



Is letter position coding when reading in L2 affected by the nature of position coding used when bilinguals read in their L1?

Huilan Yang¹ · Debra Jared² · Manuel Perea^{3,4} · Stephen J. Lupker²

Accepted: 5 December 2020 / Published online: 19 January 2021
© The Psychonomic Society, Inc. 2021

Abstract

Using four-character Chinese word targets, Yang, Chen, Spinelli, and Lupker (*Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(8), 1511–1526, 2019) and Yang, Hino et al. (*Journal of Memory and Language*, 113, 104017, 2020) demonstrated that backward primes (Roman alphabet example—*dcb* priming *ABCD*) produce large masked priming effects. This result suggests that character position information is quite imprecisely coded by Chinese readers when reading in their native language. The present question was, If Chinese readers have evolved a reading system not requiring precise position information, would Chinese–English bilinguals show more extreme transposed letter priming effects when processing English words than both English monolinguals and other types of bilinguals whose L2 is English? In Experiment 1, Chinese–English bilinguals, but not English monolinguals, showed a clear backward priming effect in a lexical decision task. In Experiment 2, the parallel backward priming effect was absent for both Spanish–English and Arabic–English bilinguals. Apparently, the orthographic coding system that Chinese–English bilinguals use when reading in their L2 leans heavily on the flexible/imprecise position coding process that they develop for reading in their L1.

Keywords Backward priming effect · Lexical decision task · Chinese · English

Considerable research effort has been expended in recent years to try to understand the orthographic coding process (Grainger, 2018). The orthographic code is defined as the code that is established early in reading a word that codes both the identities of the letters contained in the word and their positions. It is assumed that it is this code that then allows

activation of higher-level codes (e.g., lexical, phonological, semantic). In most languages, this code must contain accurate letter identity and letter position information in order to allow readers to distinguish between words like *tame* and *fame* or between words like *trial* and *trail*.

The task most commonly used to investigate the orthographic coding process is the masked priming lexical decision task (Forster & C. Davis, 1984). In this task, a forward mask, consisting of multiple hash marks (e.g., #####), is typically presented for 500 ms, followed by a lowercase prime that is presented so briefly (e.g., < 70 ms) that it is typically unavailable to consciousness, followed by an uppercase target that the participant must classify as a word or a nonword. When the prime is a nonword that is orthographically similar to the target (e.g., *honse*–*HOUSE*), responding is typically more rapid than when the prime and target are not orthographically similar (e.g., *tairn*–*HOUSE*).

In general, orthographic priming has been thought to be a lexical phenomenon. That is, the standard account of such effects (e.g., C. J. Davis, 2010; Gómez, Ratcliff, & Perea, 2008) is that the orthographic code created during prime processing activates lexical representations of words with similar orthography. As a result, access to the target's lexical representation is facilitated when the orthographic codes of the

✉ Stephen J. Lupker
lupker@uwo.ca

Huilan Yang
yanghuilanhn@hotmail.com

Debra Jared
djared@uwo.ca

Manuel Perea
Manuel.Perea@uv.es

¹ Department of Foreign Languages, Zhejiang Gongshang University, Hangzhou 310018, Zhejiang, China

² Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2, Canada

³ Departamento de Metodología, Facultad de Psicología, Universitat de València, Av. Blasco Ibáñez, 21, 46010 Valencia, Spain

⁴ Universidad Nebrija, Madrid, Spain

prime and target are similar. An alternative, sublexical account, that the effect is due to the prime facilitating target processing by activating the target's letter-level representations, has received little support since significant orthographic priming effects for nonword targets are typically nonexistent in lexical decision tasks (Forster, 1998; Forster & Davis, 1984). As will be discussed subsequently, the question of whether there was a priming effect for nonword targets in the present experiments will be considered in some detail.

The issues central to the present paper concern priming effects that emerge when the orthographically similar primes are created by transposing letters in their targets. In the simplest case, transposed letter (TL) primes are created by transposing two adjacent letters in their targets (i.e., *jugde*–*JUDGE*). The control condition typically involves primes containing all the letters that are in the same positions in the TL primes and their targets while substituting other letters for the two transposed letters (e.g., *jupte*–*JUDGE*). A faster average latency in the former condition than in the latter condition is referred to as a TL priming effect. Perea and Lupker (2003, 2004) provided some of the earliest demonstrations of TL priming effects as well as demonstrating that transpositions producing TL effects can involve nonadjacent letters (e.g., “*caniso*” is an effective TL prime for *CASINO*).¹

TL priming effects, which has been observed in many languages (e.g., in English, Perea & Lupker, 2003; in Spanish, Perea & Lupker 2004; in Chinese, Yang, Chen, Spinelli, & Lupker, 2019), indicate that letter/character position coding is not totally precise in those languages. If the position coding process were precise, “*jugde*” would be no more similar to *JUDGE* than “*jupte*” is and hence should provide no more priming for *JUDGE* than “*jupte*” does. Specifically, if position coding were precise, the existence of the *g* and the *d* in “*jugde*” would be irrelevant to the processing of *JUDGE*, because those two letters would be coded as being in different positions in (and hence being different elements in) the prime and target.

Following Perea and Lupker's (2004) demonstration of TL priming from nonadjacent transpositions, Guerrero and Forster (2008) provided an extensive investigation of how extreme a transposition can be and still produce priming. Their results showed that the system for English readers is tolerant of some rather extreme transpositions. For example, primes created by pairwise transposing the middle six letters of an eight-letter word (i.e., *sdiwelak*–*SIDEWALK*) were effective primes. However, there was also a limit, as primes created by pairwise transposing all the letters (i.e., *isedawkl*–*SIDEWALK*) or reversing the four letters in both the first half

and second half of the target (i.e., *edisklaw*–*SIDEWALK*) did not produce priming (in all cases, the control condition involved a prime sharing at most one letter with the target).

In a further examination of this issue with English readers, C. Davis, Kim, and Forster (2008) used both forward (e.g., “*face*”) and backward (e.g., “*ecaf*”) primes for backward targets (*ECAF*) in a masked priming lexical decision task. Participants had to decide whether, if the letters in the target were read backwards, the target would be a word or a nonword. C. Davis et al. found that priming only arose when the backward (word) targets were preceded by a forward prime. Hence, even with backward targets, which undoubtedly require considerable processing, backward primes, the types of primes to be investigated in the present research, were ineffective for English readers.

More recently, Yang and colleagues (Yang et al., 2019; Yang et al., 2020) have provided an examination of extreme transposition priming in Chinese. Those authors demonstrated that, using four-character Chinese words as targets, primes that were the targets written backwards (i.e., using the Roman alphabet to reflect this relationship, it would be “*dcba*” priming *ABCD*) produce 50+ ms priming effects, although those effects were still significantly smaller than the identity priming effect that Yang et al. (2019) reported (i.e., *abcd*–*ABCD* – 80 ms).

Subsequently, Yang et al. (2020) demonstrated that these backward priming effects for Chinese readers are neither phonologically nor morphologically based. That is, because a backward prime in Chinese contains all the syllables and all the morphemes of the target (in a backward order), Yang et al. (2020) suggested that Yang et al.'s (2019) backward priming effect might not be entirely orthographically based. The results of their five experiments, however, appeared to rule out any contribution of either phonologically/syllabically or morphologically/meaning-based priming. Instead, the backward priming effects that Yang and colleagues observed appear to be orthographically based.

These results, juxtaposed with Guerrero and Forster's (2008) and C. Davis et al.'s (2008) results suggest a clear difference between orthographic coding in Chinese versus English. Specifically, they indicate that, for Chinese readers, the position coding component of the orthographic coding system is much more flexible (i.e., much less precise) than it is for English readers (see also Gu, Li, & Liversedge, 2015; Lally, Taylor, Lee, & Rastle, 2020; Lerner, Armstrong, & Frost, 2014; Taft, Zhu, & Peng, 1999), although, as noted, some position information must be coded by Chinese readers because presenting the prime characters in the correct order (*abcd*–*ABCD*) does produce more priming than presenting them backwards (*dcba*–*ABCD*).

As Yang et al. (2020) note, a reasonable explanation of these data patterns is that the orthographic coding system for Chinese readers is much more focused on getting character

¹ A similar pattern of results has been found in eye-movement studies, with parafoveal previews of TL primes producing a greater facilitation of target processing than that of SL primes when evaluating a number of different viewing duration measures (e.g., first fixation duration, gaze duration) (in English, Johnson, Perea, & Rayner, 2007; in Chinese, Gu et al., 2015).

identity correct than getting character position correct than is the system for readers of most alphabetic-script languages. In any alphabetic-script language there are a limited number of letters, whereas in Chinese there are more than 30,000 characters. Therefore, there would appear to be a much larger challenge to getting identity information correct in Chinese compared with that in alphabetic-script languages. Further, there are a large number of anagrams in alphabetic-script languages (e.g., *pots*, *stop*, *opts*, *tops*), whereas there are very few in Chinese. For example, even though there are two-character words in Chinese with large neighborhoods (e.g., 大米, 大学, 大象, and 大选), 97% of two-character Chinese words are not anagrams (i.e., they do not make another word when the character positions are transposed). Longer Chinese words (e.g., the four-character words used by Yang and colleagues) do not have anagrams, and few even have neighbors. Therefore, even if a Chinese reader were to get the position information in the word 羊亡牢补 wrong, it would be obvious what the actual word was because there are no other words containing those four characters.

It makes sense, therefore, that there would be a clear difference in the nature of the orthographic coding systems of Chinese versus English readers. Indeed, Lally et al. (2020) have provided experimental evidence that the precision of a reader's position coding depends on the nature of their writing system. In their experiments, participants were trained in an artificial language that had either no anagrams (sparse orthography) or many anagrams (dense orthography). Participants learning the language with the sparse orthography produced a larger transposed letter priming effect than participants learning the language with the dense orthography, indicating that the letter position coding process that the participants evolved was less precise in the former group than in the latter.

The goal of the present research is to take this idea a step further. If Chinese readers do not have an orthographic coding system that has been trained to require as precise coding of position information as readers of English (and likely other alphabetic-script languages), one may very well be able to see the footprint of that experience in the behavior of Chinese–English bilinguals. In particular, one might expect that those individuals would show more extreme transposition priming effects when processing English words than would English monolinguals.

This issue was examined in Experiment 1. The targets in all the present experiments were four-letter or five-letter English words. The orthographically similar primes were those targets written backwards. Based on prior literature (e.g., C. Davis et al., 2008; Guerrero & Forster, 2008), English monolinguals should show no priming from these primes. In contrast, if the orthographic coding system of the Chinese–English bilinguals reflects the processes they have developed from over a decade

of reading Chinese, one would expect to see evidence of priming from these backward primes.

To look ahead, Chinese–English bilinguals did show a clear backward priming effect in Experiment 1, whereas English readers did not. Experiments 2a and 2b were then attempts to determine whether this effect also arises in other bilinguals using the same English primes and targets (i.e., to determine whether this effect truly is a Chinese L1 effect). Experiment 2a involved Spanish–English bilinguals. Potentially, the effect observed in Experiment 1 was merely an L2 effect. If so, one would expect it to show up in any bilingual whose L2 is English.

In Experiment 2b, Arabic–English bilinguals were tested. Arabic employs an Abjad script, which is quite similar to alphabetic script. That is, consonants are represented as letters although vowels are often not. Crucially, however, Arabic is written right-to-left. Therefore, Arabic–English bilinguals have experience reading words written in what is a backward direction for readers of most alphabetic-script languages. Chinese readers have a similar skill because Chinese words are written in a top-to-bottom direction as well as (rarely) a backward direction. If the reason that Chinese readers show a backward priming effect when reading in English is that they have experience reading words written in a number of different directions, Arabic–English bilinguals might also have gained a similar ability and, hence, may show a backward priming effect in Experiment 2b.

Experiment 1

Method

Participants Forty-two Chinese–English bilinguals from Western University (London, Ontario, Canada) participated in this experiment in return for receiving 0.5 credits in an introductory psychology course. All participants indicated that they are proficient in reading Simplified Chinese as well as reading English words written in uppercase Roman letters. All participants were born in China, had lived in mainland China for at least 10 years, and had Chinese as their first language. Their mean age was 19 years ($SD = 1.73$) and they moved to Canada at a mean age of 17 years ($SD = 2.48$). Thirty-three participants had completed the IELTS and had a mean score of 6.5, and six participants had completed the TOEFL and had a mean score of 107. Participants' self-ratings of the percentage of the day spent using each language and their self-rating skill in each language (from 1 = *none* to 10 = *very fluent*) are presented in Table 1. All had normal or corrected-to-normal vision, with no reading disorders.

Table 1 Language experience for Chinese–English bilingual participants

	Chinese	English
% of use in each of the following contexts		
Home	92.36	5.83
School	35.60	63.05
Other	50	42.02
Self-rating of current skill (out of 10)		
Understanding	9.67	8.12
Speaking	9.67	7.26
Reading	9.26	7.64
Writing	8.62	6.86

Forty-two English monolinguals from Western University also participated in this experiment in return for receiving 0.5 credits in an introductory psychology course. All English monolingual participants indicated that English was their first language and the only language that they used on a regular basis.²

Materials A total of 120 single-syllable English four or five letter words were selected as word targets. The words had an average SUBTLEX-US frequency per million words (Brysbaert & New, 2009) of 371.3 (range: 101.1–2691.4), an average orthographic neighborhood size (Coltheart, Davelaar, Jonasson, & Besner, 1977) of 2.6 (range: 0–5), and an average word length of 4.7 letters (range: 4–5 letters). The word targets and their associated primes are listed in Appendix 1.

In addition, 120 single-syllable four-letter or five-letter nonword targets were selected from the English lexicon project database (Balota et al., 2007). For both word and nonword targets, each target was preceded by two different types of primes, counterbalanced across participants: (1) a backward prime—that is, a nonword containing the same letters as the target with those letters being presented in a right-to-left direction (e.g., *naelc*–*CLEAN*, the backward condition); (2) an unrelated nonword prime created by selecting a different word or nonword target and presenting its letters in a backward (i.e., right-to-left) direction (e.g., *gnuoy*–*CLEAN*, the backward

unrelated condition). The unrelated primes for any given participant were based on a set of targets that were not presented to that participant (but were presented to other participants), as described below.

Each participant was presented with only 80 (of the 120) word and nonword targets, counterbalanced across participants. To accomplish the relevant counterbalancing, the word and nonword targets were each divided into three lists with 40 stimuli in each list. Two of those lists of targets were presented to each participant. The targets in one of the lists were preceded by their backward primes. The targets in the other list were preceded by unrelated primes that were the backward primes for the targets in the list not being used for that participant. The lists were rotated across participants, so that each list of targets was primed by backward primes for one third of the participants, was primed by unrelated primes for one third of the participants, and did not appear for the remaining one third of the participants. The reasons for using this procedure were that (a) we wanted to use the same primes (across participants) for the related and unrelated trials, but (b) we did not want participants to be presented with what would be the backward primes for their unrelated targets at any time in the experiment.

Procedure E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA; see Schneider, Eschman, & Zuccolotto, 2002) was used for data collection. Stimuli were presented on a computer monitor with a refresh rate of 60 Hz. Primes and targets were presented in black 35-pt Courier New typeface on a white background. Each trial began with a mask (#####) presented in the center of the screen for 500 ms, followed by a lowercase prime for 50 ms. The uppercase target then appeared, remaining on the screen until a response was made or for 3,000 ms. Participants were instructed to decide as quickly and as accurately as possible whether each presented uppercase letter string is a real English word, and to indicate their decision by pressing the button on the keyboard labelled “word” if the presented letter string is a word and the button labelled “nonword” if it is a nonword. The order of the presentation of stimuli was randomized for each participant. The experimental block included 160 trials in total. Participants completed 16 practice trials with feedback before starting the experimental block. The research reported as Experiments 1 and 2b was approved by the Western University REB (Protocol #s 108835 and 104255).

Results

Response latencies that were less than 300 ms, more than three standard deviations from the participant’s mean latency (2.1% of the data for the word targets, 1.8% of the data for the nonword targets), or from incorrect trials (3.4% of the data for the word targets, 5.4% of the data for the nonword targets) were excluded from the latency

² Because we had no a priori hypothesis concerning the size of the backward priming effect for our Chinese–English bilinguals, our initial method for determining whether we had a sample size that would produce an appropriate level of power in each of the experiments was based on Brysbaert and Stevens’s (2018) suggestion that there should be at least 1,600 trials in each condition. Once the data had been collected in Experiment 1, we did a subsequent power analysis using the *simr* package (Green & Macleod, 2016) based on the obtained (16 ms) difference in priming effect sizes for our two groups of participants in order to obtain a power estimate for the interaction. The calculated power is 79% (95% CI [77%, 82%]) for that interaction between relatedness and language group based on a simulation involving 1,000 samples with an alpha of 0.05.

analyses. Linear mixed-effects models and generalized linear mixed-effects models from the lme4 package in R were used to analyze the latency and error rate data, respectively (Baayen, 2008; Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000), with the *p* values for the latency results coming from the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). As is typical when using linear mixed-effects models to analyze latency data, the latencies were inverted in order to “normalize” the distribution in all the experiments. Doing so also helped the models converge. The *p* values for the glmer models for the error rate results were reported by default in the lme4 package. The emmeans package was used to conduct the post hoc analyses (Lenth, Singmann, Love, Buerkner, & Herve, 2020). Subjects and items were treated as random effects. Relatedness (related vs. unrelated) and language group (Chinese–English bilinguals vs. English monolinguals) were treated as fixed effects. Before running the model, R-default treatment contrasts were altered to sum-to-zero contrasts (Levy, 2014; Singmann & Kellen, 2020). The R code used in the analyses for each experiment are shown in Appendix 2. The mean latencies and percentage of errors from a subject-based analysis are shown in Table 2. The detailed information concerning the results of our analyses is reported in Table 3.

Word data In the latency analysis, the main effect of relatedness was not significant, $\beta = -0.005$, $SE = 0.004$, $t = -1.46$, $p = .144$, whereas the main effect of language group was

Table 2 Mean lexical decision latencies (RTs, in milliseconds) and percentage error rates in Experiment 1 (with standard deviations in parentheses)

Condition	RT	%E
Chinese–English bilinguals word targets		
Backward condition	798 (152)	4.0 (5)
Backward unrelated condition	812 (137)	3.9 (4)
Priming	14	-0.1
English monolinguals word targets		
Backward condition	569 (64)	2.7 (3)
Backward unrelated condition	567 (68)	2.9 (3)
Priming	-2	0.2
Chinese–English bilinguals nonword targets		
Backward condition	1,069 (256)	9.0 (7)
Backward unrelated condition	1,079 (320)	8.0 (7)
Priming	10	-1
English monolinguals nonword targets		
Backward condition	636 (89)	2.2 (3)
Backward unrelated condition	635 (87)	2.3 (3)
Priming	2	0.1

RT = reaction time; %E = percentage error rate

Table 3 Details of random effects and multicollinearity in Experiment 1, latency data

Groups name	Variance	SD	Corr
Word random effects			
Item (intercept)	0.006	0.080	
Language group/item	0.001	0.035	0.63
Participant (intercept)	0.039	0.197	
Residual	0.079	0.282	
Correlation of fixed effects			
(Intr)		Relatedness	Language group
Relatedness	0.000		
Language group	0.029	0.000	
Interaction	0.000	0.009	0.000
Nonword random effects			
Item (intercept)	0.008	0.091	
Relatedness/item	0.001	0.030	0.33
Participant (intercept)	0.061	0.247	
Residual	0.085	0.292	
Correlation of fixed effects			
(Intr)		Relatedness	Language Group
Relatedness	0.001		
Language group	0.010	0.001	
Interaction	0.001	0.033	0.001

The number of observations was 6,356 and 6,244 for word and nonword trials, respectively

significant, $\beta = 0.235$, $SE = 0.022$, $t = 10.67$, $p < .001$. Latencies were faster for English monolinguals than for bilinguals. More importantly, the interaction between relatedness and language group was significant, $\beta = -0.010$, $SE = 0.004$, $t = -2.78$, $p = .005$. In the analysis of the interaction, the 14-ms difference between the backward prime (798 ms) and the unrelated prime (812 ms) conditions was significant for the Chinese–English bilinguals, $\beta = -0.030$, $SE = 0.010$, $t = -2.99$, $p = .003$; however, the -2-ms priming effect (569 vs. 567 ms) for the English monolinguals was not, $\beta = 0.009$, $SE = 0.010$, $t = 0.94$, $p = .347$. In the error rate analysis, the main effect of language group approached significance, $\beta = -0.189$, $SE = 0.107$, $t = -1.76$, $p = .078$, whereas the main effect of relatedness and the interaction between relatedness and language group did not, both $ps > .10$.

Nonword data For the latency analysis, the model converged after a restart. The other details were the same as those for the word targets. In the latency analysis, only the main effect of language group was significant, $\beta = 0.283$, $SE = 0.027$, $t = 10.36$, $p < .001$. Latencies were faster for the English monolinguals than for the Chinese–English bilinguals. Neither the main effect of relatedness nor the interaction between relatedness and language group approached significance, all $ps > .10$. In the error rate analysis, only the main effect of language

group was significant, $\beta = -0.722$, $SE = 0.105$, $t = -6.89$, $p < .001$. Fewer errors were produced by English monolinguals. Neither the main effect of Relatedness nor the interaction between relatedness and language group approached significance, both $ps > .10$.³

Bayes factor analyses We also conducted a Bayes factor analysis for both word and nonword targets in order to quantify the statistical evidence in favor of or against the relatedness effects for both language groups. Bayes factor analyses involve calculating an estimate of the likelihood that an effect is real (i.e., that the alternative hypothesis is correct) versus the likelihood that the null hypothesis is correct, and combining those values in a ratio. The Bayes factor analysis was calculated using the *lmBF* function of the BayesFactor package with default JZS type being used to calculate the Bayes factor (Morey, Rouder, Jamil, & Morey, 2015). The analysis was based on trial-level inverted latencies. For the Chinese–English bilinguals, Model 1 (the full model with a main effect of relatedness) was compared with Model 0 (the null model with no main effect) using the formula “Model1/Model0.” For word targets, the contrast between these two models produced a BF_{10} of $2.88 \pm 1.62\%$, a result favoring the hypothesis of a relatedness effect. For nonword targets, Model 0 was preferred based on a calculated BF_{10} of $0.04 \pm 0.96\%$, strongly favoring the hypothesis that there was no relatedness effect for nonword targets.

The same procedure was used for the English monolingual data. Model 1 with a main effect of relatedness was compared with Model 0 with no main effect of relatedness. For word targets, Model 0 was strongly preferred based on a calculated BF_{10} of $0.06 \pm 1.12\%$. For nonword targets, Model 0 was again strongly preferred based on a calculated BF_{10} of $0.04 \pm 0.86\%$. Thus, for English monolinguals, there was very clear evidence for a null effect of relatedness for both words and nonwords.

Finally, we evaluated whether the significant priming effect for the Chinese–English bilinguals followed the normal pattern expected for orthographic priming effects (e.g., Taikh & Lupker, 2020)—that is, that the impact of an orthographically related prime is to shift the latency distribution by a constant due to the fact that the backward prime essentially creates a head start in target processing. To do so, we analyzed the nature of the word target priming effect for the Chinese–English bilinguals using quantile plots for each condition. The graphs of the latencies as a function of quantile are shown in Fig. 1.

In order to examine the quantile data statistically, we added quantile group as a fixed factor in an analysis of the Chinese–English bilinguals’ word data. The quantile group factor had five levels, with eight trials in each of these levels (i.e., 40 stimuli in each condition). Because not all participants

provided eight trials in both conditions in the fifth quantile (range: 1–8), we did not include the data from this quantile in our analysis, although the results for that quantile are shown in Fig. 1. The function *anova* in the *car* package (Fox & Weisberg, 2016) was used to test for significance and to provide the p values for this analysis.

In the quantile group analysis, the two main effects of relatedness and quantile group were both significant (all $ps < .01$), however, the Relatedness \times Quantile Group interaction failed to approach significance, $\chi^2(3) = 3.13$, $p = .372$. Thus, these results indicate that the overall priming effect was essentially constant across quantiles.

Discussion

As expected, English monolinguals showed no evidence of a backward priming effect. In contrast, the Chinese–English bilinguals showed a clear backward priming effect for word targets. This result nicely supports the idea that these individuals’ orthographic coding system when reading English does not involve as precise a letter position coding scheme as those developed by monolingual readers of English.

Assuming this conclusion is correct, one question that will need to be addressed is whether the effect is a lexical effect (i.e., the letters in the prime activate the relevant lexical representation, allowing more rapid lexical processing even though the letter order is backward) or a sublexical effect (i.e., the letters in the prime make the letters in the target easier to process even though they are in different orders in the two strings). As noted previously, there is little, if any, evidence for sublexical priming in prior masked priming lexical decision experiments. Hence, the general assumption is that orthographic priming is a lexically based effect. However, that literature is mainly based on priming experiments involving L1 stimuli and, with respect to Chinese–English bilinguals, the lexical versus sublexical question may be more relevant. As in virtually all masked priming experiments, the targets in Experiment 1 appeared in uppercase, an unfamiliar format in English language reading instruction in China. Therefore, it is possible that the lowercase letters appearing in the prime might have made it easier for participants to recognize the unfamiliar uppercase letters in the target (i.e., the effect in Experiment 1 arose at the sublexical level). As such, a more detailed analysis of the question of the locus of the priming effect will be undertaken following the presentation of Experiment 2.

As noted previously, the reason that one would have suspected a priming effect for Chinese–English bilinguals is that position coding when reading Chinese is almost unnecessary, as there are very few anagrams in that language. Thus, most sets of characters are uniquely identifiable as a particular word, regardless of what order those characters appear in. However, there are other possible explanations for why there was an effect for the Chinese–English bilinguals in

³ The results of an additional analyses of these data are presented in Footnote 6 in the General Discussion.

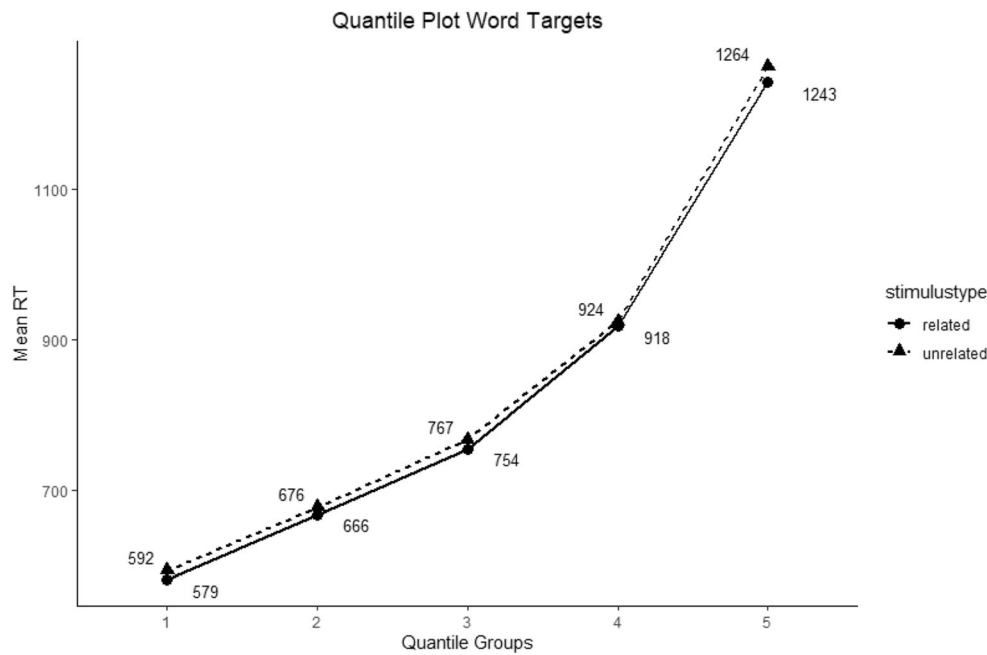


Fig. 1 Quantile plot for Chinese–English bilinguals, word targets. *Note.* The priming effects for the Quantile Groups 1 to 5 were 13 ms, 10 ms, 13 ms, 6 ms, and 21 ms, respectively

Experiment 1—explanations that have little to do with the nature of character position coding in Chinese.

One explanation is that the effect arose simply because the Chinese–English bilinguals were performing the task in a language that they were not as familiar with as they are with their L1. This idea was evaluated in Experiment 2a by determining whether there is a backward priming effect (using English stimuli) for Spanish–English bilinguals. If, as hypothesized, the effect for Chinese–English bilinguals is due to the fact that position coding is unimportant in reading Chinese but is considerably more important for readers of alphabetic-script languages, one would not expect to find a backward priming effect in Experiment 2a.

Experiment 2b allowed an examination of an additional alternative explanation of the results in Experiment 1. As noted, Chinese readers are exposed to words written in different directions. In particular, although most words are written in a left-to-right direction, some are written top-to-bottom or (rarely) right-to-left. Therefore, those individuals presumably have had considerable experience reading words in other than a left-to-right direction. Simple exposure to text presented in various directions could, therefore, be what was responsible for Chinese–English bilinguals producing a backward priming effect in Experiment 1.

Arabic script is written right-to-left. Therefore, Arabic–English bilinguals have had considerable experience reading words written in both right-to-left and left-to-right directions. If the backward priming effect observed for Chinese–English bilinguals in Experiment 1 were due to them having learned to read words written in multiple directions, Arabic–English

bilinguals, who also have had this experience, should show a backward priming effect in Experiment 2b.

On the other hand, Arabic script is unlike in Chinese script in that precise letter position information is needed for successful reading. As such, Arabic readers should have developed an orthographic coding system in which letter position information is coded in a fairly precise way. Therefore, if the reason that the Chinese–English bilinguals showed a backward priming effect is that their position coding process when reading in their L2 (English) shows the impact of their (fairly imprecise) L1 position coding system, Arabic–English bilinguals should not show a backward priming effect in Experiment 2b.

Experiment 2a

Method

Participants Forty-two Spanish–English bilinguals from Universitat de València (Valencia, Spain) participated in this experiment in return for receiving 0.5 credits in an introductory psychology course. They were recruited from a high-academic performance school in which nearly all the courses are taught in English. All Spanish–English bilingual participants indicated that they regarded themselves as proficient in reading both Spanish and English words written in uppercase Roman letters and that Spanish was their first and dominant language. They all had normal or corrected-to-normal vision and no reading disorders. Six participants had achieved a C2 level on the CEFR test (Common European Framework of

Reference for Languages), 14 had achieved a C1 level on the CEFR test, 12 had achieved a B2 level on the CEFR test, and 10 had achieved a B1 level on the CEFR test.⁴

Materials The stimuli were the same as in Experiment 1.

Procedure The procedure was virtually the same as in Experiment 1. The two minor differences were that DMDX software (Forster & Forster, 2003) was used for data collection instead of E-Prime software and that all primes and targets were presented in 14-pt Courier New typeface. This research was approved by the Universitat de València REB.

Results

Response latencies that were less than 300 ms, more than three standard deviations from the participant's mean latency (2.1% of the data for the word targets, 1.7% of the data for the nonword targets), or from incorrect trials (3.8% of the data for the word targets, 6.2% of the data for the nonword targets) were excluded from the latency analyses. Subjects and items were treated as random effects. Relatedness (related vs. unrelated) was treated as a fixed effect. The mean RTs and percentage of errors from a subject-based analysis are shown in Table 4. The detailed information concerning the results of our analyses is reported in Table 5.

Word targets The –5-ms relatedness effect was not significant in the latency analysis, $\beta = 0.009$, $SE = 0.006$, $t = 1.42$, $p = .158$. The error rate analysis showed a small –1.6% negative priming effect that did reach significance, $\beta = -0.246$, $SE = 0.104$, $z = -2.37$, $p = .018$.

Nonword targets The 11-ms relatedness effect was not significant in the latency analysis, $\beta = -0.009$, $SE = 0.006$, $t = -1.39$, $p = .166$. The 0.2% difference was not significant in the error rate analysis, $\beta = -0.071$, $SE = 0.083$, $z = -0.85$, $p = .397$.

Bayes factor analyses We also conducted a latency-based Bayes Factor analysis for both word and nonword targets in order to quantify the statistical evidence supporting the null priming effects. The details were the same as in Experiment 1. The Bayes factor value for word targets in Experiment 2a, $BF_{10} = 0.23 \pm 2.09\%$, constitutes good evidence for the null hypothesis. The Bayes factor value for nonword targets in Experiment 2a, $BF_{10} = 0.20 \pm 1.17\%$, also constitutes good evidence for the null hypothesis. Therefore, coupled with the analysis reported above, the most straightforward interpretation is that there was no priming effect for either target type.

⁴ For readers unfamiliar with this notation, C levels indicate an advanced level of ability in using a language, and B levels indicate an intermediate level of ability in using a language.

Table 4 Mean lexical decision latencies (RTs, in milliseconds) and percentage error rates in Experiment 2a (with standard deviations in parentheses)

Condition	RT	%E
Spanish–English bilinguals word targets		
Backward condition	607 (93)	4.6 (6)
Backward unrelated condition	602 (88)	3.0 (4)
Priming	–5	–1.6
Spanish–English bilinguals nonword targets		
Backward condition	740 (187)	6.5 (9)
Backward unrelated condition	751 (194)	5.8 (7)
Priming	11	–0.7

RT = reaction time; %E = percentage error rate

Experiment 2b

Method

Participants Twenty-eight Arabic–English bilinguals from Western University (London, Ontario, Canada) participated in this experiment in return for receiving \$10 (CDN). All participants indicated that they were highly proficient in reading standard Arabic and in reading English words written in uppercase Roman letters, that Arabic was their first and dominant language, and that they lived in an Arabic speaking country for at least 11 years. Their mean age was 22 years

Table 5 Details of random effects and multicollinearity in Experiment 2a, latency data

Groups name	Variance	SD	Corr
Word random effects			
Item (intercept)	0.009	0.096	
Relatedness/item	0.002	0.044	0.08
Participant (intercept)	0.051	0.226	
Residual	0.063	0.251	
Correlation of fixed effects			
	(Intr)		
Relatedness	0.013		
Nonword random effects			
Item (intercept)	0.011	0.103	
Relatedness/item	0.002	0.045	0.03
Participant (intercept)	0.092	0.303	
Residual	0.070	0.264	
Correlation of fixed effects			
	(Intr)		
Relatedness	0.004		

The number of observations was 3,163 and 3,098 for word and nonword trials, respectively

($SD = 4.81$), and they moved to Canada at a mean age of 17 years ($SD = 4.87$). Participants’ self-ratings of the percentage of the day spent using each language and their self-rating skill in each language (from 1 = *none* to 10 = *very fluent*) are presented in Table 6. They all had normal or corrected-to-normal vision and no reading disorders.

Materials The stimuli were the same as in Experiment 1. The only difference was that for the Arabic–English bilingual group, all 120 word and 120 nonword targets were presented to each participant. To do so, the word and nonword targets were divided into two lists with 60 stimuli in each list. For one-half of the participants, the 60 word targets from one of the lists and the 60 nonword targets from one of the lists were preceded by a backward prime, whereas the other 60 targets of each type were preceded by an unrelated prime. For the other half of the participants, the assignment of word and nonword targets to prime type was reversed. The unrelated prime–target pairs were created by re-pairing the backward primes and targets in the target list being used to create the unrelated condition. The reason that the prime–target counterbalancing manipulation was slightly different in Experiment 2b than in the previous experiments is because there was only a small sample of Arabic–English bilinguals at Western University, and therefore we needed to maximize the number of targets each participant would see in order to have at least 1,600 trials in each condition.

Procedure The procedure was essentially the same as in Experiment 1. The only difference was that for the Arabic–English bilingual group, the experimental block included 240 trials in total (120 word trials and 120 nonword trials).

Results

Response latencies that were less than 300 ms, more than three standard deviations from the participant’s mean latency (1.7% of the data for word targets and 1.7% of the data for nonword targets), or from incorrect trials (3.3% of the data for word targets and 4.1% of the data for nonword targets) were excluded from the latency analyses. Subjects and items were treated

Table 6 Language experience for the Arabic–English bilingual participants

	Arabic	English
% use in daily activities	63.93	31.59
Self-rating of current skill (out of 10)		
Understanding	9.30	9.15
Speaking	9.33	8.74
Reading	8.74	9.07
Writing	8.26	8.74

as random effects, and relatedness (related vs. unrelated) was treated as a fixed effect. The other details were the same as those in Experiment 2a. The mean RTs and percentage of errors from a subject-based analysis are shown in Table 7. The detailed information concerning the results of our analyses is reported in Table 8.

Word targets The 2-ms relatedness effect was not significant in the latency analysis, $\beta = 0.003$, $SE = 0.005$, $t = 0.55$, $p = .584$. The -0.1% relatedness effect was not significant in the error rate analysis, $\beta = -0.003$, $SE = 0.10$, $z = -0.03$, $p = .976$.

Nonword targets The -1 -ms relatedness effect was not significant in the latency analysis, $\beta = 0.002$, $SE = 0.005$, $t = 0.32$, $p = .751$. The 0.2% relatedness effect was not significant in the error rate analysis, $\beta = 0.025$, $SE = 0.093$, $z = 0.28$, $p = .783$.

Bayes factor analyses We also conducted a latency-based Bayes factor analysis for both word and nonword targets in order to quantify the statistical evidence supporting the null priming effects. The details were the same as in Experiment 2a. The Bayes factor in Experiment 2b for the word targets, $BF_{10} = 0.05 \pm 1.08\%$, indicates good evidence for the null hypothesis. The Bayes factor for nonword targets in Experiment 2b, $BF_{10} = 0.04 \pm 1.37\%$, also indicates good evidence for the null hypothesis. Again, coupled with the analyses reported above, the most straightforward interpretation is that there was no priming effect for either word or nonword targets for Arabic–English bilinguals.

Discussion

As there was no backward priming in either Experiment 2a or 2b, it does not appear that the backward priming effect when reading English words observed in Experiment 1 for Chinese–English bilinguals was due to either (a) the fact that the experiment was carried out in the participant’s L2 or (b) the fact that

Table 7 Mean lexical decision latencies (RTs, in milliseconds) and percentage error rates in Experiment 2b (with standard deviation in parentheses)

Condition	RT	%E
Arabic–English bilinguals word targets		
Backward condition	601 (105)	3.4 (3)
Backward unrelated condition	603 (123)	3.3 (4)
Priming	2	-0.1
Arabic–English bilinguals nonword targets		
Backward condition	716 (193)	4.0 (6)
Backward unrelated condition	715 (199)	4.2 (5)
Priming	-1	0.2

RT = reaction time; %E = percentage error rate

Table 8 Details of random effects and multicollinearity in Experiment 2b, latency data

Groups name	Variance	SD
Word random effects		
Item (intercept)	0.005	0.060
Participant (intercept)	0.050	0.225
Residual	0.078	0.280
Correlation of fixed effects		
	(Intr)	
Relatedness	0.000	
Nonword random effects		
Item (intercept)	0.006	0.080
Participant (intercept)	0.078	0.280
Residual	0.094	0.307
Correlation of fixed effects		
	(Intr)	
Relatedness	0.000	

The number of observations was 3,191 and 3,169 for word and nonword trials, respectively

Chinese–English bilinguals, like Arabic–English bilinguals, have been exposed to text written in a number of different directions.

General discussion

The proposal that there are language/script differences in how readers complete the orthographic coding process, in particular, the letter position coding component of that process, has received considerable support in recent years. As noted, it is now generally accepted that the process is differently tuned in Chinese versus English readers (Gu et al., 2015; Taft et al., 1999), with English readers (and likely readers of most alphabetic-script languages) showing less flexibility in terms of letter position coding than do Chinese readers. Other arguments consistent with this basic proposal come from research on reading in Hebrew (Velan & Frost, 2007, 2011) and in Korean (Lee & Taft, 2011; Rastle, Lally, & Lee, 2019), with the results of the relevant experiments in those languages suggesting that the letter position coding component for those readers shows even less flexibility than that for readers of most alphabetic-script languages. The general explanation of why the orthographic coding systems differ has been that the nature of any writing system, in particular, the nature of the orthographic neighborhoods created by that system, tunes the position coding process optimally for reading in the language in question (Frost, 2012; Lally et al., 2020; Lerner et al., 2014). Readers of languages with denser

neighborhoods and more anagrams (e.g., readers of Hebrew⁵) require a more precise position coding process than readers of languages with less dense neighborhoods and fewer anagrams (e.g., readers of Chinese) in order to successfully complete the reading process.

The present research takes that story one step further. In our experiments, the question investigated was whether the more flexible character position coding system that is used when reading Chinese would impact the orthographic coding process of Chinese–English bilinguals when reading English. The answer appears to be yes. In Experiment 1, Chinese–English bilinguals showed a clear backward priming effect, whereas English monolinguals did not. In follow-up experiments, neither Spanish–English nor Arabic–English bilinguals showed a backward priming effect when tested using English stimuli. Therefore, it appears that the backward priming effect when reading English words for Chinese–English bilinguals arose because the more flexible character position coding process that Chinese readers develop impacts the position coding system that Chinese–English bilinguals use when reading in their L2. That is, their orthographic coding system when reading in English is more flexible in terms of position coding than the system employed by most other readers of English (and likely other alphabetic-script languages).

Is the locus of the backward priming effect lexical or sublexical?

As noted, the backward priming effect for Chinese–English bilinguals could have either a sublexical or a lexical locus. As also noted, typically, masked priming effects in the lexical decision task are assumed to be lexically based (C. J. Davis, 2010; Forster, 1998; Forster & C. Davis, 1984). That is, the letter units activated by the prime form an orthographic code which then activates a set of lexical (i.e., word) units (i.e., units for words whose stored orthographic codes are similar to the orthographic code created by the prime). With respect to Chinese readers, the argument would be that the character identity information in the prime's orthographic code plays the main role in determining which lexical representations get activated by the prime with position information in the prime's orthographic code playing only a minimal role. As a result, even orthographic codes in which the position information in the prime and target do not match (e.g., backward primes) will still activate lexical representations when the prime and target share characters.

What is also possible, however, is that a backward priming effect could have a sublexical locus. That is, the effect might

⁵ As suggested by one of the reviewers, Hebrew's lack of vowel markers likely also makes deriving accurate position information from consonants more crucial than in other alphabetic-script languages.

be due to the (lowercase) letters in the prime activating the relevant abstract letter units, making the processing of the (uppercase) letters in the target more efficient even though the letters in the prime and target are in different positions. Therefore, in theory, any orthographic priming effect could have arisen as a result of letter level processing.

The proposal that letter level processing is facilitated in masked priming experiments makes a clear prediction, specifically, it predicts that both word and nonword targets would show facilitation. In alphabetic-script lexical decision experiments, there is virtually no evidence for orthographic priming effects for nonwords, which is why those priming effects are assumed to be lexically based. However, such may not be true in the present situation. That is, as noted, uppercase letters, the format typically used for targets in lexical decision experiments are infrequently used in English language reading instruction in China, with Chinese L1 participants often reporting that they have some difficulty reading uppercase targets. Therefore, it is not impossible that letter-level processing for Chinese–English bilinguals may be responsible for orthographic priming effects in English.

As suggested, the most straightforward way to contrast the lexical versus sublexical accounts would be to examine the relationship between word and nonword priming effects. If prime letters were activating target letters which then facilitated target processing, nonword targets should show the same priming pattern as word targets. If the primes were activating lexical representations, only word targets would show priming. Based on the analyses described above, the present data provide a reasonably convincing conclusion concerning this issue. That is, for word targets in Experiment 1, both the standard analysis and the Bayes Factor analysis indicated that Chinese–English bilinguals showed a priming effect for word targets. In contrast, neither analysis provided any evidence that those bilinguals showed a priming effect for nonword targets. However, this conclusion is somewhat compromised by the fact that the sizes of the priming effects for word and nonword targets were not substantially different (14 ms for word targets, 10 ms for nonword targets). Therefore, we undertook some additional, more detailed analyses of the word–nonword difference.⁶

If priming is acting at the same level for word and nonword targets, there should be a correlation between priming effect sizes for the two target types (across participants). That is, if

the effect is sublexical, participants showing larger word priming effects should be the ones showing larger nonword priming effects. The correlation between priming effect sizes over the 42 Chinese–English bilinguals in Experiment 1, showed no suggestion of such a correlation ($r = .088, p = .578$).

A second way of examining this issue is to compare the quantile plots for word and nonword targets. If they derive from the same process (i.e., sublexical priming), they should look similar. If the nonsignificant 10-ms nonword effect is just due to random variability, the two effects would likely look different. As noted, the plot for word targets (see Fig. 1) is consistent with the typical head-start effect found in other form priming experiments (e.g., Taikh & Lupker, 2020), an effect often assumed to be due to the prime activating the lexical representation of the target. The effect for nonword targets is somewhat (although not dramatically) less so with effect sizes in the five quantiles of 6 ms, 3 ms, 10 ms, 12 ms, and –68 ms, respectively.

A third way to evaluate this issue is to contrast the better (i.e., faster) Chinese–English bilingual participants with the weaker (i.e., slower) ones using a median split. If the priming effects come from the same process, the two groups should show parallel patterns in terms of both word and nonword priming effects. (Note also that the former group is much more comparable to the other two sets of bilinguals in terms of overall latencies, so this contrast also addresses the question of whether the backward priming effect might only arise in the weaker bilinguals.)

For the faster participants, the word target priming effect was 20 ms (685 ms on backward prime trials, 705 ms on unrelated prime trials), whereas it was 10 ms for the slower participants (907 ms vs. 917 ms). This result suggests that the priming effects for word targets were fairly similar for the two groups of participants while, more importantly, clearly showing that the priming effect for word targets was not being carried by the participants who were weaker in English. More relevant to the present discussion, the faster participants showed a 28-ms inhibition effect for nonword targets (826 ms on backward prime trials, 798 ms on unrelated prime trials), whereas the slower participants showed a 45-ms priming effect (1,307 ms vs. 1,352 ms) for nonword targets. Thus, while the word target priming effect was only slightly different in the two groups, the nonword target priming effect was being entirely carried by the participants who showed considerable difficulty with nonword targets.

Although these analyses may not be entirely conclusive, they imply that the pattern for the word and nonword priming effects are, in general, somewhat different from one another. Therefore, it seems unlikely that those effects arose from the same process (i.e., sublexical priming). Rather, it seems more likely that the word target priming effect is a lexical effect (the standard explanation for masked priming effects with word targets) while the (nonsignificant) nonword target priming

⁶ At the request of one of the reviewers, we undertook an additional analysis of the results of Experiment 1, including target type (word vs. nonword) as an additional factor in the analysis. At issue is whether there was a three-way interaction involving relatedness, language group, and target type. A three-way interaction would be consistent with the claim that significant priming was found only for Chinese–English bilinguals when presented with word targets. This interaction was marginal, $\beta = 0.005, SE = 0.003, t = 1.93, p = .054$, providing at least some additional support for the idea that the priming effects for word and nonword targets did not follow the same pattern in the Chinese–English bilingual and English monolingual groups.

effect was due simply to random variability, mainly in the results of those participants who had considerable difficulty determining that an English nonword really is a nonword.

Models of orthographic coding

The noticeable difference between the orthographic coding systems for Chinese versus English readers in terms of how they code position information raises the question of whether it is, nonetheless, possible to maintain the idea that there is a universal approach to modelling word recognition and reading. That is, could there be a single model framework involving a letter position coding process that could produce little, if any, transposed letter priming effects in some languages (e.g., Hebrew, Korean), some transposed letter priming effects in others, as long as the transpositions are not extreme (e.g., English, Spanish), and a backward priming effect in others (e.g., Chinese)? And, further, if such a model framework could do so, could that same system be able to explain how it is possible to observe a 50+ ms backward priming effect in Chinese (suggesting that the system makes little effort to produce accurate position coding), while at the same time explaining the 80-ms identity priming effect reported by Yang et al. (2019), an effect showing that there is a clear advantage of having the Chinese characters in the same order in the prime and target. Although building such a system is beyond the scope of the present paper, it may be useful to consider the nature of the orthographic coding models currently available in an attempt to get some idea of whether such a system could be built based on the structure of one of them.

Current models of orthographic coding (all of which are derived from alphabetic-script language data and typically data from experiments with monolinguals) fall into two general categories. One is what C. J. Davis and Lupker (2017) refer to as the “local context” models, or what are more typically referred to as the “open-bigram” models (Grainger, Granier, Farioli, Van Assche, & Van Heuven, 2006; Grainger & Van Heuven, 2003; Schoonbaert & Grainger, 2004; Whitney, 2001; Whitney & Marton, 2013). Those types of models are based on the idea that a set of bigram units exists as an intermediate level of representation between abstract letter units and word units. The bigram units represent ordered bigrams, and a given unit would be activated only if both letters represented in the unit are contained in the proper order in the letter string being read. For example, when reading the TL prime “judge,” the open bigrams *ju*, *jd*, *je*, *ug*, *ud*, *ue*, *gd*, *ge*, and *de* would be activated following activation of the letter units. Most of the bigrams that are relevant to processing/activating the target word JUDGE would, therefore, be activated by the TL prime “judge,” which is not the case for a substitution letter prime like “jupte.” As a result, there would be a TL priming effect.

Although local context models have had some success in explaining various priming effects, they would not be able to provide any kind of unified framework of the sort desired here, because they would have no means of explaining a backward priming effect. The reason is that none of the bigrams activated by a backward prime would be relevant to target processing. The only exception is the overlap open-bigram model (Grainger et al., 2006), which does assume that transposed open bigrams are activated; however, it also assumes that they are only activated to a minimal degree. Note also that these models would have some trouble explaining the general lack of transposition priming observed in Hebrew and Korean without making assumptions that the models currently do not make.

The other type of models is what C. J. Davis and Lupker (2017) referred to as the “noisy position” models (C. J. Davis, 2010; Gómez et al., 2008; Norris & Kinoshita, 2012; Norris, Kinoshita, & van Casteren, 2010). In these models, the precision of position coding is represented by a parameter that can be thought of as a “variance” parameter. Although the details of the models differ, the general concept of position noisiness can be thought of as follows.

When reading a letter string (e.g., a TL prime like “judge”), the system codes each letter as being most likely to be in the position it actually is in (e.g., a “j” is in Position 1, a “u” is in Position 2). However, it also codes the likelihood that each of the letters is in the other positions (i.e., the probability that the “u” may really be in Position 1, Position 3, etc., with the probability of it being in those alternative positions decreasing as a function of their distance from the actual position). The value of the relevant variance parameter is what determines how large those probabilities are. A zero value would mean that the position coding system was quite precise (i.e., the “u” must be in Position 2, as the probabilities for all the other positions are zero). A large value would mean that there is a reasonable chance that the “u” is in a position other than Position 2. The extent to which the lexical representation for a word (e.g., *judge*) is activated and hence TL priming is produced, is a function of the letter activation in the relevant positions. That is, the letter “u” will provide the most activation to words having a “u” in the second position (e.g., *judge*), less activation to words having a “u” in either Position 1 (e.g., *under*) or Position 3 (e.g., *chute*).

Models of this sort may be able to explain the data from different language experiments because, in this framework, the language differences are quantitative. That is, the precision of position coding is essentially determined by the value of any variance parameter, and that parameter is essentially a free parameter. The value it would assume for readers of any given language would presumably reflect the nature of that language and, in particular, the nature of the language’s orthographic neighborhoods (Frost, 2012). It would, therefore, be possible for such a model to explain why there is no transposed letter

priming in one language while, at the same time, there are extreme transposed letter priming effects (e.g., backward priming) in others, as well as why the position coding process in a reader’s L2 might, at least initially, show a tendency to resemble that in their L1.

Where this sort of model may run into problems, however, is in trying to find a sweet spot for the value of the variance parameter—that is, a spot that would allow it to predict all the relevant orthographic priming effects observed. For example, as noted previously, the backward priming effect for Chinese readers responding to four-character words is large (50+ ms), implying that the variance parameter would have a large value. However, the identity priming effect for those words is noticeably larger (i.e., 80 ms), suggesting that there is something important to be gained when the order of the characters in the prime and the target match. It would be quite difficult to explain such a pattern if the assumption was that the variance parameter had an extremely large value for Chinese readers.

Conclusion

The orthographic coding process, in particular, the position coding component of that process, is clearly different for Chinese readers than for readers of most other languages. Specifically, position coding is a much smaller component of that process than it is for readers of other languages. The present results indicate that when Chinese readers learn to read in English (hence, becoming Chinese–English bilinguals), the position coding component of their orthographic coding system maintains attributes of the system they have learned to use

when reading Chinese. In particular, their position coding component is not as precisely tuned as is the position coding component of English monolinguals, which allows Chinese–English bilinguals to show backward priming effects in an English lexical decision task. Such is not the case for bilinguals whose L1 is one in which accurate position coding is important. A question for future research would be whether continuing to develop their L2 skill will ultimately cause this tendency of Chinese–English bilinguals to disappear or whether they will continue to read alphabetic-script languages in a way that is different from that of a monolingual reader of those languages.

Author note This research was partially supported by Natural Sciences and Engineering Research Council of Canada Grants to Stephen J. Lupker and Debra Jared. We would like to thank Giacomo Spinelli and Zian Chi for their assistance in the data collection.

Open practices statement All materials used in these experiments are reported in the Appendix 1. The data sets analyzed in the present research and the R scripts used to carry out the analyses are available in the Open Science Framework repository (<https://osf.io/eu5s3/>). None of these experiments was preregistered.

Compliance with ethical standards

Declarations of interest None.

Appendix 1. Word targets and their primes

All stimuli below were used in Experiments 1 and 2. BA = backward primes, BA UN = backward unrelated primes.

Target	BA	BA UN	Target	BA	BA UN
WRITE	etirw	txen	MANY	ynam	yenom
KNOWN	nwonk	dnats	DROP	pord	rojam
ALONG	gnola	ecalp	SEVEN	neves	yros
CLEAN	naelc	gnuoy	VERY	yrev	cisum
CITY	ytic	yrow	TABLE	elbat	tniop
SIGN	ngis	rehto	DRIVE	evird	klaw
MOVIE	eivom	yzarc	HAIR	riah	hturt
DRINK	knird	esuac	UNTIL	litnu	kciuq
WRONG	gnorw	kcalb	EASY	ysae	elohw
ORDER	redro	eerf	ABLE	elba	kniht
ALIVE	evila	dluoc	COURT	truoc	hgih
BODY	ydob	traeh	CLEAR	raelc	truh
LEAVE	evael	hcae	EARLY	ybrae	nwot
HOTEL	letoh	worht	SHUT	tuhs	ffuts
HURRY	yrruh	eulb	IDEA	aedi	kaerb
FACT	tcaf	esuoh	STORY	yrots	daeha

(continued)

Target	BA	BA UN	Target	BA	BA UN
SENSE	esnes	tsrif	CHILD	dlihc	hcuot
BABY	ybab	retfa	PAPER	repap	enohp
GLAD	dalg	etihw	HUMAN	namuh	yenoh
SINCE	ecnis	ydaer	TRUST	tsurt	esolc
MONEY	yenom	eerht	NEXT	txen	hcum
MAJOR	rojam	teiuq	STAND	dnats	ydal
SORRY	yrros	tnega	PLACE	ecalp	dlrow
MUSIC	cisum	maerd	YOUNG	gnuoy	taerg
POINT	tniop	htrow	WORRY	yrrow	reve
WALK	klaw	lrig	OTHER	rehto	htaed
TRUTH	hturt	htuom	CRAZY	yzarc	yreve
QUICK	kciuq	nalp	CAUSE	esuac	reven
WHOLE	elohw	kcehc	BLACK	kcalb	ssik
THINK	kniht	llams	FREE	eerf	ykcul
HIGH	hgih	enola	COULD	dluoc	tfel
HURT	truh	yadot	HEART	traeh	nrtut
TOWN	nwot	flah	EACH	hcae	kaeps
STUFF	ffuts	ecnad	THROW	worht	yawa
BREAK	kaerb	gniht	BLUE	eulb	elcnu
AHEAD	daeha	namow	HOUSE	esuoh	nrael
TOUCH	hcuot	eceip	FIRST	tsrif	trats
PHONE	enohp	rednu	AFTER	retfa	tnorf
HONEY	yenoh	ysub	WHITE	etihw	nwod
CLOSE	esolc	niaga	READY	ydaer	driew
MUCH	hcum	etirw	THREE	eerht	ynam
LADY	ydal	nwonk	QUIET	teiuq	pord
WORLD	dlrow	gnola	AGENT	tnega	neves
GREAT	taerg	naelc	DREAM	maerd	yrev
EVER	reve	ytic	WORTH	htrow	elbat
DEATH	htaed	ngis	GIRL	lrig	evird
EVERY	yreve	eivom	MOUTH	htuom	riah
NEVER	reven	knird	PLAN	nalp	litnu
KISS	ssik	gnorw	CHECK	kcehc	ysae
LUCKY	ykcul	redro	SMALL	llams	elba
LEFT	tfel	evila	ALONE	enola	truoc
TURN	nrtut	ydob	TODAY	yadot	raelc
SPEAK	kaeps	evael	HALF	flah	ylrae
AWAY	yawa	letoh	DANCE	ecnad	tuhs
UNCLE	elcnu	yrroh	THING	gniht	aedi
LEARN	nrael	tcaf	WOMAN	namow	yrots
START	trats	esnes	PIECE	eceip	dlihc
FRONT	tnorf	ybab	UNDER	rednu	repap
DOWN	nwod	dalg	BUSY	ysub	namuh
WEIRD	driew	ecnis	AGAIN	niaga	tsurt

Appendix 2. R code used in the analyses

Experiment 1 main analysis:

For the latency analysis, the model, based on inverse transformed latencies in order to normalize the latency distribution, was: $RT = \text{lmer}(\text{invert } RT \sim \text{Relatedness} \times \text{Language Group} + (1 | \text{subject}) + (\text{Language Group} | \text{item}), \text{data} = \text{cleaned data}, \text{optCtrl} = \text{list}(\text{maxfun} = 1e6), \text{control} = \text{lmerControl}(\text{optimizer} = \text{"bobyqa"}))$. For the error rate analysis, the model was: $\text{Accuracy} = \text{glmer}(\text{accuracy} \sim \text{Relatedness} \times \text{Language Group} + (1 | \text{subject}) + (1 | \text{item}), \text{family} = \text{"binomial"}), \text{data} = \text{Raw data}, \text{optCtrl} = \text{list}(\text{maxfun} = 1e6))$.

Experiment 1 quantile analysis:

For the latency analysis, the Quantile Group model was: $RT = \text{lmer}(\text{invert } RT \sim \text{Relatedness} \times \text{Quantile Group} + (\text{Quantile Group} | \text{subject}) + (\text{Quantile Group} | \text{item}), \text{optCtrl} = \text{list}(\text{maxfun} = 1e6))$.

Experiment 2a:

For the latency analyses, the model was: $RT = \text{lmer}(\text{invert } RT \sim \text{Relatedness} + (1 | \text{subject}) + (\text{Relatedness} | \text{item}), \text{data} = \text{cleaned data}, \text{optCtrl} = \text{list}(\text{maxfun} = 1e6))$. For the error rate analyses, the model was: $\text{Accuracy} = \text{glmer}(\text{accuracy} \sim \text{Relatedness} + (1 | \text{subject}) + (1 | \text{item}), \text{family} = \text{"binomial"}), \text{data} = \text{Raw data}$.

Experiment 2b:

For the latency analyses, the model was: $RT = \text{lmer}(\text{invert } RT \sim \text{Relatedness} + (1 | \text{subject}) + (1 | \text{item}), \text{data} = \text{cleaned data}, \text{optCtrl} = \text{list}(\text{maxfun} = 1e6))$. For the error rate analyses, the model was: $\text{Accuracy} = \text{glmer}(\text{accuracy} \sim \text{Relatedness} + (1 | \text{subject}) + (1 | \text{item}), \text{family} = \text{"binomial"}), \text{data} = \text{Raw data}, \text{optCtrl} = \text{list}(\text{maxfun} = 1e6))$.

References

- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., ... Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, 39(3), 445–459. <https://doi.org/10.3758/BF03193014>
- Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990. <https://doi.org/10.3758/BRM.41.4.977>
- Brysbaert, M. & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), 1–20. <https://doi.org/10.5334/joc.10>
- Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Mahwah: Erlbaum.
- Davis, C., Kim, J., & Forster, K. I. (2008). Being forward not backward: Lexical limits to masked priming. *Cognition*, 107(2), 673–684. <https://doi.org/10.1016/j.cognition.2007.07.016>
- Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, 117(3), 713–758. <https://doi.org/10.1037/a0019738>
- Davis, C. J. & Lupker, S. J. (2017). A backwards glance at words: Using reversed-interior masked primes to test models of visual word identification. *PLOS One*, 12(12), e0189056. <https://doi.org/10.1371/journal.pone.0189056>
- Forster, K. I. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research: Special Issue*, 27(2), 203–233. <https://doi.org/10.1023/A:1023202116669>
- Forster, K. I. & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 680–698. <https://doi.org/10.1037/0278-7393.10.4.680>
- Forster, K. I. & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116–124. <https://doi.org/10.3758/bf03195503>
- Fox, J., & Weisberg, S. (2016). car: Companion to applied regression (R Package Version 2.1-2) [Computer software]. Retrieved from <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>
- Frost, R. (2012). A universal approach to modelling visual word recognition and reading: Not only possible, but also inevitable. *Behavioural and Brain Sciences*, 35(5), 310–329. <https://doi.org/10.1017/S0140525X12000635>
- Gómez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, 115(3), 577–600. <https://doi.org/10.1037/a0012667>
- Grainger, J. (2018). Orthographic processing: A ‘mid-level’ vision of reading: The 44th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology*, 71(2), 335–359. <https://doi.org/10.1080/17470218.2017.1314515>
- Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & Van Heuven, W. J. B. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance*, 32(4), 865–884. <https://doi.org/10.1037/0096-1523.32.4.865>
- Grainger, J., & Van Heuven, W. J. B. (2003). *Modeling letter position coding in printed word perception*. In P. Bonin (Ed.), *The mental lexicon* (pp. 1–23). Hauppauge: Nova Science.
- Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Gu, J., Li, X., & Liversedge, S. P. (2015). Character order processing in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), 127–137. <https://doi.org/10.1037/a0038639>
- Guerrera, C., & Forster, K. (2008). Masked form priming with extreme transposition. *Language and Cognitive Processes*, 23(1), 117–142. <https://doi.org/10.1080/01690960701579722>
- Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance*, 33(1), 209–229. <https://doi.org/10.1037/0096-1523.33.1.209>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13). <https://doi.org/10.18637/jss.v082.i1>
- Lally, C., Taylor, J. S. H., Lee, C. H., & Rastle, K. (2020). Shaping the precision of letter position coding by varying properties of a writing system. *Language, Cognition and Neuroscience*, 35(3), 374–382. <https://doi.org/10.1080/23273798.2019.1663222>

- Lee, C. H., & Taft, M. (2011). Subsyllabic structure reflected in letter confusability effects in Korean word recognition. *Psychonomic Bulletin & Review*, 18(1), 129–134. <https://doi.org/10.3758/s13423-010-0028-y>
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). *emmeans: Estimated Marginal Means*. R package version 1.4.8. Retrieved from <https://CRAN.R-project.org/package=emmeans>
- Lerner, I., Armstrong, B. C., & Frost, R. (2014). What can we learn from learning models about sensitivity to letter-order in visual word recognition? *Journal of Memory and Language*, 77(1), 40–58. <https://doi.org/10.1016/j.jml.2014.09.002>
- Levy, R. (2014). *Using R formulae to test for main effects in the presence of higher-order interactions*. ArXiv Preprint ArXiv:1405.2094. Retrieved from <https://arxiv.org/pdf/1405.2094.pdf>
- Morey, R. D., Rouder, J. N., Jamil, T., & Morey, M. R. D. (2015). Package ‘bayesfactor’. <https://cran.r-project.org/web/packages/BayesFactor/index.html>.
- Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there’s nothing special about the perception of orthography. *Psychological Review*, 119, 517–545. <https://doi.org/10.1037/a0028450>
- Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, 62(3), 254–271. <https://doi.org/10.1016/j.jml.2009.11.002>
- Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 97–120). Hove: Psychology Press. <https://doi.org/10.4324/9780203502846>
- Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, 51(2), 231–246. <https://doi.org/10.1016/j.jml.2004.05.005>
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.
- Rastle, K., Lally, C., & Lee, C. H. (2019). No flexibility in letter position coding in Korean. *Journal of Experimental Psychology: Human Perception and Performance*, 45(4), 458–473. <https://doi.org/10.1037/xhp0000617>
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User’s guide*. Pittsburgh: Psychology Software Tools.
- Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, 19(3), 333–367. <https://doi.org/10.1080/01690960344000198>
- Singmann, H., & Kellen, D. (2020). An introduction to mixed models for experimental psychology. In D. Spieler & E. Schumacher (Eds.), *New methods in neuroscience and cognitive psychology* (pp. 1–39). Hove: Psychology Press.
- Taft, M., Zhu, X., & Peng, D. (1999). Positional specificity of radicals in chinese character recognition. *Journal of Memory and Language*, 40(4), 498–519. <https://doi.org/10.1006/jmla.1998.2625>
- Taikh, A., & Lupker, S. J. (2020). Do visible semantic primes pre-activate lexical representations? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(8), 1533–1569. <https://doi.org/10.1037/xlm0000825>
- Velan, H., & Frost, R. (2007). Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review*, 14(5), 913–918. <https://doi.org/10.3758/BF03194121>
- Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition*, 118(2), 141–156. <https://doi.org/10.1016/j.cognition.2010.11.013>
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, 8(2), 221–243. <https://doi.org/10.3758/BF03196158>
- Whitney, C., & Marton, Y. (2013). *The SERIOL2 model of orthographic processing* (Online submission). Retrieved from <http://eric.ed.gov/?id=ED543279>
- Yang, H., Chen, J., Spinelli, G., & Lupker, S. J. (2019). The impact of text orientation on form priming effects in four-character Chinese words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(8), 1511–1526. <https://doi.org/10.1037/xlm0000655>
- Yang, H., Hino, Y., Chen, J., Yoshihara, M., Nakayama, M., Xue, J., & Lupker, S. J. (2020). The origins of backward priming effects in logographic scripts for four-character words. *Journal of Memory and Language*, 113, 104017. <https://doi.org/10.1016/j.jml.2020.104107>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.