The Word and Character Frequency Effect in Chinese Natural Reading

Longjiao Sui[1], Evy Woumans[2], Wouter Duyck[3,4], Sam Boeve[3], Nicolas Dirix[3]

[1] School of Foreign Languages, Dalian Maritime University, Dalian, China

[2] Department of Translation, Interpreting and Communication, Ghent University, Ghent, Belgium

[3] Department of Experimental Psychology, Ghent University, Ghent, Belgium

[4] The Accreditation Organisation of the Netherlands and Flanders (NVAO), Den Haag, Netherlands

Address for correspondence:

Longjiao Sui, School of Foreign Languages, Dalian Maritime University, Linghai Road 1, 116026, Dalian, China. Email: longjiao.sui@dlmu.edu.cn or longjiao.sui@gmail.com **and** Nicolas Dirix, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium. Email: Nicolas.Dirix@UGent.be

Abstract

The word frequency effect, where high-frequency words are processed faster than low-frequency ones, has been extensively studied in alphabetic writing systems. The effect has also been observed in Chinese reading, a language that differs enormously from alphabetic languages, not only in appearance but also in the nature of the words. In the present study, we investigated the word and character frequency effects in Chinese natural reading by analysing reading data from an eye-tracking corpus in which participants read an entire novel (GECO-CN). The results show that as character frequency in Chinese increases, the facilitative word frequency effect tends to flatten or even reverse, and vice versa. These findings suggest that Chinese sentence processing is influenced by the frequencies of both words and their constituent characters, indicating the importance of considering character frequencies when studying the word frequency effect. In addition, these results also provide a plausible explanation for the inconsistent character effects found previously.

Keywords: Frequency effects, Chinese reading, Eye-movements, word processing, character processing

The word frequency effect (FE) is the phenomenon where processing speed is affected by the frequency of word occurrence, with high-frequency (HF) words being processed faster than low-frequency (LF) words in both word recognition (Rayner & Duffy, 1986; Rayner & Raney, 1996) and production (Griffin & Bock, 1998; Sui et al., 2019). It is one of the most potent effects observed in language research to date (Brysbaert et al., 2018), explaining around 30%~40% of the variance in (alphabetic) language processing speed (Brysbaert et al., 2016). This effect is independent of other confounding factors, such as word length, number of phonemes, age of acquisition, and orthographic neighborhood (Brysbaert et al., 2016). However, our current understanding of FE is based primarily on findings of alphabetic languages, whose words are visually salient by spatial separation (e.g., Rayner & Duffy, 1986). Thus, a question is whether the same applies to nonalphabetic writing systems (i.e., Chinese), which differ qualitatively from alphabetic languages in both writing and pronunciation.

Early efforts have examined the FE of Chinese words using categorical designs (sets of HF vs. LF word stimuli), showing a negative correlation between processing speed and word frequency, similar to findings in alphabetic languages (Ma et al., 2015). It should be noted, however, that Chinese words are composed of a limited number of orthographically independent characters, many of which are words themselves. Moreover, Chinese words are not as salient as characters in sentences due to the lack of interspace for demarcating word boundaries. Although character properties (e.g., frequency) appear to affect word recognition (e.g., Yan et al., 2006), evidence indicates that word properties also influence Chinese reading (e.g., Li et al., 2014; Yang et al., 2012). As a result, a process that segments visually consecutive words in a sentence is necessary in Chinese reading (Li & Pollatsek, 2020).

Obviously, the peculiar characteristics of Chinese word composition and the necessity of segmenting text strings into words during sentence reading pose great challenges to the feasibility and reliability of studying the Chinese word FE separately from character effects. The absence of visually clear spaces between words in Chinese sentences complicates word processing, as a single character can either be a one-character word or part of a multi-character word with its adjacent characters. Consequently, the influence of character properties on word processing should be stronger in sentence reading than in isolated word conditions, where visually distinct word boundaries allow for the quick exclusion of other possible word combinations. This may even affect the word FE. Differentiating between and understanding Chinese character and word FEs can advance our specific knowledge of Chinese reading, inspire further research and interpretation, and refine models that have already been and are being developed to account for Chinese reading (e.g., Li & Pollatsek, 2020).

Notably, existing evidence for Chinese FEs in reading is mainly based on single words (Mattingly & Xu, 1994) and isolated controlled sentences manipulating the target word (e.g., low-constraint sentences which often have avoided word ambiguity; Cui et al., 2021; Ma et al., 2015). Yet, word identification performance under these conditions cannot adequately and fairly reflect the performance in reading coherent paragraphs of a certain length, which are the most common reading materials in everyday life (Dirix et al., 2019; Kuperman et al., 2013; Sui et al., 2023). This is important as text-level context generates semantic, grammatical, and lexical expectations that may influence the reading of subsequent words. Moreover, the difference in reading performance between experimentally controlled and natural reading conditions can be expected to be more severe in Chinese, where an additional process of segmenting text strings into words is needed, compared with in alphabetic language where word boundaries are visually and conventionally marked.

This work, therefore, aims to clarify the issues of character and word FEs in natural reading by using frequencies as continuous variables. In the following section, we will first briefly summarize the Chinese writing system. Next, we will examine whether the process of segmenting text strings into words specific to Chinese sentence reading influences word recognition and, in turn, its recognition performance. Finally, we will discuss existing findings on Chinese word and character FEs.

The Chinese writing system

Chinese is a logographic language composed of box-shaped characters constructed by a number of strokes under certain rules. Character complexity varies based on the number of strokes, with more strokes resulting in greater visual complexity and longer recognition times (Liversedge et al., 2014). Chinese words are composed of a limited number of characters (about 6,000 characters, Li et al., 2022), which can form more than 56,000 words (Li & Su, 2022). Most of these characters can form one-character words or parts of different words and appear in various positions. For example, the character 中 can be part of 264 commonly used words: 1 one-character word (i.e., 中), 145 two-character words (109 of which have this character as the initial character, e.g., 中国, 初中), 64 three-character words (appearing at the beginning of 43 words, in the middle of 18 words and at the end of 3 words), 50 four-character words, 2 five-character words, 1 six-character word, and 1 seven-character word, based on Lexicon of Common Words in Contemporary Chinese Research Team (2008). Note that the average length of Chinese words is comparatively shorter than English ones, with 97.2% of written words being one- and two-character words (Li & Pollatsek, 2020).

What further discerns Chinese and English words is the physical layout of words in sentences. Modern Chinese texts are written vertically from left to right, lacking physical cues between words (e.g., blank spaces).

In contrast, characters are spatially discrete from each other. The absence of delineated word boundaries makes the segmentation of words less apparent than characters. Moreover, the concept of words did not appear in Chinese until the twentieth century (Li et al., 2015). Even now, most dictionaries are character and not word dictionaries. The less well-defined word concepts and the lack of clear word boundaries have sometimes led Chinese readers to have divergent segmentation decisions on identical strings, resulting in disagreement on the number of characters that constitute a word (Liu et al., 2013).

In addition, some studies have shown that disrupting word processing, such as inserting blank spaces within words or between characters (Bai et al., 2008), masking one of the characters that constitute a word using the moving window paradigm (Li et al., 2013), or separating a word and placing the characters on different lines (Li et al., 2012) can interfere with reading performance. In contrast, spacing between words elicited a similar (Bai et al., 2008) or a facilitative performance in reading times (Oralova & Kuperman, 2021; Zhang et al., 2013) or in ambiguous string conditions (Hsu & Huang, 2000) compared to conventional, unspaced characters. Collectively, these findings demonstrate that efficient sentence reading requires word identification (e.g., Li et al., 2014; Yang et al., 2012), which involves the processing of their constituent characters (e.g., Chen et al., 2003; Hoosain, 1992).

Segmentation into words during Chinese Reading

As indicated earlier, Chinese words are considered as the critical units determining reading efficiency, and Chinese readers are accustomed to reading unspaced scripts (e.g., 是个上好的旧式柜，上面点缀着铜钉, which means “a fine old piece, all studded with brass nails” in English). Segmenting sentences that do not carry word boundary information into words is essential for word identification. However, this process is far more than segmenting the contiguous words with no explicitly marked boundaries into individual words, as Chinese characters can appear at various positions in words. That is, a character in text strings may act as a one-character word (e.g., 面, face) or form a word with its predecessor(s) (e.g., 上面, on top of) or successor(s) (e.g., 面点, pastry). Thus, a continuous text can sometimes be segmented into words in several ways (e.g., 上//面点//缀//着/铜钉 or 上//面//点缀//着/铜钉 or 上面//点缀//着/铜钉). Evidence shows that readers’ performance is very likely to be affected by the properties of the possible words that constitute such overlapping ambiguous strings (e.g., 上面点, Huang et al., 2021). If the activated word candidates appear more frequently, yet are implausible to the sentence context, participants are more likely to spend a longer time fixating on the ambiguous part (Ma et al., 2014). If the frequency of possible two-character words is similar, Chinese readers prefer to segment ambiguous strings into left-side style (i.e., AB-C; Huang, & Li, 2020).

These results can be well explained by the segmentation of text strings into words hypothesis proposed by the Chinese Reading Model (CRM; Li & Pollatsek, 2020). This model demonstrates that all characters within the perceptual span (e.g., 上面点) are processed in parallel, activating all the words they can compose (e.g., 上面 and 面点). Since some characters are themselves a word, single-character words are also activated (e.g., 上 and 面). The activated words then compete with each other for selection (also see Ma et al., 2014; Li et al., 2009). Multiple-character words (e.g., 上面) have advantages over the words that constitute them (e.g., 上 and 面), as they receive feedforward activation from all their characters, whereas embedded words (e.g., 上 or 面) receive activations only from themselves. Additionally, the frequency of activated words affects competition. HF words are more likely to win the competition. Once a word unit wins the competition, the segmentation of text strings into words occurs concurrently with its recognition as a target word.

Interestingly and notably, one could assume that character properties may cause multiple influences on the reading times of Chinese words based on the above hypothesis. That is, all activated characters activate and facilitate identifying their constituent words (including one-character words). The higher the frequency of a character, the stronger the activation of the candidate words, and the less time it takes to retrieve the words. Yet, the stronger the activation of a one-character word, the more it interferes with the target (multi-character) word and the longer it takes to process at this stage. Although these model assumptions are not yet appropriately empirically supported and need further verification and investigation, it could indicate that character FEs may affect word recognition due to the need for segmenting text strings into words in Chinese sentence reading.

Word and character FEs in eye movements

Eye movement measures are employed to elucidate the underlying processes in word reading. The two basic components of eye movements are saccades and fixations (Rayner, 2009). The former refers to the action of quickly moving the eyes to a point, while the latter refers to the moment when the eyes concentrate on a point. Moreover, several timed measures can be calculated, based on the duration of fixations: a) first fixation duration (FFD), the duration of the first fixation on a word; b) single fixation duration (SFD), the duration when the word is fixated only once; c) gaze duration (GD), the duration of all fixations on a word before the next (right-side) word in a sentence is fixated; and d) total reading time (TRT), the duration of all fixations on the word. Measures such as FFD and GD belonging to first-pass time are generally considered “early” measures and are supposed to reflect processes at the initial stages of word identification. In contrast, measures such as TRT that involve the second-pass time (including regressions to previous content while reading) are referred to as "late" measures (Boston et al., 2008). Additionally, there is a dichotomous measure that refers to the probability of skipping a word in the first-pass reading, the skipping probability.

Chinese word properties seem to affect reading in similar ways with other highly different scripts when looking at eye movement measures. Evidence has shown that influential factors such as word frequency (Ma et al., 2015 in Chinese; Rayner et al., 1996; Slattery et al., 2007 in English), word length (Zang et al., 2018 in Chinese; Rayner et al., 2011 in English), and predictability (Rayner et al., 2005 in Chinese; Rayner et al., 2011 in English) in Chinese have similar effects on both number and durations of fixations as in alphabetic language reading. Readers make shorter fixations on more frequent (Liu et al., 2016; Wei et al., 2013), shorter (Zang et al., 2018), and highly predictable words (Rayner et al., 2005).

However, the unique features of Chinese script could have substantial influences on reading processes. As mentioned, Chinese words are composed of a limited number of characters. The frequency of a word may be low, but its constituent characters may be highly frequent. If HF components accelerate reading times for LF words, which should surely be more than they do for HF words, then the difference in reading times between HF and LF Chinese words may be reduced, or even negligible. A crucial question is whether the robust word FE found in alphabetic languages can also be observed in a completely disparate script. Below, we will elaborate on the empirical findings with regard to word and character frequency in Chinese reading.

The word FE in Chinese reading

Early efforts have investigated the word FE in Chinese sentence reading by means of eye-tracking (e.g., Liversedge et al., 2014; Zhou et al., 2018). These studies primarily examine the performance of target (content) words embedded in isolated low-constrained sentences. Target words are often categorized into ‘HF’ and ‘LF’, and controlled for some influential factors like length and complexity (the number of strokes in a character, e.g., Yu et al., 2021, but see Li et al., 2014). These studies found significantly shorter fixation durations for HF words (Cui et al., 2021; Liu et al., 2019; Ma et al., 2015; Li et al., 2014; Wei et al., 2013; Yu et al., 2021; Yan et al., 2006; Zhou et al., 2018), but there are a few notable deviations from these results (Liversedge et al., 2014).

When reading one-character target words embedded in sentences, the word FE seems less consistent. In Liversedge et al. (2014), the main effect of word frequency was not significant in fixation durations (i.e., SFD, FFD, and GD) but did show up in the skipping probability. HF words are thus processed as quickly as LF words but are skipped more often. The interaction between word frequency and character complexity was significant in SFD and FFD, but not in GD and the skipping probability, showing that words with LF and high character complexity are processed more slowly in early stages of word recognition.

In contrast, Zang et al. (2016) observed the word FE in GD and the skipping probability, even using the same test material as Liversedge et al. (2014). Readers spent more time gazing at LF one-character words and skipped them less often. Zang et al. (2016) also found a significant interaction between word frequency and visual complexity, showing shorter fixations for HF, less complex words. Yet, the interaction was observed in GD but not in FFD and SFD, contrary to that of Liversedge et al. (2014). The discrepancy between the two studies may be due to low statistical power, with 2880 datapoints of young adults recruited in the former study compared to 5120 in the latter.

When reading a multi-character word in a sentence, it has been consistently found across eye-tracking measures that HF words are recognized more quickly than LF words (in two-character words: Liu et al., 2016; Liu et al., 2019; Ma et al., 2015; Yan et al., 2006; Yu et al., 2021; in three-character words: Zhou et al., 2018; for all word lengths: Li et al., 2014; Yu et al., 2021). Although Wei et al. (2013) failed to find a reliable word FE for two-character words in FFD (i.e., only marginally significant in the item analysis and non-significant in the subject analysis), the authors do observe it in GD.

In sum, the word FE appears to be robust in multi-character words in Chinese reading, with a pattern similar to that of alphabetic languages (e.g., Li et al., 2014). Yet, since word and character frequencies inevitably have certain collinearity, it remains to be determined whether the observed word FEs indeed rely on the frequency of the word, or if character frequency also plays a role. To study whether and how characters influence word recognition, a few studies have investigated two-character words which can tease apart the influence of character and word frequencies, in comparison with single character words, on word recognition.

The character FE in Chinese reading

Some existing research has investigated the effect of character frequency on eye movement behavior by studying two-character target words embedded in a single sentence (Yan et al., 2006; Yu et al., 2021), often using categorical designs based on word and/or character frequency (Cui et al., 2021; Cui et al., 2013; Ma et al., 2015). In addition to word frequency, these studies also manipulated first (C1) and second character (C2) frequency (Yan et al., 2006; Cui et al., 2021) or C1 frequency only (controlling for C2 frequency; Yu et al., 2021). Surprisingly, different results emerged among the few existing studies, with some observing a facilitative (Yan et al., 2006; also see Mattingly & Xu, 1994; Tse & Yap, 2018 in lexical-decision task) or an inhibitory C1

FE (Yu et al., 2021; Cui et al., 2021). Others found no influence of character frequency (Li et al., 2014; Ma et al., 2015; Cui et al., 2013).

Yan et al. (2006) observed significantly shorter fixation times in words with frequent C1s in FFD, SFD, and GD but not in TRT. The frequency of the C2 did not have an effect ($p > .05$). The interaction between word and C1 frequency was significant in FFD but not in SFD, GD, and TRT, and between word and C2 frequency was significant in GD but a hint in FFD. The character FE was negligible for HF words but appeared to have a facilitative effect for LF words. Xiong et al. (2023) also observed C1 FEs in LF words but not in HF words in sentence reading (i.e., FFD, GD, and TRT) and lexical decision tasks. However, they found inhibitory rather than facilitative effects. Notably, they observed an overall facilitative C1 FE in the word naming task but not the interaction between word and character frequencies. In contrast, Cui et al. (2021) found an inhibitory character FE when analyzing LF two-character words, showing that the higher the frequency of the C1, the longer the fixation duration on the entire word. However, the C2 frequency did not affect the fixation duration on the word, consistent with what Yan et al. (2006) observed.

Yu et al. (2021) also observed an inhibitory character FE when analyzing the carefully controlled target words (e.g., controlling for the mean frequency of words and their C2s and the mean character complexity between HF and LF C1s), with longer FFDs on words with HF C1s (experimental analysis). However, when analyzing all words in a sentence (except the first and last ones) from the same corpus and including the properties (e.g., frequency) of the current word, its preceding word, and its succeeding word (as well as character frequencies etc.) into one analysis (corpus-based)[1], a facilitative character FE emerged. The authors explained that facilitative character FE could be due to uncontrolled collinearity. Moreover, the interaction between word and character frequencies of target words was not significant, contrary to the previous results (e.g., Yan et al., 2006). Yu et al., (2021) explained that the discrepancy between the reported character FEs could be due to the predictability of upcoming words. High-constraint sentences can narrow the number of lexical candidates to those compatible with the sentence context, thereby attenuating lexical competition and leading to the facilitative character FE (Yu et al., 2021). It should be noted, however, that the cloze

[1] Note that in such an analysis, the properties of each word are analysed twice or three times as the preceding, current, and succeeding words. For example, in the sentence “他/大学/的/时候/学习/并/积累/了/大量/的/烟酒/方面/的/理论/知识/”, the word “时候” could be the next word of “的”, the current word itself, and the preceding word of “并”.

predictability in both their studies was negligibly low (M = 0.1%, estimated by 80 participants in Yu et al., 2021, and M = 1.5%, estimated by 10 participants in Yan et al., 2006).

Finally, some studies did not observe a reliable character FE (Cui et al., 2013; Li et al., 2014; Ma et al., 2015). Notably, Cui et al. (2013) show that although the C1 of a word had no main effect on word fixation duration, fixations on HF C1s were shorter than those on LF ones when sharing the same C2. Furthermore Ma et al. (2015) found that the fixations of the pre-target word decrease with the increase of the C1 frequency of the target word. In summary, the existing controlled experiments show that the word FE is only reliable in multi-character words, not in single-character words, whereas the character FE varies in two-character words.

The present study

Inspired by the existing evidence, the theoretical inferences, and the above reasons, this study aims to investigate the character and word FEs in Chinese reading. Firstly, we will clarify how character frequency affects word recognition in natural reading, introducing new empirical evidence to the existing literature. Secondly, we will investigate whether the character effect interacts with word frequency. Thirdly, we will examine whether there is a word FE, independent of character frequency, in Chinese natural reading. Investigating these questions can clarify existing inconsistencies in Chinese character FEs and contribute to refining both Chinese reading models and those aimed at explaining universal performance across writing systems.

Note that previous eye-tracking studies have primarily focused on target words embedded in isolated, controlled sentences, limited both in the number of stimuli and contextual sentence diversity. Such manipulations severely narrow the variations that occur naturally in written language and may not provide a comprehensive picture of the word and character FEs or their interaction with other word characteristics. This study, therefore, will explore FEs in natural text reading. In addition, existing research mainly consists of small-scale controlled studies, often with arbitrary categorizations of the frequency variables. That is, HF and LE words are classified based on inconsistent and theoretically unsupported frequency means and ranges. However, categorizing continuous variables can minimize the probability that observed phenomena were due to stimulus selection or reduced statistical power, reliability, or inappropriate rejection of the null hypothesis (Balota et al., 2004). Thus, this work will investigate frequencies as continuous variables, rather than using dichotomous frequency categories (also see Tse & Yap, 2018 for lexical-decision tasks).

To assess continuous effects of character and word frequency in natural reading, we will employ data from the Chinese Ghent Eye-tracking Corpus (GECO-CN; Sui et al., 2023), a high-quality corpus with over a million datapoints (9 hundred thousand datapoints in Chinese reading). It provides the statistical power to reliably detect (minimal) effects and interactions, and can help gain further insight into word and character FEs. Additionally, the corpus entails data from readers processing the content of a continuous narrative (i.e., a fiction novel). Furthermore, GECO-CN has a diverse range of word stimuli, and thus a wide range of word and character frequencies that can be investigated as continuous predictors. It can minimize the probability that observed phenomena were due to stimulus selection or reduced statistical power, reliability, or inappropriate rejection of the null hypothesis due to categorizing continuous variables (Balota et al., 2004).

Method

Participants, Materials, and procedure

The GECO-CN (Sui et al., 2023) is an eye-tracking corpus in which Chinese-English bilinguals read an entire novel. A group of 30 native Chinese speakers with an average age of 25.3 years (SD = 2.60) read half of the book in their first language and the other half in their second language. They also took a series of language proficiency tests (e.g., HSK test [Chinese Proficiency Test, n.d.]) and answered comprehension questions after each chapter. For more details on the experimental procedure, materials, and participants of the database, we refer the reader to Sui et al., 2023. As the current work aims to document the FEs of Chinese word and character and the potential interaction between them, we only used eye movement data from Chinese reading, more specifically, of the two-character content words in the corpus. The Chinese version of the novel contains 59,403 words, 36331 content words, and 4,835 unique content word types.

Analysis

Fixation durations shorter than 100 ms were not considered to reflect meaningful word processing and were therefore removed from the analysis (e.g., Sereno & Rayner, 2003). Names (e.g., Mary) (removing 4.82% of the data) or words that appear at the beginning or end of a line (removing 11.41% of the existing data) were excluded from the analysis to avoid potential contamination and wrap-up effects (e.g., Rayner et al., 1989). This left us with 212,135 data points. Reading times that differed by more than 2.5 standard deviations from the subject mean in each measure were excluded from the analysis (removing 0.93% ~1.54% of the data for the four measures).

All analyses reported here were performed using R (Version 2021.09.1-372, developed by the R Core Team). We ran linear mixed-effects models (LMMs) using the lme4 package (Version 1.1–12). Since this experiment used natural materials without any carefully experimental controls, factors affecting word processing, besides character and word frequency, were also taken into account to avoid them obscuring the results. The model includes word frequency and numbers of word repetitions of the current word (without including the properties of preceding and succeeding words to avoid repeatedly including a word property in the analysis) in Chinese sessions as word-level predictors, frequency, and complexity of each character being character-level predictors, and Chinese proficiency of participants[2]. The eye-tracking database and analysis code used in this study are freely available online[3].

All predictors are continuous variables and were centred. The word and character frequencies used in this work are Zipf frequencies, a standardized frequency measure independent of corpus size, in SUBTLEX-CH (Cai & Brysbaert, 2010). We obtained the number of strokes of each character from the *Modern Chinese Dictionary 7th Edition* (2016). Random effects of the model were participants and word tokens. The dependent variables were eye movement measures, including FFD, SFD, GD, and TRT. The dependent variable was Box-Cox transformed to normalize the distribution. Such transformation does not change the functional relationship between the dependent and predictor variables.

The model in each reading time measure starts with the same full model. The full model includes interactions between word frequency and C1 and C2 frequency, as well as complexity and language proficiency. Furthermore, the interaction between C1 and C2 frequency and the main effect of word repetition (as a covariate) were included. We first strive to maximize the random factor structure (Barr et al., 2013) and compare the incremental model with the full model using restricted maximum likelihood (REML). If the model fails to converge or is not significantly different from the full model, the previous model will be used. Ultimately, the random factor structure in all reading time measures is unchanged.

We then discover the optimal model by performing stepwise selection to remove non-significant predictor terms. The maximum likelihood is used to select the optimal model, while the REML is used for the final optimal model. In addition, to estimating multicollinearity of coefficients in regression models, the Variance Inflation Factor (VIF) was calculated for each model, using *car* package. A VIF greater than 10 is considered a problem with severe multicollinearity (Fox & Weisberg, 2010; cited from Dirix & Duyck, 2017), whereas a VIF greater than 5 is considered a moderate influence. The largest VIF in the analyses reported below is 1.958, indicating there are no multicollinearity issues (see Table 1).

[2] Full model: Word_Frequency * (C1_Frequency + C2_Frequency + Language_Proficiency + C1_ Complexity + C2_ Complexity) + C1_Frequency * C2_Frequency + Word_Repetitions + (1 | Subjects) + (1 |Word_ID)
[3] The data and materials used in this work are from the published eye-movement corpus GECO-CN (Sui et al., 2022): https://doi.org/10.3758/s13428-022-01931-3; Link to R-codes: https://osf.io/c87pa.

Results

First fixation duration

The repetition and C2 complexity effects were significant (see Table 1). Words with fewer occurrences or more complex C2s take longer to process. The word and C2 FEs were also significant, showing higher word frequency or lower C2 frequency for shorter FFDs at the reference levels (i.e., intercept). Their significant interaction indicated that the pattern and magnitude of word FEs change with increasing C2 frequency and vice versa (see Fig 1.A). HF words have shorter FFD than LF words when the C2 frequency is less than 6.19 (Zipf frequencies; also see Table 2) but do not differ significantly when greater than this value (see Fig 1.A). Likewise, the character FE exhibits a facilitative pattern (reading times decrease with increasing character frequency) when word frequency is below 2.25 but an inhibitory pattern when above 4.23. The FE disappears when word frequency is between these values.

Table 1 *Analyses of Fixation-Duration Measures*

	Predictors	Estimate	Std. Error	t value	p value	VIF
First Fixation Duration	(Intercept)	4.648868	0.012514	371.483	<.001***	
	Word Frequency	-0.003585	0.001084	-3.306	<.001***	1.685
	Second-Character Frequency	0.005209	0.001508	3.454	<.001***	1.873
	Second-Character Strokes	0.001829	0.000281	6.516	<.001***	1.182
	Repetition	-0.000126	0.000045	-2.824	.005**	1.350
	Word Frequency:Second-Character Frequency	0.004129	0.001112	3.713	<.001***	1.217
Single Fixation Duration	(Intercept)	4.6485423	0.01265	367.473	<.001***	
	Word Frequency	-0.004719	0.0011455	-4.119	<.001***	1.491
	Second-Character Frequency	0.0052607	0.0016748	3.141	.002**	1.838
	Second-Character Strokes	0.0018508	0.0003127	5.919	<.001***	1.174
	Word Frequency:Second-Character Frequency	0.0037987	0.0012122	3.134	.002**	1.139
Gaze Duration	(Intercept)	2.700000	0.003767	716.619	<.001***	
	Word Frequency	-0.001530	0.000375	-4.076	<.001***	1.932
	First-Character Frequency	-0.001115	0.000444	-2.511	.012*	1.509
	Second-Character Frequency	0.002270	0.000499	4.553	<.001***	1.958
	Second-Character Strokes	0.000520	0.000091	5.731	<.001***	1.181
	Repetition	-0.000040	0.000014	-2.806	.005**	1.351
	Word Frequency:Second-Character Frequency	0.001793	0.000360	4.985	<.001***	1.217
Total Reading Time	(Intercept)	2.158000	0.002300	938.315	<.001***	
	Word Frequency	-0.000880	0.000229	-3.840	<.001***	1.751
	First-Character Frequency	-0.000863	0.000285	-3.029	.002**	1.507
	Second-Character Frequency	0.001493	0.000318	4.696	<.001***	1.933
	Second-Character Strokes	0.000230	0.000058	3.968	<.001***	1.175
	Word Frequency:Second-Character Frequency	0.001111	0.000224	4.973	<.001***	1.141

*p < .05. **p < .01. ***p < .001 (*p* values were calculated using *lmerTest* package)

Figure 1. Plots of the effect of second-character frequency on word frequency in predicted first fixation duration (A), single fixation duration (B), gaze duration (C), and total reading times (D). The grey shadow is the confidence interval.

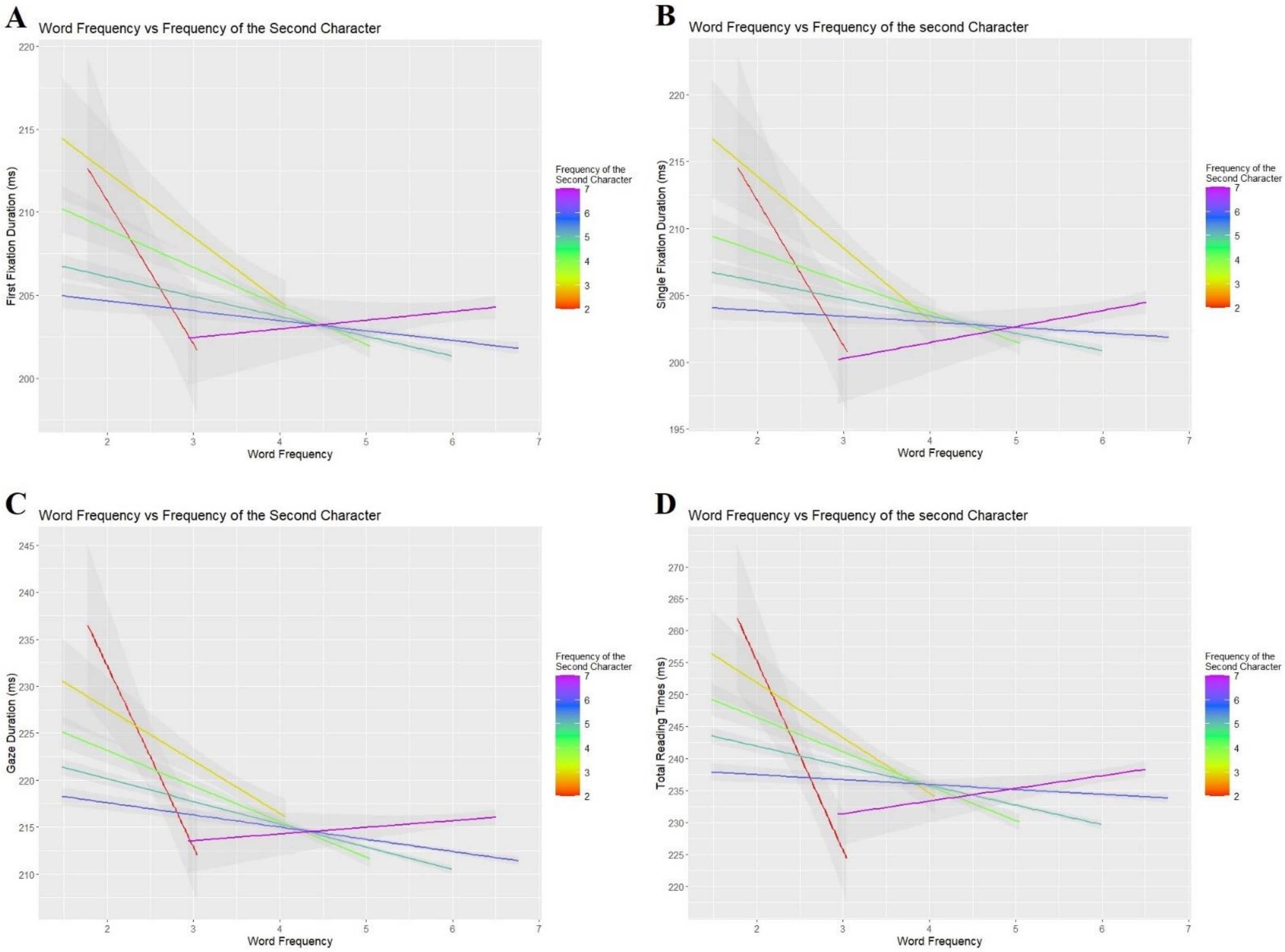

Table 2 *Tests for Linear Regression*

		Word Frequency			Second Character Frequency	
		Facilitative	**Inhibitory**		**Facilitative**	**Inhibitory**
First Fixation Duration	**Second Character Frequency**	6.1933	7.5754	**Word Frequency**	2.2461	4.2348
	Chisq	3.8458	2.6837	**Chisq**	3.8474	3.8658
	Pr(>Chisq)	0.0499	0.1014	**Pr(>Chisq)**	0.0498	0.0493
		Word Frequency			**Second Character Frequency**	
		Facilitative	**Inhibitory**		**Facilitative**	**Inhibitory**
Single Fixation Duration	**Second Character Frequency**	6.4607	7.5754	**Word Frequency**	1.4744	4.2554
	Chisq	3.8535	0.6643	**Chisq**	3.6883	3.8439
	Pr(>Chisq)	0.0496	0.4150	**Pr(>Chisq)**	0.0548	0.0499
		Word Frequency			**Second Character Frequency**	
		Facilitative	**Inhibitory**		**Facilitative**	**Inhibitory**
Gaze Duration	**Second Character Frequency**	6.2837	7.4507	**Word Frequency**	2.7251	4.0753
	Chisq	3.8595	3.8665	**Chisq**	3.8445	3.8448
	Pr(>Chisq)	0.0495	0.0493	**Pr(>Chisq)**	0.0499	0.0499
		Word Frequency			**Second Character Frequency**	
		Facilitative	**Inhibitory**		**Facilitative**	**Inhibitory**
Total Reading Time	**Second Character Frequency**	6.2387	7.3107	**Word Frequency**	2.5883	4.0234
	Chisq	3.8604	3.8477	**Chisq**	3.8610	3.8570
	Pr(>Chisq)	0.0494	0.0498	**Pr(>Chisq)**	0.0494	0.0495

*p < .05.

Single Fixation Duration

The SFD pattern was similar to that of FFD (see Table 1). Except for the repetition, the word and its C2 FEs, their interaction, and the C2 complexity effect were still present. Words with fewer strokes have shorter SFD, while the word or C2 FEs depends on the other. The word FE shows a negative pattern when the C2 is less than 6.46 and disappears when above it (see Table 2 and Fig 1.B). Similarly, an inhibitory character FE occurs when word frequency exceeds 4.26, but it disappears when below this threshold. While character frequency is visually facilitative trended when word frequency is low, it is not significant.

Gaze Duration

At the word level, frequency and repetition effects were significant (see Table 1). Repetition has an expected facilitative effect on GD. At the character level, C1 frequency, C2 frequency, and C2 complexity effects were significant. Words with higher C1 frequency or fewer C2 strokes were processed faster. The effect of word frequency on GD was influenced by C2 frequency, as their interaction was significant. GD decreased as word frequency increased when the C2 frequency was below 6.28, unaffected when between 6.28 and 7.45, and increased when above 7.45 (see Table 2 and Fig 1.C). Likewise, words with an HF C2 took less time to read than those with LF characters when the word frequency was below 2.73 but more time when it exceeded 4.08. The character effect disappeared when word frequency was between 2.73 and 4.08.

Total Reading Time

The effects of word frequency, C2 frequency, their interactions, and C2 complexity were significant, consistent with the aforementioned measures (see Table 1) and exhibited similar effect patterns. Similar to those found in GD, there was a statistically significant facilitative C1 FE. Additionally, facilitative, flattening, and inhibitory word FEs were observed when the C2 frequency was below 6.24, between 6.24 and 7.31, and above 7.31, respectively (see Table 2 and Fig 1.D). Also, the character FE appears only when word frequency is below 2.73 and above 4.08, showing facilitative and inhibitory trends, respectively.

Discussion

This study investigated the impact of word and character frequencies on eye movements in natural Chinese reading. Our first goal was to understand the effect of character frequency on visual word processing. The results showed that (first- and second-) character properties affect word processing. The C1 frequency has a facilitative effect on word processing, while the C2 frequency has an overall inhibitory effect. Our second goal was to investigate potential interactions between character and word frequencies. We found an interaction between word and C2 frequencies, with the direction of C2 FE varying with word frequency and vice versa. Our third goal was to study whether the word FE was present in natural reading after accounting for character frequency. Evidence shows that the overall facilitative word FE persists. In the next section, we address the word- and character-level results in detail and discuss their theoretical relevance for Chinese reading.

Character FE

First character

Our first key finding was the effect of C1 frequency on word processing times. Word fixation duration decreased as the C1 frequency increased. The common facilitative FE we found seems consistent with Yan et al. (2006) but contrary to what was reported by Cui et al. (2021) and Yu et al. (2021). Although previous studies obtained inconsistent FE patterns of C1, when only analyzing the target word, they all observed it in FFD but not in TRT. Contrary to previous experiments, we found the C1 frequency effects in GD and TRT but not in FFD and SFD, demonstrating that the C1 frequency affects the early and late stages of the word recognition process.

One possible reason for not having detected the C1 FE in FFD and SFD could arise from parafoveal processing. As illustrated, segmenting continuous text into words is necessary for reading in Chinese. To decide on word boundaries, readers may benefit from the parafoveal processing on the un-fixated character to the right of the fixation point (Yang et al., 2009). The C1 of a word is likely to have been processed to some extent in the last fixation, affecting its influence on the current fixation. This postulation is supported by previous findings in which the lexical properties of the subsequent word affect the current word processing (Li et al., 2014). It is also supported by additional analyses in the present study, which incorporated the frequencies of pre- and post-target characters or words. The results show that the frequency of post-target characters or words significantly affects current word processing, whereas the frequency of pre-target characters or words does not (see Table S.1 and Table S.2 in the Supplementary materials). This could explain the absence of a) the C1 FE in FFD and SFD; b) the C1 complexity effect and the interactions between word frequencies and C1 frequency or complexity throughout the time course of word processing in this work (see Table 1).

What differentiates previous work from the current study regarding C1 processing is that we studied natural reading in context. Some previous studies presented target words embedded in low-constraint sentences, which often avoid using ambiguous words (Cui et al., 2021; Cui et al., 2013), or sometimes with certain rules of preceding and following verbs and commas (Yan et al., 2006). Such experimental manipulations may reduce the need for parafoveal processing by easing the segmentation of text strings into words, either through the use of punctuation marks, which can serve as explicit word boundaries and the exclusion of ambiguous words. As a result, readers may rely less on parafoveal processing for this process than in natural reading, which we studied here.

Indeed, in our study sentences are neither controlled nor manipulated, allowing readers to read as they would in daily life. Therefore, it is not surprising that the C1 effect appeared in FFD in controlled studies but not in ours. This finding is also in line with studies that report a low degree of correspondence between paradigms in psycholinguistics (with lowest correspondence on early eye-tracking measures; see Dirix et al., 2019; Kuperman et al., 2013). The C1 FE found in our work may not reflect the complete character processing due to the parafoveal processing discussed above, which might (at least partially) explain the difference with previous findings where the C1 may have been full processed (e.g., Cui et al., 2021). Below, we will discuss possible reasons for the different results between the previous work and ours that can be ruled out.

One could argue that the facilitative character FE found in our own and Yan et al. (2006)'s work is an artifact of predictability, as Yu et al. (2021) suggested. Indeed, word anticipation could reduce competition between the target word and the candidate words (including one- and multiple-character words) whose meanings are implausible in the sentence context, resulting in target word reading facilitation. However, since no empirical evidence nor theoretical assumptions have suggested that predictability affects C1 but not C2, we expected a facilitatory effect of the C2 frequency as well. Yet, we observed facilitative as well as absent and reversed FEs of C2, which will be discussed in detail below. It is apparent that the predictability explanation account for some of our results (i.e., facilitative C1 and C2 FEs), but not all (i.e., absent and inhibitory FEs of C2).

Second character

The second remarkable finding of this work is the reliable C2 FE in all reading time measures. Previous studies only found a FE of the C1 (e.g., Yan et al., 2006; Cui et al., 2021). This work, however, is the first to observe the significant main effect of the C2, demonstrating that its frequency affects both the early- and late-stages of the word recognition process. In addition to the possible reasons for the parafoveal processing discussed above, it is worth noting that Yu et al. (2021) only manipulated the frequency of the C1 and controlled that of the C2, making an estimation of the C2 FE impossible. In Yan et al. (2006), the authors did not detect a main effect of the C2 frequency but observed a significant interaction between the C2 and word frequency, consistent with the current findings.

Before proceeding, it is important to clarify two things: (a) whether the underlying processes differ for C1 and C2 reading; (b) whether the two characters in a word influence each other. If one of (a) and (b) is fulfilled, it would be inappropriate to compare the FE of C2 either to the C1 or to studies where the C1 has been fully processed. Cui et al. (2021) suggested that since HF C1s appear more often as the initial character of words than LF C1s, by definition, and have larger morphological family sizes, they are less predictive of the C2 in two-character words. Thus, the fixation duration on the C2 should be longer, consistent with what they found.

From this, they explained that the inhibitory C1 FE they found in LF words was a carry-over effect due to morphological family size. Hence, according to Cui et al. (2021), the processing of the C2 is affected by the frequency of the C1 and may have a different underlying process than the C1. However, if C2 processing is largely constrained by the characteristics of the C1, the properties of the C2 should have minimal effects on word reading. That is, if two different C1s have similar morphological family sizes, the processing time of the accompanying C2s, regardless of their frequency and complexity, should be similar.

However, the interaction between the C1 and C2 frequencies was not reliable in the current and previous work. Instead, the current study demonstrated reliable C2 FEs in all time measures, arguing against the above inference. Additionally, Yu et al. (2021) argued on the basis of a post hoc analysis that, although statistical power was small, the morphological family size might not be responsible for the inhibitory character

FE they observed. Collectively, the frequency of the C1 seems unlikely to influence the C2 processing or to be different from its process.

Interestingly, our findings regarding the fully processed C2 frequency seem to be consistent with previous results that contradict each other (e.g., Yan et al., 2006; Yu et al., 2021). The facilitative character FE of shorter fixation duration for HF characters observed in LF words (see Table 2) was consistent with Yan et al. (2006). In contrast, the inhibitory pattern found in HF words was similar to that of Yu et al. (2021) and Cui et al. (2021), showing that higher character frequency was related to longer word fixation duration. The character FE was absent in medium-frequency words, congruent with the findings of Yan et al. (2006) and Cui et al. (2021) in their HF condition. It may explain why some studies failed to obtain the character FE (e.g., Li et al., 2014).

One potential concern is that the observed overall inhibitory FE for C2 may be due to the repeated occurrence of the same C1 paired with different C2s (e.g., "命令" (order) vs. "命运" (destiny)). To examine this, an additional factor was incorporated into models to assess whether a C2 had previously appeared with the same C1. The results show that the overall inhibitory FE on C2 remained unchanged even after controlling for this factor, ruling out this possibility (see Table S.3 in the Supplementary materials). Another concern is that the observed effects could be an artifact of predictability, given that the study used natural materials, which are arguably more influenced by predictability than experimentally designed materials. To address this concern, additional analyses were conducted by incorporating the predictability derived using the same method as in Boeve and Bogaerts (submitted). The results show that the observed character FEs persisted even after accounting for predictability, thereby arguing against this possibility (see Table S.4 in the Supplementary materials).

In sum, the variation of the character FE found in this work reiterates the importance of having a wide range of test stimuli and using them as continuous factors. Otherwise, the obtained results may reflect only part of the effect and lead to unnecessary contradiction and confusion, as shown by previous evidence. So far, there are some explanations for the character FE. Cui et al. (2021) indicated that the inhibitory character FE was due to the morphological family size, which we have discussed above, and argued against it being the main reason for the character FE. The Chinese E-Z Reader (CEZR) proposed by Yu et al. (2021) involves an interesting alternative, independent of morphological or neighborhood family size. We will explore in detail whether CEZR and CRM (Li & Pollatsek, 2020) can explain the current results after discussing the word FE.

Word frequency effect

Consistent with previous studies on Chinese (Wei et al., 2013) and alphabetic language reading (Cop et al., 2015; Rayner & Duffy, 1986), this work also observed a significant facilitatory main effect of word frequency across all measures. HF words had shorter fixation durations than LF words at the reference level. Furthermore, the word FE was found to be modulated by character frequency: the facilitative word FE decreases and even reverses as the character frequency increases. These findings suggest that word frequency affects word recognition and C2 frequency throughout the processing stage.

The facilitative word FE appears when the C2 frequency is less than 6.2 (the highest character frequency in Cai & Brysbaert (2010) is about 7.64). It is apparent that this pattern is the most frequently observed in Chinese reading (e.g., Cui et al., 2021; Wei et al., 2013). In addition, the fixation duration on words with HF characters is often shorter than those with LF characters (see Figure 1). Thus, experiments that categorize the continuous frequency variables are likely to find the facilitative word FEs only, which is also the case in literature (e.g., Ma et al., 2015; Yan et al., 2006). It also reiterates the importance of studying continuous variables to detect the full picture of the effect. When the C2 frequency exceeds about 6.2, the word FE is no longer significant due to the strong influence of C2 frequency. It could clarify the previously found seemingly counterintuitive results, where weak or absent word FE may be due to the influence of the character frequency rather than low statistical power or design flaws (e.g., Liversedge et al., 2014; Wang et al., 2018).

At extremely high C2 frequency (> 7.3), there was a tendency for an inhibitory word FE, which only occurred in GD and TRT (see Figure 1 and Table 2). Indeed, such an effect is expected because word FE should change as C2 frequency increases in a similar way to how word frequency affects character FE. Yet, the number of stimuli under this condition is limited even in this very large-size database. Further investigation is needed to verify the stability and reproducibility of the inhibitory word FE. Notably, various patterns of the word and character FEs are unlikely to be fully noticed in isolated word recognition. The predetermined word length in the isolated word condition allows for the quick exclusion of word candidates of different lengths (e.g., one-character words), thereby limiting the influence of character frequency on word retrieval. It may explain the observation of only facilitative word and character FEs in the lexical-decision task (Tse & Yap, 2018), suggesting that experimentally controlled tasks may not reflect the natural reading process fairly (also see Dirix et al., 2019; Kuperman & Van Dyke, 2013).

Our findings regarding the word FE have several theoretical implications. First, it strengthens the argument that words are indeed an important processing unit in Chinese reading, although their word boundaries are less explicit than in characters (Li et al., 2014). Their properties influence eye-movement measures during reading. Second, the interaction between word and character frequencies indicates that the word FE observed in Chinese sentence reading is influenced by the character frequencies of its constituents, exhibiting facilitative, absent, or even inhibitory patterns. Moreover, their interaction allowed us to learn the extent of their influence on each other. Given the word FE is only affected by relatively HF characters, one can infer that the influence of word frequency is greater than that of character frequency.

One might argue that, in some cases, the Chinese word processes might qualitatively differ depending on their word frequency. In fact, Cui et al. (2021) indicated that HF words could be processed as a holistic unit, and character property only plays a role in LF words. This suggestion could explain the absence of character FE in HF words found previously (Yan et al., 2006; Cui et al., 2021), it cannot account for the reversed character FE we found. Therefore, word processing is unlikely to vary with word frequency. One could also argue that word processing may differ between isolated word conditions and natural reading. Indeed, clear word boundaries in isolated word conditions could facilitate word recognition and prevent misinterpreting a part of a word (i.e., free morpheme) as an individual word or a part of another word. Such misinterpretations could adversely affect word recognition (Juhasz et al., 2005). As a result, any competition between a word and its components is likely to be quickly suppressed, limiting component interference in word recognition. In contrast, the competition among candidate words during natural reading is likely more intense due to the absence of clear word boundaries. Thus, Chinese characters may play a larger and more complex role in word processing in natural reading, where the segmentation of text strings into words is required than in the isolated word condition.

Additionally, bilingualism may be argued to affect the exploration of FEs in Chinese reading. However, previous research has shown that word FEs are affected by language proficiency rather than bilingualism, indicating no notable distinction between monolinguals and bilinguals in native language processing (Cop et al., 2015). Furthermore, no empirical or theoretical supports suggests that bilingualism influences character FEs. Additionally, since English is a compulsory course in China starting no later than junior high school, recruiting young adults who only speak Mandarin is a considerable challenge, even in China. Therefore, there is no reason to assume that the current results are qualitatively different from previous related findings. Nonetheless, future research is needed to investigate these issues.

Implications for Chinese reading models and research

A reputable Chinese reading model should be able to explain existing results. To do so, it should posit a role for character processing and incorporate its influence on word frequency into word processing. However, some existing models are designed for character identification (e.g., split-fovea model, Hsiao & Shillcock, 2004, 2005) and cannot explain the widely observed word FE (e.g., Li et al., 2014; Yu et al., 2021). Conversely, some are designed to illustrate word processing without character processes engaged (e.g., extended E-Z Reader model, Rayner et al., 2007). These are clearly at odds with the observed character FE (Yan et al., 2006). There are also models based on word identification and including the role of characters, namely the CRM (Li & Pollatsek, 2020) and the CEZR (Yu et al., 2021). While they can explain the word and character FEs found previously well, they fail to explain the novel results we observed.

In CEZR (Yu et al., 2021), word processing is related to how words are segmented. If the processing time of the first unidentified character exceeds a certain threshold, as may be the case with a LF C1 in a two-character word, the word-identification system initially infers it as a single-character word with a shorter fixation. In contrast, if the processing time does not exceed the threshold and is less than the time to process the combination of C1 and C2, for instance in HF C1 words, the word-identification system infers it as a two-character word with longer fixation. Although these hypotheses may seem counterintuitive, they can explain the obtained inhibitory character FE. The authors argued that their model predicts a facilitative character FE for non-target words (i.e., all words in the sentence except the target) while the character FE for carefully controlled target words is inhibitory. The failure to observe the facilitative character effect in the analysis of target words is due to the strong collinearity of words and character frequencies. CEZR (Yu et al., 2021) appears to be able to explain the standard word FE as well as both facilitative and inhibitory character FEs obtained in this work. However, it does not fully account for the conditions under which different character FEs occur. In addition, it seems unable to explain the C2 FE we found, as it states the role of the C1 but does not specify if and how the C2 impacts word recognition. Therefore, it is even less able to explain the interaction between the C2 and the word frequencies this study observed.

In contrast, CRM (Li & Pollatsek, 2020) is able to account for the effects of the C2 properties. Note that C1 in this study may be affected by parafoveal processing from the last fixation, and thus only C2 is discussed below. However, since CRM (Li & Pollatsek, 2020) does not elucidate the role of character frequency on word frequency, it is unable to fully explain the interactions reported here unless further assumptions are proposed. As discussed, the model assumes that character frequency affects word recognition in a hybrid manner. HF characters can greatly facilitate the retrieval of the words they constitute, but also induce more protracted interference as single-character words in lexical competition (also see Li et al., 2022). The amount of word processing time is affected by the magnitude of the character impact at different stages.

If one further hypothesizes that character frequency interacts with word frequency at the word retrieval stage, as an extension of CRM (Li & Pollatsek, 2020), then our results can be well explained. That is, when word frequency is low, word retrieval is difficult and time-consuming. It can benefit more from its HF components. The amount of facilitation from the HF character could be more influential than the prolonged time of the interference it causes. Its counteracted effect may be greater than that of LF characters, therefore demonstrating a facilitative character FE in LF words, as found in this work. Conversely, as the word frequency increases, the word is easier to access from memory and therefore benefits less from HF characters. The facilitation of HF characters may not differ much from that of LF characters, yet it still causes more interference in lexical competition. Its interference may affect word processing more than its facilitation, with a greater residual effect than for LF characters. Thus, the fixation durations of words with LF characters are shorter than those with HF characters, consistent with the inhibitory character frequency we found in HF words. Whereas in medium-frequency word, the counteracted effect of facilitation and interference of HF characters may be quantitatively similar to those of LF characters. Thus, the processing times of HF and LF characters are not statistically different, explaining the absent character FE this study found.

Likewise, LF characters cause limited facilitation and interference. Therefore, words with higher frequency are processed faster than those with lower frequency. Conversely, HF characters cause high interference but greatly facilitate word retrieval for LF words compared to HF ones. After counteracting the influence between facilitation and interference, the word FE could illustrate a flattened or even a reversed trend. These assumptions could explain the facilitative and absent word FEs we found and provide a theoretical hypothesis for the inhibitory trends when the frequencies of its characters are fairly high. In sum, the present study showed that both character and word frequency influence word reading, and influence each other, already in early stages of word recognition, until late stages.

An alternative inference that somewhat differs from the CRM (Li & Pollatsek, 2020) might also account for the current findings. While the LF multiple-character word retrieval is time-consuming, HF characters can quickly access one-character words. One-character words that are not plausible in the sentence context can be quickly excluded, resulting in reconsidering multiple-character words. That is, the segmentation of text strings into words may occur more than once. Thus, the LF multiple-character words with HF characters should be processed faster than those with LF ones. In contrast, HF multiple-character words can be quickly retrieved and compete with single-character words activated by HF characters but not with LF single-character words. Therefore, HF multiple-character words with HF characters are expected to be processed slower than those with LF characters. Future research needs to verify how exactly word and character frequencies are engaged in Chinese reading, as they are undoubtedly vital for understanding Chinese reading.

Nevertheless, the novel findings in this work can be an important new “benchmark” with significant implications for assessing theoretical assumptions of Chinese reading. Future model developments should be able to simulate the findings reported here, derived from a large corpus of natural reading. Future research on Chinese reading and aiming to seek computational models or mechanisms should pay more attention to natural reading behavior, given the unique features of the Chinese writing system. Furthermore, future studies should be cautious in categorizing continuous variables to ensure a comprehensive understanding of their effects.

Conclusion

This work, we believe, is the first to test the continuous effect of Chinese word and character frequencies across the entire ranges in text reading. In the present study, the word FE exhibited facilitative and flattening patterns with the change of character frequency, demonstrating that word frequency interacts with character frequency. Additionally, the current study observed all the seemingly contradictory character FEs (e.g., facilitative, flattening, and inhibitory) found in previous research (e.g., Li et al., 2014; Yan et al., 2006), indicating the different appearance of the character FE is due to the variation in word frequency. Finally, our findings emphasize the importance of using natural reading material to explore effects as continuous variables in statistically powerful research to reveal the full picture of effects.

References

Ashby, J., Rayner, K., & Clifton, C., Jr. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 58A(6), 1065–1086. https://doi.org/10.1080/02724980443000476

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

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual Word Recognition of Single-Syllable Words. *Journal of Experimental Psychology: General*, 133(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

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, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Boeve, S. & Bogaerts, L. (submitted). A Systematic Evaluation of Dutch Large Language Models' Surprisal Estimates in Sentence, Paragraph, and Book Reading. *Behavior Research Methods*. Preprint: https://osf.io/wr4qf/

Boston, M.F., Hale, J.T., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*. 2(1): 1-12. https://doi.org/10.16910/jemr.2.1.1

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, 27(1), 45-50. https://doi.org/10.1177/0963721417727521

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, 42(3), 441–458. https://doi.org/10.1037/xhp0000159.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, 5, e10729. https://doi.org/10.1371/journal.pone.0010729

Chen, H., Song, H., Lau, W. Y., Wong, K. F. E., & Tang, S. L. (2003). Chinese reading and comprehension: A cognitive psychology perspective. In C. McBride-Chang & H. Chen (Eds.), *Reading development in Chinese children* (pp. 157–169). Westport, CT: Praeger.

Chinese Proficiency Test. (n.d.) *HSK Level VI*. http://www.chinesetest.cn/gosign.do?id=1&lid=0#

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin & Review*, 22(5), 1216–1234. https://doi.org/10.3758/s13423-015-0819-2

Cui, L., Wang, J., Zhang, Y., Cong, F., Zhang, W., & Hyönä, J. (2021). Compound word frequency modifies the effect of character frequency in reading Chinese. *Quarterly Journal of Experimental Psychology*, 74(4), 610-633. https://doi.org/10.1177/1747021820973661

Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., & Liversedge, S. P. (2013). Processing of compound-word characters in reading Chinese: An eye-movement-contingent display change study. *Quarterly Journal of Experimental Psychology*, 66(3), 527-547. https://doi.org/10.1080/17470218.2012.667423

Dictionary Editorial Office of Linguistic Research Institute, Chinese Academy of Social Sciences (2016). *Modern Chinese Dictionary* (7th Edition). Beijing, China: The Commercial Press.

Dirix, N., & Duyck, W. (2017). The first-and second-language age of acquisition effect in first-and second-language book reading. *Journal of Memory and Language*, 97, 103-120. https://doi.org/10.1016/j.jml.2017.07.012

Dirix, N., Brysbaert, M., & Duyck, W. (2019). How well do word recognition measures correlate? Effects of language context and repeated presentations. *Behavior Research Methods*, 51(6), 2800–2816. https://doi.org/10.3758/s13428-018-1158-9

Fox, J., & Weisberg, S. (2010). *An R Companion to Applied Regression*. SAGE.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38(3), 313–338. https://doi.org/10.1006/jmla.1997.2547

Hand, C. J., Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2010). The frequency-predictability interaction in reading: it depends where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1294 –1313. https://doi.org/10.1037/a0020363

Hoosain, R. (1992). Psychological Reality of the Word in Chinese. In H. C. Chen, & O. J. L. Tzeng (Eds.), *Language Processing in Chinese* (pp. 111-130). Amsterdam, NY: North-Holland. http://dx.doi.org/10.1016/S0166-4115(08)61889-0

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

Hsiao, J. H., & Shillcock, R. (2004). Connectionist modelling of Chinese character pronunciation based on foveal splitting. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty Sixth Annual Conference of the Cognitive Science Society* (pp. 601–606). Erlbaum.

Hsiao, J. H., & Shillcock, R. (2005). Differences of split and non-split architectures emerged from modelling Chinese character pronunciation. In B. G. Bara, L. W. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the Twenty Seventh Annual Conference of the Cognitive Science Society* (pp. 989–994). Erlbaum

Huang, L., & Li, X. (2020). Early, but not overwhelming: The effect of prior context on segmenting overlapping ambiguous strings when reading Chinese. *Quarterly Journal of Experimental Psychology*, 73(9), 1382-1395. https://doi.org/10.1177/1747021820926012

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

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

Janssen, N., Bi, Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds. *Language and Cognitive Processes*. https://doi.org/10.1080/01690960802250900.

Juhasz, B. J., Inhoff, A. W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes*, *20*(1–2), 291–316. https://doi.org/10.1080/01690960444000133

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 802–823. https://doi.org/10.1037/a0030859

Lexicon of Common Words in Contemporary Chinese Research Team. (2008). *Lexicon of common words in contemporary Chinese*. Beijing, China: The Commercial Press.

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

Li, X., Gu, J., Liu, P., & Rayner, K. (2013). The advantage of word-based processing in Chinese reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 879–889. https://doi.org/10.1037/a0030337

Li, X., Huang, L., Yao, P., & Hyönä, J. (2022). Universal and specific reading mechanisms across different writing systems. *Nature Reviews Psychology*, 1-12. https://doi.org/10.1038/s44159-022-00022-6

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., 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., & Su, X., (2022). *Lexicon of common words in contemporary Chinese Second Edition*. Beijing, China: The Commercial Press.

Li, X., Zang, C., Liversedge, S. P., & Pollatsek, A. (2015). The role of words in Chinese reading. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 232–244). New York: Oxford University Press. doi:10.1093/oxfordhb/9780199324576.013.14

Li, X., Zhao, W., & Pollatsek, A. (2012). Dividing lines at the word boundary position helps reading in Chinese. *Psychonomic bulletin & review*, 19(5), 929-934. https://doi.org/10.3758/s13423-012-0270-6

Liu P-P, Li W-J, Lin N, Li X-S (2013) Do Chinese Readers Follow the National Standard Rules for Word Segmentation during Reading? *PLoS ONE*, 8(2): e55440. doi:10.1371/journal.pone.0055440

Liu, Y., Reichle, E.D., & Li, X. (2016). The effect of word frequency and parafoveal preview on saccade length during the reading of Chinese. *Journal of Experimental Psychology: Human Perception and Performance*, 42, 1008–1025. https://doi.org/10.1037/xhp0000190

Liu, Y., Yu, L., Fu, L., Li, W., Duan, Z., & Reichle, E. D. (2019). The effects of parafoveal word frequency and segmentation on saccade targeting during Chinese reading. *Psychonomic bulletin & review*, 26(4), 1367-1376. https://doi.org/10.3758/s13423-019-01577-x

Liversedge, S. P., Drieghe, D., Li, X., Yan, G., Bai, X., & Hyönä, J. (2016). Universality in eye movements and reading: A trilingual investigation. *Cognition*, 147, 1-20. https://doi.org/10.1016/j.cognition.2015.10.013

Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during Chinese reading. *Visual Cognition*, 22(3-4), 441–457. https://doi.org/10.1080/13506285.2014.889260

Lu, Z.W., (1964). *The Word Formation of Chinese*. Science Press.

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(3), 1046–1059. https://doi.org/10.1037/a0035389.

Ma, G., Li, X., & Rayner, K. (2015). Readers extract character frequency information from nonfixated-target word at long pre-target fixations during Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, 41(5), 1409-1419. https://doi.org/10.1037/xhp0000072

Mattingly, I. G., & Xu, Y. (1994). Word superiority in Chinese. *Advances in the Study of Chinese Language Processing*, 1, 101–111.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375–407. https://doi.org/10.1037/0033-295X.88.5.375

Oralova G and Kuperman V. (2021). Effects of Spacing on Sentence Reading in Chinese. *Frontiers in Psychology*. 12:765335. https://doi.org/10.3389/fpsyg.2021.765335

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & cognition*, 14(3), 191-201. https://doi.org/10.3758/BF03197692

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, 3(2), 245-248. https://doi.org/10.3758/BF03212426

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The Effects of Frequency and Predictability on Eye Fixations in Reading: Implications for the E-Z Reader Model. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 720–732. https://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E-Z reader model of eye movement control to Chinese readers. *Cognitive Science*, 31(6), 1021–1033. https://doi.org/10.1080/03640210701703824

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review*, 12(6), 1089-1093. https://doi.org/10.3758/BF03206448

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, 22(5), 1188–1200. https://doi.org/10.1037/0096-1523.22.5.1188

Rayner, K., Sereno, S. C.,Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(special issue), 21–49. https://doi.org/10.1080/01690968908406362

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 514–528. https://doi.org/10.1037/a0020990

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences*, 7(11), 489–493. https://doi.org/10.1016/j.tics.2003.09.010

Slattery, T.J., Pollatsek, A., & Rayner, K. (2007). The effect of the frequencies of three consecutive content words on eye movements during reading. *Memory & Cognition*, 35, 1283-1292. https://doi.org/10.3758/BF03193601

Sui, L., Dirix, N., Woumans, E., & Duyck, W. (2023). GECO-CN: Ghent Eye-tracking COrpus of sentence reading for Chinese-English bilinguals. *Behavior Research Methods*, 55, 2743–2763. https://doi.org/10.3758/s13428-022-01931-3

Sui, L., Kruger, H., & Slatyer, H. (2019). The central processing bottleneck during word production: Comparing simultaneous interpreters, bilinguals and monolinguals. *Bilingualism: Language and Cognition*, 22(5), 968-985. https://doi.org/10.1017/S1366728918000871

Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 761–775. https://doi.org/10.1037/0278-7393.23.3.761

Tse, C. S., & Yap, M. J. (2018). The role of lexical variables in the visual recognition of two-character Chinese compound words: A megastudy analysis. *Quarterly Journal of Experimental Psychology*, 71(9), 2022-2038. https://doi.org/10.1080/17470218.2014.985234

Wang, J., Li, L., Li, S., Xie, F., Liversedge, S. P., & Paterson, K. B. (2018). Effects of aging and text-stimulus quality on the word-frequency effect during Chinese reading. *Psychology and Aging*, 33(4), 693–712. https://doi.org/10.1037/pag0000259

Wei, W., Li, X., & Pollatsek, A. (2013). Word properties of a fixated region affect outgoing saccade length in Chinese reading. *Vision Research*, 80, 1-6. https://doi.org/10.1016/j.visres.2012.11.015

Xiong, J., Yu, L., Veldre, A., Reichle, E. D., & Andrews, S. (2023). A multitask comparison of word- and character-frequency effects in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 49(5), 793–811. https://doi.org/10.1037/xlm0001192

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(2), 259–268. https://doi.org/10.1348/000712605X70066

Yang, J., Staub, A., Li, N., Wang, S., & Rayner, K. (2012). Plausibility effects when reading one- and two-character words in Chinese: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(6), 1801–1809. https://doi.org/10.1037/a0028478

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word n + 2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 35(4), 1192–1204. https://doi.org/10.1037/a0013554

Yu, L., Liu, Y., & Reichle, E. D. (2021). 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. https://doi.org/10.1037/xge0001014

Zang, C., Fu, Y., Bai, X., Yan, G., & Liversedge, S. P. (2018). Investigating word length effects in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, 44(12), 1831–1841. https://doi.org/10.1037/xhp0000589

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

Zang, C., Zhang, M., Bai, X., Yan, G., Paterson, K. B., & Liversedge, S. P. (2016). Effects of word frequency and visual complexity on eye movements of young and older Chinese readers. *Quarterly Journal of Experimental Psychology*, 69(7), 1409–1425. https://doi.org/10.1080/17470218.2015.1083594

Zhou, J., Ma, G., Li, X., & Taft, M. (2018). The time course of incremental word processing during Chinese reading. *Reading and Writing*, 31(3), 607-625. https://doi.org/10.1007/s11145-017-9800-y