The Transient Scaffold: A Longitudinal Eye-Tracking Study on the Dynamic Role of Interword Spacing in Chinese L2 Reading

Authors: Yu, Miao, Wang, Tianlin, Sun, Hui, Long,Jiaxin, Yu, Miao, Tianlin Wang

Date: 2025-11-28T10:41:00+00:00

Abstract

While inserting spaces between words facilitates reading in Chinese as a Second Language (CSL) learners, it remains unclear whether this benefit is short-lived or enduring throughout the learning process. Using a longitudinal eye-tracking design, we tracked L1-English CSL learners across four stages of a one-year intensive program in China. The results revealed a dissociable pattern: spacing facilitated global sentence reading only during the initial stage (S1), whereas it supported local lexical processing throughout the first three stages (S1–S3), with benefits disappearing only at the highest proficiency level (S4). These findings indicate that the functional role of interword spacing is highly dynamic and proficiency-dependent. It acts as a scaffold that is gradually withdrawn as reading skills develop, ultimately becoming unnecessary as learners internalize robust orthographic and lexical representations. These results underscore the importance of stage-sensitive pedagogical implementation: spacing can be strategically introduced in beginning instruction to reduce cognitive load, but should be phased out in intermediate and advanced stages to promote adaptive reading of authentic, unspaced Chinese texts.

Full Text

The Transient Scaffold: A Longitudinal Eye-Tracking Study on the Dynamic Role of Interword Spacing in Chinese L2 Reading

Miao Yu[1], Tianlin Wang[2], Hui Sun[1], Jiaxin Long[1]

[1]Key Research Base of Humanities and Social Sciences of the Ministry of Education, Academy of Psychology and Behavior, Faculty of Psychology, Tianjin

Normal University, 393 Binshui West Road, Xiqing District, Tianjin 300387, China

[2]Educational & Counseling Psychology, University at Albany, State University of New York, New York, United States of America

Abstract

While inserting spaces between words facilitates reading in Chinese as a Second Language (CSL) learners, it remains unclear whether this benefit is short-lived or enduring throughout the learning process. Using a longitudinal eye-tracking design, we tracked L1-English CSL learners across four stages of a one-year intensive program in China. The results revealed a dissociable pattern: spacing facilitated global sentence reading only during the initial stage (S1), whereas it supported local lexical processing throughout the first three stages (S1–S3), with benefits disappearing only at the highest proficiency level (S4). These findings indicate that the functional role of interword spacing is highly dynamic and proficiency-dependent. It acts as a scaffold that is gradually withdrawn as reading skills develop, ultimately becoming unnecessary as learners internalize robust orthographic and lexical representations. These results underscore the importance of stage-sensitive pedagogical implementation: spacing can be strategically introduced in beginning instruction to reduce cognitive load, but should be phased out in intermediate and advanced stages to promote adaptive reading of authentic, unspaced Chinese texts.

Keywords: Chinese as a Second Language (CSL); interword spacing; eye-tracking; longitudinal study; reading

Introduction

Lexical access is central to reading comprehension, as readers process information largely in terms of word units (Bai et al., 2008; Li et al., 2014; Radach & Kennedy, 2004; Shen et al., 2010; Winskel et al., 2009). In Latin-script orthographies, interword spaces provide overt visual cues that facilitate word identification and segmentation. By contrast, written Chinese lacks explicit word boundary markers, posing a significant challenge for Chinese as a Second Language (CSL) learners, particularly those from alphabetic backgrounds. Although previous cross-sectional studies have examined the effects of interword spacing on CSL learners' reading, the long-term impact remains unclear. Key questions persist regarding whether the benefits of spacing endure throughout extended learning. Moreover, longitudinal evidence is scarce on how L2 learners visually and cognitively adapt to spaceless texts, and how such adaptation interacts with increasing proficiency. To address these gaps, the present study tracks CSL learners across four phases over the course of one academic year to investigate the evolving role of word spacing in reading development.

In unspaced languages (e.g., Chinese), it is necessary to segment words from

the continuous stream of text, and therefore word segmentation plays a crucial role in word recognition (Li & Pollatsek, 2020; Li et al., 2009). Identifying word boundaries is much less arduous in writing systems that employ interword spacing, since the spaces between words mark explicitly where words begin and end. As an overt visuospatial indicator, interword spaces are helpful for readers in carrying out word identification (Bai et al., 2008). Indeed, many studies of English reading have found that reading is disrupted and reading speed decreases when spaces between words are removed or filled with other characters (Rayner & Pollatsek, 1996; Rayner et al., 1998; Winskel et al., 2009).

A fundamental difference between Chinese and alphabetic scripts like English is the absence of interword spaces in the former. The lack of this overt word boundary information raises a key psycholinguistic question: how do readers segment a continuous string of characters into meaningful words? To explore this, numerous studies have employed interword spacing manipulations, presenting artificially spaced texts to examine the manipulation' s facilitative or inhibitory effects on sentence reading. Collectively, these studies show null spacing benefits for proficient native readers, who are highly adept at segmenting words without overt cues (Bassetti, 2009; Bassetti & Lu, 2016; Shen et al., 2010), but consistent facilitation for younger elementary school children (Li et al., 2021; Song et al., 2021) and CSL learners (Shen et al., 2012). Some have therefore argued that the influence of added spacing is questionable. Indeed, since words in Chinese are typically short (the majority being comprised of just one or two characters) and there is less uncertainty in their length distribution, readers can relatively easily identify word boundaries without spaces (Huang et al., 2024). As a result, proficient native Chinese speakers rarely encounter any challenge with word segmentation. However, for CSL learners, especially those whose first language is alphabetic, the absence of spacing poses a substantial segmentation challenge (Gao & Jiang, 2015), as their native orthography routinely relies on low-level visuospatial cues for lexical parsing. This lifelong experience with spaced alphabetic systems means that alphabetic-L1 learners approach Chinese reading with pre-existing segmentation strategies that are poorly suited to an unspaced script.

Therefore, it is unsurprising that inserting spaces into Chinese text facilitates reading for CSL learners, particularly those from alphabetic L1 backgrounds. A critical question remains, however, concerning the durability of this facilitation: is it long-lasting and stable, resembling native readers' entrenched processing strategies, or is it merely transient, diminishing as learners gain proficiency and adapt to unspaced text?

Two competing theoretical possibilities may explain the persistence of spacing effects in second language Chinese reading. One account posits that L1 reliance on interword spacing induces positive transfer, thereby sustaining long-term benefits in L2 Chinese reading. From this perspective, spacing facilitates core reading processes such as word segmentation by visually marking word boundaries to reduce cognitive load (Zang et al., 2013), and therefore attenuates lat-

eral interference from adjacent characters (Bai et al., 2008). Moreover, spaces help guide saccadic targeting toward the optimal viewing position within words (Rayner, 2009; Rayner et al., 1998; Winskel et al., 2009), thereby enhancing lexical identification and overall reading efficiency. If spacing supports these fundamental perceptual and lexical processes in this manner, its benefits are likely to persist throughout learning.

An alternative account suggests that spacing functions as a temporary scaffold during acquisition, mainly by aiding word segmentation and lexical consolidation. Since Chinese words are composed of character sequences, learners must actively parse continuous text into discrete word units to access meaning. Interword spacing provides salient visual cues that make this process easier. However, as learners develop stronger lexical representations and become increasingly proficient in segmenting unspaced text, the reliance on spacing is hypothesized to diminish (Bassetti & Lu, 2016). According to this view, spacing benefits are not permanent but will gradually diminish with increasing proficiency, potentially disappearing once a sufficient level of reading skill is achieved.

Given these competing theoretical possibilities, the issue of whether interword spacing offers temporary or enduring benefits remains empirically unresolved. While several studies have examined this question, their methodological differences limit clear conclusions.

Certain findings suggest that spacing may support sustained improvements in reading proficiency. For instance, Chang (2002) conducted an 8-week intervention with beginner CSL learners and observed that those trained with spaced texts showed significantly greater gains in reading comprehension and writing than a control group using traditional unspaced texts. These results indicate that prolonged exposure to spaced input may facilitate the acquisition of linguistic representations. Further supporting the potential for lasting effects, Bai et al. (2013) demonstrated that CSL learners (American university students) read novel words more quickly during a test without spacing if they had first been exposed to spaced texts during a prior learning phase. Similarly, Gao and Jiang (2015) reported facilitative effects of spacing on the reading comprehension of long Chinese texts for not only beginning-level but also intermediate-level CSL learners whose native language is alphabetic. Together, these findings suggest that interword spacing may promote the formation of robust lexical representations that persist beyond initial exposure.

In contrast, other studies point to the transient nature of spacing benefits. Bassetti and Lu (2016) found that interword spacing increased reading rates among L1-English CSL learners, but this advantage diminished with increasing proficiency. Higher-proficiency readers derived no significant benefit, indicating that spacing may serve as a scaffold that becomes less necessary as reading skill develops. Corroborating this view, Yao (2011) found that intermediate-to-advanced CSL learners, all of whom had native languages that use spaced writing systems, derived no benefit from interword spaces in their reading speed or comprehension. Collectively, these findings suggest the utility of spacing may be confined

to the earlier stages of L2 Chinese reading development.

However, notable limitations across these studies prevent firm conclusions regarding the long-term utility of spacing. Chang' s (2002) intervention did not include intermediate assessments to trace the emergence of benefits over time, and Bai et al.' s (2013) learning phase spanned only two days and focused narrowly on vocabulary acquisition. Thus, while existing research confirms that spacing facilitates CSL learners' reading, its sustained usefulness across developmental stages remains an open question. The present study addresses this gap through a longitudinal design that tracks the evolution of spacing effects across multiple stages of learning.

The question of whether the effect of word spacing in Chinese second language learning is temporary or enduring has important implications for Chinese language instruction. Some studies suggest that spacing can facilitate reading for CSL learners and may therefore warrant inclusion in early instructional texts (Gao & Jiang, 2015; Peng & Su, 2009; Shen et al., 2012). However, if research reveals that spacing effects are limited to the earliest stages of learning, then its role should not be overemphasized in curriculum design.

Most existing research on the effects of interword spacing in Chinese L2 reading has relied predominantly on cross-sectional designs, which limit insight into how such effects evolve with prolonged learning. In addition, prior studies often involved learners with varying levels of pre-existing Chinese knowledge, making it difficult to isolate the role of spacing in the initial formation of lexical representations. To address these limitations, the present study employed a longitudinal eye-tracking approach to examine spacing effects over a one-year intensive program in L1-English learners with no prior exposure to Chinese. This design allows us to trace the trajectory of spacing effects from the very onset of learning.

Based on previous findings that spacing facilitates reading in CSL learners (Gao & Jiang, 2015; Shen et al., 2012), we expected to observe initial benefits among our participants. The central question, however, was whether these benefits would persist or diminish with increasing proficiency. If spacing offers a general facilitatory influence on reading processes such as word segmentation and identification, its benefits should remain stable. Alternatively, if spacing functions primarily as a temporary scaffold, its benefits should decline as learners develop more robust mental lexical representations and become able to segment words without visual cues. We predicted that spacing would prove most beneficial during early stages, with facilitation attenuating as proficiency grows.

Method

Participants. Nineteen participants (11 females; Mage = 19 years, SD = 0.21) from Scotland participated in the experiment. All participants had zero background in Chinese and were true novice learners. They were right-handed native speakers of English with normal or corrected-to-normal vision. All participants

progressed through the Immersion Chinese Language Program during their one year of stay in China. Power analysis (Faul et al., 2007) indicated that with a significance level of $\alpha = 0.05$ and a medium effect size of f = 0.25, the sample size (n = 19) in this study achieved statistical power of 0.89. Moreover, based on the results from Yu et al.' s (2015) meta-analysis, the effect size of interword spacing on local fixation duration for CSL learners is 1.37. Given this effect size, a sample of 19 participants would provide statistical power of 1.

Design and Materials. The experiment employed a longitudinal within-subject design spanning a full academic year, with two independent variables: Spacing (spaced vs. unspaced text) and Learning Stage (four stages: S1 to S4). Data were collected at four stages aligned with the participants' midterm and final examinations: S1 (mid-first semester), S2 (end of first semester), S3 (mid-second semester), and S4 (end of second semester). Each learning stage represented an instructional period of approximately eight weeks. This alignment was intentionally chosen to situate the experiment within participants' ongoing reading development assessments, thereby enhancing both engagement and ecological validity. Participants were explicitly informed, however, that their eye-movement data would not contribute to their academic evaluation. The approval for this research was granted by the relevant committee at the university' s Research Ethics Committee.

We constructed four sets of test sentences based on the vocabulary from the first ten lessons of the course textbooks (*The Basic Course of Experiencing Chinese (I)* and *The Oral Course of Experiencing Chinese (I)*). Each set contained 64 sentences ranging from 6 to 12 Chinese characters in length (M = 8.86, SD = 1.48). The words in each sentence differed across the four testing stages, but the number of words per sentence remained the same in each test set, and the sentences were broadly matched on syntactic structure (e.g., subject-verb-object word order). To ensure consistency in difficulty and naturalness across the four stimulus sets, twenty native Chinese-speaking college students rated the naturalness of each sentence on a 5-point scale (1 = very unnatural to 5 = very natural). Additionally, twenty Chinese language teachers currently teaching CSL students provided difficulty ratings on a scale from 1 (very easy) to 5 (very difficult). Test sentences comprising each set were then selected to ensure comparable levels of difficulty and naturalness across the four stimulus sets.

Table 1 presents the means and standard deviations of naturalness and difficulty ratings for the experimental sentences. Statistical analysis indicated no significant differences in naturalness ($F(3, 252) = 0.173$, $p = 0.91$, $^2 = 0.002$) or difficulty ratings ($F(3, 252) = 0.07$, $p = 0.98$, $^2 = 0.0001$) among the four groups of sentences. In each testing stage, two experimental files were constructed for the text spacing manipulation: a spaced condition and an unspaced condition. There were 32 sentences in each condition, and conditions were rotated across files according to a Latin Square design. Sentences in each condition were presented randomly in a blocked format. Table 2 shows an example of experimental

sentences.

Apparatus. All stimuli were presented in a single line on a 21-inch CRT screen at a viewing distance of 70 cm, with a resolution of 1024 × 768 pixels and a frame rate of 120 Hz. Chinese characters were presented in 21-point Songti font, rendered in black, with each character subtending a 0.9° visual angle against a white background. Participants' eye movements were recorded using an SR Research EyeLink 1000 eye tracker with a sampling rate of 1000 Hz. Viewing was binocular, but only the movements of the right eye were recorded.

Procedure. All participants were tested individually. They were instructed to read the sentences silently at their own pace for comprehension. A standard 3-point grid calibration procedure was conducted before the experiment. After successful calibration, sentences appeared on the screen one at a time. Participants were recalibrated if tracker loss occurred or after each break. Upon completing a sentence, participants pressed the space bar to indicate comprehension and terminate the display. Following this action, on approximately one-third of trials they were presented with a yes/no comprehension question. Participants read 8 practice trials before the formal experiment, which lasted about 20 minutes.

Results

The average comprehension accuracy for all participants across the four testing stages was 95.26%, 94.12%, 96.05%, and 95.79%, respectively, indicating that participants fully understood the sentences. All fixation durations longer than 1200 ms or shorter than 80 ms were excluded. Trials with eye movement measures exceeding three standard deviations from the mean were also removed (6.97% of the total data).

Two sets of analyses of the eye movement data were conducted. First, we computed local eye-movement measures to evaluate the time course of spacing effects during sentence reading. Referencing Li et al. (2021) and Shen et al. (2020), we focused on two-character words in both spaced and unspaced conditions to investigate the effect of spacing on CSL learners' reading across the four learning stages. The local measures included first fixation duration (FFD: the duration of the first fixation on a word), gaze duration (GD: the sum of all fixation durations on a word before moving to another word), refixation probability (the probability of making a second fixation on a word before moving to another word), total fixation duration (TFD: the sum of all fixation durations on a word), regression time (the sum of all fixation durations after a regression to the current word), and fixation count (FC: total number of fixations on a word). FFD and GD on target words usually reflect early-stage processing, while regression time, FC, and TFD typically reflect later cognitive processing activities (Rayner, 1998, 2009). Refixation probability indicates the degree to which cognitive processing of the target word interrupts the regular sequence of word-to-word saccades during normal reading (Shen et al., 2012).

Second, we computed global eye movement measures to assess participants' overall reading performance, including total sentence reading time (the sum of all fixation durations on a sentence), total number of fixations (the number of fixations on a sentence), average fixation duration (the average duration of all fixations on a sentence), and average saccade amplitude (the average length of all saccades on a sentence).

Under the R language environment (R Development Core Team, 2020), linear mixed-effects models (LMMs) and generalized mixed-effects models (GLMMs) were used for data analysis with the lme4 package (Bates et al., 2023). Time-based eye movement measures were log-transformed. The model included learning stages (S1, S2, S3, and S4) and spacing conditions (spaced vs. unspaced) as fixed factors, with subjects and items as crossed random effects. All analyses started with the maximum random effects structure (Barr et al., 2013). If the model failed to converge, item correlations, then item slopes, followed by subject correlations and slopes were sequentially removed until the model converged successfully.

The present study employed both global (four measures) and local (six measures) eye-movement analyses, which increased the likelihood of Type I error (von der Malsburg & Angele, 2017). To control for multiple comparisons, we applied the Holm–Bonferroni correction separately to the global and local measures. Additionally, for learning-stage effects (comparisons across S1-S4), we applied correction to adjust p-values for all pairwise comparisons. All adjusted p-values are reported in Tables 4 and 6 .

Local Analyses

The grand means for the local measures are shown in Table 3 , and Table 4 presents the fixed-effect estimates. No significant differences between the word spaced and unspaced conditions were observed for FFD and GD ($|t|s < 1.48$, $ps > 0.15$). FFD in S4 was significantly shorter compared to S1 ($t = -3.19$, $p < 0.01$), and GD in S4 was significantly shorter compared to S1 and S2 ($|t|s > 3.11$, $ps < 0.01$). The interaction between spacing condition and learning stage was significant for GD. Further analysis revealed that in stage one, GD under the spaced condition was significantly longer than under the unspaced condition ($t = -3.26$, $p < 0.01$). However, in S2, S3, and S4, no significant differences were found between the two spacing conditions ($|t|s < 0.63$, $ps > 0.53$). No significant interaction effect was observed for FFD ($|t|s < 1.25$, $ps > 0.21$).

No significant differences were found for spacing conditions ($z = -0.24$, $p = 0.81$) or learning stages ($|z|s < 1.83$, $ps > 0.07$) on refixation probability. However, the interaction between learning stages and spacing conditions was significant. Further analysis revealed that refixation probability was significantly higher for spaced text compared to unspaced text at S1 ($t = -2.55$, $p = 0.01$). In S2, S3, and S4, there were no significant differences between the two spacing conditions ($|z|s < 1.33$, $ps > 0.18$).

TFD and regression time were shorter, and FC was fewer in the spaced condition compared to the unspaced condition (|t|s > 4.49, ps < 0.001). Both TFD and FC decreased sequentially across the four learning stages (|t|s > 3.79, ps < 0.001). Regression time decreased sequentially from S1 to S3 (|t|s > 4.49, ps < 0.001), with no significant difference between S3 and S4 (t = -0.50, p = 0.61). The interaction between spacing condition and learning stage was significant. Further analysis showed that in the first three stages, both TFD and regression time were significantly shorter for the spaced condition than the unspaced condition (|t|s > 2.53, ps < 0.05). In stages two and three, FC was significantly fewer for the spaced condition than the unspaced condition (|t|s > 3.42, ps < 0.001). However, by the fourth stage, differences between spacing conditions were no longer significant for TFD, FC, and regression time (|t|s < 1.41, ps > 0.16). The means for the two spacing conditions and four learning stages are shown in Figure 1 [Figure 1: see original paper], using TFD as an example.

The local analysis revealed two distinct patterns. For early-stage processing measures—including FFD, GD, and refixation probability—spacing had no significant facilitative effect on word recognition among CSL learners. A limited inhibitory effect was observed only at the first learning stage (S1), where spaced text led to longer GD and higher refixation probability. In contrast, spacing consistently facilitated later cognitive processes, as reflected in TFD, FC, and regression time, with significant benefits observed during the first three learning stages (S1–S3).

Global Analyses

The grand means for the global measures are shown in Table 5 , and Table 6 presents the fixed-effect estimates. Total sentence reading time is a key measure that reflects the overall duration readers spent processing sentences across spacing conditions (Bai et al., 2008). The means for the two spacing conditions and four learning stages are shown in Figure 2 [Figure 2: see original paper]. Total sentence reading time in the spaced condition was slightly shorter than in the unspaced condition (t = 1.83, p = 0.068). Total sentence reading time decreased sequentially across the four stages (|t|s > 3.36, ps < 0.01). Additionally, there was a significant interaction between spacing conditions and learning stage. Further analysis revealed that in S1, learners had significantly shorter total sentence reading time in the spaced condition compared to the unspaced condition (t = 2.99, p < 0.01). However, this trend was not found in S2, S3, or S4 (|t|s < 1.60, ps > 0.05). These findings suggest that the facilitative role of interword spacing is most pronounced during the initial stage of Chinese L2 reading but gradually diminishes as learners accumulate experience and proficiency.

Average fixation duration was significantly shorter for spaced than unspaced text (t = 9.75, p < 0.001). Average fixation durations decreased steadily from S1 to S3 (|t|s > 4.05, ps < 0.01), but no significant difference emerged between S3 and S4 (t = -0.99, p = 0.34). The interaction between spacing condition

and learning stage was significant. Further analysis revealed that spaced text yielded significantly shorter average fixation durations than unspaced text at all learning stages ($|t|s > 2.94$, $ps < 0.01$), although the magnitude of this effect declined numerically over time (S1 difference = 22 ms; S2 difference = 19 ms; S3 difference = 10 ms; S4 difference = 8 ms). Thus, although spacing benefits persisted across all stages, they weakened with increasing proficiency.

Average saccade amplitude was significantly shorter in the unspaced condition than in the spaced condition ($t = -27.24$, $p < 0.001$). S1 amplitude was significantly shorter than at later stages ($|t|s > 11.21$, $ps < 0.001$). A significant interaction between spacing and learning stage was observed. Further analysis confirmed that saccade amplitudes were consistently shorter for unspaced than spaced text across all stages ($|t|s > 12.15$, $ps < 0.001$), though the magnitude of spacing effects varied across time (S1 difference = 0.45; S2 difference = 0.64; S3 difference = 0.57; S4 difference = 0.5). It is important to note that average saccade amplitude is influenced by information density. Inserting spaces reduces the horizontal density of characters, thereby decreasing the amount of information available per unit of visual span. This may prompt readers to make longer saccades to compensate for reduced information availability (Shen et al., 2010). For this reason, and because average fixation duration is similarly sensitive to visual-textual confounds, neither measure was treated as a primary indicator of processing effort in this study.

No significant difference was observed in the total number of fixations between the spaced and unspaced conditions ($t = -0.38$, $p = 0.70$). However, the total number of fixations decreased sequentially across the four learning stages ($|t|s > 2.45$, $ps < 0.05$), indicating improved processing efficiency with increasing experience, regardless of text spacing. A significant interaction between spacing and learning stage was also observed. Further analysis revealed that the number of fixations did not differ between spaced and unspaced sentences at S1, S2, or S3 ($|t|s < 0.86$, $ps > 0.39$). At S4, however, readers made significantly more fixations for spaced than unspaced text ($t = -2.04$, $p = 0.042$).

Discussion

Employing a longitudinal eye-tracking design, the present study is the first to directly investigate how interword spacing influences Chinese L2 reading across four progressive stages of learning. The results revealed a robust main effect of learning stage: as participants gained proficiency through immersive instruction, their reading efficiency improved consistently, as reflected in shortened fixation durations and fewer fixation counts over time. Critically, the facilitative effect of spacing attenuated with increasing proficiency, though the time course differed across processing levels. Global reading measures, such as total sentence reading time, showed a spacing benefit only at the earliest stage (S1), which then diminished. In contrast, local measures of lexical processing, including total fixation duration, fixation count, and regression time, showed that spacing supported word identification through the first three stages, with

benefits disappearing only at the final stage (S4). These findings suggest that interword spacing serves as a scaffold that is gradually withdrawn as reading skills develop.

Our longitudinal findings reveal that the facilitative effect of interword spacing is transient yet multifaceted, its time course subject to differential constraints from both the time scale of cognitive activity (early vs. late eye-movement measures) and the level of cognitive processing (lexical vs. sentence). This dynamic pattern is particularly evident in lexical processing, where our results chart a clear trajectory of learners' progression from initial disruption to efficient utilization and eventual independence from the scaffold.

At Stage 1 (8 weeks), spacing introduced a temporary cost to early visual processing, as evidenced by longer gaze duration (GD) and higher refixation probability, indicating an inhibitory effect on early word identification. One plausible explanation lies in perceptual learning mechanisms related to visual crowding (Chiu et al., 2023; Zhang et al., 2009). For L1-English learners, the abrupt insertion of spaces alters the familiar visual gestalt of densely packed Chinese character strings, temporarily interfering with their newly developing visual-processing routines. This interference, however, was short-lived. From S2 to S4, early-stage measures (FFD, GD, refixation probability) showed no further facilitative or inhibitory effects, suggesting rapid adaptation to the altered visual input and dissipation of initial disruption.

Concurrently, and despite the early visual disruption, spacing exerted a persistent facilitative effect on deeper, post-lexical cognitive processes. This was evidenced by consistent reductions in late-stage measures, including total fixation duration (TFD), fixation count (FC), and regression time across the first three learning stages (S1–S3). This pattern suggests that as learners accumulated experience, they quickly learned to capitalize on spacing as a reliable word-boundary cue. This sustained benefit aligns with cognitive load theory (Sweller et al., 2019): once learners recognized spacing as a segmentation cue, it offloaded the cognitively demanding process of word boundary identification, thereby reducing processing burden and freeing resources for higher-level linguistic operations such as syntactic parsing and semantic integration.

Critically, this facilitative scaffold was gradually withdrawn as proficiency increased. By the fourth and final learning stage (S4), the facilitative effect of spacing on word identification disappeared, as shown by the convergence of late-stage measures (TFD, FC, regression time) across conditions. This indicates that as CSL learners developed more robust orthographic and lexical representations, their dependence on low-level visual cues diminished. The disappearance of both the initial cost and the later benefit highlights that learners ultimately internalized word segmentation skills, achieving efficient reading of unspaced text.

At the sentence level, interword spacing facilitated reading only during the initial learning stage (S1), with no observable effects at later stages (S2–S4). This

indicates that the benefit of spacing for global text integration is short-lived. For alphabetic-L1 learners, added spaces provide a temporary advantage specifically during the earliest phase of exposure to Chinese text. The pattern that spacing benefits diminish with increasing proficiency is well-established across languages and learner groups, with evidence from Thai reading (Kasisopa, 2011), proficient CSL learners (Bassetti, 2009; Yao, 2011), and Chinese children, where spacing aids lower-grade but not higher-grade readers (Li et al., 2021; Song et al., 2021). This converging evidence strongly indicates that the utility of spacing declines as reading skill develops.

This transient sentence-level benefit contrasts with the more persistent lexical-level facilitation. The dissociation between these effects highlights a core feature of L2 reading development: different subskills follow distinct developmental trajectories. The finding that spacing aided global sentence comprehension only at the earliest stage (S1) suggests that learners rapidly shifted toward higher-order predictive processing skills. As their proficiency grew, they increasingly relied on established lexical representations and top-down cues, such as syntactic knowledge, semantic context, and grammatical predictions, for sentence integration, thereby reducing their reliance on visual boundaries for global comprehension.

In contrast, the sustained facilitation of spacing on lexical processing across three stages (S1–S3) indicates that word identification continues to rely heavily on visual cues for an extended period. This prolonged benefit likely arises because spacing aids in recognizing less familiar words and resolving segmentation ambiguities, challenges that persist even as sentence-level skills advance. This narrowing functional role of spacing, from supporting both local and global processing early on to primarily aiding lexical access later, reflects the asynchronous development of reading subskills. It further suggests that the influence of spacing diminishes first at higher cognitive levels, where predictive and contextual mechanisms dominate, and only later at the word level, where robust orthographic representations take longer to solidify. Thus, spacing operates not as a uniformly beneficial cue, but as a compensatory aid whose utility varies across processing stages and proficiency levels.

Overall, while the persistence of spacing facilitation differed between lexical and sentence-level processing, its role in both domains demonstrates a clear developmental trajectory aligned with increasing L2 proficiency. The present findings illuminate a process of progressive adaptation: learners transition from initially relying on spacing as an external scaffold across multiple levels of processing, to gradually internalizing word segmentation skills and employing higher-order linguistic knowledge, ultimately achieving efficient reading of unspaced text. This dynamic reflects a fundamental principle of reading acquisition: the shift from reliance on surface visual cues to the efficient use of internalized lexical, syntactic, and semantic representations.

The present study provides a nuanced theoretical contribution by distinguishing the differential impacts of interword spacing on lexical versus sentence-level processing in Chinese L2 reading. Our longitudinal design uniquely captures

the dynamic, stage-dependent nature of these effects, revealing that spacing serves as a transient scaffold for global integration but offers prolonged support for lexical access. These findings advance our understanding of the cognitive architecture of reading development in non-spaced orthographies by illustrating how learners' reliance on visual cues evolves with increasing proficiency. Specifically, the results illuminate the asynchronous development of subskills such as visual decoding, word segmentation, and sentence integration, offering a developmental framework for how visual formatting influences reading processes across different stages of acquisition.

From a practical perspective, our results support the adoption of a pedagogical approach in CSL instruction that is both stage-sensitive and spatially targeted. Specifically, spaced text provided a transient benefit for sentence reading confined to the initial stage (S1), while its facilitative effect on word identification persisted into later stages (S2–S3). These findings suggest that spacing can be strategically introduced to ease lexical segmentation and reduce cognitive load for beginners, particularly when dealing with unfamiliar or complex word structures. However, given that spacing effects diminish rapidly for sentence reading and disappear entirely by the final stage (S4), instructors should avoid prolonged or blanket use of spaced text. Instead, spacing should be gradually phased out as learners reach intermediate proficiency to encourage adaptation to authentic, unspaced Chinese text and prevent visual overreliance.

Importantly, standard beginner textbooks already employ word-unit segmentation in pinyin, a practice that serves a similar cueing function, thereby further reducing the added value of interword spacing in character-based text. Accordingly, we recommend that teachers emphasize pinyin-based word segmentation instruction in the early stages. This approach effectively scaffolds word recognition without introducing artificial visual formatting. Interword spacing should therefore be used strategically and reserved primarily for the initial phase of character-based reading acquisition. Its application should be limited to targeted learning objectives, such as resolving segmentation ambiguities or clarifying word boundaries in syntactically challenging sentences. Beyond the beginner stage, educators should transition learners to fully unspaced text to promote adaptation to authentic reading materials. In the long term, the limited and transient benefits of spacing do not justify its large-scale incorporation into curricular design, particularly given the added spatial and digital overhead it introduces.

However, the present study has several limitations. First, to cleanly assess developmental changes in spacing sensitivity, we used texts of consistent difficulty across all learning stages. While this controlled for text-related confounds and isolated proficiency effects, it may have reduced the ecological validity of later sessions, as advanced learners read simpler materials. Spacing effects may interact with text complexity, and future studies should incorporate proficiency-matched texts at each stage to clarify how spacing supports reading under varied cognitive demands. Second, all participants were native readers of English, an

alphabetic writing system with explicit interword spaces. Whether the observed trajectory of spacing effects generalizes to learners from writing systems without word spacing (e.g., Thai, Japanese) remains an open question. Future research involving such learners would determine the extent to which these patterns reflect universal processes of visual adaptation in L2 reading. Finally, while this study focused on spacing effects in continuous reading, prior research suggests that spacing may also facilitate the reading of specific text types such as ambiguous sentences (Hsu & Huang, 2000a), dynamically scrolling texts (Shieh et al., 2005), and highly complex texts (Hsu & Huang, 2000b) even among native readers. Further work could examine whether CSL learners show similar facilitation under these conditions, thereby extending the functional profile of spacing beyond general reading ease.

Conclusion

The present longitudinal eye-tracking study provides a fine-grained account of how interword spacing dynamically influences Chinese L2 reading acquisition across learning stages. Our core finding is that spacing acts as a transient and dissociable scaffold: it facilitated global text integration only at the earliest stage of learning, whereas it supported local lexical processing throughout the first three stages, with benefits fully attenuating only at the highest proficiency level. This indicates that the utility of spacing narrows over time, shifting from broadly aiding comprehension to selectively assisting lexical access, before becoming redundant as learners develop robust orthographic and lexical representations.

These findings illuminate the dynamic interplay between visual support and reading skill development. More importantly, they offer an empirically grounded basis for implementing stage-informed pedagogical practices in CSL instruction. Educators may consider incorporating spaced texts as a temporary instructional scaffold during initial learning phases, while gradually phasing them out as learners progress, thereby promoting adaptation to authentic Chinese text.

Data availability statement. Data and analysis scripts are available at the following OSF repository: https://osf.io/2ze4w/?view_{only}=63680f1f852b49fcb5e0267b80470185

References

Bai, X., Liang, F., Blythe, H. I., Zang, C., Yan, G., & Liversedge, S. P. (2013). Interword spacing effects on the acquisition of new vocabulary for readers of Chinese as a second language. *Journal of Research in Reading, 36*(S1), S4-S17. https://doi.org/10.1111/j.1467-9817.2013.01554.x

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1277-1287. https://doi.org/10.1037/0096-1523.34.5.1277

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

Bassetti, B. (2009). Effects of adding interword spacing on Chinese reading: A comparison of Chinese native readers and English readers of Chinese as a second language. *Applied Psycholinguistics, 30*(4), 757-775. https://doi.org/10.1017/S0142716409990105

Bassetti, B., & Lu, M. (2016). Effects of interword spacing on native English readers of Chinese as a second language. *International Review of Applied Linguistics in Language Teaching, 54*(1), 1-22. https://doi.org/10.1515/iral-2016-0014

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2023). *lme4: Linear mixed-effects models using* '*Eigen*' *and S4*. https://cran.r-project.org/web/packages/lme4/index.html

Chang, C. P. (2002). *Marking text boundaries and learning the Chinese language* (Doctoral dissertation, University of Southern California).

Chiu, T. Y., & Drieghe, D. (2023). The role of visual crowding in eye movements during reading: Effects of spacing. *Attention, Perception, & Psychophysics, 85*(8). https://doi.org/10.3758/s13414-023-02787-1

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power *3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences.* Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/BF03193146

Gao, S., & Jiang, X. (2015). The effect of word boundaries on Chinese reading for second language learners. *Language Teaching and Linguistic Studies,* (4), 8-16.

Hsu, S. H., & Huang, K. C. (2000a). Interword spacing in Chinese text layout. *Perceptual & Motor Skills, 91*(2), 355–365. https://doi.org/10.2466/pms.2000.91.2.355

Hsu, S. H., & Huang, K. C. (2000b). Effects of word spacing on reading Chinese text from a video display terminal. *Perceptual & Motor Skills, 90*(1), 81–92. https://doi.org/10.2466/PMS.90.1.81-92

Huang, L., Reichle, E. D., & Li, X. (2024). Comparative analyses of the information content of letters, characters, and inter-word spaces across writing systems. *Annals of the New York Academy of Sciences, 1537*(1), 129-139. https://doi.org/10.1111/nyas.15178

Kasisopa, B. (2011). *Reading without spaces between words: Eye movements in reading Thai* (Doctoral dissertation, University of Western Sydney).

Li, S., Wang, Y., Lan, Z., Yuan, X., Zhang, L., & Yan, G. (2021). Effects of word spacing on children' s reading: Evidence from eye movements. *Reading and Writing, 35*(4), 1019-1033. https://doi.org/10.1007/s11145-021-10215-9

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58*(4), 525-552. https://doi.org/10.1016/j.cogpsych.2009.02.003

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., Bicknell, K., Liu, P., Wei, W., & Rayner, K. (2014). Reading is fundamentally similar across disparate writing systems: A systematic characterization of how words and characters influence eye movements in Chinese reading. *Journal of Experimental Psychology: General, 143*(2), 895–913. https://doi.org/10.1037/a0033580

Peng, Y., & Su, L. (2009). A study on eye movement of Korean students reading Chinese texts with or without marks for word boundaries. *International Conference on Asian Language Processing: Recent Advances in Asian Language Processing*, 332–336. https://doi.org/10.1109/IALP.2009.83

R Development Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology, 16*(1-2), 3-26. https://doi.org/10.1080/09541440340000295

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506. https://doi.org/10.1080/17470210902816461

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*(8), 1129-1144. https://doi.org/10.1016/S0042-6989(97)00274-5

Rayner, K., & Pollatsek, A. (1996). Reading unspaced text is not easy: Comments on the implications of Epelboim et al.' s (1994) study for models of eye movement control in reading. *Vision Research, 36*(3), 461-465. https://doi.org/10.1016/0042-6989(95)00132-8

Shen, D. L., Bai, X., Zang, C., Yan, G., Feng, B., & Fan, X. (2010). Effect of word segmentation on beginners'reading: Evidence from eye movements. *Acta Psychologica Sinica, 42*(2), 159-172. https://doi.org/10.3724/SP.J.1041.2010.00159

Shen, D., Liversedge, S. P., Tian, J., Zang, C., Cui, L., & Bai, X., et al. (2012). Eye movements of second language learners when reading spaced and unspaced

Chinese text. *Journal of Experimental Psychology: Applied, 18*(2), 192-202. https://doi.org/10.1037/a0027485

Shieh, K. J., Hsu, S.-H., & Liu, Y. C. (2005). Dynamic Chinese text on a single-line display: Effects of presentation mode. *Perceptual and Motor Skills*. https://doi.org/10.2466/pms.100.3c.1021-1035

Song, Z., Liang, X., Wang, Y., & Yan, G. (2021). Effect of alternating-color words on oral reading in grades 2–5 Chinese children: Evidence from eye movements. *Reading and Writing, 34*(10), 2627-2643. https://doi.org/10.1007/s11145-021-10164-3

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2). https://doi.org/10.1007/s10648-019-09465-5

von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language, 94*, 119–133. https://doi.org/10.1016/j.jml.2016.10.003

Winskel, H., Radach, R., & Luksaneeyanawin, S. (2009). Eye movements when reading spaced and unspaced Thai and English: A comparison of Thai-English bilinguals and English monolinguals. *Journal of Memory and Language, 61*, 339–351. https://doi.org/10.1080/17586801.2016.1225562

Yao, Y. (2011). Interword spacing effects on reading Mandarin Chinese as a second language. *Writing Systems Research, 3*(1), 23-40. https://doi.org/10.1093/wsr/wsr009

Yu, X., Ren, Z., & Ye, Y. (2015). The effects of interword spacing in Chinese text reading: A meta-analysis based on eye movements research. *Advances in Psychological Science, 23*(11), 1894-1909. https://doi.org/10.3724/SP.J.1042.2015.01894

Zang, C., Liang, F., Bai, X., Yan, G., & Liversedge, S. P. (2013). Interword spacing and landing position effects during Chinese reading in children and adults. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 720-734. https://doi.org/10.1037/a0030097

Zhang, J. Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2009). Legibility of Chinese characters in peripheral vision and top-down influences on crowding. *Vision Research, 49*(1). https://doi.org/10.1016/j.visres.2008.09.021

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