye movements were recorded using an SR Research EyeLink 2000 (1,000 Hz) eye tracker tracking the right eye. We used a Dell CRT presentation monitor with 1024 × 768 pixel resolution and the monitor's refresh rate was 160 Hz. Sentences were presented in Song font in black on white background. Each character was 1.1 × 1.1 cm² in size at a viewing distance of 60 cm.

Materials and design

We used a 2 (predictability: high vs. low) × 2 (preview: identity vs. pseudocharacter) within-subject repeated measures design. We selected 72 compound words with a high- or low-predictable second constituent and the same first constituent. The second constituents were matched on the number of strokes and character frequency; there was no significant difference in the frequency of the compound words ($p_s > .10$; see Table 1).

Table 1. Lexical-statistical properties for compound words.

	Target compound word with high-predictable second constituent	Target compound word with low-predictable second constituent
Frequency of second character	290 (603)	209 (464)
Strokes of second character	6.81 (2.53)	7.44 (3.71)
Word frequency	62 (111)	38 (197)

Word frequency is measured as words per million using the *Chinese Daily Word Frequency Dictionary* (1998). Character frequency is measured as characters per million (Cai & Brysbaert, 2010).

(1) Sentence frame containing a two-character compound target word in which the second character has high predictability

老师正在使用**胶符**往墙上贴小红花。

老师正在使用**胶水**往墙上贴小红花。

The teacher is using **glue** to paste a small red flower on the wall.

(2) Sentence frame containing a two-character target compound word in which the second character has low predictability

老师正在使用**胶符**往墙上贴小红花。

老师正在使用**胶布**往墙上贴小红花。

The teacher is using **adhesive tape** to paste a small red flower on the wall.

Figure 1. Example Chinese stimuli for the different experimental conditions. The target words were compound words in which the second character was high predictable, or low predictable. These are shown in bold in the figure, but appeared normal in the experiment. The location of the invisible boundary is indicated by a vertical line. The preview of the second character of the two-character target was either an identity character (“水” or “布”) or a dissimilar pseudocharacter (e.g., “符”). When the reader’s point of fixation crossed the invisible boundary location, the preview was replaced by the target character (“水” or “布”).

Pseudocharacter previews were created using True Font software. They very closely resembled real characters but were completely meaningless. Sentence frames were identical up to the target word and content differences after the target were minimised. The sentences appeared on a single line (maximum 17 characters) and the target never appeared as the initial or final word.

We adopted the boundary paradigm (Rayner, 1975). The invisible boundary was placed between the two characters of the compound words (see Figure 1).

Rating of materials

Predictability of the second character of the target word based on preceding sentential context and local predictability. Forty participants completed a list of sentence fragments that included all of the words and characters prior to the second character of the target compound word. Completions showed that predictability of the target character was

higher for compound words under high ($M=62\%$, $SD=0.22$) than low ($M=4.2\%$, $SD=0.05$) constraint conditions, $t(71)=21.33$, $p<.001$.

There is an important aspect of the off-line predictability rating data that we acquired prior to conducting the eye movement experiment, which we must not neglect. In these rating studies, we found that the local predictability prescreen results, where participants were presented with only the first character of the compound word, were almost identical to the results from the prescreen study in which participants received the sentential context and the local predictability cue. On this basis, one might initially assume that the local predictability influences were modulated very little by predictability cues from sentential context. In other words, local predictability, not sentential predictability, would drive our effects. However, any such assumption would be incorrect. In fact, the two sources of predictability do both exert an influence and produce a joint constraint on the likelihood of the target character. The fact that the

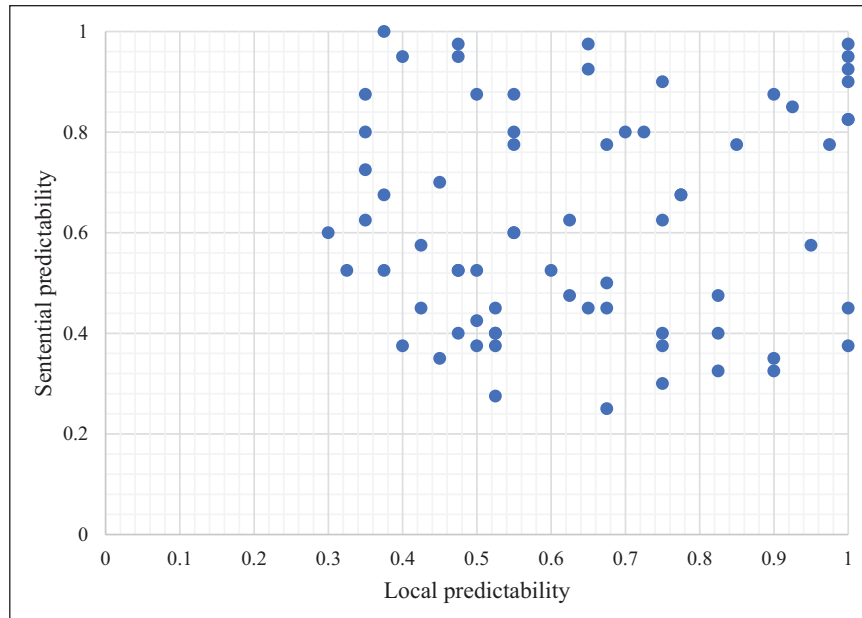


Figure 2. Local predictability ratings plotted against sentential predictability ratings for each of the experimental stimuli.

correct target word completion rates in the high-predictable condition (62% and 64%) were similar may not be of particular significance and to assume that it reflects no change in the predictability influences across the local and sentential predictability prescreens would likely be erroneous. To make this absolutely clear, if all of the predictability constraint in both the prescreens derived entirely from influences of local predictability, then target completion rates for each item (in each condition) across the two prescreen tests should be identical (because the sentential context should not exert any influence beyond that of the influence of local predictability). This possibility was something that we were able to test directly by plotting the two predictability scores (one from each pretest) in the high predictability conditions for each item against each other. As can be seen from Figure 2, it is very clearly the case that the scores from each pretest were not highly correlated, demonstrating that each predictability cue made its own contribution.

Beyond these data, and as we noted earlier, the combination of two constraints, for example, the joint constraint of preceding sentential context and the constraint imposed by local predictability, may produce effects that are either increased or reduced relative to the influence of one, or the other, alone. To be clear, the modulatory influence of contextual constraint over constraints of local predictability is not unidirectional in nature (see Figure 2).

To avoid any question regarding the pattern of effects in our predictability completion data, we undertook a replication and required two new groups of 40 participants to respectively undertake the two original prescreen studies (the prescreen with the local predictability cue alone and the prescreen with both the local predictability

cue and the cue from sentential context). The results of these studies were very similar to the original results that we obtained: Local predictability completions were higher for compound words under high ($M=64%$, $SD=0.21$) than low ($M=4.2%$, $SD=0.05$) constraint conditions, $t(71)=23.17$, $p<.001$; sentential context and local predictability correct completions were higher for compound words under high ($M=69%$, $SD=0.22$) than low ($M=4.2%$, $SD=0.05$) constraint conditions, $t(71)=23.79$, $p<.001$. Given the comparability of the effects between the first prescreens and the replication prescreens, we also combined the two data sets together to allow for a set of analyses with very substantial statistical power. Unsurprisingly, the combined analyses showed that the proportion of correct completions was approximately similar in both prescreens and that there was a substantial influence of predictability, local predictability completions, 64% and 4%; $t(71)=23.13$, $p<.001$; sentential context and local predictability completions, 66% and 4%; $t(71)=23.71$, $p<.001$. Overall, these analyses demonstrate the reliability of the original pattern of prescreen effects and show that despite the similarity of the mean correct completion rates across the studies, it was clearly the case that both sources of constraint (sentential context and local predictability) each exerted an influence on completions such that the pattern across items was differentially influenced across each prescreen. Thus, more generally, our prescreen results demonstrate that the joint constraint of preceding sentential context and the constraint imposed by local predictability may be either increased or reduced relative to the influence of one, or the other, alone.

Second constituent predictability from first character (local predictability). To obtain an index of transitional probability, 40 participants created two-character words by adding the first character that came to mind to a list formed from each first character of each target compound word. There were significantly more high-predictable ($M=64%$, $SD=0.21$) than low-predictable ($M=4.1%$, $SD=0.05$) word completions, $t(71)=23.09$, $p < .001$.

Plausibility ratings. Thirty participants rated the target sentences for their plausibility, using a 5-point scale (1 = *very plausible*, 5 = *very implausible*). Besides the 72 experimental sentences, we added 64 implausible filler sentences. There were no significant differences in plausibility for sentences with high- ($M=2.00$, $SD=0.51$) and low-predictable ($M=2.09$, $SD=0.48$) second characters, $t(71) < 1.33$.

Procedure

Prior to the experiment, participants were read a set of instructions. Then, a 3-point calibration was performed and its accuracy checked and rechecked before each sentence. Recalibration was performed whenever necessary. Participants were told to read sentences for comprehension at their own rate. The items were counterbalanced using a Latin square design such that participants saw each sentence and target word only once. After every three sentences, a comprehension question was asked about the preceding sentence. The participants answered the questions by pressing a Yes or No key on a gamepad. After the experiment, participants were asked whether they experienced anything unusual during reading. A small number of subjects reported seeing something flicker on the screen on only one or two trials. No participant was able to report exactly what it was that they had seen.

In total, participants read 114 sentences: 72 experimental sentences randomly intermingled with 36 filler sentences, preceded by six practice sentences. Including 5 min for the initial calibration, the whole experiment lasted about 30 min.

Results

Eight participants were discarded because more than 25% of the display changes occurred during a fixation. For the 36 participants included in the analyses, trials in which the display change occurred during a fixation due to drift were excluded. Fixations less than 60 ms or greater than 800 ms were also excluded, in total 10.2% of the data were excluded (including track losses). All readers scored over 75% correct on the questions; the mean comprehension accuracy was 89.6% for all of participants who were included in the analysis.

We estimated the statistical power of the experiment based on an effect size estimated from the most comparable existing study in the literature (Cui et al., 2013b). We used PANGEA (v0.2) (Westfall et al., 2014). The power of our current design is 0.911 for an effect size of $d=.47$, which was calculated based on means and standard deviations in the study by Cui et al. (2013b). This value is greater than the recommended level of 0.8 and this demonstrates that our study has good power to establish an effect of the type we expected.

Linear mixed models (LMMs) were conducted using the lme4 package (version 1.1-7) in R (R Development Core Team, 2014). As fixed factors, we included the predictability and preview conditions and their interaction. A “full” random model including intercepts and slopes for the main effects and their interactions with participants and items as random factors was run, and if it either did not converge for the dependent measures due to missing values or had perfect correlations in the random structure (a sign of overparameterisation), then a model was conducted with intercepts and where possible slopes for the main effects with participants and items as random factors. Furthermore, two contrasts were programmed to test for preview effects in the two predictability conditions (Drieghe et al., 2019). The first contrast compared the identical and pseudocharacter previews in the high predictability condition, and the second contrast compared the identical and pseudocharacter previews in the low predictability condition. The fixation times were analysed using log-transformed data and the skipping probability was analysed using logistic LMMs. Fixation time measures and skipping probability measures averaged across participants are presented in Table 2 and fixed effect estimations for the fixation times and skipping probability measures are shown in Table 3.

Eye fixation measures for the first constituent

Significant effects of preview were observed for all fixation time measures and skipping probability with longer fixation times and reduced skipping when the dissimilar preview was presented compared with when the identical preview was presented. Clearly, participants were sensitive to the presence of a pseudocharacter relative to the character itself prior to the boundary change. The main effect of predictability was never close to significant in any of the measures. However, the interaction between preview and word type was significant in FFD and gaze duration (GD) at this point in the sentence, although note that it was not significant for single fixation duration, go-past time, and skipping probability. The patterns of effects for FFD and GD here are interesting. Indeed, as we predicted, in the identity conditions, these times were shorter in the high

Table 2. Eye movement measures for the first constituent, the second constituent, and the whole two-character word region.

	High predictability		Low predictability	
	Identical preview	Pseudocharacter preview	Identical preview	Pseudocharacter preview
First character				
FFD	239 (76)	267 (91)	246 (77)	254 (84)
SFD	241 (76)	267 (90)	245 (76)	256 (84)
Gaze	247 (85)	279 (102)	249 (81)	265 (94)
Go-past	279 (124)	322 (155)	294 (141)	315 (156)
SP	.53 (.50)	.44 (.50)	.48 (.50)	.45 (.50)
Second character				
FFD	246 (85)	290 (98)	245 (81)	285 (105)
SFD	247 (85)	292 (96)	245 (80)	291 (103)
Gaze	250 (88)	309 (112)	254 (93)	307 (114)
Go-past	292 (159)	412 (198)	298 (156)	439 (211)
SP	.52 (.50)	.30 (.46)	.46 (.50)	.34 (.47)
The whole word				
FFD	243 (81)	272 (90)	247 (79)	268 (88)
SFD	243 (81)	276 (89)	245 (78)	271 (87)
Gaze	289 (141)	415 (208)	300 (142)	396 (198)
Go-past	345 (217)	525 (299)	392 (264)	536 (316)
SP	.20 (.40)	.10 (.29)	.14 (.35)	.13 (.33)

FFD: first fixation duration (the duration of the first fixation on the word, regardless of how many fixations are made); SFD: single fixation duration (duration of the only fixation on a word during the first pass); Gaze: gaze duration (sum of all fixation durations on a word during the first pass); Go-past: go-past time (sum of all fixations on a word and any words to the left of it before going past it to the right); SP: skipping probability (rate of not fixating a word during the first pass); I: identical; D: dissimilar. Standard deviations are provided in parentheses.

Table 3. Fixed effects estimates from the LME models for all measures across all target regions.

Region	Effect	Measure														
		FFD			SFD			Gaze			Go-past			SP		
		<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>z</i>
First constituent	Predictability	-0.00	0.02	-0.15	-0.00	0.02	-0.20	-0.01	0.02	-0.55	0.03	0.02	1.07	-0.08	0.08	-0.96
	Preview	0.07	0.02	3.78	0.07	0.02	4.25	0.08	0.02	3.75	0.08	0.04	2.20	-0.29	0.10	-2.77
	Interaction	-0.08	0.03	-2.35	-0.06	0.03	-1.81	-0.07	0.03	-2.00	-0.04	0.04	-0.96	0.24	0.18	1.31
Second constituent	Predictability	-0.02	0.02	-1.01	-0.01	0.02	-0.57	-0.00	0.02	-0.07	0.04	0.02	1.61	-0.02	0.08	-0.29
	Preview	0.14	0.02	8.18	0.16	0.02	7.55	0.19	0.02	8.28	0.36	0.02	15.9	-0.76	0.08	-9.19
	Interaction	-0.03	0.03	-0.88	-0.02	0.03	-0.60	-0.03	0.04	-0.89	0.03	0.04	0.70	0.40	0.17	2.41
Whole word	Predictability	0.00	0.01	0.29	0.00	0.02	0.05	-0.00	0.02	-0.25	0.05	0.02	2.53	-0.07	0.12	-0.56
	Preview	0.09	0.01	6.95	0.12	0.02	6.23	0.29	0.03	10.9	0.36	0.04	9.70	-0.53	0.12	-4.52
	Interaction	-0.04	0.03	-1.45	-0.04	0.03	-1.42	-0.08	0.04	-2.28	-0.10	0.04	-2.41	0.80	0.24	3.39

LME: linear mixed-effects; FFD: first fixation duration; SFD: single fixation duration; SP: skipping probability; SE: standard error. Significant terms are presented in bold.

than the low predictability condition (FFD=7 ms effect, GD=2 ms effect). These differences suggest a sensitivity to the identity of the second character of the target prior to its direct fixation, although it must be acknowledged that any such sensitivity is limited given the modest differences that we observed. However, recall that at this point in the sentence, the eyes had not yet crossed the boundary, and therefore, the text remained identical in the high and low predictability pseudocharacter conditions. Despite this, we obtained

greater numerical differences for the pseudocharacters across the predictability conditions than was the case for the identity stimuli. Note also that reading times were actually longer for the high than the low predictability pseudocharacter conditions (FFD = 13 ms effect, GD = 14 ms effect). It is not at all evident why such a pattern of effects should occur for the pseudocharacter previews at this point in the sentence, and for this reason, we feel that this aspect of our results should be treated with caution.

Eye fixation measures for the second constituent

At the second constituent, preview effects were similar to those observed for the first constituent. Significant effects of preview were observed for skipping probability and all fixation time measures with reduced skipping and longer fixation times when the preview was dissimilar relative to when it was identical to the target. There were no main effects of predictability. No significant interactions between preview and predictability were observed for any of the fixation time measures but there were significant effects for the skipping probability with larger preview benefit in the high (0.22) compared with the low predictability condition (0.12). Note also that this effect reflects processing on the preboundary region (i.e., prior to the boundary change) because the decision to skip the second constituent was almost certainly made during the fixation prior to the saccade across the boundary.

Eye fixation measures for the whole compound

The analyses of the whole region can provide an indication of the nature of processing over the whole target compound word, although reading times for this region may include fixations both before and after the boundary change. Taking this region as a whole allows us to provide early indices (skipping, FFD, and single fixation duration) of disruption or facilitation associated with the preview, as well as somewhat later measures that reflect processing that occurs prior to a reader making a fixation on words away from the target (GD), or fixations on new words to the right (go-past time). These later measures represent how the word is processed in its entirety in the context of the sentence. The main effect of predictability was not close to significant in any of the measures. However, there was a significant main effect of predictability for go-past time such that times on the compound word with a low-predictable second constituent were 58 ms longer than in the high predictability condition. Significant effects of preview were observed for skipping probability and all fixation time measures with reduced skipping and longer fixation times when the dissimilar preview was presented compared with when the identical preview was presented. The interaction between preview and word type was significant in GD, go-past time, and skipping probability but was not significant in FFD and single fixation duration. The interaction was due to the preview effect being larger for the high predictability condition (gaze=126 ms; go-past=180 ms; skipping probability=.20) compared with the low predictability condition (gaze=96 ms; go-past=144 ms; skipping probability=.01). Note, as with the skipping effect for the second constituent, the skipping probability effect here reflects processing on the pretarget

region (i.e., an effect prior to the boundary change) because the decision to skip the whole target word was made during the fixation before crossing the boundary.

Discussion

We investigated how predictability information, both in relation to sentential context and local predictability, and perceptual information concerning the upcoming constituent character of a word jointly affect lexical identification and eye movement behaviour during natural Chinese reading. We used two kinds of two-character compound target words that shared the same first character, and we manipulated whether the second constituent character of the target word was either highly predictable (based on sentential context and local predictability cues) or low predictable. We also manipulated the availability of reliable preview information for the second constituent character of the target word using the boundary paradigm with the boundary positioned between the two constituent characters of the target.

Let us first deal with the basic preview effects that we obtained. We found that when the preview of the second constituent of the target was a pseudocharacter, all reading time measures were longer and skipping probability was reduced relative to the identity preview condition. These effects were very robust and occurred for the first constituent, second constituent, and the whole target word. The results reflect standard preview benefit effects, that is, when the target word was available in the parafovea, then processing of it occurred more rapidly when it was fixated compared with when it was unavailable. A further aspect of these results that deserves comment concerns the effects that occurred at the first constituent. Recall that the boundary was positioned after the first constituent, and therefore, the effects observed at the first constituent reflect a sensitivity to the pseudocharacter in the parafovea. To be clear, the effects observed at this point reflect disruption to processing due to the presence of an orthographically illegal upcoming character that was detected prior to its fixation. This is clearly a parafoveal-on-foveal effect that is very likely to be orthographically, or perhaps morphologically, based.

The second aspect of our results that requires discussion concerns the predictability effect that we obtained. Somewhat surprisingly, only one reading time measure, go-past time, for one of our regions, the whole target region, showed an effect of predictability. Based on our prescreen data, it is clear that our stimuli induced the appropriate expectations in our participants, and of course, the predictability effect that we did observe for the go-past reading time does indicate that our manipulation in this regard was effective. Nonetheless, it is slightly surprising that we did not obtain more evidence of predictability influences on other eye movement measures, although we

believe that the position of the boundary within our target word and the presentation of a pseudocharacter preview of the second constituent on half the trials might have contributed to the relatively weak predictability effects that we obtained.

The most important results in our study, however, were the interactive effects of predictability and preview. Interactions occurred in skipping for the whole target word and the second constituent, for FFD and GD for the first constituent and GD and go-past time for the whole target word. Recall that we suggested earlier that the FFD and GD effects that we observed for the first constituent of the target word should be treated with caution. While there were modest mean differences between the high and low predictability identity conditions in a direction that may suggest a sensitivity to the predictability of the target word prior to the boundary change (and its direct fixation), it is almost certainly the case that the mean differences between the pseudocharacter previews for the high and low predictability conditions also contributed to this robust effect. Recall that in these conditions, the stimuli were identical up to this point in processing. Consequently, it is unclear why such differences would have occurred. For this reason, we will not discuss this result further.

In dealing with the remaining interactions, it is helpful to categorise the effects into those that reflect processing prior to the boundary change and those that reflect processing both before and after the boundary change. The interactive effects for skipping for the whole target word and the second constituent both reflect processing commitments made prior to the boundary change. The decisions to skip the second constituent or to skip the whole target word were made during the fixation prior to the saccade that crossed the boundary. To this extent, these effects reflect a processing sensitivity to the preview manipulation, and entirely consistent with this suggestion, the difference in skipping rate between the identity and pseudocharacter preview conditions for both the second constituent and the entire target was greater in the high than the low predictability conditions. These results demonstrate that the extent to which a word was predictable influenced how effectively information regarding the identity of the second constituent character of the target (positioned to the right of fixation) was processed. Furthermore, in relation to the whole target word region skipping results, these influences were apparent even when the character manipulated for preview was at least two characters distant from fixation. We think that these effects are very likely due to the strong predictability cue provided by the first constituent of the target in relation to the identity of the second character. Note also that this effect is consistent with the numerical differences observed between the high and low predictability identity conditions for FFD and GD on the first constituent of the target word. It appears that readers were treating the first and second constituent as a

single orthographic unit for lexical identification (cf. Zang, 2019; Zang et al., 2021). These interactive effects illustrate that predictability and preview jointly constrained the word identification process, reducing potential candidates likely to be the target word. At a more general level, this result demonstrates that Chinese readers use perceptual information about the upcoming constituent characters of a word jointly with expectations of a word's likelihood generated from sentence context based on real-world knowledge and stored, local predictability metrics. It appears that these expectations constrain lexical identification online during natural reading. To be clear, we have demonstrated that both higher order sources of semantic information (obtained from text to the left of fixation) and lower level perceptual and relatively shallow linguistic sources of information (obtained from upcoming text to the right of fixation) exert a direct and relatively immediate influence on word identification as we read (Chang, Zhang, et al., 2020; Staub & Goddard, 2019). Our results are a clear illustration of concurrent top-down and bottom-up constraints in operation during sentence processing.

The final results that we must consider are the interactive effects that we obtained in GD and go-past time for the whole target word. Recall that readers took much less time to process the whole target word in the identity compared with the pseudocharacter preview conditions when it was high relative to low predictable. Importantly, these effects reflect fixations that occur both before and after the boundary change occurred. That is to say, the fixations that occurred prior to the boundary change reflect processing of predictability in relation to the target character preview, whereas those made after the boundary change reflect processing of predictability in relation to the target character itself. For this reason, we undertook some exploratory analyses in which we split the data for the GD and go-past measures for the whole target region to consider separately the fixations that occurred before (GD=253 ms, 41.3% of fixations, and go-past=297 ms, 38.0% of fixations) and after (GD=324 ms, 58.7% of fixations, and go-past=407 ms, 62.0% of fixations) the boundary change. It is clear that readers spent longer processing the target word as a whole after the boundary change (in both measures) than they did processing the word (under preview conditions) before the boundary change, that is, they spent less time processing the first constituent and the preview of the second constituent relative to the time they spent processing the whole word in its identity form. It is difficult to form a clear interpretation of the whole word reading times due to them capturing effects associated with preview processing and effects associated with the integration of the whole target word with sentence context. Teasing these influences apart is not straightforward. However, what we can be assured of is that the majority of the fixations and time that these measures capture were associated with processing the word after the boundary had changed, and presumably, a significant

proportion of that time would reflect sentence integration processes (rather than processing of the preview). If this suggestion is correct, then we might tentatively interpret the interactive effects for GD and go-past time for the whole word, at least in part, to reflect integrative processing.

One final issue that is somewhat difficult to disentangle on the basis of the present results concerns the representational level at which these effects occur. One possibility is that the preview influences derive from aspects of orthographic processing, such that basic visual characteristics of a valid preview work with predictability to constrain lexical candidates. Alternatively, these effects may operate at a morphological level, with the initial morpheme of the target word activating the second morpheme of the compound at a lexical level. We see this, to some degree at least, as a distinction between perception (orthographic preview) and prediction (morphological priming). To be clear, the effects here may reflect a contextually based expectation for a particular orthographic form (which may or may not be available contingent on the nature of the preview) or instead (or even also) may reflect a morphologically based expectation derived from the identity of the initial character of the compound word. Future experimentation is required to discriminate between these possibilities. What we can say with some certainty on the basis of the present set of results, however, is that there are concurrent bottom-up influences derived from visual characteristics of the upcoming character, and top-down expectations based on sentential context that constrain lexical identification during natural Chinese reading.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: L.C. was supported by a grant from the Humanities and Social Sciences of Chinese Ministry of Education (18YJC190001), a grant from the Social Science Foundation of Shandong (19BYSJ46), a grant from the Natural Science Foundation of Shandong (ZR2020MC222), and a Research Scholarship from the China Scholarship Council. S.P.L. and C.Z. acknowledge support from ESRC Grant (ES/R003386/1).

ORCID iD

Lei Cui  <https://orcid.org/0000-0003-3584-8477>

Data accessibility statement



The data and materials from the present experiment are publicly available at the Open Science Framework website: <https://osf.io/awxp7/>

References

- Abbott, M. J., Angele, B., Ahn, Y. D., & Rayner, K. (2015). Skipping syntactically illegal the previews: The role of predictability. *Journal of Experimental Psychology Learning Memory & Cognition*, *41*(6), 1703–1714.
- Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477–492.
- Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *58*(6), 1065–1086.
- Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*, 364–390.
- Bélanger, N. N., & Rayner, K. (2013). Frequency and predictability effects in eye fixations for skilled and less-skilled deaf readers. *Visual Cognition*, *21*, 477–497.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLOS ONE*, *5*(6), Article e10729.
- Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *Spanish Journal of Psychology*, *5*(1), 66–77.
- Chang, M., Hao, L., Zhao, S., Li, L., Paterson, K. B., & Wang, J. (2020). Flexible parafoveal encoding of character order supports word predictability effects in Chinese reading: Evidence from eye movements. *Attention, Perception & Psychophysics*, *82*(6), 2793–2801.
- Chang, M., Zhang, K., Hao, L., Zhao, S., McGowan, V. A., Warrington, K. L., Paterson, K. B., Wang, J., & Gunn, S. C. (2020). Word predictability depends on parafoveal preview validity in Chinese reading. *Visual Cognition*, *28*, 33–40.
- Chinese Daily Word Frequency Dictionary*. (1998). Beijing Language and Culture University.
- Choi, W., Lowder, M. W., Ferreira, F., Swaab, T. Y., & Henderson, J. M. (2017). Effects of word predictability and preview lexicality on eye movements during reading: A comparison between young and older adults. *Psychology and Aging*, *32*, 232–242.
- Cui, L., Drieghe, D., Yan, G., Bai, X., Chi, H., & Livsedge, S. P. (2013a). Parafoveal processing across different lexical constituents in Chinese reading. *Quarterly Journal of Experimental Psychology*, *66*(2), 403–416.
- Cui, L., Yan, G., Bai, X. J., Hyönä, J., Wang, S., & Livsedge, S. P. (2013b). Processing of compound-word characters in reading Chinese: An eye movement contingent display change study. *Quarterly Journal of Experimental Psychology*, *66*(3), 527–547.
- Cutter, M. G., Drieghe, D., & Livsedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1778–1786.
- Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, G. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology*, *16*, 79–103.

- Drieghe, D., Veldre, A., Fitzsimmons, G., Ashby, J., & Andrews, S. (2019). The influence of number of syllables on word skipping during reading revisited. *Psychonomic Bulletin & Review*, 26(2), 616–621.
- Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20, 641–655.
- Fitzsimmons, G., & Drieghe, D. (2013). How fast can predictability influence word skipping during reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(4), 1054–1063.
- Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, 140, 186–209.
- Hand, C. J., Mielle, S., O'Donnell, P. J., & Sereno, S. C. (2010). The frequency-predictability interaction in reading: It depends where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1294–1313.
- Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & Cognition*, 32, 523–532.
- Kretzschmar, F., Schlesewsky, M., & Staub, A. (2015). Dissociating word frequency and predictability effects in reading: Evidence from coregistration of eye movements and EEG. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 1648–1662.
- Lee, Y. A., Binder, K. S., Kim, J. O., Pollatsek, A., & Rayner, K. (1999). Activation of phonological codes during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 948–964.
- Lexicon of common words in contemporary Chinese*. (2008). The Commercial Press.
- Li, N., Wang, S., Mo, L., & Kliegl, R. (2018). Contextual constraint and preview time modulate the semantic preview effect: Evidence from Chinese sentence reading. *Quarterly Journal of Experimental Psychology*, 71(1), 241–249.
- Liversedge, S. P., Drieghe, D., Li, X., Yan, G., Bai, X., & Hyönä, J. (2016). Universality in eye movements and reading: A trilingual investigation. *Cognition*, 147, 1–20.
- McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, 14, 648–652.
- McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, 43, 1735–1751.
- Mielle, S., Sparrow, L., & Sereno, S. C. (2007). Word frequency and predictability effects in reading French: An evaluation of the E-Z Reader model. *Psychonomic Bulletin & Review*, 14, 762–769.
- R Development Core Team. (2014). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65–81.
- Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology Human Perception & Performance*, 30(4), 720–732.
- Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology Human Perception & Performance*, 37, 514–528.
- Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3, 504–509.
- Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review*, 18, 371–376.
- Staub, A., & Goddard, K. (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45, 110–127.
- Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5), 2020–2045.
- White, S. J., Rayner, K., & Liversedge, S. P. (2005). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review*, 12(3), 466–471.
- 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. (2021). Foveal and parafoveal processing of Chinese three-character idioms in reading. *Journal of Memory and Language*, 119, Article 104243.
- Zang, C., Wang, Y., Bai, X., Yan, G., Drieghe, D., & Liversedge, S. P. (2016). The use of probabilistic lexicality cues for word segmentation in Chinese reading. *The Quarterly Journal of Experimental Psychology*, 69(3), 548–560.
- Zola, D. (1984). Redundancy and word perception during reading. *Perception & Psychophysics*, 36, 277–284.