Psychonomic Bulletin & Review
https://doi.org/10.3758/s13423-024-02614-0

BRIEF REPORT

Readers may not integrate words strictly in the order in which they appear in Chinese reading

Hui Zhao[1,2] · Linjieqiong Huang[1] · Xingshan Li[1,2]

Accepted: 30 October 2024

Abstract

The current study investigated whether word integration follows a strictly sequential order during natural Chinese reading. Chinese readers' eye movements were recorded when they read sentences containing a three-character string (ABC), where BC was always a two-character word and AB was also a two-character word in the overlapping condition but not a word in the non-overlapping condition. We manipulated the extent to which word BC was plausible as an immediate continuation following prior context (cross-word plausibility); the string AB was always implausible given the prior context, and the sentence continued in a manner that was compatible with A-BC. The results showed that there were longer second-pass reading times on the string ABC region in the cross-word plausible condition than those in the cross-word implausible condition in both the overlapping condition and the non-overlapping condition. These results imply that readers do not always integrate words strictly in the order in which they appear in Chinese reading.

Keywords Chinese reading · Eye movements · Prior context · Word integration

Introduction

Due to constraints in retinal visual acuity and visual attention, readers can perceive new information from a limited region within a fixation called the perceptual span (Rayner, 1975). How readers perceive new visual information within the perceptual span and how they integrate newly acquired information have been interesting questions in the last decades for both alphabetic and logographic writing systems (Reichle, Liversedge et al., 2009; Snell & Grainger, 2019a).

One interesting question regarding how words are processed within the perceptual span is whether words are integrated in a strictly sequential order. While some theories suggest words are integrated sequentially as presented (Reichle, Warren et al., 2009), others assume words are not integrated into context in a strictly sequential order (Gibson et al., 2013; Huang & Staub, 2021, 2023). Integrating words in correct order is important for accurate language comprehension. For example, if readers do not integrate words in the order in which they appear, they cannot differentiate "a dog bites a man" from "a man bites a dog." This fact supports the argument that readers integrate words in a strict sequential order. However, some studies showed that readers did not always do so. For example, Gibson et al. (2013) showed that readers might ignore word N (the currently fixated word) and integrate word N+1 (the right word of the currently fixated word) directly into context. Recent studies on the transposed-word effect also suggested that readers did not always integrate words sequentially, and word N+1 might be integrated earlier than word N (Huang & Staub, 2021; Mirault et al., 2018; Snell & Grainger, 2019a, 2019b). These studies showed that readers sometimes failed to notice word transposition errors, reporting an ungrammatical sentence with two transposed words (e.g., The white *was cat* big) to be grammatical. The transposed-word effect was found even when words were presented serially (Hossain & White, 2023; Huang & Staub, 2023; Liu et al., 2022; Milledge et al., 2023; Mirault et al., 2022; but see Snell & Nogueira-Melo, 2024).

The question regarding how words within the perceptual span are processed is more challenging and interesting in Chinese, because it is a distinct writing system with several

✉ Xingshan Li
lixs@psych.ac.cn

1 CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

2 Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

Published online: 18 November 2024

unique properties (Li et al., 2022). Chinese text is written by characters located within equal-sized boxes, without spaces explicitly marking word boundaries. For this reason, Chinese readers cannot easily segment words using low-level visual information provided by inter-word spaces as their English counterparts do. These unique properties might cause Chinese readers to integrate information differently while reading.

Using overlapping ambiguous strings (OASs) as stimuli, Huang et al. (2021) have showed that Chinese readers may not always integrate words strictly sequentially into context. An OAS usually consists of three characters (ABC, denoting characters from left to right), where the middle character can form different words with characters on its left (word AB) and its right (word BC). For example, for the OAS "带头发", the left two characters form a word "带头" (meaning *take the lead*) and the right two characters form another word "头发" (meaning *hair*). Huang et al. manipulated the extent to which the right-side word (BC) was plausible as an immediate continuation following pre-OAS context, which is referred to as cross-word plausibility hereafter (see Table 1 for examples). They found longer gaze durations and higher regression-in probabilities on the ABC region when word BC was more plausible given sentence context prior to A, indicating readers sometimes ignored word A and integrated word BC directly.

The findings of Huang et al. (2021) were discussed in the framework of the Chinese Reading Model (CRM; Li & Pollatsek, 2020). CRM assumes that all words in the perceptual span are activated during reading, and words that overlap in space compete for a winner. Factors such as word frequency and left-side advantage affect word competition. Once a word wins the competition, it is identified and segmented from the text simultaneously. Therefore, CRM assumes that word segmentation and word identification is a unified process. Huang et al. suggested that prior context might also affect word competition when processing OASs. If word BC is plausible give prior-A context, word BC is more likely to be segmented as a word than when word BC is implausible. It should be noted that CRM did not implement a syntactic or a semantic component, thus it cannot simulate how sentence context affects word segmentation. The results of Huang et al. suggest that the syntactic and semantic components should be implemented in the future. In Huang et al.'s experiments, word AB was always plausible because the correct segmentation of sentences was AB-C. Compared with the less cross-word plausible condition where only word AB was plausible, the competition between words AB and BC was stronger in the more cross-word plausible condition because both words AB and BC were plausible, which led to longer gaze durations. They inferred that word BC won the competition more frequently when its cross-word plausibility was high. When word BC won the competition, readers might ignore word A and directly integrate word BC into context.

Notably, Huang et al. (2021) used OASs as stimuli, leaving it unclear whether their findings apply to text that does not include OASs. For most character strings, neighboring words do not overlap spatially and CRM assumes they do not compete with each other. If words N and N+1 do not overlap, do Chinese readers still integrate words out of order sometimes? This is the question we asked in the present study. The answer to this question can help us to understand at what stage the cross-word plausibility effects occur. Huang et al. suggested that prior context might influence the word segmentation stage when the target strings are OASs. In the present study, we also used non-overlapping strings in which word boundaries were not ambiguous. Therefore, prior context does not influence the outcome of word segmentation, and it can exert an effect only after the word segmentation stage, likely during the word integration stage. Therefore, if we find that Chinese readers sometimes integrate words out of order for words without ambiguous

Table 1 Examples of stimuli in Experiment 1 in Huang et al. (2021)

Condition	Introduction	Example/Translation
More cross-word plausible	Stimuli	张铭说他有**带头-发**传单给路人
	The whole sentence	Zhang Ming said he had *taken the lead* in *handing out* leaflets to passers-by.
	Prior context + word AB	Zhang Ming said he had *taken the lead*…
	Prior context + word BC	Zhang Ming said he had *hair*...
Less cross-word plausible	Stimuli	有经验的张铭**带头-发**传单给路人
	The whole sentence	Zhang Ming who was experienced *took the lead* in *handing out* leaflets to passers-by.
	Prior context + word AB	Zhang Ming who was experienced *took the lead*...
	Prior context + word BC	Zhang Ming who was experienced *hair*...

The overlapping ambiguous strings (OASs) are in bold and the hyphens are added for illustrative purposes, but the characters were not bolded or segmented in the experiment. The OAS, word AB and word BC in the corresponding translation are in italics boundaries, it suggests that the cross-word plausibility effect can happen after the word segmentation stage.

There are two possible answers to our question. One is that word integration occurs strictly in a sequential manner, where readers integrate each word sequentially in the direction of reading. This aligns with the assumptions of some computational models of reading, such as the E-Z Reader model (Reichle, Warren et al., 2009).

The other possibility is that Chinese readers do not always integrate words strictly in the order they appear; instead, later words may sometimes be integrated into context before earlier ones. The noisy channel model makes similar assumptions (Gibson et al., 2013). According to this model, readers sometimes unconsciously correct errors in text based on their own experience and knowledge, making non-literal but reasonable interpretations. For example, when an extra word is inserted into text, readers may automatically ignore the "strange" word and directly integrate the subsequent text into the sentence. However, few studies explored whether readers also ignore individual characters in totally correct text and integrate later words directly into the sentence.

In this study, we report an eye-tracking experiment that examined whether word integration is strictly sequential in Chinese reading. We embedded a three-character string ABC in sentences to investigate whether readers sometimes ignore word A and directly integrate word BC into sentence context. We manipulated the plausibility of word BC given preceding context before word A (cross-word plausibility) to investigate whether readers integrate word BC into context while ignoring word A in some situations. We also manipulated whether characters A and B constitute a two-character word to investigate whether the integration process is different when the string ABC is an OAS or not. Because BC was always a word, the string ABC formed an OAS if characters A and B constituted a word, but the string was not an OAS if characters A and B did not constitute a word.

The two possible answers above make different predictions regarding second-pass reading times on the string ABC region. If Chinese readers integrate words in a strictly sequential order, second-pass reading times on the string ABC region will be comparable between two different cross-word plausibility conditions. Conversely, if Chinese readers do not integrate words in a strictly sequential order, when word BC is plausible given the pre-A context, readers might mistakenly integrate BC into context. However, because ignoring word A will make sentences hard to interpret when reading the subsequent text, readers might need to make a regressive eye movement to fix the mistake, which leads to longer second-pass reading times. In contrast, when word BC is implausible given pre-A context, readers integrate it into context only after integrating word A during the first-pass reading. Therefore, readers will encounter less difficulties when they read the subsequent text and do not need to go back to the string ABC region to reread, leading to shorter second-pass reading times. In summary, if Chinese readers do not integrate words strictly in the order in which they appear, there will be longer second-pass reading times on the string ABC region in the cross-word plausible condition than those in the cross-word implausible condition.

The manipulation of whether the string ABC is an OAS or not can help us to understand whether Chinese readers integrate an OAS differently from other regular character strings. Because Huang et al. (2021) have found the cross-word plausibility effect on OASs, if whether the three-character string is an OAS or not does not affect the integration process, we expect similar results for two overlapping type conditions. Otherwise, the cross-word plausibility effect only occurs when the string ABC is an OAS, but not when it is not an OAS.

Methods

Participants

One hundred and fifteen college students (88 females; age range 18–30 years, M = 23.34 years, SE = .20) participated in the experiment. Given the number of trials in each condition, this yielded 1,610 total observations per condition, which is comparable to the recommendation of Brysbaert and Stevens (2018) for well-powered within-subjects designs. All participants were native Chinese speakers with normal or corrected-to-normal vision, and no self-reported history of reading or language disorders. Informed consent was obtained for each participant.

Materials and design

A total of 56 words were chosen as target words (BC), which were medium-frequency words (Lexicon of Common Words in Contemporary Chinese Research Team, 2008). Each word BC was paired with a character (character A) to form a three-character string ABC. In the overlapping condition, character A was a single-character word and characters A and B also constituted a two-character word (word AB). For example, in the string "跟前任", character A "跟" and character B "前" formed a word "跟前" (meaning *in front of*), and character B "前" and character C "任" formed another word "前任" (meaning *ex*). The word frequency (AB: M = 10.92 occurrences per million, SE = 1.31, ranging between 1.50 and 44.51; BC: M = 13.94 occurrences per million, SE = 1.37, ranging between 1.48 and 43.42, $t(55) = -1.58$, p = .120) and stroke number (AB: M = 14.23, SE = .54, ranging between 7 and 26; BC: M = 15.11, SE = .59, ranging between 9 and 27, $t(55) = -1.56$, p = .124) were matched between words

Table 2 Properties of the stimuli

	Overlapping		Non-overlapping		Overlapping type	Plausibility	Interaction
	Plausible	Implausible	Plausible	Implausible			
Plausibility of the whole sentence	5.46 (0.08) Range: 4.0–6.6	5.57 (0.08) Range: 4.3–6.7	5.50 (0.09) Range: 4.1–6.6	5.59 (0.08) Range: 4.0–6.6	$F(1,55) = 0.12$, $p = .731$	$F(1,55) = 1.85$, $p = .180$	$F(1,55) = 0.02$, $p = .879$
Plausibility: prior context + string ABC	4.65 (0.14) Range: 2.4–6.7	4.66 (0.14) Range: 2.5–6.8	4.80 (0.15) Range: 1.8–6.5	4.86 (0.16) Range: 1.7–6.4	$F(1,55) = 1.82$, $p = .183$	$F(1,55) = 0.06$, $p = .805$	$F(1,55) = 0.07$, $p = .798$
Plausibility: prior context + word AB	3.22 (0.07) Range: 1.8–3.9	3.09 (0.08) Range: 1.6–3.9	/	/		$t(110) = 1.22$, $p = .225$	
Plausibility: prior context + word BC	5.27 (0.09) Range: 4.2–6.8	2.72 (0.08) Range: 1.5–3.9	5.30 (0.09) Range: 4.3–6.8	2.73 (0.08) Range: 1.5–3.9	$F(1,55) = 0.88$, $p = .354$	$F(1,55) = 474.69$, $p < .001$	$F(1,55) = 0.21$, $p = .647$
Sentence length	18.29 (0.31)	18.18 (0.32)	18.29 (0.30)	18.29 (0.32)	$F(1,55) = 0.42$, $p = .518$	$F(1,55) = 0.22$, $p = .644$	$F(1,55) = 3.11$, $p = .083$
Word frequency of character A	1,752.36 (247.45)		1,423.92 (176.79)		$t(110) = -1.08$, $p = .283$		
Character frequency of character A	2,211.34 (298.37)		1,672.56 (205.33)		$t(110) = -1.49$, $p = .140$		
Stroke number of character A	7.25 (0.36)		7.02 (0.37)		$t(110) = -0.45$, $p = .652$		

Standard errors are given in parentheses. The units of word frequency and character frequency are the number of occurrences per million

AB and BC. In the non-overlapping condition, character A was also a single-character word but characters A and B could not constitute a word. For example, in the string "和前任", character A "和" and character B "前" could not form a word, but character B "前" and character C "任" formed a word "前任" (meaning *ex*). The character and word frequency and stroke number of character A were not significantly different between the overlapping and non-overlapping conditions (see Table 2 for details).

Each three-character string ABC was embedded into four sentences (see Table 3 for examples). In the sentences, the correct segmentation of the string ABC was always A-BC. We manipulated the extent to which word BC was plausible given the preceding context before word A, in the case of ignoring word A. This results in a cross-word plausible condition and a cross-word implausible condition. For example, in the cross-word plausible condition, given a preceding context "她下决心删掉" (meaning *She was determined to delete*), word BC "前任" (meaning *her ex*) was plausible. In contrast, for the cross-word implausible condition, given a preceding context "她因为冲动地" (meaning *She impulsively*), word BC "前任" (meaning *her ex*) was implausible. Therefore, the design was a 2 (overlapping type: overlapping vs. non-overlapping) × 2 (plausibility: plausible vs. implausible) within-participant and within-item design. We made every effort to minimize differences in stimuli across conditions. The target strings (ABC) were the same in the plausible and implausible conditions, and the sentence frames were identical for the overlapping and non-overlapping conditions.

We conducted four norming studies to assess the plausibility of word BC, strings AB and ABC and whole sentences on a 7-point scale (1 = very implausible, 7 = very plausible). A total of 120 native Chinese speakers who did not participate in the main experiment were recruited. Forty participants were assigned to one of four counterbalanced lists and asked to rate the plausibility of string ABC, and another 40 participants were asked to rate the plausibility of whole sentences. Twenty participants were assigned to one of two counterbalanced lists and were asked to rate the plausibility of word AB, and another 20 participants were asked to rate the cross-word plausibility of word BC. When participants rated the cross-word plausibility of word BC, they were only presented with "prior context + word BC" without word A in the sentence fragment (e.g., "她下决心删掉前任"/"她因为冲动地前任"). The instructions received by the participants were: "*All sentences you see are incomplete, please rate the plausibility of sentences so far*." The results showed that the cross-word plausibility of the right-side words (BC) was rated significantly higher in the plausible condition than that in the implausible condition. Within each item, the plausibility value in the plausible condition was greater than 4 and that in the implausible

Table 3 Examples of stimuli in the Experiment

Overlapping type	Cross-word plausibility	Introduction	Example/Translation
Overlapping	Plausible	Stimuli	她下决心删掉跟-前任一起拍的所有亲密照片
		The whole sentence	She was determined to delete all intimate photos she had taken *with her ex*
		Prior context + word AB	She was determined to delete *in front of her*...
		Prior context + word BC	She was determined to delete *her ex*…
		Prior context + string ABC	She was determined to delete *with her ex* …
	Implausible	Stimuli	她因为冲动地跟-前任说了要复合而后悔不已
		The whole sentence	She regretted impulsively telling *her ex* that she wanted to get back together
		Prior context + word AB	She impulsively *in front of her*...
		Prior context + word BC	She impulsively *her ex*...
		Prior context + string ABC	She impulsively *with her ex* …
Non-overlapping	Plausible	Stimuli	她下决心删掉和-前任一起拍的所有亲密照片
		The whole sentence	She was determined to delete all intimate photos she had taken *with her ex*
		Prior context + string AB	She was determined to delete *and before*…
		Prior context + word BC	She was determined to delete *her ex*…
		Prior context + string ABC	She was determined to delete *with her ex* …
	Implausible	Stimuli	她因为冲动地和-前任说了要复合而后悔不已
		The whole sentence	She regretted impulsively telling *her ex* that she wanted to get back together
		Prior context + string AB	She impulsively *and before*…
		Prior context + word BC	She impulsively *her ex*...
		Prior context + string ABC	She impulsively *with her ex* …

The strings ABC are in bold and the hyphens are added for illustrative purposes, but the characters were not bolded or segmented in the experiment. The string ABC, word/string AB and word BC in the corresponding translation are in italics

condition was smaller than 4. Because characters A and B could not form a word in the non-overlapping condition, the plausibility of string AB was low. To ensure consistency, we controlled the plausibility of word AB in the overlapping condition, maintaining values below 4, which showed no significant difference between the plausible and implausible conditions. The plausibility of whole sentences and string ABC and sentence lengths were comparable across conditions (see Table 2 for details). To make sure readers do not predict any part of the target strings (ABC), we also assigned 20 participants who did not participate in the main experiment to one of two counterbalanced lists and asked them to write down the words they predicted after reading the preceding context. The predictability of words A, AB, or BC in the string ABC was close to zero (0.01, 0, 0.01 in the overlapping plausible condition; 0.03, 0, 0 in the overlapping implausible condition; 0.01, 0, 0.01 in the non-overlapping plausible condition; 0.01, 0, 0 in the non-overlapping implausible condition). Given our effort to control the material properties cross conditions, any discrepancies observed in the results should be primarily attributed to the independent variables that we manipulated.

Apparatus

Participants' eye movements were recorded using an SR Research Eyelink 2000 eye-tracking system with a sampling rate of 1,000 Hz. The materials were presented on a 21-in. cathode-ray tube monitor (resolution: 1,024 × 768 pixels; refresh rate: 150 Hz) connected to a Dell personal computer. Each sentence was displayed on a single line in Song 20-pt font, and the characters were shown in white (RGB: 255, 255, 255) on a black background (RGB: 0, 0, 0). A chin rest and forehead rest were employed to minimize head movement during the entire experiment. Participants were seated 58 cm from the computer screen; at this distance, one character subtended a visual angle of approximately 0.7°. For each participant, the viewing was binocular, but only the right eye was monitored. The experiment was programmed with the *Eye-Track* software developed by the UMASS Eye Tracking Lab.

Procedure

When participants came into the lab, they were given the experimental instructions and a brief description of the apparatus. The eye tracker was calibrated at the beginning of the experiment and again during the experiment as needed. A three-point calibration and validation procedure were used, and the maximal error of validation was below 0.5° visual angle. Each sentence appeared after participants fixated on a character-sized box at the location of the first character of each sentence. Next, each participant read six sentences for practice, followed by 56 experimental sentences and 56 filler sentences in a random order. Participants were asked to read the sentences silently and to answer comprehension questions following one third of the sentences. After reading each sentence, they pressed a response button to start the next trial.

Data analysis

We conducted separate analyses on the string ABC region and the word BC region. The reason why we conducted the analysis on the BC regions was that BC consistently forms a word across different conditions; therefore, any observed differences can only be attributed to whether the string ABC region has spatially overlapping words. We analyzed the following eye-movement measures: (1) *first fixation duration* is the duration of the first fixation on the target region during first-pass reading; (2) *gaze duration* is the summed duration of all first-pass fixations on the target region before moving on to other words; (3) *go-past time* is the summed duration starting when entering the target region until this region's right boundary is crossed; (4) *second-pass reading time* is the summed duration of all fixations on the target region following the first-pass reading (including zero times when the target region is not fixated, see Clifton et al., 2007), (5) *regression-in probability* is the percentage of regressions made back to the target region after leaving it.

Data were analyzed using *linear mixed-effects models* (LMMs) for continuous variables and *generalized linear mixed-effects models* (GLMMs) for binary variables. We used the (generalized) linear mixed-effects models to directly test theoretically motivated hypotheses using the following three customized contrasts (Schad et al., 2020): (1) assessing the effect of overlapping type (overlapping condition vs. non-overlapping condition), (2) testing whether the two levels of cross-word plausibility differ significantly for the overlapping condition, and (3) testing whether the two levels of cross-word plausibility differ significantly for the non-overlapping condition. We specified the participants and items as crossed random effects, including intercepts and slopes (Baayen et al., 2008). Following Barr et al. (2013), we used the maximal model that could converge. We first constructed a model with a maximal random factor structure. When the maximal model failed to converge, we used a zero-correlation parameter model and dropped the random components that generated the smallest variances (see Appendix Tables 8 and 9 for the final random components). The *lmer* and *glmer* functions from the *lme4* package (Bates et al., 2015) were used. We report the regression coefficients (*b*s, which estimate the effect size), standard errors (*SE*s), *t*-values (for durations), *z*-values (for binary dependent variables), and corresponding *p*-values. We estimated and reported the *p*-values for the effects using the summary function of the *lmerTest* package (Kuznetsova et al., 2017). Fixation duration measures were log-transformed, except for second-pass reading times, which included values of zero.

The mean accuracy of the comprehension questions was 95%, indicating that the participants understood the sentences well. Trials in which participants made more than three blinks while reading the entire sentence or made one or more blinks on the string ABC region were excluded from analyses, resulting in the exclusion of 4.94% of the trials. One subject was excluded because the accuracy rate was too low (63%) and five subjects were excluded because more than one third trials were deleted due to making too many blinks. Fixations with durations longer than 1,000 ms or shorter than 80 ms (approximately 1.75%) were also excluded from analysis.

Results

The string ABC region

Second-pass reading times were longer in the overlapping condition (*M* = 369 ms, *SE* = 20) compared to the non-overlapping condition (*M* = 277 ms, *SE* = 16). Second-pass reading times were significantly longer in the cross-word plausible condition than the cross-word implausible condition for both the overlapping and non-overlapping strings (overlapping and plausible condition: *M* = 401 ms, *SE* = 24; overlapping and implausible condition: *M* = 336 ms, *SE* = 19; non-overlapping and plausible condition: *M* = 299 ms, *SE* = 18; non-overlapping and implausible condition: *M* = 256 ms, *SE* = 17). First fixation durations and gaze durations showed no significant difference between the overlapping and non-overlapping conditions, nor between the cross-word plausible and implausible conditions in either condition. Go-past times were longer in the overlapping condition (*M* = 688 ms, *SE* = 23) compared to the non-overlapping condition (*M* = 620 ms, *SE* = 19), despite the lack of a significant cross-word plausibility effect. However, the difference in regression-in probabilities showed a trend of difference between the overlapping condition (*M* = .43, *SE* = .02) and the non-overlapping condition (*M* = .39, *SE* =

.02). Additionally, in the non-overlapping condition, regression-in probabilities were higher in the cross-word plausible condition compared to the cross-word implausible condition, while in the overlapping condition, regression-in probabilities were comparable between the cross-word plausible and implausible conditions. Detailed eye-movement measures and fixed-effect estimates from the (G)LMMs for all measures on the string ABC region are shown in Tables 4 and 5.

The word BC region

All measures were larger in the overlapping condition (first fixation durations: M = 296 ms, SE = 5; gaze durations: M = 362 ms, SE = 9; go-past times: M = 572 ms, SE = 17; second-pass reading times: M = 269 ms, SE = 15; regression-in probabilities: M = .36, SE = .01) than those in the non-overlapping condition (first fixation durations: M = 281 ms, SE = 4; gaze durations: M = 331 ms, SE = 7; go-past times: M = 501 ms, SE = 15; second-pass reading times: M = 195 ms, SE = 12; regression-in probabilities: M = .32, SE = .02). First fixation durations, gaze durations, and go-past times were comparable between the cross-word plausible and implausible conditions in both the overlapping and non-overlapping conditions. Second-pass reading times were significantly longer in the cross-word plausible condition than the implausible condition in the non-overlapping condition, while the difference was marginally significant in the overlapping condition. Regression-in probabilities showed a marginally significant difference between the cross-word plausible and implausible conditions in the non-overlapping condition, while they were comparable between the cross-word plausible and implausible conditions in the overlapping condition. Detailed eye-movement measures and fixed-effect estimates from the (G)LMMs for all measures on the word BC region are shown in Tables 6 and 7.

Table 4 Eye-movement measures on the string ABC region

Measures	Overlapping		Non-overlapping	
	Plausible	Implausible	Plausible	Implausible
First fixation duration (ms)	286 (5)	288 (5)	280 (5)	283 (4)
Gaze duration (ms)	490 (16)	481 (14)	460 (14)	459 (14)
Go-past time (ms)	696 (26)	678 (25)	623 (22)	615 (21)
Second-pass reading time (ms)	401 (24)	336 (19)	299 (18)	256 (17)
Regression-in probability	.44 (.02)	.42 (.02)	.42 (.02)	.37 (.02)

Table 5 Results of the (generalized) linear mixed-effects models on the string ABC region

Measures	Fixed effect	*Estimate*	*SE*	*t/z*	*p*
First fixation duration	Overlapping type	-0.01	0.01	-1.23	.223
	Non-overlapping: Plausible vs. Implausible	-0.01	0.01	-0.79	.431
	Overlapping: Plausible vs. Implausible	-0.01	0.01	-0.56	.580
Gaze duration	Overlapping type	-0.04	0.03	-1.54	.129
	Non-overlapping: Plausible vs. Implausible	0.01	0.02	0.49	.624
	Overlapping: Plausible vs. Implausible	0.001	0.02	0.06	.955
Go-past time	**Overlapping type**	**-0.08**	**0.04**	**-2.10**	**.040**
	Non-overlapping: Plausible vs. Implausible	0.02	0.04	0.52	.607
	Overlapping: Plausible vs. Implausible	0.01	0.04	0.20	.841
Second-pass reading time	**Overlapping type**	**-94.58**	**30.51**	**-3.10**	**.003**
	Non-overlapping: Plausible vs. Implausible	**43.76**	**21.55**	**2.03**	**.047**
	Overlapping: Plausible vs. Implausible	**64.49**	**31.24**	**2.06**	**.044**
Regression-in probability	Overlapping type	-0.20	0.10	-1.93	.054
	Non-overlapping: Plausible vs. Implausible	**0.22**	**0.09**	**2.38**	**.017**
	Overlapping: Plausible vs. Implausible	0.09	0.10	0.86	.389

Significant effects are indicated in bold

Table 6 Eye-movement measures on the word BC region

Measures	Overlapping		Non-overlapping	
	Plausible	Implausible	Plausible	Implausible
First fixation duration (ms)	295 (6)	296 (5)	280 (5)	283 (4)
Gaze duration (ms)	365 (11)	359 (9)	331 (8)	329 (7)
Go-past time (ms)	590 (23)	555 (18)	496 (17)	506 (17)
Second-pass reading time (ms)	293 (18)	246 (14)	211 (13)	180 (12)
Regression-in probability	.37 (.02)	.35 (.02)	.33 (.02)	.30 (.02)

Table 7 Results of the (generalized) linear mixed-effects models on the word BC region

Measures	Fixed effect	*Estimate*	*SE*	*t/z*	*p*
First fixation duration	**Overlapping type**	**-0.04**	**0.01**	**-4.26**	**<.001**
	Non-overlapping: Plausible vs. Implausible	-0.02	0.01	-1.33	.184
	Overlapping: Plausible vs. Implausible	-0.01	0.02	-0.40	.694
Gaze duration	**Overlapping type**	**-0.08**	**0.02**	**-4.46**	**<.001**
	Non-overlapping: Plausible vs. Implausible	-0.001	0.02	-0.06	.951
	Overlapping: Plausible vs. Implausible	0.002	0.03	0.06	.950
Go-past time	**Overlapping type**	**-0.11**	**0.04**	**-2.71**	**.009**
	Non-overlapping: Plausible vs. Implausible	-0.01	0.04	-0.18	.856
	Overlapping: Plausible vs. Implausible	0.04	0.04	1.01	.319
Second-pass reading time	**Overlapping type**	**-76.99**	**23.06**	**-3.34**	**.002**
	Non-overlapping: Plausible vs. Implausible	**31.93**	**14.35**	**2.23**	**.030**
	Overlapping: Plausible vs. Implausible	46.88	23.83	1.97	.054
Regression-in probability	**Overlapping type**	**-0.20**	**0.10**	**-2.07**	**.039**
	Non-overlapping: Plausible vs. Implausible	0.16	0.09	1.87	.061
	Overlapping: Plausible vs. Implausible	0.09	0.11	0.87	.386

Significant effects are indicated in bold

Discussion

In the present study, an eye-tracking experiment was conducted to investigate whether word integration is strictly in the order of presentation in natural Chinese reading. We found that the cross-word plausibility of word BC affected second-pass reading times. Second-pass reading times in the cross-word plausible condition were longer than those in the implausible condition on both the word BC and the string ABC regions, in both the overlapping and the non-overlapping conditions. We also found that whether the key three-character string is an overlapping ambiguous string (OAS) affected all eye-movement measures on the word BC region, and affected second-pass reading times and go-past times on the string ABC region. Reading times were longer when the three-character string was an OAS compared with when it was not.

The cross-word plausibility effect observed in both the overlapping and the non-overlapping conditions suggests that Chinese readers sometimes do not integrate words strictly in the order of presentation, allowing later words to be integrated before earlier ones. Put another way, word BC can be integrated into context directly without integrating word A in the non-overlapping condition, and without integrating either word A or word AB in the overlapping condition. When word BC was cross-word plausible, readers might integrate word BC into context and this would cause an error. However, when word BC was cross-word implausible, it was easy for readers to integrate word A into context because they found word BC was implausible as an immediate continuation following prior context. As a result, there were longer second-pass reading times when word BC was cross-word plausible than when it was implausible.

The cross-word plausibility effect in the overlapping condition in our experiment aligns with Huang et al. (2021), which showed longer reading times in the more cross-word plausible condition compared to the less cross-word plausible condition. In Huang et al.'s study, all of the stimuli were

OASs, thus their study could not distinguish whether the results were caused by the specific properties of OASs. In the present study, we found the cross-word plausibility effect in both overlapping and non-overlapping conditions. These results suggest the cross-word plausibility effect observed by Huang et al. was not totally caused by the specific properties of OASs. Thus, word integration may be not strictly sequential during Chinese reading, regardless of whether sentences contain an OAS or not.

These findings help us to understand at what stage the cross-word plausibility effects occur. Huang et al. (2021) suggested that prior context influenced the word segmentation stage when the target strings are OASs in which words AB compete with words BC. However, since there were no word boundary ambiguities in the non-overlapping strings used in the present study, prior context can exert an effect only after the word segmentation stage. Therefore, the cross-word plausibility observed in the non-overlapping condition suggests that this effect can happen after the word segmentation stage.

Why do Chinese readers not integrate words in a strictly sequential order? The noisy channel model offers one account (Gibson et al., 2013). According to this model, readers use both bottom-up and top-down information during reading, making non-literal but more plausible inferences based on the world knowledge and linguistic knowledge. When word BC was cross-word plausible, readers might infer that ignoring word A might be more reasonable for sentence comprehension, so that they directly integrate word BC into sentence context. However, this may lead to comprehension difficulties when reading the subsequent text, resulting in regressive eye movements for correction. In contrast, when word BC was cross-word implausible, readers might rely less on inference and were less likely to ignore word A. In this case, readers would integrate word A first and then word BC, leading to correct integration order without the need for corrective eye movements.

Another possibility is that words are integrated to sentence context immediately upon identification, regardless of presentation order. According to the Chinese Reading Model (CRM; Li & Pollatsek, 2020), Chinese characters within the perceptual span are processed in parallel, all words composed of these Chinese characters will be activated, and any word will be identified as long as the activation degree exceeds the threshold. In most situations, words are identified in the order they appear. However, in some special situations, word N+1 might be identified before word N, leading to earlier integration of subsequent words. Both accounts stated above can explain the results of the present study, warranting further studies to distinguish these possibilities.

The OB1-reader model provides an explanation on how word order is encoded during reading (Snell et al., 2017, 2018; Snell & Grainger, 2019a). The model assumes that multiple words are processed in parallel, and the activated words are mapped onto a spatiotopic sentence-level representation to keep track of word order using word length information. The OB1-reader model relies heavily on inter-word spaces to determine word length, which are used to map identified words into proper position in the sentence. But it should be noted that it is very hard to extend OB1-reader model into Chinese reading because Chinese script does not use inter-word spaces to demarcate words, and readers cannot perceive word length with low-level vision.

While both Huang et al. (2021) and the present study showed words are not always integrated in the order they appear, the results are not identical. Huang et al found longer gaze durations when word BC was more cross-word plausible than those when word BC was less plausible. However, in the present study, we did not find the cross-word plausibility effect on gaze durations. This difference may be attributed to different properties of stimuli between the two experiments. Huang et al.'s experiments manipulated the frequency of word BC to be higher or lower than word AB, whereas in our experiment, words AB and BC were medium-frequency words, and the word frequencies were comparable between these two words. It is possible that the cross-word plausibility can only exert an effect on the word segmentation stage when the activation passes a certain threshold. In Huang et al., the frequency of word BC was high enough for its plausibility to influence the competition between words AB and BC. However, in the present study, the frequency of word BC was not high enough. As a result, the cross-word plausibility effect was small and the present study did not have enough power to detect it on gaze durations. This argument aligned with Huang et al.'s observation that the cross-word plausibility effect of gaze durations diminished when the word frequency of BC was low.

One limitation of the present study is that we did not distinguish between syntactic plausibility and semantic plausibility. It is unknown whether the cross-word plausibility effects due to syntactic plausibility and those due to semantic plausibility are different. Future studies are needed to investigate this question.

We found reading times in the overlapping condition were significantly longer than those in the non-overlapping condition, reflected by first fixation durations and gaze durations on the word BC region. One possible reason is that both words AB and BC are activated in the overlapping condition, while only word BC is activated in the non-overlapping condition. Activation of word AB results in longer reading times, which aligns with previous studies (Inhoff & Wu, 2005; Perfetti & Tan, 1999). In the overlapping condition, words AB and BC compete with each other, as character B can only belong to one word, whereas in the non-overlapping condition, only word BC is activated without competition. Word competition may result in longer reading times, which is consistent with the assumptions of the CRM.

Conclusion

To summarize, the results of the present study suggest that Chinese readers sometimes do not integrate words strictly in the order in which they appear. The present study extends the findings of Huang et al. (2021), suggesting that word integration is not strictly sequential even when words are non-overlapping. Furthermore, the cross-word plausibility effect can happen after the word segmentation stage.

Appendix

Table 8 Final (Generalized) Linear Mixed-effects Models on the String ABC Region

Measures	Formula
First fixation duration	log(DV) ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + overlapping type + overlapping: plausible vs. implausible ǁ subject) + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǁ item)
First pass reading time	log(DV) ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + non-overlapping: plausible vs. implausible ǁ subject) + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǁ item)
Go-past time	log(DV) ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǀ subject) + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǀ item)
Second-pass reading time	DV ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + overlapping type ǁ subject) + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǁ item)
Regression-in probability	DV ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + non-overlapping: plausible vs. implausible ǁ subject) + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǁ item)

DV, dependent variable

Table 9 Final (Generalized) Linear Mixed-effects Models on the Word BC Region

Measures	Formula
First fixation duration	log(DV) ~ 1s + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + overlapping: plausible vs. implausible ǁ subject) + (1 + overlapping type + overlapping: plausible vs. implausible ǁ item)
Gaze duration	log(DV) ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + overlapping type + overlapping: plausible vs. implausible ǁ subject) + (1 + overlapping type + overlapping: plausible vs. implausible ǁ item)
Go-past time	log(DV) ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǁ subject) + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǁ item)
Second-pass reading time	DV ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 + overlapping type ǁ subject) + (1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible ǁ item)
Regression-in probability	DV ~ 1 + overlapping type + non-overlapping: plausible vs. implausible + overlapping: plausible vs. implausible + (1 ǀ subject) + (1 + overlapping type + overlapping: plausible vs. implausible ǁ item)

DV, dependent variable

Authors' contributions Hui Zhao: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – Original draft, Writing – Reviewing and Editing.

Linjieqiong Huang: Writing – Reviewing and Editing.

Xingshan Li: Conceptualization, Methodology, Formal analysis, Visualization, Supervision, Writing – Reviewing and Editing, Funding acquisition.

Funding This research was supported by a grant from the National Natural Science Foundation of China (31970992).

Data availability The data and materials of this study are publicly available at Science Data Bank and can be accessed at https://www.scidb.cn/s/6bmyIr.

Code availability The code of this study is publicly available at Science Data Bank and can be accessed at https://www.scidb.cn/s/6bmyIr.

Declarations

Conflicts of interest None.

Ethics approval The study was approved by the ethics committee of the Institute of Psychology, Chinese Academy of Sciences.

Consent to participate All participants gave their informed consent prior to their inclusion in the study.

Consent for publication All participants have consented to the submission of their data to the journal.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

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. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 9. https://doi.org/10.5334/joc.10

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. *Eye movements* (pp. 341–371). Elsevier. https://doi.org/10.1016/B978-008044980-7/50017-3

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences, 110*(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

Hossain, J., & White, A. L. (2023). The transposed word effect is consistent with serial word recognition and varies with reading speed. *Cognition*, *238*, 105512. https://doi.org/10.1016/j.cognition.2023.105512

Huang, K. J., & Staub, A. (2021). Using eye tracking to investigate failure to notice word transpositions in reading. *Cognition, 216*, 104846. https://doi.org/10.1016/j.cognition.2021.104846

Huang, K. J., & Staub, A. (2023). The transposed-word effect does not require parallel word processing: Failure to notice transpositions with serial presentation of words. *Psychonomic Bulletin & Review, 30*(1), 393–400. https://doi.org/10.3758/s13423-022-02150-9

Huang, L., Staub, A., & Li, X. (2021). Prior context influences lexical competition when segmenting Chinese overlapping ambiguous strings. *Journal of Memory and Language, 118*, 104218. https://doi.org/10.1016/j.jml.2021.104218

Inhoff, A. W., & Wu, C. (2005). Eye movements and the identification of spatially ambiguous words during Chinese sentence reading. *Memory & Cognition, 33*(8), 1345–1356. https://doi.org/10.3758/BF03193367

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lexicon of Common Words in Contemporary Chinese Research Team. (2008). *Lexicon of common words in contemporary Chinese*. Commercial 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. https://doi.org/10.1037/rev0000248

Li, X., Huang, L., Yao, P., & Hyönä, J. (2022). Universal and specific reading mechanisms across different writing systems. *Nature Reviews Psychology, 1*(3), 133–144. https://doi.org/10.1038/s44159-022-00022-6

Liu, Z., Li, Y., Cutter, M. G., Paterson, K. B., & Wang, J. (2022). A transposed-word effect across space and time: Evidence from Chinese. *Cognition*, *218*, 104922. https://doi.org/10.1016/j.cognition.2021.104922

Milledge, S. V., Bhatia, N., Mensah-Mcleod, L., Raghvani, P., A. McGowan, V., Elsherif, M. M., Cutter, M. G., Wang, J., Liu, Z., & Paterson, K. B. (2023). The transposed-word effect provides no unequivocal evidence for parallel processing. *Attention, Perception, & Psychophysics*, *85*(8), 2538–2546. https://doi.org/10.3758/s13414-023-02721-5

Mirault, J., Snell, J., & Grainger, J. (2018). You that read wrong again! A transposed-word effect in grammaticality judgments. *Psychological Science, 29*(12), 1922–1929. https://doi.org/10.1177/0956797618806296

Mirault, J., Vandendaele, A., Pegado, F., & Grainger, J. (2022). Transposed-word effects when reading serially. *PLOS ONE*, *17*(11), e0277116. https://doi.org/10.1371/journal.pone.0277116

Perfetti, C. A., & Tan, L. H. (1999). The constituency model of Chinese word identification. In J. Wang, A. W. Inhoff, & H. Chen (Eds.), *Reading Chinese Script: A cognitive analysis* (pp. 115–134). Lawrence Erlbaum Associates.

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

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences, 13*(3), 115–119. https://doi.org/10.1016/j.tics.2008.12.002

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*(1), 1–21. https://doi.org/10.3758/PBR.16.1.1

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language, 110*, 104038. https://doi.org/10.1016/j.jml.2019.104038

Snell, J., & Grainger, J. (2019a). Readers are parallel processors. *Trends in Cognitive Sciences, 23*(7), 537–546. https://doi.org/10.1016/j.tics.2019.04.006

Snell, J., & Grainger, J. (2019b). Word position coding in reading is noisy. *Psychonomic Bulletin & Review, 26*(2), 609–615. https://doi.org/10.3758/s13423-019-01574-0

Snell, J., & Nogueira-Melo, A. (2024). Do love you me? Failure to notice word transpositions is induced by parallel word processing. *Journal of Cognition, 7*(1), 21. https://doi.org/10.5334/joc.335

Snell, J., Meeter, M., & Grainger, J. (2017). Evidence for simultaneous syntactic processing of multiple words during reading. *PLOS ONE, 12*(3), e0173720. https://doi.org/10.1371/journal.pone.0173720

Snell, J., Van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-reader: A model of word recognition and eye movements in text reading. *Psychological Review, 125*(6), 969–984. https://doi.org/10.1037/rev0000119