Bilingualism: Language and Cognition

cambridge.org/bil

Processing strategy or representation difference? Investigating the word segmentation difficulty of second language learners of Chinese

Panpan Yao[1], Huangxia Li[1], Xinwei Chen[1] and Xingshan Li[2,3]

[1]Cognitive Science and Allied Health School, Institute of Life and Health Sciences, Key Laboratory of Language and Cognitive Science (Ministry of Education), Beijing Language and Culture University, Beijing, China; [2]Institute of Psychology, Chinese Academy of Sciences, Beijing, China and [3]School of Psychology, University of Chinese Academy of Sciences, Beijing, China

Research Article

Cite this article: Yao, P., Li, H., Chen, X. and Li, X. (2025). Processing strategy or representation difference? Investigating the word segmentation difficulty of second language learners of Chinese. *Bilingualism: Language and Cognition* 1–13. https://doi.org/10.1017/S1366728925100643

Received: 12 December 2024
Revised: 3 September 2025
Accepted: 22 September 2025

Keywords:
word segmentation; L2-Chinese processing; lexical representation; processing strategy; eye-tracking

Corresponding author:
Panpan Yao;
Email: yaopp@blcu.edu.cn

This research article was awarded Open Data badge for transparent practices. See the Data Availability Statement for details.

Abstract

Due to the lack of explicit word boundary markers, L2-Chinese learners have shown some difficulties in Chinese word segmentation. This study aimed to tackle the possible reasons of L2-Chinese learners' difficulties in word segmentation: L1-biased processing strategy or developing mental representations of Chinese compound words, or both. In an eye-tracking experiment, high-frequency two-character Chinese compound words were used as targets. These compound words were embedded in sentences where their first component characters with prior verbs were manipulated to be either plausible or implausible, while the whole compound words were always plausible. Sentences were presented in character-spaced or word-spaced style. High-proficiency L2-Chinese learners and native Chinese speakers participated. Results revealed non-native-like patterns of L2-Chinese learners: they holistically processed compound words only in the word-spaced condition, while native speakers did so regardless how sentences were presented. The findings indicated that high-proficiency L2-Chinese learners' difficulty in word segmentation is predominantly caused by their L1-biased processing strategy.

Highlights

1. High-proficiency L2-Chinese learners have acquired holistic representations of Chinese compound words.
2. L2-Chinese learners preferred to use spaces to segment continuous character strings into words.
3. The meanings of constituent characters of compound words were activated for L2-Chinese learners.
4. Compound words were holistically processed for native speakers.
5. Text presentation styles affected L2-Chinese learners' reading but not native speakers'.

1. Introduction

Previous studies revealed that words play an important role in reading in both alphabetic languages such as English and logographic languages such as Chinese (Inhoff et al., 2000; Li et al., 2009; Perea & Acha, 2009; Rayner, 1998, 2009; Reichle et al., 1999; Yan et al., 2006). In alphabetic languages, inter-word spaces mark word boundaries and provide salient visual cue for the start and end of a word. Different from alphabetic languages, there are no explicit visual word boundary markers in Chinese text (Hoosain, 1992; Tsai & McConkie, 2003; Tsang & Chen, 2008). Chinese characters occupy the same unit of space, with little space between every two neighboring characters but not words. One primary and important task in Chinese reading is to determine which continuous characters belong to a word (i.e., word segmentation). Previous studies found that people who learn Chinese as a second language (hereinafter, L2-Chinese learners) find it difficult to segment continuous character strings into words when reading Chinese on-line (Bai et al., 2013; Cui, 2023; Shen et al., 2012; Yang, 2021; Yao et al., 2023). The present study aimed to investigate the possible reasons for the word segmentation difficulty of high-proficiency L2-Chinese learners.

1.1. Processing ambiguity of Chinese compound words

There are more than 6,000 characters in Chinese, and most characters represent a morpheme (Standardization Administration of China, 1980). Chinese words vary with their lengths (the number of constituent characters): there are single-character words (approximately 20%), two-character words (approximately 70%) and words composed by three or more characters

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

(approximately 10%, Lexicon of Common Words in Contemporary Chinese, 2009). Most of the multicharacter words are compound words (80%, Yuan, 1990), and the constituent characters could be independent words and contribute to the meanings of the whole words. For example, the constituent characters of the compound word *笔记* (meaning "notebook") are *笔* ("pen") and *记* ("record"), which are independent and single-character words by themselves. In other words, most compound words and their constituent components are all lexicalized words that can be found in Chinese dictionaries as independent entries.

Since Chinese words vary in length, compound words and their components could all be meaningful and independent units and there is no salient word boundary marker, Chinese readers might experience temporary word boundary ambiguities during reading (Ma et al., 2014; Zhou et al., 2017). For example, the character string *花生长* could be segmented into two different ways: *花生–长* (meaning "peanut-grow") versus *花–生长* (meaning "flower-grow"). Different segmenting methods lead to different meanings of the character string and, consequently, the whole sentence. For successful text comprehension, Chinese readers must rely on context and world knowledge to decide whether each incoming character is an independent word or a constituent morpheme of a compound word (Chen, 1996, 1999). In general, the script features of Chinese are expected to cause temporary ambiguities and reanalysis processes, since the grammatical properties and meanings of each incoming character are controversial and dependent on the context (Liu et al., 2013; Yao et al., 2021).

Previous studies found that Chinese readers can successfully segment compound words from continuous character strings without being affected by the semantic meanings of constituent morphemes. Yang et al. (2012) used two-character Chinese compound words to explore whether Chinese readers take each incoming character as a processing unit and integrate it with the previous context incrementally. The plausibility of the first constituent of target words was manipulated as plausible versus implausible by varying the prior verbs, while the target words were always plausible in the context (see examples in (1a) and (1b)).

(1a) 围观的人看着他踢打门卫却无动于衷。
People were inattentive when he kicked the **gatekeeper**.
(1b) 围观的人看着他哀求门卫却无动于衷。
People were inattentive when he entreated the **gatekeeper**.

If characters were taken as processing units, when reading the sentences in (1), *门* ("door") would be treated as an independent word and integrated with the prior context. Readers would realize this analysis error when reaching *卫* ("guard," which cannot be an independent word in this context and has to be combined with *门* ["door"] to form a compound word), abandon their original analysis, reanalyze the character string *门卫* ("gatekeeper") as a whole word and combine it with the prior verb *踢打* ("kick")/*哀求* ("entreat"). At this point, it would be easier to give up an implausible one (*哀求门* "entreat door") than a plausible one (*踢打门* "kick door"), which would lead to longer processing times in (1a) than in (1b). However, the results revealed that the processing of *门卫* ("gatekeeper") was not significantly changed between (1a) and (1b), indicating that Chinese readers' processing of compound nouns was not affected by whether the constituent component was plausible with prior verbs or not. The authors argued that native Chinese readers' insensitivity to the plausibility between the component morphemes and the preceding verbs indicated that during on-line reading, the semantic meanings of the component morphemes in compound words are either not activated or insufficiently strong to influence lexical processing. In other words, Chinese compound words are processed holistically, with the entire compound word taking priority over its constituent morphemes in semantic activation. Similar findings were reported by Zhou and Li (2021), who found that when processing three-character compound words, the plausibility of embedded words that were composed by the first two characters did not affect the processing of whole compound words.

These observations indicated a "long word advantage" in Chinese lexical recognition: when multiple-character strings can form a meaningful word, the long word is activated first, rather than its individual components. These findings revealed two interesting facts: (1) native Chinese speakers take whole words as processing units, rather than morphemes/characters; and (2) these compound words are represented holistically by native Chinese speakers, and the meanings of compound words can be accessed directly as a whole, without activating and combining the meanings of each component morpheme.

1.2. L2-Chinese learners' segmentation difficulty

Previous studies consistently show that L2-Chinese learners exhibited difficulties or non-native-like patterns in Chinese word segmentation. For example, artificially inserting spaces between words did not significantly impact native Chinese speakers' reading (Bai et al., 2015; Li et al., 2011; Yan et al., 2010; Zang et al., 2013), while it could facilitate L2-Chinese learners' processing of Chinese. Bai et al. (2013) investigated how the visual form of input information (continuous character strings versus words with inter-word spaces) affects the learning and processing of new words for L2-Chinese learners. In the learning phase, participants were asked to learn new words through two types of text: word-spaced versus unspaced. In the test phase, participants read unspaced sentences that contained the newly learned words in the learning phase. The results revealed that in the learning phase, participants read the new words faster in the word-spaced condition than in the unspaced condition, and this space-facilitated effect was maintained in the subsequent test phase. The authors attributed this benefit to the stronger connections made between the constituent characters due to the introduction of inter-word spaces. Further evidence for the space-facilitated effect indexed by shorter reading times, fewer fixations and lower regression probabilities was reported (Cui, 2023; Shen et al., 2012). These findings suggest that L2-Chinese learners have problems in processing unspaced Chinese scripts, particularly in segmenting character strings. Inter-word spaces could serve as salient visual cues, helping to segment continuous character strings into meaning-conveying units and consequently ease sentence processing by reducing the uncertainty of what characters should be grouped together to form a word.

More relevant to the present study, Yao et al. (2023) found that, unlike native Chinese speakers who activate whole-word meanings during on-line reading, high-proficiency L2-Chinese learners rely more on the semantic meanings of characters/morphemes. For example, when processing sentences like (1a) and (1b), the plausibility status between the verb *踢打* ("kick")/*哀求* ("entreat") and the first constituent word *门* ("door") in *门卫* ("gatekeeper") affected L2-Chinese learners' reading significantly: compared to the implausible condition (*哀求门* "entreat door"), they spent longer time staring at *门卫* ("gatekeeper") in the plausible condition (*踢打门* "kick door"). This indicates that they initially take each incoming character as a processing unit, rather than as a compound word. This finding suggests that unlike native Chinese speakers who segment words from character strings and prioritize the activation of whole words over their component morphemes,

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

L2-Chinese learners face challenges in determining what characters should be grouped into words, even with a high level of Chinese proficiency.

Overall, previous studies highlight the critical role of word segmentation in Chinese reading for L2-Chinese learners, revealing their difficulties in segmenting words from continuous character strings without the aid of inter-word spaces. However, it remains unclear what causes L2-Chinese learners' difficulty in Chinese word segmentation. Multiple factors may affect L2-Chinese learners' word segmentation: their on-line processing strategies, insufficient mental representations of L2 words, L2 proficiency level, age of L2 acquisition, L1–L2 distance and so on. Unfortunately, the existing findings are comparatively limited, and it is difficult to get a clear picture of what factors play dominant roles in L2-Chinese learners' word segmentation difficulty. The critical research question of the present study is the following: For high-proficiency L2-Chinese learners who start to learn Chinese after their puberty and have inter-word spaces as word boundary demarcations in their native languages, what is/are the critical factors for their word segmentation difficulty – their on-line processing strategies and/or insufficient L2 lexical representations?

It should be noted that the procedure of Chinese word segmentation is modulated by the combined effects of processing strategy and lexicon representations. The former decides how readers segment strings of input and incrementally process sentences based on what units, while the latter determines how the semantic meanings of the units are stored and accessed. The natural coexistence of these two factors in language processing makes it difficult to tease apart their independent contributions. The present study aimed to investigate which of these two factors plays a dominant role in high-proficiency L2-Chinese learners' word segmentation difficulty.

1.3. Script difference and processing strategies

Differences in scripts are accompanied by differences in readers' cognitive mechanisms and processing strategies. Specifically, readers develop different strategies of word segmentation adapted to their languages (Li et al., 2022). As a typical alphabetic language, English has inter-word spaces to mark word boundaries. Letters that form a word are grouped by spaces explicitly, and readers process each chunk of letters (word) during reading. Previous studies found that (1) removing inter-word spaces would hamper alphabetic language processing (Perea & Acha, 2009; Rayner, 1998); (2) preferred viewing location is formed based on the inter-word spaces, which can be perceived by parafoveal vision and benefit word identification (O'Regan & Jacobs, 1992; Rayner, 1979); (3) when processing compound words, regardless of whether the lexical representation is familiar (such as "mountain lion") or novel (such as "cafeteria manager"), the processing unit is the chunk of letters that are grouped by spaces (Staub et al., 2007). These findings indicate that readers of alphabetic languages rely on inter-word spaces heavily during reading. They are very sensitive to the space cue and would take the chunk of letters grouped by spaces as processing units, no matter whether it is a word or a component morpheme of a compound word.

Logographic languages such as Chinese, on the other hand, do not have explicit word boundary demarcations. Readers of Chinese need to segment certain characters from a continuous character string to form a word to realize word identification. Inserting inter-word spaces would not affect native Chinese speakers' reading performance (Bai et al., 2015; Li et al., 2011; Yan et al., 2010; Zang et al., 2013), indicating that different from readers of alphabetic languages, Chinese readers do not use inter-word space as a cue for word boundary marking. Instead, they use a processing-based strategy to segment words. Li and Pollatsek (2020) constructed a Chinese reading model (CRM) to simulate how Chinese readers segment words without the aid of inter-word spaces. According to the CRM, all the characters within the perceptual span are activated, and these activated characters further facilitate the activation of the corresponding word nodes (the words that can be formed by these characters). Word identification is the result of the competition among these activated word nodes. Many linguistic factors affect this competition among word nodes: words with higher frequency, less strokes and higher predictability would reach the activation threshold faster and then got identified quicker. In addition, in the competition, if the constituent character is a free morpheme and thus has a corresponding word node (such as 门 "door" in 门卫 "gatekeeper"), this constituent word node (门 "door") would compete with the long-word node (门卫 "gatekeeper") at the word level. Since the long-word node receives bottom-up facilitation from two or more component character nodes and the constituent word node only receives bottom-up facilitation from one character node, the activation of the long-word node would be stronger than that of the constituent word node and would exceed the threshold first. Thus, the long words always have activation advantages over their constituent morphemes. As soon as a word has been identified, it is segmented from the character strings. In other words, word segmentation and identification are a unified procedure in Chinese reading. Readers of Chinese use this processing-based strategy to segment and identify words from continuous character strings repeatedly until sentence processing accomplished.

In general, distinct features of scripts result in script-specific strategies of word segmentation. For L2-Chinese learners whose native languages are alphabetic languages, there is no need to intentionally segment words in their native language reading. In other words, the processing-based word segmentation strategy is unique to L2 to them (Bertram et al., 2004). In addition, this processing strategy can only be implicitly acquired (if possible) through immersive language experience. Would L2-Chinese learners be able to implicitly acquire this unique-to-L2 word segmentation strategy when reached a high-level of Chinese proficiency? Is it possible that their difficulties with Chinese word segmentation were predominantly caused by their L1-biased processing strategy?

For high-proficiency L2-Chinese learners, it could be the case that they have not acquired the processing-based strategy. Specifically, they are sensitive to the inter-word spaces as cues for processing units in their L1, transfer this L1-specific processing strategy to L2 processing and take each character (separated by little spaces and thus presented as a salient orthographic unit) as a processing unit and thus integrate morphemes instead of words with prior context. From this perspective, within the framework of CRM, there would be no competition between the component word nodes and the whole word nodes at the word level for L2-Chinese learners, since the whole words would not be initially chunked and treated as processing units and the constituent morphemes would be the first to be activated and integrated with prior context. Consequently, no "long word advantage" would be expected on L2-Chinese learners, and the plausibility status of the first constituents with prior verbs would inevitably affect the processing of whole compound words.

1.4. Developing mental representations in L2

In addition to the processing strategy, L2-Chinese learners' mental representations of Chinese words may also affect their Chinese word segmentation. Existing findings posit that Chinese compound

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

words (especially those with high frequency, low semantic transparency and short word length) are represented and processed in a holistic way (Cui et al., 2020; Hyönä et al., 2020; Yao et al., 2021; Cf. Yu et al., 2020). Very few previous studies examined L2-Chinese learners' representation and processing of Chinese compound words. Using a repetition priming task, Gao et al. (2022) found that with increased Chinese proficiency, L2-Chinese learners' representations of compound words developed from decompositional to holistic. For intermediate L2-Chinese learners, both the whole-word prime (天气–天气 "weather") and the morpheme prime (天 "sky"-天气 "weather") significantly primed target words' activation when compared to the symbol prime (****-天气 "weather"), with no significant difference between the whole-word and morpheme priming conditions. For advanced L2-Chinese learners, both the whole-word and morpheme primes evoked shorter reaction times on target words than the symbol prime, and the priming effect was bigger for the whole-word prime than for the morpheme prime. For native Chinese speakers, the whole-word prime evoked a bigger priming effect than the morpheme and symbol primes, with no significant difference between the latter two. These findings indicated that even though advanced L2-Chinese learners rely more on whole-word representations instead of morphemes than intermediate L2-Chinese learners, they still process words differently from native speakers: the representations of both whole words and morphemes were activated for advanced L2-Chinese learners, but only whole-word representations were activated for native speakers. Similarly, using a semantic priming test, Hong and Feng (2010) found that compared to native Chinese speakers who demonstrated a holistic fashion of compound word representation, L2-Chinese learners were more easily affected by morpheme-level information. These findings indicated that even with advanced L2 proficiency levels, L2-Chinese learners may have non-native-like representations of Chinese words.

These observed non-native-like patterns of L2-Chinese learners can be explained by the ontogenesis model, which was recently proposed by Bordag and colleagues (Bordag et al., 2021) to analyze the development of L2 word representation with an emphasis on the inherent ambiguity in L2 lexicons compared to L1 lexicons. According to the ontogenesis model, the acquisition of words is not an all-or-nothing process, but rather an ontogenetic process. Specifically, in the mental lexicon, all lexical representations dynamically move along ontogenetic curves. There is an optimum in each dimension (orthography, phonology and semantic) for each word, and the representation is properly encoded and fully specified at this optimum. While most representations of L1 words can reach their optimum and stay around it, most L2 representations do not reach their optimum – they stay fuzzy and move along the ontogenetic curve dynamically with the development of L2 proficiency. In other words, the central property of L2 lexicon is fuzziness, referring to the inexact or ambiguous encoding of different components of lexical representations. L2 representation fuzziness can be caused by several factors, including script-specific mappings from orthography to semantics, problems with L2 phonological encoding, changes in neural plasticity and so on. In general, the ontogenesis model argues that compared to L1 representations, L2 representations are fuzzy, instable and incomplete, but can gradually approach the optimum (which is the same for L1 and L2 speakers of the same dialect) with increasing L2 proficiency.

Previous studies and the ontogenesis model all indicated that L2-Chinese learners' representations of Chinese words are developing – move along the ontogenetic curves. But it is difficult to identify what is the status of lexicon representations when they reach the advanced level of Chinese proficiency. Can high-proficiency L2-Chinese learners ultimately acquire native-like representations (in a holistic fashion) for Chinese words (have reached the optimum) or still have developing representations (in a decompositional fashion, stay fuzzy but gradually approaching the optimum)?

L2-Chinese learners' difficulties in word segmentation could be caused by their developing mental representations of Chinese compound words. Different from native Chinese speakers who represent and process Chinese compound words in a holistic way, L2-Chinese learners may represent compound words decompositionally and need to activate the meanings of each constituent to access the meaning of the whole compound word. This way, it is reasonable to observe the plausibility effect of the first constituents on the processing of whole compound words.

1.5. Hypotheses of the present study

The present study aimed to address the question of what is/are the dominant factors of high-proficiency L2-Chinese learners' difficulty in Chinese word segmentation: L1-biased processing strategy or developing representations of L2 words, or both. Specifically, we investigated whether inserting spaces between words or characters would affect high-proficiency L2-Chinese learners' processing of high-frequency two-character compound words, with a focus on their sensitivity to the plausibility status of the first constituents of compound words with preceding verbs.

If high-proficiency L2-Chinese learners have acquired the unique-to-L2 processing strategy and native-like mental representations of Chinese compound words, native-like behaviors would be expected. Specifically, in both character-spaced and word-spaced conditions, no plausibility effect would be expected. This is because just like native speakers, L2-Chinese learners have holistic representations of high-frequency Chinese compound words and read Chinese text with a processing-based strategy. The two characters of target compound words would be perceived by one fixation in either character-spaced or word-spaced conditions. Using a processing-based strategy, both the constituent morpheme node and the whole word node would be activated and compete with each other, and the whole word node would finally win the competition and get identified and segmented. Since these compound words are represented holistically, the semantic meanings of their constituent morphemes would not be sufficiently activated to affect the accessing of the entire words. Consequently, no plausibility effect would be expected in either character- or word-spaced condition.

If high-proficiency L2-Chinese learners' word segmentation difficulty was dominantly caused by their L1-biased processing strategy but not by developing mental representations, inserting spaces between words would facilitate their segmentation of compound words and then evoke holistic processing of compound words. In this case, no plausibility effect would be expected in the word-spaced condition. In the character-spaced condition, on the other hand, inter-character spaces would be taken as cues to segmenting processing units following their L1 processing habits, and L2-Chinese learners would exhibit the plausibility effect of the first constituents with prior verbs on compound words even with holistic compound word representations, since they identify and integrate each individual character that is separated by spaces as processing units incrementally.

If high-proficiency L2-Chinese learners' word segmentation difficulty was dominantly caused by their developing mental representations but not by L1-biased processing strategy, the plausibility effect of the first constituent morpheme on compound word processing

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

Table 1. Hypotheses: Whether or not to expect a plausibility effect in a certain condition among L2-Chinese learners

		Holistic representation	Decompositional representation
L1-biased processing strategy (space-sensitive)	Character-spaced	Yes	Yes
	Word-spaced	No	Yes
L2-adaptive processing strategy (processing-based)	Character-spaced	No	Yes
	Word-spaced	No	Yes

would be expected in both character-spaced and word-spaced conditions. This is because L2-Chinese learners' mental representations of Chinese compound words are decompositional even for high-frequency words. No matter how the text was presented, L2-Chinese learners need to access the meanings of each constituent to activate the whole word's meaning. The semantic plausibility effect of the first constituent is inevitable in this process.

If high-proficiency L2-Chinese learners' word segmentation difficulty was caused by both their L1-biased processing strategy and developing mental representations, the plausibility effect of the first constituent morpheme would be expected in both character-spaced and word-spaced conditions. No matter what the processing unit is, compound words need to be accessed through the activation of all the constituent morphemes. In other words, the semantic information of the first constituent has to be activated and play a role in the processing of compound words. Consequently, the plausibility effect of the first constituent morpheme is inevitable.

To sum up, as illustrated in Table 1, for high-proficiency L2-Chinese learners, there are three possibilities of the processing patterns in the character-spaced and word-spaced conditions based on the processing strategy hypothesis and mental representation hypothesis. The reason for setting the word-spaced condition was to explore how L2-Chinese learners represent Chinese compound words: holistically or decompositionally. With explicit visual cues of word boundary (inter-word spaces), the plausibility effect of the first constituent on the processing of compound words would be taken as an index of the decompositional fashion of representations. In addition, a direct comparison between word- and character-spaced conditions would reveal the predominant reason for L2-Chinese learners' word segmentation difficulty.

2. Method

2.1. Participants

Forty L2-Chinese learners (27 female, aged 17 to 35 years, M_{age} = 23.2) and 40 native Chinese speakers (29 female, aged 19 to 30 years, M_{age} = 22.5) participated in this study. L2 participants were students in a Chinese-learning major. They all started learning Chinese after puberty. All L2 participants had passed the HSK5/6 (Hanyu Shuiping Kaoshi-advanced level, the standard Chinese language proficiency test for non-native speakers administered by the Ministry of Education of the People's Republic of China), indicating that they were high-proficiency L2-Chinese learners. Their Chinese-learning years within mainland China ranged from two to ten years ($M_{Chinese\text{-}learning\ years}$ = 4.78 years). These L2 participants have different native languages (Korean, Spanish, Russian, etc.), but all of these languages have inter-word spaces explicitly marking word boundaries. The number of L2 participants was decided based on the results of the power analysis on a pilot study of 20 L2-Chinese participants with the exact same materials and procedure. The total reading times for compound words in the character-spaced condition of these 20 participants were analyzed using a linear mixed-effects (LME) model focusing on the effect of plausibility (Yao et al., 2023). The powerSim function (Judd et al., 2012) was used to test the statistical power of this model. Results showed that the power would reach .90 with 32 participants. Following Yao et al. (2023), 40 high-proficiency L2-Chinese learners were recruited as L2 participants.

2.2. Materials

Thirty-one sentence frames with high-frequency two-character compound words as targets were constructed. Examples are illustrated in Table 2. For each sentence frame, two types of sentences were developed by manipulating the verb prior to the target: the plausible condition as in (2), in which the preceding verb is plausible with both the compound word and its first constituent component; and the implausible condition as in (3), in which the preceding verb is plausible with only the compound word but not with its first constituent component.

The mean frequency of the pre-target verbs was 42 per million in the plausible condition and 41 per million in the implausible condition, with no significant difference between these two types of verbs, $p > .5$. The mean frequency of target words was 81.28 per million, with a high semantic transparency score (mean = 4.43) in a five-point-scale test conducted on 20 native Mandarin speakers (who did not participate in the eye-tracking experiment). All the materials (including all the characters, words and sentences) were familiar and acceptable to high-proficiency L2-Chinese learners through a familiarity test and an understandability and acceptability test, which were conducted on 20 high-proficiency L2-Chinese learners (all passed the HSK5/6 and did not participate in the main eye-tracking experiment). The plausibility of the target words and their first characters with the preceding verbs was assessed by

Table 2. Material examples

	Plausibility condition	Spacing condition	Examples
(2a)	Plausible	Character-spaced	弟 弟 知 道 爱 惜 **书 包** 让 人 觉 得 他 懂 事 了。
(2b)	Plausible	Word-spaced	弟弟 知道 爱惜 **书包** 让 人 觉得 他 懂 事 了。 Brother know cherish **bag** let people feel him sensible LE "My younger brother knows to take care of his bag to let people feel that he is responsible."
(3a)	Implausible	Character-spaced	弟 弟 知 道 清 洗 **书 包** 让 人 觉 得 他 懂 事 了。
(3b)	Implausible	Word-spaced	弟弟 知道 清洗 **书包** 让 人 觉得 他 懂 事 了。 Brother know wash **bag** let people feel him sensible LE "My younger brother knows to wash his bag to let people feel that he is responsible."

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

another 20 high-proficiency L2-Chinese learners (who did not participate in the familiarity test and the understandability and acceptability test, nor the main eye-tracking experiment). These L2 participants were asked to rate the plausibility of the verb-noun/character phrases on a five-point scale (1 = very implausible; 5 = very plausible). The mean plausibility ratings of the target compound words in the plausible and implausible conditions were 4.92 and 4.93, respectively, and the mean plausibility ratings of the first constituent characters in the plausible and implausible conditions were 4.89 and 1.12, respectively. Plausibility score of the first characters in the implausible condition was significantly lower than that in the plausible condition, $p < .001$. There was no significant difference in the plausibility score of the target compound words between these two conditions.

Two spacing conditions were developed for each sentence: character-spaced condition (a single space was inserted between every two neighboring characters), as in (2a) and (3a), and word-spaced condition (a single space was inserted between every two neighboring words), as in (2b) and (3b). To test the validity of the word segmentation in the word-spaced condition, 20 native Chinese speakers (who did not participate in the eye-tracking experiment) were required to indicate the word boundaries in unspaced sentences. The agreement reached 94.6% (ranging from 85% to 100%) among participants.

In total, there were four conditions of sentences: character-spaced & plausible, character-spaced & implausible, word-spaced & plausible and word-spaced & implausible. There were 31 items in each condition. These experimental sentences were divided into four lists, each list containing 31 experimental sentences as well as 31 filler sentences. The plausibility condition was a within-participant variable, while the spacing style was a between-participant variable. Specifically, list 1 contained 16 character-spaced & plausible sentences and 15 character-spaced & implausible sentences, and list 2 contained 16 character-spaced & implausible sentences and 15 character-spaced & plausible sentences. The same manipulations were used for list 3 and list 4 for word-spaced sentences. This is to avoid participants' extra attention paid to the spacing styles caused by the shift between word-spaced and character-spaced conditions in one list. The presentation order was randomized and counterbalanced. Each of the four conditions of the experimental sentence sets appeared once across the four lists. None of the fillers involved implausibility.

2.3. Apparatus

Participants' eye movements were recorded using an EyeLink 1000 eye tracker with a sampling rate of 1000 Hz. Each sentence was presented in one line in the middle of the screen. Participants were seated 60 cm away from the monitor. Following a three-point horizontal calibration and validation, the maximum gaze-position error was less than 0.5°. Eye movements were recorded from the right eye, but viewing was binocular.

2.4. Procedure

Participants were tested individually. The experiment started with a brief instruction and a standard horizontal three-point grid calibration and validation. Then, five practice trials were run to ensure that participants understood the task and were familiar with the apparatus. Critical experimental trials were run after the practice trials. Participants were required to read sentences on the screen silently. At the beginning of each trial, a drift check was conducted. Each sentence appeared after participants fixated on a character-sized box at the location of the first character of each sentence. After finishing reading one sentence, participants were asked to press a button so that the original sentence disappeared and was replaced by a meaning-related question, to which participants then responded by pressing a button. The entire experiment lasted around 30 mins.

2.5. Results

The critical region of interest (ROI) is the target word region. To match the length of ROI between the character-spaced and word-spaced conditions, each ROI includes the two constituent characters of the compound word and the space between (in the character-spaced condition) or after (in the word-spaced condition) the two characters (as illustrated by the yellow markings in Table 2), so that the physical lengths and linguistic information in each ROI in these two conditions were exactly the same. Three eye movement indices were analyzed. Gaze duration (GD) is the sum of the fixation durations before the eyes first move out of a region, reflecting the early stage of processing. Go-past time (GP) is the duration from first entering a region to first going past it forwardly, reflecting the processing of error detection and reanalysis. Total reading time (TT) is the sum of the durations of all fixations in a region, reflecting the late stage of processing and general processing difficulties (Rayner, 1998, 2009; Rayner et al., 1998). Following the hypotheses raised in Table 1, we predict that for L2 participants, the possible plausibility effect may be observed on both early (GD) and late (TT) measures and especially on the measure that reflects the reanalysis procedure (GP).

Three L2 participants were excluded from the final analysis due to their low accuracy (<80%) on the comprehension questions. Thus, the data of 37 L2 participants and 40 native participants were used in the final analysis. The mean accuracy of these 37 L2 participants was 86%, ranging from 80% to 98%. The mean accuracy of the 40 native participants was 96%, ranging from 80% to 100%. Trials in which participants blinked once or more in the ROI were excluded from the analysis. This approach resulted in a loss of 2.9% of the trials. Fixations shorter than 80 ms or longer than 1000 ms (approximately 2.5% of all fixations) were removed.

LME models (Baayen et al., 2008) were constructed for GD, GP and TT. Each model included plausibility conditions (plausible versus implausible, −.5 versus .5), text spacing conditions (character-spaced versus word-spaced, −.5 versus .5) and speaker groups (native participants versus L2 participants, .5 versus −.5) as fixed effects, specifying participants and items as crossed random effects, including intercepts and slopes. To directly test the possible interactions between plausibility conditions and speaker groups in different spacing conditions, following Schad et al. (2020), we used the following customized contrasts: (a) a contrast that tests the main effect of text spacing conditions (character-spaced versus word-spaced); (b) a contrast that tests the main effect of speaker groups (native participants versus L2 participants); (c) a contrast that tests the plausibility effect in the character-spaced trials; (d) a contrast that tests the plausibility effect in the word-spaced trials; (e) a contrast that tests the interaction between spacing and speaker groups; (f) a contrast that test the interaction between plausibility effect and speaker groups in the character-spaced trials; (g) a contrast that test the interaction between plausibility effect and speaker groups in the word-spaced trials.

The statistical procedure was conducted using *lmer* function (Bates et al., 2011) and *lmerTest* function (Kuznetsova et al., 2017) in R (version 3.5.1; R Core Team, 2018). Simple effects were tested

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

Table 3. The means and SEs (in parentheses) of each eye movement index in the ROI of compound words

	Native Chinese speakers		L2-Chinese learners	
	Character-spaced	Word-spaced	Character-spaced	Word-spaced
Gaze duration				
Implausible	258(28.8)	274(27.2)	554(26.3)	610(26.4)
Plausible	263(29.1)	260(25.9)	716(26.0)	604(25.3)
Go-past time				
Implausible	375(57.5)	375(55.2)	778(53.7)	748(53.5)
Plausible	396(58.6)	388(52.4)	1066(53.3)	812(51.5)
Total reading time				
Implausible	375(74.2)	387(73.6)	1107(76.4)	1146(75.1)
Plausible	401(74.9)	420(72.5)	1401(75.4)	1192(73.6)

Note. The data in boldface are those with significant differences between plausible and implausible conditions.

using *emmeans* function (Lenth, 2018). The means and SEs for each eye movement index in ROI are summarized in Table 3. The log-transformed data of fixation measures yield the same patterns of statistical significance as the analysis based on the raw data. In the interest of transparency of effect sizes, we report the analysis of the untransformed data here, with a focus on *beta*, *SE*, *t* and *p values*, which are summarized in Table 4. All the data and analysis code can be found at https://osf.io/4mjpy/.

The main effect of text spacing was not significant on any of the three indexes (GD, GP and TT). The main effect of speaker group was significant on all three indexes, with L2 participants exhibiting significantly longer reading times than native participants. The plausibility effect was significant on GD, GP and TT in the character-spaced condition, with longer durations in the plausible condition compared to the implausible condition. On the other hand, no significant plausibility effect was observed in the word-spaced condition on any index. The interaction between spacing and speaker groups was not significant on any index. The interaction between plausibility and speaker groups was significant in the character-spaced condition, but not in the word-spaced condition, and this pattern was consistently reflected in all three indexes.

Further simple effects analysis revealed that, for native speakers, the plausibility effect was not significant in either character-spaced or word-spaced condition on any index, while for L2 participants, the plausibility effect was significant in the character-spaced condition but not in the word-spaced condition, and this pattern was consistently reflected in all three indexes. The results of the simple effects analysis are summarized in Table 5.

Bayes factor analysis was further conducted on log-transformed GD, GP and TT to determine the strength of evidence for the null/significant plausibility effect in the character-/word-spaced conditions for native participants and L2 participants separately. Bayes factors were computed using the *lmBF* function from the *BayesFactor* package in R (Morey et al., 2015). In all analyses, we assumed the default Cauchy prior for effect size (Abbott & Staub, 2015; Staub & Goddard, 2019). In each condition, Bayes factors were calculated to compare a model that included subject and item intercepts and the plausibility effect to a null model that included only subject and item intercepts. The results of the Bayes factor analysis are summarized in supplementary materials in Table S1.

Table 4. Summary of LME analysis results in the ROI of compound words

	Estimate	SE	t	p
Gaze duration				
(Intercept)	442.42	10.71	41.30	<.001
Space	−10.91	21.42	−0.51	.61
Speaker	−357.26	20.40	−17.51	<.001
Space_character: Plausibility	−84.10	25.37	−3.32	<.001
Space_word: Plausibility	9.75	24.03	0.41	.69
Space × Speaker	34.13	40.80	0.84	.41
Space_character: Plausibility×Speaker	157.08	47.23	3.33	<.001
Space_word: Plausibility×Speaker	8.23	44.25	0.19	.85
Go-past time				
(Intercept)	617.19	22.31	27.67	<.001
Space	−73.00	44.62	−1.64	.11
Speaker	−466.99	44.39	−10.52	<.001
Space_character: Plausibility	−154.55	46.03	−3.36	<.001
Space_word: Plausibility	−38.76	43.12	−0.90	.37
Space × Speaker	137.63	88.77	1.55	.13
Space_character: Plausibility×Speaker	266.85	91.13	2.93	<.001
Space_word: Plausibility×Speaker	50.08	85.20	0.59	.56
Total reading time				
(Intercept)	803.80	34.25	23.47	<.001
Space	−34.66	68.49	−0.51	.61
Speaker	−815.58	66.92	−12.19	<.001
Space_character: Plausibility	−159.65	47.11	−3.39	<.001
Space_word: Plausibility	−39.27	45.84	−0.86	.39
Space × Speaker	100.22	133.83	0.75	.46
Space_character: Plausibility×Speaker	267.31	83.37	3.21	<.001
Space_word: Plausibility×Speaker	13.12	80.44	0.16	.87

Bayes analysis favored the null model for all three indexes in both character-spaced and word-spaced conditions for native participants, which delivered evidence against the plausibility effect. For L2 participants, in the character-spaced condition, the Bayes factors for GD, GP and TT were bigger than 10, offering supportive evidence for the observed plausibility effect. In the word-spaced condition, however, the Bayes analysis for all three indexes favored the null model.

2.5.1. The first character of compound words as ROI

In addition to the compound words, analyses were conducted in the ROI of the first constituent character of the compound words to further explore whether and how the presentation style affected participants' processing of the same characters when embedded in plausible versus implausible conditions.

Similar to the ROI of compound words, LME models including plausibility conditions (plausible versus implausible, −.5 versus .5),

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

Table 5. Summary of the plausibility effects in the ROI of compound words

Speaker group	Text spacing	Estimate	SE	t	p
Gaze duration					
Native participants	Character-spaced	−5.56	37.10	−0.15	.88
	Word-spaced	13.86	33.40	0.42	.68
L2 participants	Character-spaced	−162.64	32.30	−5.04	<.001
	Word-spaced	5.63	32.10	0.18	.86
Go-past time					
Native participants	Character-spaced	−21.10	69.40	−0.31	.76
	Word-spaced	−13.70	62.20	−0.22	.83
L2 participants	Character-spaced	−288.00	60.30	−4.78	<.001
	Word-spaced	−63.80	59.40	−1.07	.28
Total reading time					
Native participants	Character-spaced	−26.00	64.30	−0.40	.69
	Word-spaced	−32.70	60.80	−0.54	.59
L2 participants	Character-spaced	−293.30	61.70	−4.76	<.001
	Word-spaced	−45.80	61.30	−0.75	.46

text spacing conditions (character-spaced versus word-spaced, −.5 versus .5) and speaker groups (native participants versus L2 participants, .5 versus -.5) as fixed effects were constructed for GD, GP and TT. The means and SEs for each eye movement index in the ROI of the first character of compound words are summarized in Supplementary Table S2, and the results of the LME analyses are summarized in Supplementary Table S3.

The main effect of text spacing was significant only on GD, but not on the other two indexes. The main effect of speaker group was significant on all three indexes, with L2 participants exhibiting significantly longer reading times than native participants. In the character-spaced condition, the plausibility effect was marginally significant only on GD, with longer GD in the implausible condition compared to the plausible condition. No significant plausibility effect was observed in the word-spaced condition on any index. The interaction between spacing and speaker group was only marginally significant on GD. The interaction between plausibility and speaker group was marginally significant in the character-spaced condition on GD, but not in the word-spaced condition on any of the three indexes. Further simple effects analysis (results of which are summarized in Supplementary Table S4) revealed that the marginal interaction between plausibility and speaker group in the character-spaced condition was caused by the contrast between native and L2 participants: L2 participants exhibited significant plausibility effect in the character-spaced condition, with longer GD on the first character in the implausible condition than in the plausible condition, while the difference was not significant for native participants. This significant plausibility effect in the character-spaced condition for L2 participants was supported by the results of Bayes factor analysis, which are summarized in Supplementary Table S5.

2.5.2. Chinese-learning years as a covariate for L2 participants

Although all L2 participants in the present study had passed HSK5/6 and reached a high level of Chinese proficiency, they had varied Chinese-learning durations within mainland China (ranging from 2 to 10 years, mean = 4.78 years). Further analyses were conducted to explore whether and how prolonged Chinese exposure times affected L2 participants' on-line reading of compound words. Specifically, we aimed to address the question of whether high-proficiency L2 participants' processing strategy and mental representations of Chinese compound words would change with extended Chinese-learning years.

In the ROI of compound words, LME models were conducted on L2 participants, containing plausibility conditions (plausible versus implausible, −.5 versus .5) and text spacing conditions (character-spaced versus word-spaced, −.5 versus .5) as fixed effects, specifying participants and items as crossed random effects, including intercepts and slopes, and Chinese-learning years within mainland Chinese as a covariate. To directly test the plausibility effect in different spacing conditions and the possible effects of Chinese-learning years, we used the following four customized contrasts: (a) a contrast that tests the main effect of text spacing; (b) a contrast that tests the main effect of Chinese-learning years; (c) a contrast that tests the plausibility effect in the character-spaced trials; and (d) a contrast that tests the plausibility effect in the word-spaced trials. The results are summarized in Supplementary Table S6.

The main effects of Chinese-learning years were significant on all three indexes, with prolonged Chinese-learning years facilitating L2 processing. In addition, L2 participants were sensitive to the spacing conditions: the plausibility effect was significant only in the character-spaced condition, but not in the word-spaced condition. This pattern was consistently observed on all the three indexes.

Further analyses were conducted to explore how plausibility and Chinese-learning years mutually affect L2 participants' on-line reading of compound words in character- and word-spaced conditions separately. Results are summarized in Supplementary Table S7 and visualized in Figure 1. The findings showed that L2 participants exhibited consistent patterns on the three indexes in the word-spaced condition: only the main effect of Chinese-learning years was significant, with increased Chinese-learning years speeding up L2 participants' processing of compound words at both early (GD) and late stages (GP and TT). Neither the main effect of plausibility nor the interaction between plausibility and Chinese-learning years was significant.

In the character-spaced condition, on the other hand, different patterns were illustrated among the three indexes. No effect reached significance on GD. On GP, however, the main effect of plausibility was significant, and the interaction between plausibility and Chinese-learning years was marginally significant. Further analysis indicated that prolonged Chinese-learning years significantly decreased GP in the plausible condition ($t = -2.48$, $p < .05$), but had a null effect in the implausible condition ($t = 0.47$, $p = .64$). As for TT, the main effects of plausibility and Chinese-learning years were significant, but not their interaction.

3. Discussion

The present study conducted an eye-tracking experiment to explore the dominant reasons for high-proficiency L2-Chinese learners' difficulties in Chinese word segmentation: L1-biased processing strategy versus developing representations of Chinese compound words. Specific attentions were focused on whether or not to observe the plausibility effect in the character- or word-spaced condition. In the subsequent part, critical findings will be discussed systematically. Each discovery will be critically evaluated in relation to the research hypotheses, previous studies, implications and potential limits.

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

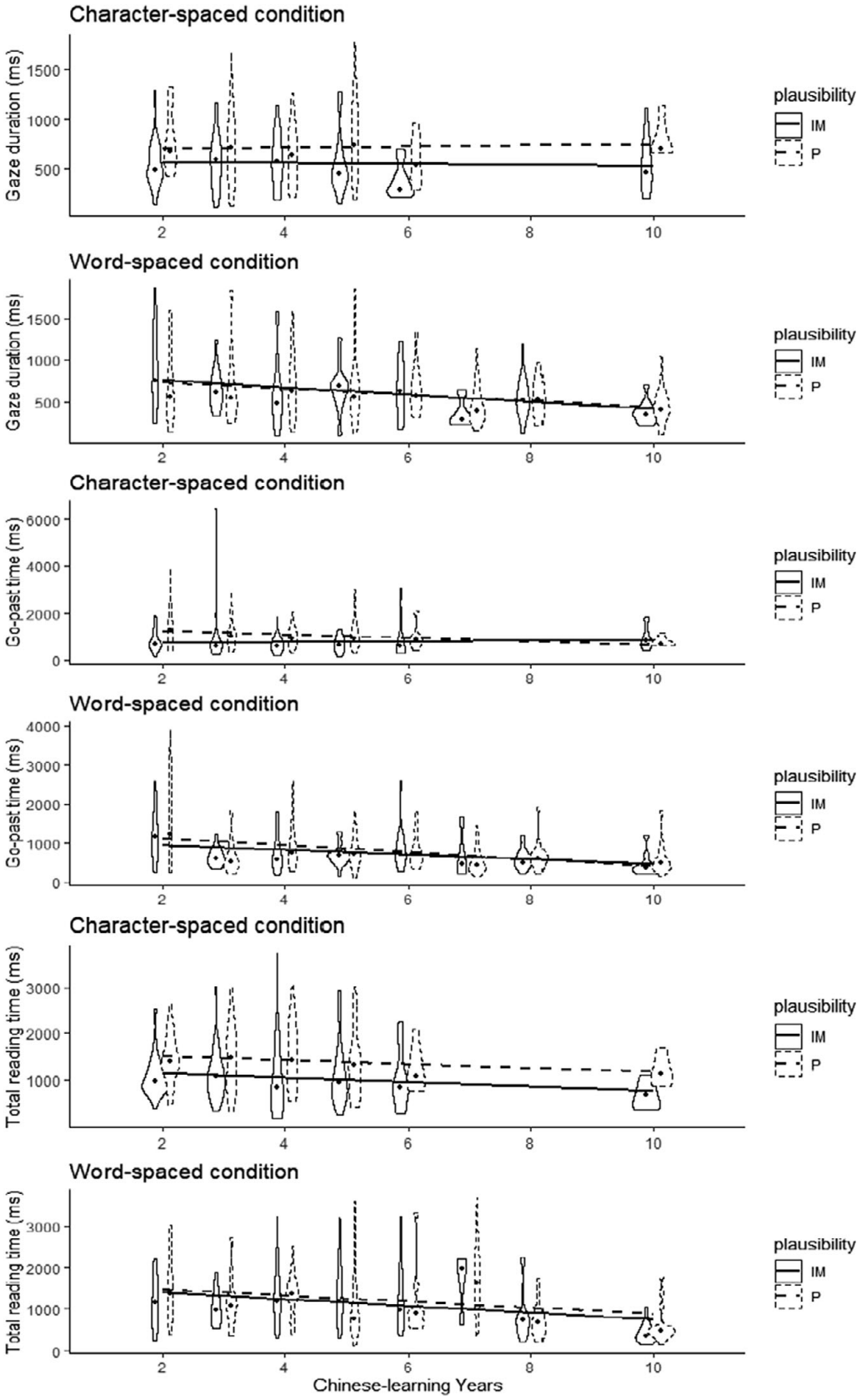


Figure 1. Violin plots illustrating how plausibility and Chinese-learning years mutually affect L2 participants' processing of compound words.

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

3.1. L2 participants' non-native-like behaviors

Critical to the research aim of the present study, some non-native-like patterns were observed among high-proficiency L2 participants: the plausibility effect was found in the region of target compound words in the character-spaced condition, with longer reading times (including all three indexes) in the plausible condition than in the implausible condition. This plausibility effect was diminished on all measures in the word-spaced condition. Native participants, on the other hand, did not exhibit a significant plausibility effect on any measures. In addition, in the region of the first constituent character of compound words, the significant plausibility effect was observed only on L2 participants in the character-spaced condition on GD, with longer GD in the implausible condition than in the plausible condition.

The significant plausibility effect in the character-spaced condition, and the opposite patterns of this plausibility effect in the ROI of compound words and the ROI of the first character of compound words, indicated that L2 participants were sensitive to the semantic information of the first constituent morphemes of compound words in the character-spaced condition. Different from native Chinese speakers who processed and activated compound words as a whole, L2-Chinese learners took each character/morpheme as a processing unit and integrated it incrementally with the preceding verbs. For example, during on-line reading sentences in the character-spaced condition such as (2a) & (3a), L2 participants originally treated 书 ("book") as an independent word and integrated it with the prior verb 爱惜 ("cherish")/清洗 ("wash"). At this point, the combination 爱惜书 ("cherish book") in (2a) is plausible, while the combination 清洗书("wash book") in (3a) is not, leading to longer GD in the ROI of the first character in the implausible condition than in the plausible condition. As the reading continued, L2 participants realized their original analysis was not correct when reaching 包 ("bag"), which cannot be an independent word in this case and needs to combine with 书 ("book") to form a compound word 书包 ("bag"). At this point, they need to give up their original analysis and reanalyze and resegment the character strings into 爱惜 ("cherish")/清洗 ("wash") -书包 ("bag"). It is reasonable to observe longer reading times in the ROI of compound words in the plausible condition than in the implausible condition, since it is easier to give up an implausible original analysis than a plausible one. These observations were similar to the findings from the study of Yao et al. (2023), where sentences were presented in a natural way without any artificial insertions of spaces.

More importantly, for L2 participants, this plausibility effect was significant only in the character-spaced condition, but not in the word-spaced condition. This difference between character- versus word-spaced conditions indicated that L2-Chinese learners already acquired native-like holistic representations of familiar high-frequency two-character compound words and that their word segmentation difficulty was predominantly caused by their L1-biased processing strategy. Specifically, in the word-spaced condition where word boundaries were explicitly marked by inter-word spaces, these artificially inserted spaces between words were more salient visual cues than the little spaces naturally existing between neighboring characters. By using their L1-biased space-sensitive strategy, L2-Chinese learners processed sentences based on the chunked characters (words) instead of individual characters. The null plausibility effect in this condition indicated that high-proficiency L2 participants have already acquired a holistic fashion of compound words representations: the meanings of compound words were accessed directly and holistically without activating the meanings of each constituent. For example, in reading sentences such as (2b) & (3b), the presegmented compound word 书包 ("bag") was activated holistically without strong intervention from the constituent characters 书 ("book") and 包 ("bag"). Since both the combination 爱惜-书包 ("cherish bag") in (2b) and the combination 清洗-书包 ("wash bag") in (3b) are grammatical and plausible, the processing times of 书包 ("bag") were comparable between these two conditions.

The developing representation possibility predicts that even with high proficiency, L2-Chinese learners may still have non-native-like and decompositional representations of Chinese compound words and thus would expect to observe the plausibility effect in both the character- and word-spaced conditions. However, this is not what we observed. The null plausibility effect in both the ROI of compound words and the ROI of the first character in the word-spaced condition indicated that just like native speakers, high-proficiency L2 participants have already acquired a holistic fashion of compound word representations: the meanings of compound words were accessed holistically without activating the meanings of each constituent when characters belonging to a word were grouped together primarily. According to the ontogenesis model, with developing proficiency, lexical representations gradually approximate an optimum (for each dimension) along ontogenetic curves. The native-like holistic processing of compound words in the word-spaced condition indicated that the representations of these words might have already reached the optimum (at least for the orthography and semantic dimensions, and their interface) for L2 participants: they were properly encoded and can be automatically activated holistically, and the mappings from orthography to semantics were strong and solid. The target compound words in the present study are high-frequency commonly used two-character words in daily life. It is reasonable for these high-proficiency L2-Chinese learners who have years of immersive Chinese-speaking living experience to develop native-like representations for these words.

With native-like holistic representations of Chinese compound words, L2 participants still exhibited some non-native-like patterns, leading to the speculation that L2-Chinese learners' word segmentation difficulty was predominantly caused by their L1-biased processing strategy. Specifically, due to their intensive experience of native languages, during on-line reading, they preferred to rely on inter-word spaces as cues for word boundary demarcation (Yao, 2011). In other words, they would take any unit that is separated by spaces as a processing unit. In the word-spaced condition, spaces were inserted between every two neighboring words. This primary segmentation of words led participants to process each word (chunk of characters grouped by spaces) as a whole. With native-like holistic compound word representations, there was no plausibility effect. However, in the character-spaced condition, spaces were inserted between every two neighboring characters. Due to the visual salience of characters as orthographic units, L2 participants preferred to process one character a time incrementally. In this case, the plausibility effect was inevitable. In general, the plausibility effect caused by the temporarily ambiguous processing was observed in the character-spaced condition but not in the word-spaced condition, indicating that L2-Chinese learners' word segmentation difficulty was predominantly caused by their L1-biased processing strategy.

Some caveats should be noted: although the observations in the present study indicated that high-proficiency L2-Chinese learners have already acquired holistic representations of high-frequency and transparent Chinese compound words and their word segmentation difficulty was predominantly caused by their L1-biased

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

processing strategy, a strong argument cannot be made that L2-Chinese learners' word segmentation was not affected by their L2 lexicon representations. As mentioned before, word segmentation is modulated by the combined effects of processing strategy and lexicon representations. The dominant role of L2-Chinese learners' L1-biased processing strategy in their word segmentation difficulty can only be admitted with the prerequisite that they have already acquired the native-like holistic representations of Chinese compound words, which was reflected by the null plausibility effect in the word-spaced condition in the present study. The independent contributions of processing strategy and L2 lexicon representations to L2-Chinese learners' word segmentation difficulties should be further explored.

Consistent with previous studies (Yang et al., 2012), the plausibility effect of the first constituent character on the processing of compound words was not observed among native Chinese speakers. In either character-spaced condition or word-spaced condition, the processing of compound words was not affected by the plausibility status of the first constituent character with the preceding verbs. This finding confirmed that native Chinese speakers process compound words holistically during on-line reading, and inserting spaces between characters or words did not affect compound words' holistic representation and processing.

To sum up, the present study found that different from native Chinese speakers whose on-line reading of compound words was not affected by the plausibility status of the first constituents with preceding verbs in either character-spaced condition or word-spaced condition, high-proficiency L2-Chinese learners were sensitive to how the text was presented: in the character-spaced condition, L2-Chinese learners exhibited a significant plausibility effect on compound words, while this plausibility effect vanished in the word-spaced condition. These patterns of L2-Chinese learners suggest that they have already acquired the holistic representations of high-frequency Chinese compound words and that their difficulty in word segmentation is mainly caused by their L1-biased language processing strategy.

3.2. The effects of L2 participants' Chinese-learning years within mainland China

Further analyses were conducted to explore the question of whether high-proficiency L2 participants' processing strategy and mental representations of Chinese compound words would change with extended Chinese exposure durations. Some interesting findings were revealed.

First of all, when including Chinese-learning years as a covariate, LME analyses revealed that Chinese-learning years had a significant main effect on L2 participants' processing of compound words: prolonged Chinese-learning years significantly decreased their processing times at both early (GD) and late stages (GP and TT). In addition, with the significant main effect of Chinese-learning years, L2 participants still exhibited a significant plausibility effect in the character-spaced condition. These findings revealed that regardless of whether or not including Chinese-learning years as a covariate, the plausibility effect in the character-spaced condition remained, indicating it is a salient and stable effect.

Furthermore, simple effects analyses revealed that in the word-spaced condition, the main effect of Chinese-learning years was significant on all three indexes, while the main effect of plausibility was not significant, nor their interaction. These findings indicated that high-proficiency L2 participants have already acquired holistic representations of compound words (as indexed by the null effect of plausibility). Prolonged Chinese exposure durations only increased L2 participants' processing efficiency in general, but did not affect their holistic representations of compound words.

More importantly, in the character-spaced condition, on GP, the main effect of plausibility was significant, and the interaction between plausibility and Chinese-learning years was marginally significant. Specifically, increased Chinese-learning years significantly shortened GP on compound words embedded in the plausible condition, but had no effect in the implausible condition. GP reflects the procedure of error detection, reanalysis and recovery from temporary ambiguous structures. The significant effect of Chinese-learning years in the plausible condition revealed that the increase in Chinese exposure times led to shorter reanalysis/recovery times for L2 participants. However, in the implausible condition, the ungrammatical combinations of the first characters of compound words with the prior verbs blocked the possibility of building temporary ambiguous structures; thus, little or no reanalysis would be needed in this case. Consequently, GP was not affected by Chinese-learning years in the implausible condition. Taken together, these findings indicated that in the character-spaced condition, L2 participants processed compound words in a character-by-character fashion, which is indexed by the significant main effect of plausibility. Additionally, their efficiency of reanalysis and recovery from temporary ambiguous structures was increased with prolonged Chinese exposure times.

In addition, in the character-spaced condition, the main effects of both plausibility and Chinese-learning years were significant on TT, but not their interaction. Prolonged Chinese exposure durations decreased TT for compound words to the same extent, no matter they were embedded in plausible or implausible conditions. Even with shortened reanalysis time (reflected in the patterns of GP), L2 participants still found it harder to process compound words in temporary ambiguous situations than in an unambiguous context. This pattern indicated that L2 participants' difficulty in word segmentation remained even with prolonged Chinese-learning years. Extended Chinese exposure durations only enhanced their processing efficiency in general, but did not help them develop a native-like processing strategy. Specifically, the facilitation from immersive linguistic environments may be attributed to strengthened lexical access pathways, accelerated syntactic pattern recognition and optimized allocation of cognitive resources through repeated contextualized practice, but had no effect on processing strategies. Even with a high proficiency in Chinese and long durations of Chinese exposure, L2-Chinese learners still use L1-biased processing strategy in word segmentation. It is extremely difficult (if possible at all) for L2-Chinese learners to implicitly develop a unique-to-L2 processing strategy.

In general, when taking L2 participants' Chinese-learning years into consideration, the analyses revealed some interesting findings: (1) high-proficiency L2 participants have already acquired holistic representations of Chinese compound words even with short Chinese exposure durations; (2) prolonged Chinese exposure times increased L2 participants' efficiency of words processing and reanalysis; and (3) even with increased Chinese exposure times, L2 participants' word segmentation difficulty remained. These findings led to the speculation that their word segmentation difficulty was predominantly caused by their L1-biased processing strategy, which is extremely difficult to implicitly acquire.

The findings of the present study revealed that high-proficiency L2-Chinese learners represented high-frequency two-character compound words in a holistic fashion: in the word-spaced condition, the meanings of the whole words can be directly accessed

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press

without activating each constituent. Additionally, these holistic representations of compound words were not affected by extended Chinese-learning years. This is different from the observations from the study of Gao et al. (2022), which indicated that L2-Chinese learners' representations of compound words were developed from a compositional to a holistic fashion with increased Chinese proficiency and that constituent morphemes affected the recognition of compound words even for high-proficiency L2-Chinese learners. We speculate that the discrepancies between the findings of the present study and previous ones were caused by different experimental tasks. In Gao et al. (2022), participants were required to accomplish a lexical decision task with a priming paradigm, while in the present study, participants were asked to read sentences silently for comprehension. To accomplish a lexical decision task, participants need to judge the acceptability of the combinations of the two constituent morphemes. However, in a reading task where all words are real, attention should be paid to the meaning activation and integration of each word. From this perspective, constituent morphemes are expected to play a bigger role in the lexical decision task than in the reading task. Thus, it is reasonable to observe the effects of constituent morphemes in the study of Gao et al. (2022), but not in the present study.

Note that the target words in the present study are all high-frequency and semantic transparent words, which are familiar to L2-Chinese learners. How lexical frequency, semantic transparency and L2 proficiency may affect L2 mental representations of Chinese words is out of the research scope of the present study. Some follow-up studies are needed to further investigate the possible factors of L2-Chinese learners' mental representations of Chinese compound words. Additionally, these studies would provide important insights into the independent influence of L2 lexicon representations on word segmentation difficulty.

The findings of the present study have some pedagogical implications. Since cross-language differences may encourage language learners to use different processing strategies, more attention should be paid to the layout of Chinese texts for L2-Chinese learners. Specifically, inter-word spaces should be employed in elementary Chinese reading materials designed for learners with a word-spaced L1 background. This would help them adjust to a typologically different script when learning new orthographic representations and grammatical rules. Stronger connections between orthographic forms and semantic representations would be made with the introduction of inter-word spaces.

4. Conclusion

In conclusion, the present study found that text presentation style affected L2-Chinese learners' on-line processing of Chinese compound words: high-proficiency L2-Chinese learners processed compound words holistically in the word-spaced condition, but decompositionally in the character-spaced condition. The findings indicated that L2-Chinese learners have acquired holistic mental representations of high-frequency and transparent Chinese compound words. Their difficulty in word segmentation was predominantly caused by their L1-biased processing strategy developed with their extensive experience in native languages.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S1366728925100643.

Data availability statement. The data that support the findings of this study are openly available at https://osf.io/4mjpy/.

Consent to participate. Informed consent was obtained from all individual participants included in the study.

Funding statement. This research was supported by Beijing Social Science Funds (22YYC016).

Competing interests. The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics approval. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

Abbott, M. J., & **Staub, A.** (2015). The effect of plausibility on eye movements in reading: Testing E-Z reader's null predictions. *Journal of Memory and Language*, **85**, 76–87. https://doi.org/10.1016/j.jml.2015.07.002.

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.

Bai, X., **Li, X.**, & **Yan, G.** (2015). Eye movement control in Chinese reading: A summary of 20-years research. *Psychology Development Education*, **31**, 85–91.

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**, S4–S17. https://doi.org/10.1111/j.1467-9817.2013.01554.x.

Bates, D., **Mächler, M.**, & **Bolker, B. M.** (2011). LME4: Linear mixed-effects models using S4 classes. R Package Version 0.999375–39.

Bertram, R., **Pollatsek, A, and Hyönä, J.** (2004). Morphological parsing and the use of segmentation cues in reading Finnish compounds. *Journal of Memory and Language*, **51**: 325–345.

Bordag, D., **Gor, K.**, & **Opitz, A.** (2021). Ontogenesis model of the L2 lexical representation. *Bilingualism: Language and Cognition*, 1–17. https://doi.org/10.1017/S1366728921000250.

Chen, H. C. (1996). Chinese reading and comprehension: A cognitive psychology perspective. In M. H. Bond (Ed.), *Handbook of Chinese psychology* (pp. 43–62). Hong Kong, China: Oxford University Press.

Chen, H. C. (1999). How do readers of Chinese process words during reading for comprehension? In J. Wang, A. W. Inhoff, & H. C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 257–278). Mahwah, NJ: Erlbaum.

Cui, L., **Wang, J.**, **Zhang, Y.**, **Cong, F.**, **Zhang, W.**, & **Hyönä, J.** (2020). Compound word frequency modifies the effect of character frequency in reading Chinese. *Quarterly Journal of Experimental Psychology*, **74**, 610–633.

Cui, Y. (2023). Eye movements of second language learners when reading spaced and unspaced Chinese texts. *Frontiers of Psychology*, **14**, 783960. https://doi.org/10.3389/fpsyg.2023.783960.

Gao, F., **Wang, J.**, **Zhao, C.**, & **Yuan, Z.** (2022). Word or morpheme? Investigating the representation units of L1 and L2 Chinese compound words in mental lexicon using a repetition priming paradigm. *International Journal of Bilingual Education and Bilingualism*, **25**, 2382–2396. https://doi.org/10.1080/13670050.2021.1913984.

Hong, W., & **Feng, C.** (2010). A study of the differences between CSL learners and Chinese native speakers in Chinese two-character compound word recognition. *Modern Foreign Languages*, **33**, 387–394.

Hoosain, R. (1992). Psychological reality of the word in Chinese. In H. C. Chen & O. Tzeng (eds), *Language processing in Chinese* (pp. 111–130). New York: North-Holland Elsevier Science Publisher.

Hyönä, J., **Pollatsek, A.**, **Koski, M.**, & **Olkoniemi, H.** (2020). An eye-tracking study of reading long and short novel and lexicalized compound words. *Journal of Eye Movement Research*, **13**. https://doi.org/10.16910/jemr.13.4.3.

Inhoff, A. W., **Radach, R.**, & **Heller, D.** (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory & Language*, **42**, 23–50. https://doi.org/10.1006/jmla.1999.2666.

Judd, C. M., **Westfall, J.**, & **Kenny, D. A.** (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, **103**, 54–69.

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.

Lenth, R. (2018). Emmeans: Estimated marginal means, aka least-squares means (R package version 1.1) [Computer software]. https://CRAN.R-project.org/package=emmeans

Lexicon of Common Words in Contemporary Chinese Research Team. (2009). *Lexicon of common words in contemporary Chinese*. Beijing: Commercial Press.

Li, X., & **Pollatsek, A.** (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review*, **127**, 1139–1162.

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

Li, X., **Liu, P.**, & **Ma, G.** (2011). Advances in cognitive mechanisms of word segmentation during Chinese reading. [中文阅读中词切分的认知机理述评]. *Advanced Psychological Science*, **19**, 459–470.

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

Liu, P., **Li, W.**, **Lin, N.**, & **Li, X.** (2013). Do Chinese readers follow the national standard rules for word segmentation during reading? *PLoS One*, **8**(2), e55440. https://doi.org/10.1371/journal.pone.0055440.

Ma, G., **Li, X.**, & **Rayner, K.** (2014). Word segmentation of overlapping ambiguous strings during Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, **40**, 1046.

Morey, R. D., **Rouder, J. N.**, & **Jamil, X.** (2015). BayesFactor: Computation of Bayes factors for common designs (Version 0.9.11) [Computer software]. http://bayesfactorpcl.r-forge.r-project.org

O'Regan, J. K., & **Jacobs, A. M.** (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 185–197.

Perea, M., & **Acha, J.** (2009). Space information is important for reading. *Vision Research*, **49**, 1994–2000. https://doi.org/10.1016/j.visres.2009.05.009.

R Core Team. (2018). R: A language and environment for statistical computing (version 3.5.0.). https://www.r-project.org

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception*, **8**, 21–30.

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. *Quarterly Journal of Experimental Psychology*, **62**, 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**, 1129–1144. https://doi.org/10.1016/s0042-6989(97)00274-5.

Reichle, E. D., **Rayner, K.**, & **Pollatsek, A.** (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z reader model. *Vision Research*, **39**, 4403–4411. https://doi.org/10.1016/s0042-6989(99)00152-2.

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.

Shen, D., **Liversedge, S.**, **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*, **18**, 192–202. https://doi.org/10.1037/a0027485.

Standardization Administration of China. (1980). *GB 2312–1980: Information Technology — Chinese Ideogram Coded Character Set for Information Interchange (Basic Set)*.

Staub, A., & **Goddard, K.** (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **45**, 110–127. https://doi.org/10.1037/xlm0000561.

Staub, A., **Rayner, K.**, **Pollatsek, A.**, **Hyönä, J.**, & **Majewski, H.** (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. Journal of experimental psychology. *Learning, Memory, and Cognition*, **33**(6), 1162–1169. https://doi.org/10.1037/0278-7393.33.6.1162.

Tsai, J. L., & **McConkie, G. W.** (2003). Where do Chinese readers send their eyes? In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 159–176). Oxford: Elsevier.

Tsang, Y. K., & **Chen, H. C.** (2008). Eye movements in reading Chinese. In K. Rayner, D. Shen, X. Bai, & G. Yan (Eds.), *Cognitive and cultural influences on eye movements* (pp. 235–254). Tianjin, China: Tianjin People's Publishing House.

Yan, G., **Tian, H.**, **Bai, X.**, & **Rayner, K.** (2006). The effect of word and character frequency on the eye movements of chinese readers. *British Journal of Psychology*, **97**, 259–268. https://doi.org/10.1348/000712605x70066.

Yan, M., **Kliegl, R.**, **Richter, E. M.**, **Nuthmann, A.**, & **Shu, H.** (2010). Flexible saccade-target selection in Chinese reading. *Quarterly Journal of Experimental Psychology*, **63**, 705–725. https://doi.org/10.1080/17470210903114858.

Yang, J., **Staub, A.**, **Li, N.**, **Wang, S.**, & **Rayner, K.** (2012). Plausibility effects when reading one- and 2-character words in Chinese: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **38**(6), 1801–1809.

Yang, S. (2021). Diagnosing reading problems for low-level Chinese as second language learners. *System*, **97**, 102433. https://doi.org/10.1016/j.system.2020.102433.

Yao, P., **Alkhammash, R.**, & **Li, X.** (2021). Plausibility and syntactic reanalysis in processing novel noun-noun combinations during Chinese Reading: Evidence from native and non-native speakers. *Scientific Studies of Reading*, **26**(5), 390–408.

Yao, P., **Jiang, X.**, **Chen, X.**, & **Li, X.** (2023). Explore the processing unit of L2 Chinese learners in on-line Chinese Reading. *Second Language Research*. https://doi.org/10.1177/02676583231202606.

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

Yu, L., **Liu, Y.**, & **Reichle, E. D.** (2020). A corpus-based versus experimental examination of word-and character-frequency effects in Chinese reading: Theoretical implications for models of reading. *Journal of Experimental Psychology: General*, **150**(8), 1612–1641.

Yuan, L. (1990). *Dictionary of usage frequency of modern Chinese words*. Beijing: Publishing House of Astronavigation.

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 & Performance*, **39**, 720–734. https://doi.org/10.1037/a0030097.

Zhou, J., & **Li, X.** (2021). On the segmentation of Chinese incremental words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **47**(8), 1353–1368.

Zhou, J., **Ma, G.**, **Li, X.**, & **Taft, M.** (2017). The time course of incremental word processing during Chinese reading. *Reading and Writing*, **31**, 607–625. https://doi.org/10.1007/s11145-017-9800-y.

https://doi.org/10.1017/S1366728925100643 Published online by Cambridge University Press