*Applied Psycholinguistics* (2024), **45**, 1147–1164
doi:10.1017/S0142716424000432

ORIGINAL ARTICLE

# The effects of L2 proficiency on L2 word reading strategies: evidence from Chinese–English bilinguals

Rong Zhao, Hang Wei and Zhao Yao

School of Foreign Studies, Xi'an Jiaotong University, Xi'an, China
**Corresponding author:** Zhao Yao; Email: yaozhao@xjtu.edu.cn

(Received 23 January 2024; revised 7 November 2024; accepted 10 November 2024; first published online 20 January 2025)

**Abstract**

The dual route cascaded (DRC) model proposes that the mapping from orthography to phonology occurs through two pathways: the sublexical and lexical routes. Cross-linguistic studies have found that Chinese character reading relies more on the lexical pathway, whereas English word reading relies more on the sublexical pathway. However, it remains unclear how these two pathways collaborate in the L2 word reading of Chinese–English bilinguals and whether their reading strategies are influenced by L2 proficiency. In the current study, 72 Chinese–English bilinguals with varying levels of L2 proficiency were tested. They were asked to name English words that varied in frequency and spelling-sound consistency. The results showed that participants with lower L2 proficiency were more sensitive to frequency, indicating a greater reliance on lexical processing in L2 word reading. In contrast, participants with higher L2 proficiency were more sensitive to consistency, suggesting a greater reliance on sublexical processing. These findings suggest that L2 word reading strategies vary as a function of L2 proficiency. As L2 proficiency increases, Chinese–English bilinguals' reading strategies may shift from primarily relying on lexical to sublexical processing. This study provides evidence from L2 readers for the DRC model, helping to broaden the explanatory scope of the model.

**Keywords:** Chinese–English bilinguals; lexical processing; L2 proficiency; L2 word reading; sublexical processing

## Introduction

Mapping orthography to phonology is a core process of word reading. According to the dual route cascaded (DRC) model (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), the phonological access of visual words is realized through two pathways: the sublexical route, which relies on grapheme-to-phoneme conversion rules for decoding, and the lexical route, which directly retrieves the phonology of visual word forms from the mental lexicon or via the semantic system. Cross-linguistic research demonstrates that native speakers of languages with a transparent orthography (e.g., alphabetic languages) rely more on the sublexical pathway, whereas native speakers of languages with an opaque orthography (e.g., logographic languages) rely more on the lexical pathway in word reading (Bhide, 2015; Cao et al., 2017). Previous research has mainly examined word reading strategies in monolinguals within the framework of the DRC model. The number of bilinguals has been increasing, with more than half of the world's population speaking more than one language (Grosjean, 2010). However, for bilinguals whose first language (L1) and second language (L2) have completely different orthographies (e.g., Chinese–English bilinguals, whose L1 is logographic and L2 is alphabetic), how the sublexical and lexical pathways collaborate during L2 word reading remains largely underexplored.

Recent studies have investigated L2 word reading strategies of Chinese–English bilinguals within the framework of the DRC model (Akamatsu, 2002; Ben-Yehudah et al., 2019). These studies have yielded mixed results. Some studies found that Chinese–English bilinguals rely more on the lexical pathway in L2 word reading (Ben-Yehudah et al., 2019; Pae, Sun, Mano, & Kwon, 2017), while other studies observed a greater reliance on the sublexical pathway (Akamatsu, 2002). More recently, studies have begun to examine the influence of L2 proficiency on L2 word reading strategies (Botezatu, 2023; Botezatu, Misra, & Kroll, 2022). For example, Botezatu (2023) found that Chinese–English bilinguals with lower L2 proficiency tended to rely more on the sublexical pathway. It must be noted, however, that this finding was observed only in ERP responses, with no effects of L2 proficiency detected in behavioral results. Additionally, some of the results from Botezatu (2023) and Botezatu et al. (2022) are conflicting. Therefore, further research is needed to examine whether Chinese–English bilinguals rely more on the sublexical or lexical pathway when reading L2 words, and whether their reading strategies vary as a function of L2 proficiency.

### Chinese character reading vs. English word reading

Chinese and English are of different orthographic types. English belongs to the alphabetic writing system, and its words are composed of linear strings of letters. Chinese is a logographic language, and its words consist of one or more square-shaped characters, with each character consisting of one or more strokes. In addition to the different visual features between English and Chinese words, word reading also differs across the two languages. Compared with Chinese, the mapping between the orthography and phonology of English words is relatively consistent, such that native English speakers rely more on sublexical processing in word reading. They use printed graphemes as visual bases and map them into the phonemes of spoken language according to the grapheme-to-phoneme conversion rules. They mainly read out English words by assembling fine-grained phonemic units (Tan et al., 2005). That is, native English speakers tend to rely more on the sublexical pathway in word reading.

In contrast, due to the lack of correspondence rules between orthography and pronunciation of Chinese characters, native Chinese readers tend to rely more heavily on lexical processing in Chinese character reading. They read Chinese characters in two ways (Pasquarella et al., 2014). Specifically, when a character consists of a single component (e.g., 叉, meaning *fork* and pronounced /cha1/), readers can only pronounce it by mapping the whole character onto its pronunciation. If a character is a compound and contains a phonetic component, readers might pronounce it drawing on the information provided by the phonetic component. For example, 笼 means *cage* and is pronounced /long2/, which has the same pronunciation as its phonetic component 龙. Although most of Chinese characters have a phonetic component, only 28% of the phonetic components can accurately represent the pronunciation of the whole character (Tan et al., 2005). Additionally, the phonetic components in Chinese characters do not correspond to the subsyllabic phonology in the way of mapping letters into phonemes in alphabetic languages (Perfetti, Liu, & Tan, 2005; Tan et al., 2005). As a consequence, instead of using the grapheme-to-phoneme conversion rules of alphabetic languages, native Chinese readers retrieve the phonological information of a character directly from their mental lexicon (Tan et al., 2005), mainly through manipulating lexical level information (Ben-Yehudah et al., 2019; Bhide, 2015).

### *Sublexical vs. lexical processing in L2 word reading*

The orthographic differences between Chinese and English have motivated researchers to investigate the L2 word reading strategies among Chinese–English bilinguals, focusing on whether they rely more on sublexical or lexical processing in L2 word reading (Akamatsu, 2002; Ben-Yehudah et al., 2019; Pae et al., 2017). Sublexical processing was typically measured by manipulating spelling-to-sound consistency/regularity (e.g., consistency effects: consistent vs. inconsistent words), variations in letter size and shape (e.g., siZe, *s*hap$^{e}$), etc. (Ben-Yehudah et al., 2019; Bhide, 2015; Botezatu, 2023; Pae et al., 2017). Taking consistency effects as an example, reading spelling-to-sound consistent words relies more on grapheme-to-phoneme conversion rules compared with reading inconsistent words, and the differences in reading these two types of words reflect the extent of the reader's reliance on sublexical processing. In contrast, lexical processing was usually measured by manipulating word frequency (high- vs. low-frequency words), scrambling the letters within a word (e.g., close→colse), etc (Ben-Yehudah et al., 2019; Bhide, 2015; Pae et al., 2017). Taking frequency effects as an example, readers typically extract the pronunciation of high-frequency words directly from their mental lexicon, thus relying more on lexical processing when reading high-frequency words compared with low-frequency words. The greater the frequency effects, the more L2 learners rely on lexical processing.

Previous studies have yielded mixed findings regarding the L2 word reading strategies of Chinese–English bilinguals. One set of studies show that Chinese–English bilinguals rely more on lexical processing in L2 word reading (Ben-Yehudah et al., 2019; Pae & Lee, 2015; Pae et al., 2017). Specifically, the L2 word reading strategies may be influenced by the bilinguals' L1 backgrounds, with L2 readers whose L1 is a logographic language relying more on lexical processing than those with alphabetic languages as their L1. In contrast, L2 readers with alphabetic languages as their L1 show more reliance on sublexical processing as compared to those with logographic languages as their L1. For example, Ben-Yehudah et al. (2019) compared the L2 English word reading of native Chinese and Korean readers and found that native Chinese readers showed stronger sensitivity to frequency effects, whereas native Korean readers were more sensitive to consistency effects. Another study, conducted by Pae et al. (2017), used L2 English words with scrambled letters as well as words with letters of varied sizes and shapes, revealing similar findings. Pae et al. (2017) found that compared with native Korean readers, native Chinese readers read scrambled words more efficiently at the lexical level, but were less efficient in naming atypical words at the sublexical level. Taken together, these findings suggest that native Chinese readers rely more on lexical processing, whereas native Korean readers rely more on sublexical processing in L2 English word reading.

In contrast, another set of studies found that Chinese–English bilinguals show a greater reliance on sublexical processing in L2 word reading (Akamatsu, 2002; Koda, 2000). For instance, Akamatsu (2002) tested fluent L2 English readers with different L1 orthographic backgrounds (Chinese, Japanese, and Persian) and found that the way these bilinguals read L2 words was highly similar. Specifically, all participants read high-frequency regular words and high-frequency exception words (words whose pronunciations do not follow the grapheme–phoneme correspondence rules) at the same speed, but read low-frequency regular words faster than low-frequency exception words. These findings suggest that L2 readers are more sensitive to sublexical processing. Based on this, Akamatsu argued that the English word reading strategies used by L2 learners might be fundamentally the same as those of native English speakers. In addition, Koda (2000) found that Chinese–English and Korean–English bilinguals used similar strategies in L2 word recognition, both exhibiting intraword structural sensitivity (sublexical level) to a similar extent. Note that studies that found greater reliance on sublexical processing (e.g., Akamatsu, 2002) tend to involve bilinguals with more advanced L2 proficiency than studies that found greater reliance on lexical processing (e.g., Ben-Yehudah et al., 2019; Pae et al., 2017), a point to which we now turn.

### *The effects of L2 proficiency on L2 word reading*

Previous studies have shown that L2 reading strategies are influenced by L2 proficiency (Chikamatsu, 2006; Matsumoto, 2013). Chikamatsu (2006) tested English–Japanese bilinguals to explore whether L2 word recognition strategies underwent restructuring as L2 proficiency increased. The author found that bilinguals with lower L2 proficiency tended to rely more on phonological coding, whereas those with higher L2 proficiency exhibited a greater reliance on visual coding. Additionally, Matsumoto (2013) tested three groups of L2 Japanese readers. The first group's L1 is an alphabetic language (English), with a beginner level in L2. The second group's L1 is a logographic language (Chinese), also with a beginner level in L2. The third group's L1 is an alphabetic language (English), with an intermediate level in L2. The results revealed that the intermediate L2 readers (the third group) showed very similar performance to that of the L2 readers at the beginning level with L1 kanji (logographic characters used in the Japanese writing system, originally borrowed from Chinese characters) knowledge (the second group). This suggests that intermediate L2 readers have begun to restructure their L2 word recognition process, probably due to their increased L2 proficiency.

More recently, researchers have begun to directly investigate the influence of L2 proficiency on sublexical processing in L2 word reading (Botezatu, 2023; Botezatu et al., 2022). Botezatu et al. (2022) examined Spanish–English bilinguals whose L1 and L2 orthographies are similar. Spanish and English both use alphabetic orthography, but the spelling-sound correspondence in Spanish is more consistent than in English. They explored the effects of L2 proficiency on sublexical processing and found that consistency effects decreased with the improvement of L2 proficiency. That is, Spanish–English bilinguals with higher L2 proficiency are less sensitive to sublexical processing in L2 English word reading. Botezatu (2023) further tested Chinese–English and Spanish–English bilinguals in L2 English word reading. The results showed that Chinese–English bilinguals with lower L2 proficiency showed larger consistency effects, whereas Spanish–English bilinguals with higher L2 proficiency exhibited stronger consistency effects. It needs to be noted, however, that the result patterns of the Spanish–English bilinguals in the two aforementioned studies are conflicting. Additionally, Botezatu (2023) found the effects of L2 proficiency on L2 word reading strategies only in ERP data, but not in behavioral data. Thus, the effects of L2 proficiency on L2 word reading strategies require further investigation.

Additionally, differences in the L2 proficiency of the participants may be a major reason for the conflicting findings among Chinese–English bilinguals (Ben-Yehudah et al., 2019; Pae et al., 2017; Akamatsu, 2002). Due to the lack of specific measurement and manipulation of L2 proficiency in these studies, we can only make some rough conjectures. Compared with studies that observed Chinese–English bilinguals relying more on lexical processing in L2 word reading, studies that observed them relying more on sublexical processing tended to involve participants with relatively higher L2 proficiency. For instance, Akamatsu (2002) tested skilled L2 English readers whose L1s were Chinese, Japanese, and Persian and found that the English word reading strategies of these participants were similar to those of native English speakers, with all participants being more sensitive to sublexical processing. Although this study did not measure the participants' L2 proficiency, most of their accuracy rates in the TOEFL vocabulary and reading comprehension sections were above 90%, indicating relatively higher L2 proficiency. In contrast, studies that included participants with relatively lower L2 proficiency found that they relied more on lexical processing. For example, participants in Ben-Yehudah et al. (2019) were international students from China and Korea in the United States, and some of them came from language courses focused on learning and improving English, suggesting that their English proficiency levels were more varied (compared with the participants in Akamatsu, 2002). The participants in Pae et al. (2017) were from local universities in Hong Kong and South Korea, with more limited experience in L2 English. From the above comparison, it can be inferred that as L2 proficiency increases, the L2 word reading strategies of Chinese–English bilinguals may undergo restructuring. When L2 proficiency is lower, bilinguals are more inclined to use lexical processing in L2 word reading. As their L2 proficiency increases, they tend to rely more on sublexical processing. Nevertheless, these assumptions need to be tested through empirical studies.

### *The present study*

The present research focused on bilingual speakers of Chinese and English, two languages that differ significantly in their orthography, to explore the influence of L2 proficiency on sublexical and lexical processing in L2 word reading. Lexical processing was operationalized as word frequency effects, which have been widely used to measure the influence of L2 proficiency at the lexical level (Ben-Yehudah et al., 2019; Bhide, 2015). Sublexical processing was operationalized as consistency effects, which are typically used to measure the influence of L2 proficiency at the sublexical level (Ben-Yehudah et al., 2019; Bhide, 2015; Botezatu, 2023). L2 proficiency was measured using Oxford Quick Placement Test (OQPT), which is a valid and reliable method for measuring L2 proficiency in Chinese–English bilinguals (Zhang, Zhang, Zheng, & Yang, 2020).

On the basis of previous studies, we hypothesize that as L2 proficiency increases, the L2 word reading strategies of Chinese–English bilinguals may undergo a shift from predominantly relying on the lexical level to the sublexical level. Specifically, bilinguals with limited L2 proficiency may rely more on lexical processing and less on sublexical processing. Consequently, they will show greater sensitivity to frequency effects and less sensitivity to consistency effects in L2 word reading. In contrast, bilinguals with higher L2 proficiency may rely more on sublexical processing and less on lexical processing. Consequently, they will exhibit greater sensitivity to consistency effects and less sensitivity to frequency effects in L2 word reading.

## Methods

### *Participants*

Seventy-two native Chinese speakers (45 females, mean age = 20.96 years old, *SD* = 2.05 years old) participated in this study. All participants learned English as the second language. They were undergraduate or graduate students from Xi'an Jiaotong University with normal or corrected-to-normal vision. All participants provided written informed consent and were paid for their participation.

### *Measures*

All participants completed the L2 proficiency test (OQPT) before the experiment. The paper and pen version of the OQPT was used in this study and it took about 30 minutes to complete. This test consists of 60 questions, and the total score is 60 points.

The test scores (mean = 40.26, *SD* = 8.21, range = 24–54) showed that the participants' L2 proficiency levels varied from elementary to advanced (Geranpayeh, 2003). Specifically, 12 participants were at the elementary level (18–29 points), 19 at the lower-intermediate level (30–39 points), 25 at the upper-intermediate level (40–47 points), and 16 at the advanced level (48–54 points).

### *Materials*

The materials consisted of 80 English monosyllabic words (see https://osf.io/3a6c4/), which were the same as the stimuli used in the Experiments 1 and 2 of Jared (1997). The frequency of these words varies, with half having lower frequency and the other half having higher frequency. The detailed frequency information of each word is provided in the supplementary materials. In addition, half of the words (including both low and high frequency) are spelling–sound inconsistent words, and the other half are spelling–sound consistent words. As in prior research (Chee, Chow, Yap, & Goh, 2020; Glushko, 1979), consistency is defined as the mapping from spelling to sound in the current study. The pronunciation of an inconsistent word (e.g., pint, in which "i" is pronounced as /ai/) is different from most words with the same spelling for medial vowels and terminal consonants (e.g., hint, lint, and mint, in which "i" is pronounced as /i/). In contrast, consistent words are pronounced in the same way (e.g., face, race, and pace). Hence, the words used in the current study belong to one of 4 conditions: low-frequency inconsistent (20 items, e.g., steak), low-frequency consistent (20 items, e.g., peer), high-frequency inconsistent (20 items, e.g., break), high-frequency consistent (20 items, e.g., beer). These four conditions were matched for mean number and mean summed frequency of friends, number of letters, and mean bigram frequency according to the norms of Solso and Juel (1980) (see Jared (1997) for details).

To exclude words that participants were unfamiliar with, we recruited 30 additional participants from the same population to rate the familiarity of the 80 English words mentioned above. Their L2 proficiency was similar to those participants in the experiment. The scores of OQPT revealed that 12 individuals were at the upper-intermediate level or above and 18 were at the lower-intermediate level or below. A five-point scale was used in the familiarity rating, with 5 = "very familiar" and 1 = "very unfamiliar". The results showed that the average familiarity rating for each word exceeded 2, with a total average score of 3.79. In line with previous studies (Lin, Cheng, & Wang, 2018; Sun-Alperin & Wang, 2011), all 80 words are suitable to serve as the experimental materials.[1]

### *Procedure*

Participants were individually tested in a quiet laboratory room, seated at an appropriate distance from the computer monitor. The stimuli were presented using E-prime 2.0 (Schneider, Eschman, & Zuccolotto, 2012). During the experiment, all words were presented in white on the black background using Times New Roman (36 pt) in the center of the monitor screen. Participants were asked to read the words aloud as quickly and accurately as possible.

In each trial, a fixation cross first appeared on the screen for 1000 ms, followed by a blank lasting 120 ms. Then, the English word was presented for 3000 ms maximum or until the participants responded. Following previous studies, all stimuli were presented twice (in two sessions, with each word appearing once per session) to obtain enough trials (Christoffels, Timmer, Ganushchak, & La Heij, 2016; Clahsen & Jessen, 2019).[2] The experimental materials were presented pseudo-randomly within each session. The presentation order of the two sessions was counterbalanced across participants. Each session included two runs, and each run consisted of 40 trials. The experiment began with 12 practice trials to help participants better adapt to the task. Reaction times were recorded using a serial response box with voice key, referring to the time from when each stimulus appeared on the screen to when it was overtly pronounced. Meanwhile, accuracy was recorded by the experimenter.

### Data analysis

The reaction time and accuracy of pronunciation for each word were treated as the dependent variables. Based on previous research (Ben-Yehudah et al., 2019), trials with reaction times below 250 ms were considered as outliers, as they might result from the anticipation of words that have not yet appeared on the screen. To examine how L2 proficiency influenced L2 word reading strategies, mixed-effects models were built using the lme4 package (Bates, Mächler, Bolker, & Walker, 2014) and lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) in R version 4.3.2.

#### *Reaction time analysis*

The maximal possible linear mixed-effects model was constructed on reaction time data, using L2 proficiency (continuous variable, OQPT scores), frequency (continuous variable), consistency (categorical variable: consistent vs. inconsistent words), and their interactions as fixed factors. The frequency counts were log transformed using the Zipf-scale (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). The categorical variable was contrast coded (inconsistent $= -.5$, consistent $= .5$), and the continuous variables were centered around their mean values to reduce collinearity (Cunnings, 2012). For random effects, the random intercepts for participants and items, by-participant random slope for the interaction between frequency and consistency, and by-item random slope for L2 proficiency were included. The by-participant random slopes for frequency and consistency, which contributed the least amount of variance, were excluded from the full model due to singular convergence (Linck & Cunnings, 2015).[3]

Next, we adopted the backward model selection method to obtain the best fitting model using the step function of the lmerTest package (Kuznetsova et al., 2017). The step function automatically eliminates the non-significant random and fixed effects by comparing the Akaike information criterion (AIC) (Kuznetsova et al., 2017; Nodari, Celata, & Nagy, 2019). Finally, the best-fitting model includes the fixed effects of L2 proficiency, frequency and their interaction; the random intercepts for participants and items, the by-participant random slope for the interaction between frequency and consistency, and the by-item random slope for L2 proficiency.

Since the interaction between continuous variables is difficult to interpret, we converted L2 proficiency and frequency to categorical variables. Specifically, the participants were divided into a lower L2 proficiency group ($n = 36$) and a higher L2 proficiency group ($n = 36$) using the median of the OQPT scores (41.5) as the cutoff point (Wang & Treffers-Daller, 2017; Whitford & Titone, 2012). Likewise, the words were divided into lower frequency words and higher frequency words using

https://doi.org/10.1017/S0142716424000432 Published online by Cambridge University Press

**Table 1.** The average response time and accuracy for reading four types of words among higher and lower L2 proficiency groups

| | | HFC | | HFINC | | LFC | | LFINC | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | *SD* | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| **RT** (ms) | Higher L2 | 583 | 131 | 607 | 172 | 617 | 149 | 633 | 181 |
| | Lower L2 | 730 | 239 | 741 | 286 | 783 | 301 | 800 | 316 |
| **ACC** | Higher L2 | .96 | .19 | .91 | .29 | .94 | .24 | .81 | .39 |
| | Lower L2 | .83 | .37 | .77 | .42 | .75 | .44 | .59 | .49 |

*Note:* RT = reaction time. ACC = accuracy. *SD* = standard deviation. HFC = high-frequency consistent words. HFINC = high-frequency inconsistent words. LFC = low-frequency consistent words. LFINC = low-frequency inconsistent words.

the median word frequency (18.5) as the cutoff point. The average reaction times and accuracy for each group under each condition are shown in Table 1.

Subsequently, the maximal possible linear mixed-effects model was built using reaction time data, with the same three factors as categorical variables, all of which were contrast coded (e.g., lower level group = −.5, higher level group = .5). Apart from this point, all model building and selection methods were the same as above. At last, the optimal model includes the fixed effects of L2 proficiency, frequency and their interaction; the random intercepts for participants and items, the by-participant random slope for frequency, and the by-item random slope for L2 proficiency. The emmeans package was used to conduct multiple comparisons for significant interaction (Kwon, 2021; Lenth et al., 2022).

### *Accuracy analysis*

Similarly, the maximal possible logistic mixed-effects model was built on accuracy data, including the same fixed factors and random effects as in the maximal linear mixed-effects model described above. This time, L2 proficiency, and frequency were used as continuous variables, and consistency was used as a categorical variable. They were coded the same way as mentioned above. We then used backward model comparison to determine the best-fitting random and fixed effects successively by comparing the AIC of each model (Kuznetsova et al., 2017; Nodari et al., 2019). Finally, the best-fitting model includes the fixed effects of L2 proficiency, consistency, and frequency; the random intercepts for participants and items, the by-participant random slope for consistency, the by-participant random slope for the interaction between frequency and consistency, and the by-item random slope for L2 proficiency.

Next, the maximal possible logistic mixed-effects model was constructed, with all three factors as categorical variables (contrast coding: e.g., lower level group = −.5, higher level group = .5). The methods for model building and model comparison were the same as above. Ultimately, the optimal model includes the fixed effects of L2 proficiency, consistency, their interaction and frequency; the random intercepts for participants and items, the by-participant random slope for consistency, the

https://doi.org/10.1017/S0142716424000432 Published online by Cambridge University Press

**Table 2.** The results of the optimal linear mixed-effects model (with L2 proficiency and frequency as continuous variables)

| Fixed effects | β | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 699.67 | 17.80 | 39.30 | $<$ .001*** |
| L2 proficiency | −12.34 | 2.06 | −6.00 | $<$ .001*** |
| Frequency | −30.76 | 8.99 | −3.42 | .001 |
| L2 proficiency × Frequency | 1.77 | .35 | 4.99 | $<$ .001*** |

*Note:* ****p* $<$ .001, ***p* $<$ .01, **p* $<$ .05.

**Table 3.** The results of the optimal linear mixed-effects model (with all factors as categorical variables)

| Fixed effects | β | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 700.23 | 18.98 | 36.89 | $<$ .001*** |
| L2 proficiency | −174.09 | 36.42 | −4.78 | $<$ .001*** |
| Frequency | −43.70 | 12.93 | −3.38 | .001** |
| L2 proficiency × Frequency | 29.99 | 14.54 | 2.06 | .042* |

*Note:* ****p* $<$ .001, ***p* $<$ .01, **p* $<$ .05.

by-participant random slope for the interaction between frequency and consistency, and the by-item random slope for L2 proficiency.

## Results

### *Reaction time*

When using L2 proficiency and frequency as continuous variables and consistency as a categorical variable, the results of the best-fitting model (see Table 2) showed that the fixed effects of L2 proficiency ($\beta = -12.34$, $SE = 2.06$, $t = -6.00$, $p < .001$), frequency ($\beta = -30.76$, $SE = 8.99$, $t = -3.42$, $p = .001$), and their interaction ($\beta = 1.77$, $SE = .35$, $t = 4.99$, $p < .001$) were significant.

Since both L2 proficiency and frequency were continuous variables, their interaction was difficult to interpret. Consequently, we used all three factors as categorical variables to construct models. The results of the optimal model (see Table 3) revealed that the fixed effect of L2 proficiency was significant ($\beta = -174.09$, $SE = 36.42$, $t = -4.78$, $p < .001$). The higher proficiency group read English words faster than the lower proficiency group. The fixed effect of frequency was significant ($\beta = -43.70$, $SE = 12.93$, $t = -3.38$, $p = .001$), suggesting that participants read high-frequency words faster than low-frequency words. The interaction between L2 proficiency and frequency was also significant ($\beta = 29.99$, $SE = 14.54$, $t = 2.06$, $p = .042$). The speed differences between naming high-frequency words and low-frequency words were larger for the lower proficiency group than for the higher proficiency group (see Figure 1). The post hoc pairwise comparisons were performed on the interaction using the emmeans function

https://doi.org/10.1017/S0142716424000432 Published online by Cambridge University Press

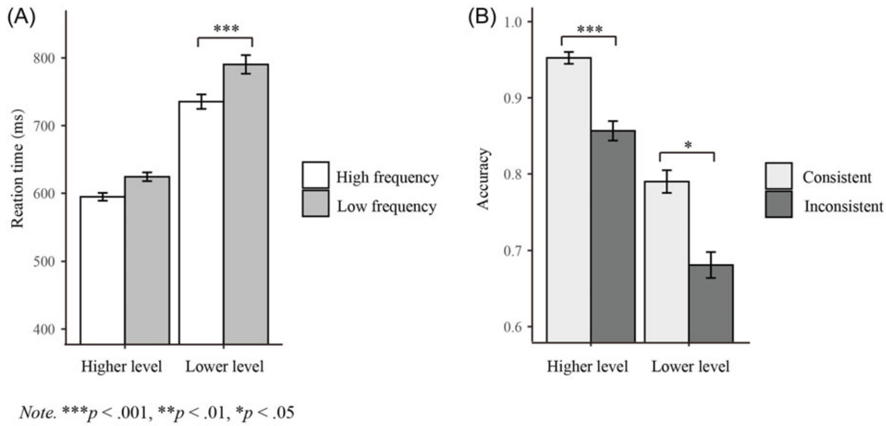


**Figure 1.** Bilinguals with different levels of L2 proficiency showed distinct sensitivities to frequency (A) and consistency (B). The lower level group demonstrated a greater sensitivity to frequency compared to the higher level group, whereas the higher level group showed a stronger sensitivity to consistency compared to the lower level group.

**Table 4.** The results of the optimal logistic mixed-effects model (with L2 proficiency and frequency as continuous variables)

| Fixed effects | β | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept | 3.00 | .22 | 13.57 | < .001*** |
| L2 proficiency | .14 | .01 | 10.32 | < .001*** |
| Consistency | 1.32 | .38 | 3.44 | < .001*** |
| Frequency | 1.15 | .28 | 4.02 | < .001*** |

*Note:* ***$p < .001$, **$p < .01$, *$p < .05$.

(Kwon, 2021; Lenth et al., 2022). The results revealed that the lower level group read high-frequency words faster than low-frequency words ($p < .001$). For the higher level group, the difference in the speed of reading low- and high-frequency words was marginally significant ($p = .053$). These findings suggest that the lower L2 proficiency group is more sensitive to frequency effects than the higher L2 proficiency group. That is, participants with lower L2 proficiency demonstrate a greater reliance on lexical processing in L2 word reading.

### *Accuracy*

The results of the best-fitting model (see Table 4), which included L2 proficiency and frequency as continuous variables and consistency as a categorical variable, revealed that the fixed effects of L2 proficiency ($\beta = .14$, $SE = .01$, $z = 10.32$, $p <$ .001), consistency ($\beta = 1.32$, $SE = .38$, $z = 3.44$, $p < .001$), and frequency ($\beta = 1.15$, $SE = .28$, $z = 4.02$, $p < .001$) were significant.

Subsequently, models were constructed with all three factors as categorical variables. The results of the optimal model (see Table 5) showed that the fixed effect

https://doi.org/10.1017/S0142716424000432 Published online by Cambridge University Press

**Table 5.** The results of the optimal logistic mixed-effects model (with all factors as categorical variables)

| Fixed effects | β | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept | 2.48 | .19 | 13.10 | $< .001^{***}$ |
| L2 proficiency | 1.75 | .22 | 7.99 | $< .001^{***}$ |
| Consistency | 1.14 | .34 | 3.39 | $< .001^{***}$ |
| Frequency | 1.12 | .33 | 3.41 | $< .001^{***}$ |
| L2 proficiency × Consistency | .51 | .27 | 1.90 | .058 |

*Note:* $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

of L2 proficiency was significant (β = 1.75, $SE$ = .22, $z = 7.99$, $p < .001$). The higher level group read English words more accurately than the lower level group. The fixed effect of consistency was significant (β = 1.14, $SE$ = .34, $z = 3.39$, $p < .001$), indicating that participants read consistent words more accurately than inconsistent words. The fixed effect of frequency was significant (β = 1.12, $SE$ = .33, $z = 3.41$, $p < .001$), suggesting that participants read high-frequency words more accurately than low-frequency words. Finally, the interaction effect between L2 proficiency and consistency was marginally significant (β = .51, $SE$ = .27, $z = 1.90$, $p = .058$). The difference in naming accuracy between consistent and inconsistent words was smaller for the lower level group than for the higher level group (see Figure 1). The interaction was further analyzed using the emmeans package (Kwon, 2021; Lenth et al., 2022). The results showed that both the lower level group ($p = .012$) and higher level group ($p < .001$) read consistent words more accurately than inconsistent words. In addition, the higher L2 proficiency group is more sensitive to consistency effects than the lower L2 proficiency group. In other words, participants with higher L2 proficiency show more reliance on sublexical processing in L2 word reading.

## Discussion

This study examined whether the L2 word reading strategies of Chinese–English bilinguals, whose L1 and L2 orthographies are significantly different, were influenced by their L2 proficiency within the DRC model (Coltheart et al., 1993; Coltheart et al., 2001). Lexical and sublexical processing were, respectively, operationalized as frequency and consistency effects in L2 word reading. Our results showed that both lower and higher L2 proficiency bilinguals exhibited significant frequency effects and consistency effects. In addition, bilinguals with lower L2 proficiency were more sensitive to frequency effects, whereas those with higher L2 proficiency were more sensitive to consistency effects. These findings indicate that Chinese–English bilinguals rely on both lexical and sublexical pathways in their L2 word reading and that the collaboration between the two pathways is influenced by L2 proficiency. Bilinguals with lower L2 proficiency rely more on the lexical pathway, while those with higher L2 proficiency rely more on the sublexical pathway. As L2 proficiency increases, their L2 word reading strategies may undergo restructuring.

https://doi.org/10.1017/S0142716424000432 Published online by Cambridge University Press

First, compared with bilinguals with higher L2 proficiency, those with lower L2 proficiency show a greater reliance on lexical processing. Specifically, there was a significant interaction between word frequency and L2 proficiency on reaction time. Bilinguals with lower L2 proficiency showed larger frequency effects than those with higher L2 proficiency. That is, they appear to be more sensitive to frequency effects and rely more on lexical processing in L2 word reading. This result is consistent with previous studies (Ben-Yehudah et al., 2019; Pae & Lee, 2015; Pae et al., 2017). Although lexical processing was manipulated differently (e.g., word frequency effects, scrambled words) in these studies, they all found that native Chinese readers were more sensitive to the lexical level processing in L2 word reading compared with native Korean readers. While these previous studies compared bilinguals with different L1 backgrounds (e.g., Chinese vs. Korean), the present study focused on bilinguals with the same L1 (but different L2 proficiency levels). We found that Chinese–English bilinguals with lower L2 proficiency tend to rely more on lexical processing.

Second, compared with bilinguals with lower L2 proficiency, those with higher L2 proficiency rely more on sublexical processing in L2 word reading. This is reflected in the significant interaction between L2 proficiency and Consistency in accuracy, with consistency effects being larger in bilinguals with higher L2 proficiency than in those with lower L2 proficiency. These results indicate that Chinese–English bilinguals with higher L2 proficiency rely more on sublexical processing in L2 word reading. Our findings are consistent with Akamatsu (2002). Akamatsu (2002) compared bilinguals with different L1 backgrounds and found that all bilingual participants rely more on sublexical processing regardless of their L1 backgrounds. Although Akamatsu did not directly measure the participants' L2 proficiency, the analysis of their TOEFL scores indicated that their L2 proficiency was relatively high.

Taken together, the L2 word reading strategies used by Chinese–English bilinguals seem to be modulated by their L2 proficiency. Bilinguals with lower L2 proficiency rely more on lexical processing, while those with higher L2 proficiency show greater reliance on sublexical processing. These findings suggest that Chinese–English bilinguals' L2 word reading strategies may undergo restructuring as their L2 proficiency improves. Studies on other bilingual populations have also found that L2 reading strategies are influenced by L2 proficiency (Chikamatsu, 2006; Matsumoto, 2013). For example, Chikamatsu (2006) focused on English–Japanese bilinguals and found that participants with lower L2 proficiency showed more phonological reliance, whereas those with higher L2 proficiency exhibited more visual reliance. Another study on L2 Japanese readers also found a restructuring of L2 word reading strategies (Matsumoto, 2013). The current study further demonstrates the restructuring of L2 word reading strategies at the sublexical and lexical processing levels.

To the best of our knowledge, two studies have focused on the influence of L2 proficiency on sublexical processing (consistency effects) in L2 word reading (Botezatu, 2023; Botezatu et al., 2022). However, results from these two studies are somewhat contradictory (see the Introduction section for details). Moreover, Botezatu (2023) found an influence of L2 proficiency on L2 reading strategies in ERP data but not in behavioral data. One reason could be the small sample size of

https://doi.org/10.1017/S0142716424000432 Published online by Cambridge University Press

participants (about 20 in each group), which may not be powerful enough to detect the influence of L2 proficiency. Another possible reason is related to the lexical decision task used in Botezatu (2023), which may not necessarily require access to phonology. The current study tested 72 Chinese–English bilinguals with different L2 proficiency levels. They were instructed to name L2 words aloud, which can more effectively tap into the process of mapping orthography to phonology. Hence, our study provides a more reliable measure of L2 word reading strategies among Chinese–English bilinguals.

Furthermore, this study sheds new light on the DRC model. The DRC model was originally based on English word reading (Coltheart et al., 1993; Coltheart et al., 2001), and researchers have since extended it to Chinese character reading (e.g., Lee et al., 2004). Recent research within this theoretical framework has examined the L2 word learning mechanisms of Chinese–English bilinguals (Cao et al. 2017; Fu et al. 2023). Following this line of research, the current study investigated L2 word reading among Chinese–English bilinguals, revealing that bilinguals with different levels of L2 proficiency use different reading strategies. Bilinguals with lower L2 proficiency rely more on the lexical pathway, while those with higher L2 proficiency rely more on the sublexical pathway. That is, the collaboration between the sublexical and lexical pathways in L2 word reading among Chinese–English bilinguals is modulated by L2 proficiency. The current findings help to expand the explanatory scope of the DRC model, demonstrating that Chinese–English bilinguals' L2 word reading strategies may undergo restructuring as their L2 proficiency increases.

## Limitations and future directions

This study has two main limitations. First, as pointed out by one reviewer, although we tested bilinguals with different levels of L2 proficiency, we used a cross-sectional design. Therefore, we cannot causally reveal the development of L2 word reading strategies in Chinese–English bilinguals as their L2 proficiency improves. Future research should use a longitudinal design to examine whether the L2 word reading strategies used by Chinese–English bilinguals undergo restructuring with the increase of their L2 proficiency.

Second, this study observed that Chinese–English bilinguals' L2 word reading strategies are influenced by their L2 proficiency, shifting from primarily relying on lexical to sublexical processing. However, the reasons underlying such a change remain unclear. We speculate that bilinguals with lower L2 proficiency may employ reading strategies in their L2 that are similar to those used in their L1 Chinese reading, relying more on lexical processing. In contrast, bilinguals with higher L2 proficiency may adopt English word reading procedures similar to those of native English speakers, relying more on sublexical processing. The findings from previous studies lend some support to our hypotheses (Ke & Chan, 2017; Akamatsu, 2002). For example, Ke and Chan (2017) focused on L2 Chinese learners with different proficiency levels. The participants were either from the Chinese cultural sphere (CCS), in which people are influenced by the Chinese orthography and culture (e.g., Korea and Japan), or from the non-Chinese cultural sphere (NCCS). They found that readers from CCS showed an advantage in character decoding strategies compared with NCCS readers at the elementary level. Intriguingly, with an increase in L2 proficiency, this advantage disappeared. Their findings suggest that L2 Chinese readers with lower L2 proficiency are susceptible to the features from their native language in the L2 reading. The current study found that compared with Chinese–English bilinguals with higher L2 proficiency, those with lower L2 proficiency use reading strategies in their L2 that are closer to those in their L1, suggesting that they might be influenced by their native language. Additionally, Akamatsu (2002) tested bilinguals from different L1 backgrounds (Chinese, Japanese, and Persian) with higher L2 (English) proficiency and found that their English word reading strategies were similar to those of native English speakers, demonstrating more reliance on sublexical processing. Our study observed that compared with Chinese–English bilinguals with lower L2 proficiency, those with higher L2 proficiency rely more on sublexical processing, indicating that their English word reading strategies might be closer to those of native English speakers.

Future studies should include two control groups: one consisting of bilinguals with a more transparent L1, such as Spanish–English bilinguals, to reveal the influence of L1 backgrounds on L2 word reading strategies. Specifically, by comparing the L2 word reading strategies of Chinese–English and Spanish–English bilinguals with lower L2 proficiency, we can determine whether these L2 readers tend to use strategies similar to those used in their L1 reading (i.e., Chinese–English bilinguals with lower L2 proficiency rely more on lexical processing, while Spanish–English bilinguals with lower L2 proficiency rely more on sublexical processing), or they use similar strategies regardless of their L1 backgrounds. The other group would be English monolinguals, which could help us determine whether Chinese–English bilinguals with higher L2 proficiency use similar English word reading strategies as English monolinguals. With these two control groups, we might observe that Spanish–English bilinguals with lower L2 proficiency rely more on sublexical processing, similar to the findings of Botezatu et al. (2022). In contrast, Spanish–English bilinguals with higher L2 proficiency might rely more on lexical processing in L2 word reading.

## Conclusion

This study focused on the influence of L2 proficiency on L2 word reading strategies in Chinese–English bilinguals whose L1 and L2 orthographic systems are markedly different. The results show that bilinguals with lower L2 proficiency rely more on lexical processing, while bilinguals with higher L2 proficiency show a greater reliance on sublexical processing. As their L2 proficiency improves, Chinese–English bilinguals' L2 word reading strategies might undergo restructuring. These findings help to expand the explanatory scope of the DRC model and reveal the mechanisms underlying L2 word reading in bilinguals.

**Replication package.** The supplementary materials including stimuli, data, analysis code are available here: https://osf.io/3a6c4/.

**Acknowledgments.** This research was supported by the Postdoctoral Fellowship Program of China Postdoctoral Science Foundation (GZC20232117), the China Postdoctoral Science Foundation (2024M752548), and the Social Science Foundation of Shaanxi Province (2021K013).

**Competing interests.** The authors declare none.

## Notes

**1** We have also explored the influence of word familiarity on the results and found that familiarity significantly improved the fit of each optimal model (*ps* $< .05$, see Table S1 in the supplementary materials), except for one model that failed to converge. However, word familiarity and word frequency were highly correlated ($r = .79$, $p < .001$). In previous research, lexical processing was typically operationalized as frequency effects. Therefore, we chose word frequency as the predictor for data analysis.

**2** All stimuli were presented twice, once in each session. To examine whether stimulus repetition influenced the results, we conducted additional data analysis. First, when constructing a model with stimulus repetition as a fixed factor, its fixed effect was not significant (*ps* $> .05$, see Table S2 in the supplementary materials), indicating that the repetition of stimuli did not induce a learning or practice effect. Second, when stimulus repetition was included as a covariate in the optimal models, it did not improve the model fit (*ps* $> .05$, see Table S3 in the supplementary materials). These suggest that repeating the stimulus twice did not significantly influence the results of this study.

**3** When we attempted to fit the maximal possible linear mixed-effects model, a warning message appeared: "boundary (singular) fit: see help ('isSingular')". This warning indicates that the fitted model is on the boundary of the parameter space, which means that one or more variance components are estimated to be zero or near zero. Therefore, we simplified the model by removing random effects that had variances estimated to be zero or near zero (the by-participant random slopes for frequency and consistency).

## References

**Akamatsu, N.** (2002). A similarity in word-recognition procedures among second language readers with different first language backgrounds. *Applied Psycholinguistics*, **23**(1), 117–133.

**Bates, D., Mächler, M., Bolker, B., & Walker, S.** (2014). Fitting linear mixed-effects models using lme4. arXiv preprint.

**Ben-Yehudah, G., Hirshorn, E. A., Simcox, T., Perfetti, C. A., & Fiez, J. A.** (2019). Chinese-English bilinguals transfer L1 lexical reading procedures and holistic orthographic coding to L2 English. *Journal of Neurolinguistics*, **50**, 136–148.

**Bhide, A.** (2015). Early literacy experiences constrain L1 and L2 reading procedures. *Frontiers in Psychology*, **6**, 1446.

**Botezatu, M. R.** (2023). The impact of L1 orthographic depth and L2 proficiency on mapping orthography to phonology in L2-English: an ERP investigation. *Applied Psycholinguistics*, **44**(2), 237–263.

**Botezatu, M. R., Misra, M., & Kroll, J. F.** (2022). Proficiency in a second language influences processing of print-to-sound mappings. *Linguistic Approaches to Bilingualism*, **14**, 285–309

**Cao, F., Sussman, B. L., Rios, V., Yan, X., Wang, Z., Spray, G. J., & Mack, R. M.** (2017). Different mechanisms in learning different second languages: Evidence from English speakers learning Chinese and Spanish. *Neuroimage*, **148**, 284–295.

**Chee, Q. W., Chow, K. J., Yap, M. J., & Goh, W. D.** (2020). Consistency norms for 37,677 english words. *Behavior Research Methods*, **52**(6), 2535–2555.

**Chikamatsu, N.** (2006). Developmental word recognition: a study of L1 English readers of L2 Japanese. *The Modern Language Journal*, **90**(1), 67–85.

**Christoffels, I., Timmer, K., Ganushchak, L., & La Heij, W.** (2016). On the production of interlingual homophones: delayed naming and increased N400. *Language, Cognition and Neuroscience*, **31**(5), 628–638.

**Clahsen, H., & Jessen, A.** (2019). Do bilingual children lag behind? A study of morphological encoding using ERPs. *Journal of Child Language*, **46**(5), 955–979.

**Coltheart, M., Curtis, B., Atkins, P., & Haller, M.** (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, **100**(4), 589–608.

**Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J.** (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, **108**(1), 204.

**Cunnings, I.** (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, **28**(3), 369–382.

**Fu, Y., Beltrán, D., Bermúdez-Margaretto, B., Dominguez, A., Fu, Y., & Huili, W.** (2023). Language proficiency modulates L2 orthographic learning mechanism: Evidence from event-related brain potentials in overt naming. *Studies in Second Language Acquisition*, 1–22.

**Geranpayeh, A.** (2003). A quick review of the English Quick Placement Test. *Research Notes*, **12**(3), 8–10.

**Glushko, R. J.** (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, **5**(4), 674.

**Grosjean, F.** (2010). *Bilingual: Life and Reality*. Cambridge, Mass: Harvard University Press.

**Jared, D.** (1997). Spelling–sound consistency affects the naming of high-frequency words. *Journal of Memory and Language*, **36**(4), 505–529.

**Ke, S., & Chan, S.-d.** (2017). Strategy use in L2 Chinese reading: The effect of L1 background and L2 proficiency. *System*, **66**, 27–38.

**Koda, K.** (2000). Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics*, **21**(3), 297–320.

**Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B.** (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, **82**(13), 1–26.

**Kwon, H.** (2021). A non-contrastive cue in spontaneous imitation: Comparing mono- and bilingual imitators. *Journal of Phonetics*, **88**, 101083.

**Lee, C.-Y., Tsai, J.-L., Kuo, W.-J., Yeh, T.-C., Wu, Y.-T., Ho, L.-T., … Hsieh, J.-C.** (2004). Neuronal correlates of consistency and frequency effects on Chinese character naming: an event-related fMRI study. *Neuroimage*, **23**(4), 1235–1245.

**Lenth, R., Buerkner, P., Herve, M., Love, J., Miguez, F., Riebl, H., & Singmann, H.** (2022). Emmeans: Estimated marginal means, aka least-squares means. *R package version*, **1**(1), 3.

**Lin, C. Y., Cheng, C., & Wang, M.** (2018). The contribution of phonological and morphological awareness in Chinese–English bilingual reading acquisition. *Reading and Writing*, **31**(1), 99–132.

**Linck, J. A., & Cunnings, I.** (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, **65**(S1), 185–207.

**Matsumoto, K.** (2013). Kanji recognition by second language learners: Exploring effects of first language writing systems and second language exposure. *The Modern Language Journal*, **97**(1), 161–177.

**Nodari, R., Celata, C., & Nagy, N.** (2019). Socio-indexical phonetic features in the heritage language context: Voiceless stop aspiration in the Calabrian community in Toronto. *Journal of Phonetics*, **73**, 91–112.

**Pae, H. K., & Lee, Y.-W.** (2015). The resolution of visual noise in word recognition. *Journal of Psycholinguistic Research*, **44**(3), 337–358.

**Pae, H. K., Sun, A. K., Mano, Q. R., & Kwon, Y.-J.** (2017). Sublexical and lexical processing of the English orthography among native speakers of Chinese and Korean. *Reading and Writing*, **30**(1), 1–24.

**Pasquarella, A., Chen, X., Gottardo, A., & Geva, E.** (2014). Cross-language transfer of word reading accuracy and word reading fluency in Spanish-English and Chinese-English bilinguals: Script-universal and script-specific processes. *Journal of Educational Psychology*, **107**(1), 96.

**Perfetti, C. A., Liu, Y., & Tan, L. H.** (2005). The lexical constituency model: some implications of research on Chinese for general theories of reading. *Psychological Review*, **112**(1), 43.

**Schneider, W., Eschman, A., & Zuccolotto, A.** (2012). *E-Prime User's Guide*. Psychology Software Tools.

**Sun-Alperin, M. K., & Wang, M.** (2011). Cross-language transfer of phonological and orthographic processing skills from Spanish L1 to English L2. *Reading and Writing*, **24**(5), 591–614.

**Tan, L. H., Laird, A. R., Li, K., & Fox, P. T.** (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. *Human Brain Mapping*, **25**(1), 83–91.

**Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M.** (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, **67**(6), 1176–1190.

**Wang, Y., & Treffers-Daller, J.** (2017). Explaining listening comprehension among L2 learners of english: the contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. *System*, **65**, 139–150.

**Whitford, V., & Titone, D.** (2012). Second-language experience modulates first-and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review*, **19**(1), 73–80.

**Zhang, C., Zhang, H., Zheng, X., & Yang, J.** (2020). A comparative study of English proficiency testing in Chinese-English bilinguals. *Foreign Language Teaching and Research*, **52**(5), 701–712.

**Cite this article:** Zhao, R., Wei, H., & Yao, Z. (2024). The effects of L2 proficiency on L2 word reading strategies: evidence from Chinese–English bilinguals. *Applied Psycholinguistics* **45**, 1147–1164. https://doi.org/10.1017/S0142716424000432