In press in *Quarterly Journal of Experimental Psychology*

Revisiting the Influence of Phonological Similarity on Cognate Processing: Evidence from Cantonese-Japanese Bilinguals

Brian W. L. WONG[a,b], Shawn HEMELSTRAND[a], and Tomohiro INOUE[a]

[a] Department of Psychology, The Chinese University of Hong Kong, Hong Kong, China

[b] BCBL, Basque Center on Brain, Language and Cognition, Donostia-San Sebastian, Spain

Author Note

Brian W. L. WONG https://orcid.org/0000-0002-0519-6116

Shawn HEMELSTRAND https://orcid.org/0000-0002-9627-1571

Tomohiro INOUE https://orcid.org/0000-0001-8163-087X

Correspondence concerning this article should be addressed to Tomohiro Inoue, Rm 352, Sino Building, Department of Psychology, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong. Email: tinoue@cuhk.edu.hk

Abstract

The influences of shared orthography, semantics, and phonology on bilingual cognate processing have been investigated extensively. However, mixed results have been found regarding the effects of phonological similarity on L2 cognate processing. In addition, most existing studies examining the influence of phonological similarity on cognate processing have been conducted on alphabetic scripts, in which phonology and orthography are always associated. Hence, in this study, we recruited Cantonese-Japanese bilinguals who used two logographic scripts, traditional Chinese and Japanese Kanji, to examine the influence of phonological similarity on L2 cognate lexical decision. Importantly, these scripts allow the manipulation of phonological similarity using identical characters across both languages. Additionally, we examined how word frequency and L2 proficiency modulate cognate processing. Results showed that although word frequency and L2 proficiency played important roles in cognate processing, there was minimal overall influence of phonological similarity on cognate lexical decision. The latter finding suggests that theoretical models of bilingual word recognition may need to be refined to enhance our understanding of cognate processing regarding the role of phonology among diverse bilingual populations.

Keywords: Cognate, phonology, Chinese-Japanese bilinguals, cognate facilitation effect

Revisiting the Influence of Phonological Similarity on Cognate Processing: Evidence from Cantonese-Japanese Bilinguals

What are the factors that affect lexical representations in bilinguals' minds? To examine this question, researchers have historically focused on the utility of cognates, which are semantic equivalents in two languages that share phonological and/or orthographic similarities (Liu et al., 2023; Xiong et al., 2020). This sharing of semantics, orthography, and/or phonology allows researchers to investigate the processing of word recognition in bilinguals directly. In turn, this allows us to enhance our understanding of established bilingual models of word recognition (e.g., Dijkstra & Van Heuven, 2002; Dijkstra et al., 2019).

Repeated research in this area has demonstrated what we call the cognate facilitation effect, a phenomenon wherein bilinguals recognize a cognate faster or more accurately compared to a non-cognate (Peeters et al., 2013). This effect has been replicated in a number of previous studies (e.g., Comesaña et al., 2012; Dijkstra et al., 1999; Xiong et al., 2020), though it may not be universal, as at least one study showed a reverse effect in a mixed-language context (Hsieh et al., 2021). Essentially, the cognate facilitation effect suggests that bilingual lexical access is non-selective, which means that word candidates from both language systems are activated in bilinguals' minds due to the language input involved (de Groot et al., 1999; Dijkstra et al., 2019). However, while extensive studies have shown the presence of this effect, there remains an important gap in research on cognate processing.

So far, a variety of studies have primarily focused on the phonology aspect of cognate or homophone processing in diverse languages (e.g., Allen & Conklin, 2013; Allen et al., 2021; Brysbaert et al., 1999; Gollan et al., 1997; Haigh & Jared, 2007; Kim & Davis, 2003; Lupker et al., 2015; Miwa et al., 2014; Nakayama et al., 2012, 2013, 2014; Pexman et al., 2001; Timmer et al., 2014; Voga & Grainger, 2007; Zhang et al., 2019; Zhou et al., 2010). In

addition, most previous studies on cognate lexical decision have been conducted on participants who use alphabetic scripts (e.g., Carrasco-Ortiz et al., 2021; Dijkstra et al., 1999, 2010; Frances et al., 2021). While these studies provide a wealth of evidence for the cognate facilitation effect, it is difficult to examine the influences of orthography and phonology on word processing separately because orthography and phonology are always closely associated in alphabetic languages. Only a few studies have adopted two logographic scripts, such as Chinese and Japanese Kanji characters (for exceptions, see Hsieh et al., 2021; Liu et al., 2023; Nakayama, 2002; Xiong et al., 2020). Therefore, how cognates are processed and represented in the mental lexicon of bilinguals who speak non-alphabetic languages remains unclear, especially in terms of the role of phonology in cognate processing. Examining lexical decision in Cantonese-Japanese bilinguals, the focus of the present study, may shed light on this issue and thus extend previous theories of bilingual representations.

Theoretical Frameworks of Bilingual Representations and Relevant Findings

Two of the most prominent models of bilingual representations are the Bilingual Interactive Activation + (BIA+) model (Dijkstra & Van Heuven, 2002) and its extended version called Multilink (Dijkstra et al., 2019). Both models postulate that the representations shared across two languages can be co-activated, even when bilinguals are performing a unilingual task, such as an L2 lexical decision task (Dijkstra & Van Heuven, 2002).

These models stipulate that cognate facilitation is caused by the orthographic overlap of cognates, leading to co-activation in both languages, which maps onto shared semantics. The co-activated orthography and shared semantics activate linked phonological representations, resulting in the cognate facilitation effect. Therefore, the two models explain the effect via the combination of shared orthography, semantics, and phonological representations. Additionally, the Multilink model considers the lexical activation spreading from orthographic to phonological representations as the fourth explanation (Dijkstra et al.,

2019). Hence, cognate processing can be facilitated by semantic, orthographic, and phonological similarities through the co-activations of shared lexical representations across languages. Numerous studies have provided evidence supporting these models and demonstrating the facilitative influences of shared orthography and semantics on cognate processing in bilinguals (see Dijkstra & Van Heuven, 2002, for a review).

However, lexical decision studies, which have investigated the influence of phonological similarity on cognate processing, have yielded mixed findings (e.g., Carrasco-Ortiz et al., 2021; Dijkstra et al., 1999, 2010; Frances et al., 2021). For example, Dijkstra et al. (2010) found that phonological overlap facilitated visual lexical decision of identical cognates (i.e., shorter reaction time [RT]) in Dutch-English bilinguals. Likewise, Carrasco-Ortiz et al. (2021) revealed that both phonological and orthographic similarities had facilitative effects on L2 English lexical decision among Spanish-English bilinguals. Similar facilitative effects have also been found in other studies with bilinguals of other sets of alphabetic languages (e.g., Haigh & Jared, 2007; Lemhöfer & Dijkstra, 2004).

In contrast, other studies have found inhibitory effects of phonological similarity on lexical decision. Dijkstra et al. (1999) showed that L2 English lexical decision was facilitated by orthographic and semantic similarities of cognates in Dutch-English bilinguals, but it was inhibited by phonological overlap (i.e., longer RT). Additionally, Frances et al. (2021) examined the effects of orthographic and phonological similarities on L2 English lexical decision in Spanish-English bilinguals. They showed that RT was faster when a cognate in L2 English was orthographically similar to its counterpart in L1 Spanish. However, higher phonological similarity led to a slower RT. These findings, showing the inhibitory effect of phonological similarity on cognate processing, call into question the basic premise of the theoretical models of bilingual representations and suggest that at least some modifications of these theories may be necessary.

These contradictory findings have led some researchers to propose alternative explanations for the relationship between phonology and cognate processing. For example, Frances et al. (2021) attributed the negative effect of phonological similarity on cognate processing to differences in the orthographic transparency of L1 Spanish and L2 English. Since English has an opaque orthography and Spanish has a transparent orthography (Frost, 2005), the mismatch in the grapheme-phoneme correspondences between the two languages may bring confusion and spread activation to non-target words. They also proposed that if the orthographic consistency of two languages is comparable (i.e., both transparent or both opaque), phonological similarity should lead to facilitation, as proposed by the BIA+ model.

Regardless of whether orthographic consistency plays a role in the facilitation of phonological similarity, examining the relationship between phonology and cognate processing in alphabetic scripts is challenging. Admittedly, it is difficult to separate the effects of cognate processing attributed to phonology from that of the orthography itself because, in alphabetic scripts, these two features are intrinsically interconnected (Hsieh et al., 2021). Therefore, creating stimuli with identical graphemes but with low phonological similarity is almost impossible in such scripts (for relevant discussions, see Frances et al., 2021).

To circumvent this problem, some previous studies have been conducted with bilinguals using different scripts (e.g., L1 Chinese or Japanese and L2 English) that have no orthographic overlap and varying phonological similarity. For example, Allen and Conklin (2013) conducted an L2 English visual lexical decision experiment in Japanese-English bilinguals. They found that increased phonological similarity of cognates led to faster RTs (for similar results, see Allen et al., 2021; Miwa et al., 2014). Importantly, since Japanese and English do not overlap in orthography, the observed effect of phonological similarity on cognate facilitation in these studies is independent of the effect of orthography. Thus, these

results provide supporting evidence for the BIA+ and Multilink models and suggest that the phonological similarity of cognates may determine the extent to which their counterparts in another language are activated in bilingual word recognition.

It should be noted, however, that these studies often involved English, an alphabetic language in which orthography and phonology are closely related. Indeed, since phonological activation during word recognition is relatively automatic in alphabetic languages (Berent & Perfetti, 1995; Perfetti & Bell, 1991), it is unknown whether the results can be generalized to bilinguals reading two logographic languages, where the mappings between orthography and phonology are more opaque. Therefore, the present study examined cognate lexical decision in two logographic scripts, namely traditional Chinese and Japanese Kanji characters. This comparison may deepen our understanding of the influence of phonology on cognate processing because, in many cases, the pronunciation of the same symbol (character) is very different in these two languages (Nakayama, 2002; Taylor & Taylor, 2014). Using these two languages allows researchers to examine the role of phonology in cognate processing by varying the phonological similarity of cognate words that share identical graphic symbols.

Characteristics of Cantonese and Japanese

Despite many shared characteristics of traditional Chinese and Japanese Kanji characters, there are some important differences. Although there are some variations in both the spoken and written forms of Chinese, we base our description below on Cantonese and traditional Chinese as commonly used in Hong Kong, where this study was conducted.

Most Chinese characters are composed of two distinct graphic components: a phonetic radical and a semantic radical (Shu & Anderson, 1997; Taylor & Taylor, 2014). The former usually represents the pronunciation of a character, while the latter offers a clue to the meaning of the character. For example, the left semantic radical of the character 抽 /cau1/ (the number denotes the tone in Cantonese) “extract” indicates that the meaning of the word

is related to “hand,” as 扌 is the radical that indicates the literal depiction of a hand. The right phonetic radical 由 /jau4/ represents the pronunciation of the character.

However, in many instances, there is no one-to-one correspondence between a phonetic radical and its pronunciation. For example, 迪 is pronounced as /dik6/ despite the same phonetic radical 由, as shown in the example above. In other words, several pronunciations are often possible for a character with the same phonetic radical, which makes Chinese orthographically opaque.

Chinese is also considered a *morphosyllabic* script because each character represents not only a syllable but also a morpheme (e.g., 大 /daai6/ “big”, 人 /jan4/ “people”, 大人 /daai6.jan4/ “adult” [the dot indicates the separation of each character]; McBride-Chang et al., 2003; Perfetti & Dunlap, 2008; Verhoeven & Perfetti, 2022).

Similar to Chinese, Japanese Kanji characters also have inconsistent character-sound mappings. Many Kanji have two different types of reading, namely *Kun-reading* and *On-reading*. Kun-reading is of Japanese origin, while On-reading is derived from Chinese pronunciation (Koda, 2017; Verdonschot et al., 2013). Therefore, the phonological similarity of Japanese words with their Cantonese counterparts is generally higher for On-reading than for Kun-reading. For example, the two-character word 態度 is pronounced as /tai.do/ in On-reading in Japanese and /taai3.dou6/ in Cantonese. In contrast, most Kun-reading words have a lower phonological similarity to their Cantonese counterparts than On-reading words. For example, the word 入口 is pronounced as /iri.guchi/ in Japanese and /jap6.hau2/ in Cantonese.

Additionally, unlike Cantonese, each Japanese Kanji character usually represents multi-syllabic morphemes and words (e.g., 車 /kuruma/ “car”, 頂 /itadaki/ “peak”). Taken together, the phonological similarity between Cantonese and Japanese words can vary

dramatically, mainly depending on the pronunciation of Japanese words (see Verdonschot et al., 2013, for more details on the characteristics of Japanese Kanji).

The characteristics of the Chinese and Japanese Kanji described above provide a unique research opportunity because one can directly examine the effects of phonological similarity on cognate processing in bilinguals of the two languages. Thus, the present study administered an L2 lexical decision task using Japanese and Cantonese cognates. Most importantly, orthography and semantics were identical for each word, but phonological similarity differed systematically.

Moderators of Cognate Processing in Bilinguals

Several studies with bilinguals of alphabetic languages have suggested that cognate processing may be influenced not only by phonological similarity (Carrasco-Ortiz et al., 2021; Dijkstra et al., 1999, 2010; Frances et al., 2021) but also by some other stimulus- and participant-level factors, such as word frequency (Dijkstra et al., 2010; Peeters et al., 2013; Xiong et al., 2020) and L2 proficiency (Carrasco-Ortiz et al., 2021; Midgley et al., 2011; Van Hell & Dijkstra, 2002). The present study aimed to extend this literature by investigating the moderating role of these three factors on cognate processing in bilinguals of logographic scripts (i.e., Chinese and Japanese Kanji).

This study operationalized the phonological similarity of cognate words with two measures. First, Levenshtein distance was used as an objective measure of phonological similarity. It is an index that quantifies the phonological difference between a pair of words. It is calculated as the number of insertions, deletions, and substitutions of sounds at the grapheme level required to transform one sound string into another. Thus, smaller values indicate higher phonological similarity (for more details, see Carrasco-Ortiz et al., 2021; Schepens et al., 2012; Wieling et al., 2012). Levenshtein distance has been used in many

previous bilingual studies on cognate processing (Carrasco-Ortiz et al., 2021; Peeters et al., 2013).

However, some researchers argue that Levenshtein distance may not be a perfect measure of phonological similarity (e.g., Frances et al., 2021) because, for example, some phonemes that have the same pronunciation may be represented by different graphemes between languages (e.g., /j/ in Cantonese and /y/ in Japanese are pronounced the same). In addition, suprasegmental information, such as lexical tones and accents, is not considered when calculating the distance between the phonologies of languages.

Therefore, to address the limitations of Levenshtein distance, we also evaluated the phonological similarity between the Chinese and Japanese cognates using participants' ratings. A previous study with Mandarin-Japanese bilingual adults has shown that the correlation between Levenshtein distance (reverse-coded) and subjective phonological similarity rated by participants was as high as .49 (Hayakawa et al., 2017). This moderate correlation indicates that the two measures may reflect at least partly different aspects of phonological similarity and thus may have different influences on cognate lexical decision. However, to our knowledge, no study has examined the effects of the two phonological similarity measures on cognate processing simultaneously.

Besides the effect of phonological similarity on cognate lexical decision, the relationship between word frequency and the cognate facilitation effect was also examined in this study. According to the BIA+ and Multilink models, visual word input triggers lexical (orthographic or phonological) representations in both L1 and L2 (Dijkstra & Van Heuven, 2002; Dijkstra et al., 2019). The activation of representations depends on their resting level activation, which is affected by word frequency. Specifically, words with higher frequency, regardless of their cognate status, should facilitate processing because the resting level activation of such words should be higher than those of low-frequency words. Put simply,

high-frequency words are activated more easily than low-frequency words in one's mental lexicon (see Figure 2 of Dijkstra & Van Heuven, 2002 and Figure 1 of Dijkstra et al., 2019, for graphic illustrations).

Several studies have tested this prediction in bilinguals of alphabetic (e.g., Dijkstra et al., 2010; Peeters et al., 2013) and logographic scripts (e.g., Xiong et al., 2020), which resulted in mixed findings. For example, in a study with French-English bilinguals, Peeters et al. (2013) found that participants' RTs were faster for cognates with high frequency in either L1 French or L2 English. In contrast, in a study with Mandarin-Japanese bilinguals, Xiong et al. (2020) found no effect of L1 Mandarin or L2 Japanese word frequencies on cognate lexical decision, which is inconsistent with the prediction of the BIA+ and Multilink models. Moreover, the cognate facilitation effect was smaller for high-frequency L2 Japanese words than for low-frequency words. Arguably, the lexical information of low-frequency L2 words is less accessible to bilinguals. Thus, they may rely more on L1 lexical information in cognate processing.

Xiong et al. (2020) suggested some potential reasons for this discrepancy, which were addressed in the present study. Specifically, the researchers warned of insufficient statistical power due to the limited number of trials and the relatively high frequency of cognate words, even for stimuli in the low-frequency condition.[1] In the present study, we recruited more participants ($N = 55$ compared to $N = 24$ in the study by Xiong et al., 2020) to have sufficient power to detect effects. Moreover, stimuli with a wider frequency range were included. To this end, stimuli with different L1 Cantonese and L2 Japanese word frequencies were used to examine whether the cognate facilitation effect, if present, was modulated by L2 Japanese word frequency. In addition, the effects of L1 Cantonese and L2 Japanese word frequencies on cognate processing were examined.

Finally, L2 proficiency is another important factor that can modulate cognate processing. The BIA+ and Multilink models postulate that resting level activation increases along with L2 proficiency, allowing for faster activation of L2 representations (i.e., higher proficiency readers are generally faster in reading L2 words, including cognates and non-cognates). Regarding cognate facilitation, previous studies in alphabetic languages have found that lower proficiency readers show larger cognate facilitation effects than higher proficiency readers (e.g., Andras et al., 2022; Bultena et al., 2014; Casaponsa et al., 2015; Krogh, 2022; Rosselli et al., 2014). For example, Krogh (2022) found a cognate facilitation effect in L2 lexical decision only for Danish-English bilinguals with lower L2 proficiency. Krogh (2022) explained this finding in terms of the BIA+ model. Since the co-activation of L1 and L2 orthographic representations is more pronounced in lower proficiency readers, their lexical representations may reach the recognition threshold for cognates faster than that for non-cognates. In contrast, the gap between L1 and L2 resting level activations is minimized in highly proficient readers. Consequently, the potential to benefit from converging activation of L1 and L2 lexical representations may also be reduced.

The Present Study

In this pre-registered study, we examined the effects of phonological similarity, word frequency, and L2 proficiency on cognate lexical decision in late bilinguals of two languages using logographic scripts, traditional Chinese characters (L1) and Japanese Kanji (L2). Specifically, we aimed to answer the following three research questions.

Our first research question of the present study was whether the cognate facilitation effect exists in logographic scripts, as well as its relationship to L2 word frequency and L2 proficiency. We hypothesized that the cognate facilitation effect would be present due to non-selective bilingual lexical access, as found in previous studies (e.g., Dijkstra et al., 2010; Peeters et al., 2013; Xiong et al., 2020). The cognate facilitation effect was expected to be

stronger for low-frequency Japanese cognates than for high-frequency Japanese cognates based on the findings from Xiong et al. (2020). In addition, a stronger cognate facilitation effect was expected in participants with lower L2 Japanese proficiency based on the results from Krogh (2022).

The second research question, and the main focus of the present study, was to examine the influence of phonological similarity on cognate lexical decision in Cantonese-Japanese bilinguals. We also examined whether L2 proficiency could modulate the influence of phonological similarity. We hypothesized that phonological similarity would facilitate cognate lexical decision based on previous findings in English lexical decision studies (Allen & Conklin, 2013; Allen et al., 2021; Miwa et al., 2014) and the fact that both Chinese and Japanese Kanji are orthographically opaque (Dong et al., 2021; Frost, 2005; Inoue et al., 2023; Wydell & Butterworth, 1999). The role of L2 proficiency in the relationship between phonological similarity and cognate lexical decision was exploratory as, to our knowledge, no study has been conducted on this issue before.

Finally, our third research question concerned the effects of word frequency in each language and L2 proficiency on cognate lexical decision. We predicted that cognates with higher word frequency in either L1 Cantonese or L2 Japanese would have a shorter RT (Peeters et al., 2013). Although Xiong et al. (2020) did not find this effect, it was expected in the present study given the increased statistical power and a wider range of word frequency among our stimuli. Failure to find an effect, on the other hand, would support a true difference in cognate processing between alphabetic and logographic scripts. In addition, faster lexical decision of cognates was expected for participants with higher L2 proficiency.

Method

The study's theoretical rationale, aims, research questions, hypotheses, experimental design, sample size estimation, and data analysis plans were all pre-registered on the Open

Science Framework (https://osf.io/jyeh9). All stimuli, data, and analysis code are available at https://osf.io/9bft6/.

Participants

Fifty-eight Cantonese-Japanese late bilingual adults were recruited. Most previous studies that found a facilitatory effect of phonological similarity on cognate processing in Japanese-English bilinguals recruited around 20 participants with a trial number similar to or lower than that of the present study (e.g., Allen & Conklin, 2013; Allen et al., 2021; Miwa et al., 2014). To ensure sufficient statistical power to detect effects, and to take into account the the difficulty of the lexical decision task with respect to the average Japanese proficiency among Cantonese-Japanese bilinguals in Hong Kong, we aimed to recruit more participants. Two participants were excluded from the analysis because they had scored below 50% on the Japanese reading task described below. Another participant was omitted due to skipping many trials in the reading task. Therefore, 55 participants (*mean age*: 24; *age range*: 18–56; 42 females) remained for our analyses. All participants received a small honorarium of HK$50 as compensation after the experiment. Written consent from each participant had been obtained prior to testing. This study was approved by the Survey and Behavioural Research Ethics Committee (SBREC) of The Chinese University of Hong Kong (CUHK).

The participants' Cantonese and Japanese language backgrounds are shown in Table 1, and their English and Mandarin language backgrounds are shown in Table S1 in Supplementary Material I. Of the total 55 participants, 37 had passed the N1 level and 18 had passed the N2 level of the Japanese-Language Proficiency Test (JLPT; Japan Foundation & Japan Educational Exchanges and Services, 2009). Most participants perceived Japanese as their fourth language because the first language of most Hongkongers is typically Cantonese, with English and Mandarin being the second or third.

Table 1

The Means and SDs (in Parentheses) of Variables Related to Participants' Language Background of Cantonese and Japanese (N = 55)

Variable	Cantonese (L1)	Japanese (L2)	
Listening age of acquisition (AoA)	0.6 (1.1)	15.1 (6.3)	
Speaking AoA	0.9 (0.9)	16.4 (5.8)	
Reading AoA	1.6 (1.5)	16.6 (5.7)	
Writing AoA	1.8 (1.5)	17.1 (5.6)	
Years of use	23.9 (7.3)	8.4 (5.9)	
Self-perceived listening proficiency [a]	7.0 (0.3)	5.1 (0.8)	
Self-perceived speaking proficiency [a]	6.9 (0.4)	4.6 (0.8)	
Self-perceived reading proficiency [a]	6.9 (0.4)	5.3 (0.7)	
Self-perceived writing proficiency [a]	6.8 (0.5)	4.5 (0.9)	
Percentage of daily use currently	66.9 (19.5)	12.9 (11.4)	
JLPT [b]		N1 (N = 34)	N2 (N = 18)
Vocabulary/grammar (max: 60)	–	44.9 (10.0)	47.6 (10.3)
Reading (max: 60)	–	47.8 (11.1)	49.9 (12.2)
Listening (max: 60)	–	48.4 (10.1)	49.4 (9.4)
Total score (max: 180)	–	141.1 (24.4)	146.9 (27.4)

Note.

[a] Self-perceived language proficiency was measured on a 7-point scale ranging from 1 = *very poor* to 7 = *very good*.

[b] The scores of JLPT are from data that includes 52 participants only, as three participants took the old version of JLPT with a different sum score for each component. Participants reported the JLPT score of the highest level that they had passed (N1 or N2).

Stimuli

The stimulus design of the lexical decision task was adapted from Xiong et al.'s study (2020) with Mandarin-Japanese bilinguals. However, while most of the stimuli in the non-cognate condition were adapted from their study, most of the stimuli in the cognate condition were modified given the different research purposes and language contexts between the studies. Specifically, a new stimulus list was created due to the differences in frequency, number of strokes, and word usage between Cantonese with traditional Chinese characters and Mandarin with simplified Chinese characters. A full list of stimuli can be found in Table S2 in Supplementary Material II or https://osf.io/9bft6/.

There were a total of 480 stimuli: 240 words and 240 non-words. Of the 240 words, half of them were Chinese-Japanese cognates with identical orthography and meaning, while the other half were Japanese words that did not exist in Chinese (non-cognates). Among the 120 Chinese-Japanese cognates, half had relatively higher Japanese word frequency, and the other half had lower Japanese word frequency (hereafter referred to as HJ and LJ, respectively). In addition, within each HJ and LJ set of cognates, 30 words had higher Cantonese word frequency (HC), and the other 30 had lower Cantonese word frequency (LC). Thus, there were four sets of cognates, each with 30 words with different frequencies in both languages, namely HJHC, HJLC, LJHC, and LJLC.[2] This stimulus selection procedure was designed to ensure that our stimuli covered a wide range of word frequencies in each language. In addition, this procedure ensured that the correlation between the two frequencies was nearly zero, $r(118) = .01$, $p = .903$, thereby allowing us to enter both Japanese and Cantonese word frequencies as continuous variables in the same linear mixed-effects model without encountering any multicollinearity issues.

The Japanese word frequency was based on the Kanji database (Tamaoka et al., 2017; http://www.kanjidatabase.com/), and the Cantonese word frequency was extracted from Cifu,

a lexical database of Cantonese (Lai & Winterstein, 2020). The frequency of Japanese words (including both cognates and non-cognates) ranged from 153 to 80,850 out of 299,695,840 tokens (M = 11,384), while the frequency of Cantonese cognates ranged from 1 to 1,783 out of million tokens (M = 95). All stimuli were two-character nouns in Chinese and Japanese. The non-words were formed by combining the first character of each word in the word condition with the second character of another word to ensure that the orthographic, phonological, and semantic properties of the stimuli were the same between the word and non-word conditions. This approach is similar to those used in previous English studies, which modified some letters or phonemes of real words (e.g., Frances et al., 2021; Peeters et al., 2013). Therefore, any differences in processing between words and non-words could be attributed solely to their lexical status rather than to other reasons. Examples for each stimulus set are shown in Table 2.

Table 2

Examples for Each Stimulus Set in the Japanese Lexical Decision Task

Stimulus Set	Stimulus Example	Meaning	Japanese word log frequency of the stimulus example	Cantonese word log frequency of the stimulus example
Non-word	送素	-	-	-
HJHC	習慣	Habit	8.54	5.54
HJLC	火災	Conflagration	9.17	0.69
LJHC	元素	Element	6.22	3.99
LJLC	水路	Waterway	6.96	1.10
HJ non-cognates	番組	Program	10.31	-
LJ non-cognates	別荘	Villa	7.00	-

Note. HJ = high Japanese word frequency. HC = high Cantonese word frequency. LJ = low Japanese word frequency. LC = low Cantonese word frequency.

The Levenshtein distance, serving as a proxy for objective phonological similarity among cognates, was calculated in three steps. First, based on the methods introduced by Hayakawa et al. (2017), the Japanese and Chinese cognates were transformed into the phonetic transcriptions used in each language (i.e., Japanese romaji and Cantonese pinyin without tone numbers, respectively).

Next, the objective phonological similarity for each cognate was calculated using the Levenshtein Distance Calculator (https://phiresky.github.io/levenshtein-demo/). Following the calculation method of Yu et al. (2018), each insertion or deletion was counted as one point, while the replacement was counted as two points. For example, the Japanese romaji and Cantonese pinyin of 人生 are /jinsei/ and /jansang/, respectively. This word requires three replacements (replacing /i/ with /a/ at position 2, replacing /e/ with /a/ at position 5, and replacing /i/ with /n/ at position 6) and one insertion (inserting /g/ at position 7) to convert the former into the latter. Therefore, the Levenshtein distance of this word is 7. The average Levenshtein distance of the 120 stimuli was 7.67 (*SD* = 2.71; range = 2 to 15).

Finally, the Levenshtein distance was reverse-coded (subtracted from 16, the maximum Levenshtein distance plus one) to form objective phonological similarity and to align its interpretation with subjective phonological similarity (see below), where the larger the number, the more phonologically similar the words were. Examples of stimuli with different levels of objective phonological similarity are shown in Table 3.

Table 3

Three Examples of Stimuli with Different Objective Phonological Similarity and Subjective Phonological Similarity used in the Japanese Lexical Decision Task

Cognate	Cantonese pronunciation	Japanese pronunciation	Objective phonological similarity	Subjective phonological similarity
入口	/jap6.hau2/	/iri.guchi/	4 (12)	1.56
人生	/jan4.sang1/	/jin.sei/	9 (7)	4.28
態度	/taai3.dou6/	/tai.do/	14 (2)	6.13

Note. Objective phonological similarity was measured by Levenshtein distance, and subjective phonological similarity was rated by the participants (1 = *very different*; 7 = *very similar*). Levenshtein distance (shown in parentheses) was reverse-coded (subtracted from 16, the maximum Levenshtein distance plus one) to form objective phonological similarity to align its interpretation with subjective phonological similarity (i.e., larger values indicate higher phonological similarity). Subjective phonological similarity was averaged across participants' ratings for each item.

We ensured that the stimuli in the four stimulus sets did not differ significantly from a statistical standpoint in terms of phonological similarity, number of strokes, or number of homophones in each language. Relevant details can be found in Supplementary Material III.

Procedure

All experiments were conducted by trained experimenters in a laboratory at the university. First, an online self-report questionnaire asking about demographic variables and JLPT scores in each category (i.e., vocabulary/grammar, reading, listening) was sent to the participants prior to the experiments. When they arrived at the laboratory, they were given

another questionnaire about their language background. This questionnaire asked about the age of acquisition (AoA) and self-perceived proficiency in reading, writing, listening, and speaking in all the languages they spoke (see Tables 1 and S1). Participants were also asked to report the proportion of time they spent using each language.

After that, participants proceeded to complete three tasks in a specified sequence: they started with the Japanese lexical decision task, followed by the Japanese reading task. The final task was the phonological similarity rating task. Details about each task are provided below.

Japanese Lexical Decision Task

E-prime 3.0 (Psychology Software Tools, Pittsburgh, PA) was used to program the experiment. All stimuli were presented on a 23-inch monitor in black on a white background in Courier New font (28-point size). The distance between the participant and the screen was 80 cm. Each trial began with a fixation cross (+) appearing at the centre of the screen for 200 ms, followed by a 200-ms blank screen. After that, a two-character word in Japanese Kanji was presented at the centre for 2,000 ms. Participants were asked to decide and indicate as quickly and as accurately as possible whether each word was a real Japanese word or not by pressing the “f” key on the keyboard for real words and the “j” key for non-words. Each stimulus appeared for 2,000 ms, regardless of whether participants responded. The RT was defined by the duration from the stimulus onset to the participants’ response. After a 200-ms blank interval, the next trial started.

Participants first completed 16 practice trials (with feedback) to familiarize themselves with the task. Stimuli in the practice session were not used in the main experiment. The practice session was followed by the main experiment, which consisted of 480 stimuli equally divided into eight blocks. Hence, each block contained 60 stimuli. Between each block, participants were allowed to take a break if needed. The stimulus

sequence was uniquely pseudo-randomized for each participant, with the restriction that no more than two words or non-words were presented consecutively to avoid repetition effects.

Japanese Reading Task

A Japanese reading task was used to assess participants' Japanese proficiency. We did not use the JLPT score to infer their Japanese proficiency because some participants passed the N1 level while some participants passed the N2 level, making their scores incomparable. In addition, our preliminary analysis showed that the correlation between word reading accuracy and JLPT score (r = .29) was relatively higher than that between self-rated proficiency and JLPT score (r = .12) for the participants who passed N1. These correlations suggest that word reading accuracy is likely to be a better indicator of language proficiency than participants' self-ratings.

The task consisted of the 240 words used in the lexical decision task. The words were printed on an A4 sheet using MS P Gothic font (11-point size), with the same randomized order for each participant. Participants were instructed to read the words one at a time in Japanese, and they could tell the experimenter if they did not know how to read the words. The experimenter recorded the accuracy for each stimulus. Accuracy was coded as 1 only if participants pronounced both characters in each word correctly, without considering the Japanese pitch accent. Otherwise, accuracy was coded as 0.

In addition, the participants' performance on the Japanese reading task was used to ensure that they knew the Japanese pronunciations of the words (both cognates and non-cognates) used in the experiment. This procedure was necessary because the participants had to know the pronunciation of the Japanese words before the relationship between phonology and cognate lexical decision, which was the focus of the present study (our second research question), could be examined. If a participant did not know the correct pronunciation of

Japanese words, the corresponding trials in the lexical decision task were excluded from the analyses for this research question.

Phonological Similarity Rating Task

The subjective phonological similarity between the Cantonese and Japanese cognates was rated by the participants after the Japanese reading task. They were asked to listen to the Cantonese and Japanese pronunciations of each word and to rate the phonological similarity by pressing a button from one to seven on the keyboard (1 = *very different*; 7 = *very similar*). The Cantonese and Japanese pronunciations were created using text-to-speech synthesizers for each language (https://text-to-speech.imtranslator.net/speech.asp for Cantonese and https://wideo.co/text-to-speech/ for Japanese). The naturalness of the pronunciations was confirmed or adjusted where necessary by the first author, who is proficient in both languages, using Audacity (Audacity Team, 2012). All data from this task were included in the analysis, even if participants did not know how to read a particular word, because participants had listened to the recordings before rating the similarity.

The apparatus and setup of the task were the same as for the lexical decision task, except that two speakers were placed next to the monitor to present the recordings. Each trial began with a one-second pause, followed by the Cantonese recording presented by the speakers. Then, after a one-second interval, the Japanese recording was presented. A seven-point Likert scale labeled “1: very different” and “7: very similar” was then shown on the screen. Participants were given five seconds to rate the phonological similarity by pressing the corresponding buttons. After pressing a button or after five seconds without a response, the next trial started immediately. The stimuli were presented in a random order. A break was allowed in the middle of the task (after 60 stimuli had been presented).

Data Analyses

All pre-registered analyses were performed using linear mixed effects models (LMEMs) with the lmerTest package (Bates et al., 2015; Kuznetsova et al., 2017) in R Version 4.2.2 (R Core Team, 2022). The LMEMs were estimated using the maximum likelihood method. An inverse RT (-1,000/raw RT) was used as the dependent variable to address the positive skewness of the raw RT (Brysbaert & Stevens, 2018). The *marginal* R^2 and the *conditional* R^2 of the model, which represent the variance explained by the fixed factors alone and the variance attributed to both fixed and random factors, respectively (Nakagawa & Schielzeth, 2013), were calculated using the MuMIn package (Barton, 2019).

Besides estimating the *p*-values for each predictor in the model, effect sizes for individual predictors also provide important information about the magnitude of their effects (Funder & Ozer, 2019; Sullivan & Feinn, 2012). Therefore, we calculated the predictor effect size, f^2, based on the *marginal* R^2 with this formula: $f^2 = (R_2^2 - R_1^2) / (1 - R_1^2)$, where R_2^2 represents the variance explained for a model with the given effect and R_1^2 signifies the variance explained for a model without the given effect (Lorah, 2018). This measure can be interpreted as the proportion of variance explained by the given effect relative to the proportion of variance that is unexplained (Aiken & West, 1991). Importantly, this measure informs us about how large the effect size is: f^2 is considered small at a value of 0.02, medium at a value of 0.15, and large at a value of 0.35 (Cohen, 1992). The R script for all data analyses is available at https://osf.io/9bft6/.

Before running the LMEMs, observations were excluded if they met the following criteria. First, for all analyses, incorrect trials (12.90% of the complete data) and trials with an RT exceeding 2.5 *SD* (4.30% of the complete data) in the lexical decision task were excluded. Since the first research question was about the relationship between L2 word frequency, L2 proficiency, and cognate facilitation effect, both cognate and non-cognate trials (50% of the complete data) were included in the analysis. The inverse RT of Japanese

words (including both cognates and non-cognates) was analyzed using word type (cognates = 0.5 vs. non-cognates = -0.5), log Japanese word frequency (continuous), L2 proficiency (continuous), and the interactions of these three variables as fixed effects. Due to outliers and incorrect responses, 16.71% of the data from both cognate and non-cognate trials were excluded, leaving 10,994 out of 13,200 observations remaining.

Only data from cognate trials (50% of the word trials) were included to examine our second research question about whether and how phonological similarity influences cognate processing. Non-cognate trials were not included in this analysis because phonological similarity cannot be measured for non-cognates. Due to the exclusion criteria described above (i.e., incorrect trials and outliers), 15.73% of the cognate data were excluded. In addition, data were removed when participants did not know the correct pronunciation of Japanese words in the reading task (11.15% of the remaining cognate data). Seven items with fewer than ten observations (0.72% of the remaining cognate data) were discarded because too few observations in each item could hinder model convergence (Brauer & Curtin, 2018). Furthermore, trials with no phonological similarity rating (0.83% of the remaining cognate data), where participants did not respond within the time limit, were discarded. Thus, 28.44% of the cognate data were excluded, leaving 4,723 out of 6,600 observations.

Two models were tested to examine our main research question about the effects of objective phonological similarity (research question 2.1) and subjective phonological similarity (research question 2.2) on cognate lexical decision. In the first model, the inverse RT of cognates was examined with objective phonological similarity (continuous), L2 proficiency as inferred by accuracy in the Japanese reading task (continuous), and the interaction of these two variables as fixed effects. Similarly, in the second model, the inverse RT of cognates was examined with subjective phonological similarity averaged across participants' ratings for each item (continuous), L2 proficiency (continuous), and the

interaction of these two variables as fixed effects. Objective phonological similarity and subjective phonological similarity were centred to the corresponding item grand means, while L2 proficiency was centred to the subject grand mean.

Finally, a model was constructed to test the word frequency effect on cognate processing for our third research question. The inverse RT of cognates was analyzed using log Japanese word frequency (continuous), log Cantonese word frequency (continuous), L2 proficiency (continuous), and the interactions of these three variables as the fixed effects. Only cognate trials (50% of the word trials) were included in the analysis. Due to outliers and incorrect responses, 15.73% of the cognate data were excluded from this analysis, leaving 5,562 out of 6,600 observations.

We did not exclude any trials based on accuracy in the reading task for the first and the third research questions due to two reasons. First, participants could judge whether a word was a Japanese word or not without knowing the pronunciation. Additionally, the effects of word frequency on cognate lexical decision may be due to shared orthography and semantics (Midgley et al., 2011; Voga & Grainger, 2007; but see below for additional analyses related to this issue).

The random intercepts for participants and items were included in all analyses. In all three initial models, the within-subject and within-item variables were used as by-subject and by-item random slopes respectively to construct a maximal random effect structure (Barr et al., 2013). However, these random slopes were dropped because the models did not converge or a singular fit was observed even after changing the optimizer to “bobyqa” and increasing the maximum iterations (Brauer & Curtin, 2018; Matuschek et al., 2017). Therefore, the models without random slopes were considered the final models (see Table S5 in Supplementary Material IV for the initial and final models for each research question).

Results

Descriptive statistics for the three tasks were analyzed prior to fitting each mixed model. For the lexical decision task, the average accuracy was 87.10% (*SD* = 6.52%). Specifically, the average accuracy rates for cognates, non-cognates, and non-words were 87.79% (*SD* = 8.00%), 85.08% (*SD* = 7.76%), and 87.77% (*SD* = 11.94%), respectively. In addition, the average RT of the correct lexical decision was 806.01 ms (*SD* = 116.41 ms). The accuracy rates and RTs in correct trials, with their *SD*s in different frequency conditions, are shown in Table 4. The average accuracy of the Japanese reading task was 79.56% (*SD* = 10.99%). This accuracy level partly reflects the fact that our Japanese reading task covered a wider frequency range than previous studies (e.g., Xiong et al., 2020) and included more infrequent words. The average phonological similarity rating across items was 4.21 out of 7.00 (*SD* = 0.87). Some examples of stimuli with different phonological similarity ratings are listed in Table 3.

Table 4

Average Accuracy Rates and RTs (in Correct Trials only) with their SDs (in Parentheses) for each Word and Non-word Condition in the Lexical Decision Task

Condition	Accuracy Rate (%)	RT (ms)
Non-word	87.77 (11.94)	844.84 (163.78)
HJHC	96.06 (4.93)	715.79 (127.09)
HJLC	93.15 (7.60)	744.95 (126.63)
LJHC	81.45 (11.93)	769.67 (137.57)
LJLC	80.48 (12.38)	818.86 (144.12)
HJ non-cognates	87.64 (6.81)	750.37 (92.62)
LJ non-cognates	82.52 (9.61)	798.52 (106.47)

Our first research question examined the effect of word frequency on the cognate facilitation effect. The final model (Table 5) showed that the effect of word type was

statistically significant, $f^2 = .011$, $p < .001$. This finding aligns with the cognate facilitation effect observed in our data: the average RT of cognates ($M = 724.86$ ms, $SD = 206.50$ ms) was shorter than that of non-cognates ($M = 746.19$ ms, $SD = 191.47$ ms). In addition, the effect of Japanese word frequency was significant, $f^2 = .029$, $p < .001$, with a negative relationship between Japanese word frequency and RT. The same relationship was shown in the significant effect of L2 proficiency, $f^2 = .047$, $p < .001$. Thus, higher Japanese word frequency and higher L2 proficiency were associated with a shorter lexical decision time. There was also a statistically significant interaction between word type and L2 proficiency. As shown in Figure 1, the interaction indicates that the cognate facilitation effect (i.e., shorter RT on cognate trials than on non-cognate trials) was larger for participants with higher L2 proficiency than for those with lower proficiency. The other interactions were not statistically significant (all $p > .071$).

Since we hypothesized an interaction between word type and Japanese word frequency, we conducted a Bayes Factor analysis to provide evidence for the null effect. To obtain evidence for the null hypothesis, we used Bayes Factors derived from the Bayesian information criterion (BIC) (Kass & Rafferty, 1995; Wagenmakers, 2007) from fitted mixed models in lmer. To do this, we compared the original model with the interaction and another model without the interaction. The Bayes Factor was 0.001, which means that the no-interaction model was favored 736 times (1/0.001) over the interaction model. This result suggests evidence for the null effect of the interaction between word type and Japanese word frequency.

Table 5

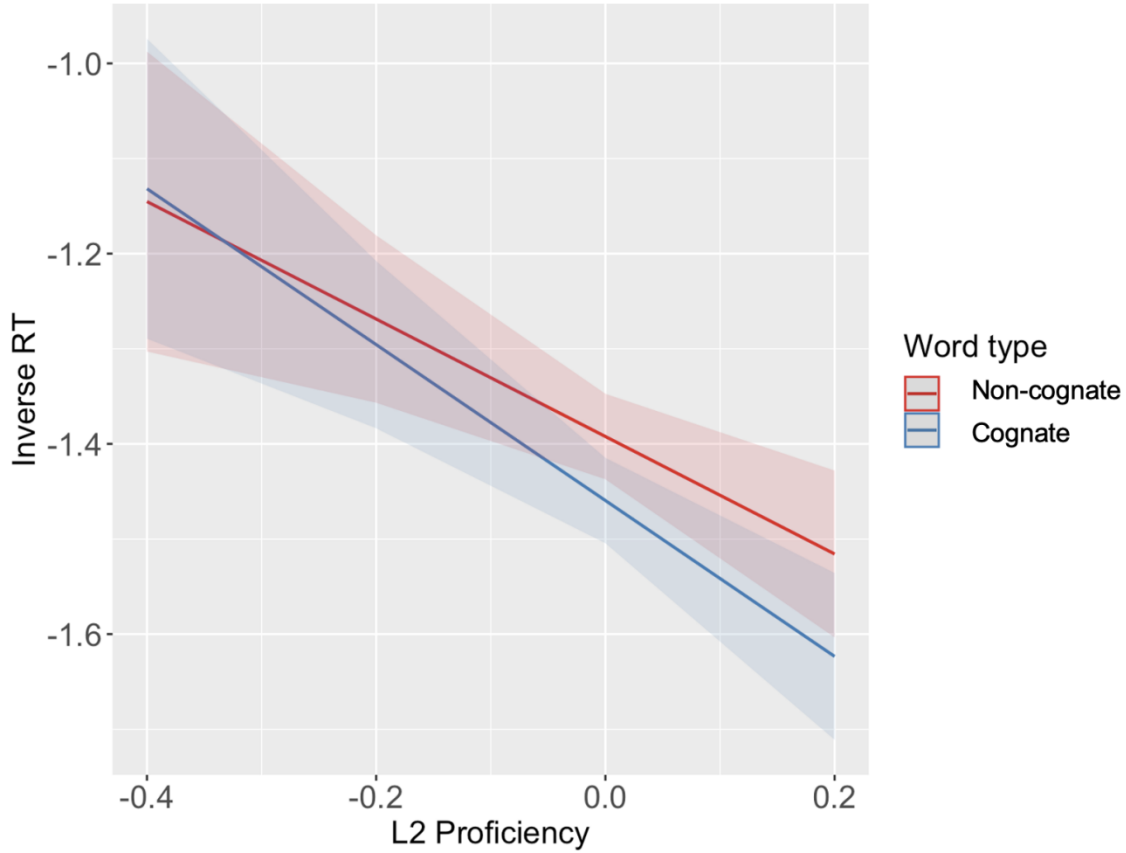
The Results of Fixed Effects in the Model of Inverse RT, Word Type, Japanese Word Frequency, and L2 Proficiency Measured by Japanese Word Reading Accuracy

	Estimate	*SE*	*t*	*p* (> |*t*|)	f^2
(Intercept)	-1.42	0.02	-65.21	< .001	–
Word type	-0.07***	0.01	-4.91	< .001	.011
Japanese word frequency	-0.05***	0.01	-8.30	< .001	.029
L2 proficiency	-0.72***	0.19	-3.76	< .001	.047
Word type x Japanese word frequency	0.02	0.01	1.81	.072	.001
Word type x L2 proficiency	-0.20***	0.05	-3.78	< .001	.001
Japanese word frequency x L2 proficiency	0.00	0.02	0.38	.701	.000
Word type x Japanese word frequency x L2 proficiency	0.00	0.04	0.75	.451	.000

Note. ***$p < .001$. *marginal* $R^2 = .08$. *conditional* $R^2 = .36$.

Figure 1

The Interaction of Word Type (Cognate vs. Non-cognate) and L2 Proficiency Measured by Japanese Word Reading Accuracy on Inverse RT



Note. The y-axis shows the inverse RT, with a more negative value indicating a faster RT. L2 proficiency was a continuous variable and was centred to the subject grand mean.

The first part of our second research question was about whether and how objective phonological similarity influences cognate lexical decision. The results (Table 6) showed that the fixed effect of objective phonological similarity on RT was statistically nonsignificant, $f^2 = .001$, $p = .348$. However, the fixed effect of L2 proficiency was statistically significant, $f^2 = .051$, $p < .001$, indicating that participants with higher L2 proficiency showed shorter RT. The interaction between objective phonological similarity and L2 proficiency was not statistically significant, $f^2 = .029$, $p = .352$.

Contrary to our prediction based on previous findings in English lexical decision studies (Allen & Conklin, 2013; Allen et al., 2021; Miwa et al., 2014), we failed to reject the null hypothesis for the effect of phonological similarity. However, this result alone does not "prove" that there is indeed a null effect (Wasserstein & Lazar, 2016). Therefore, again we conducted a Bayesian analysis by comparing a model that did not include objective phonological similarity with a model that did. The resulting Bayes Factor was 0.0002, which means that the model without the objective phonological similarity was favored 4,614 times (1/0.0002) over the model with objective phonological similarity. This result suggests that objective phonological similarity does not play a role in the RT of lexical decision.

Similarly, the second part of the primary research question, which was about whether and how subjective phonological similarity influences cognate lexical decision, was examined (Table 7). The fixed effect of subjective phonological similarity was not statistically significant, $f^2 = .001$, $p = .174$, but the fixed effect of L2 proficiency was statistically significant, $f^2 = .050$, $p < .001$, indicating that the RTs of participants with higher L2 proficiency were faster than those of participants with lower L2 proficiency. The interaction between subjective phonological similarity and L2 proficiency was not statistically significant, $f^2 = .028$, $p = .916$.

The Bayes Factor, generated by comparing the model without subjective phonological similarity with the model that included this variable, was 0.001. This value means that the model without subjective phonological similarity was favored 998 times (1/0.001) over the model with subjective phonological similarity. This result provides evidence that subjective phonological similarity did not affect lexical decision speed.

Table 6

The Results of Fixed Effects in the Model of Inverse RT, Objective Phonological Similarity, and L2 Proficiency Measured by Japanese Word Reading Accuracy

	Estimate	*SE*	*t*	*p* (> |*t*|)	f^2
(Intercept)	-1.46	0.02	-56.32	< .001	-
Objective phonological similarity	0.00	0.00	0.94	.348	.001
L2 proficiency	-0.80***	0.22	-3.59	< .001	.051
Objective phonological similarity x L2 proficiency	-0.02	0.02	-0.93	.352	.000

Note. ***$p < .001$. *marginal* $R^2 = .05$. *conditional* $R^2 = .35$.

Table 7

The Results of Fixed Effects in the Model of Inverse RT, Subjective Phonological Similarity, and L2 Proficiency Measured by Japanese Word Reading Accuracy

	Estimate	*SE*	*t*	*p* (> |*t*|)	f^2
(Intercept)	-1.46	0.03	-56.29	< .001	-
Subjective phonological similarity	-0.01	0.01	-1.37	.174	.001
L2 proficiency	-0.80***	0.22	-3.60	< .001	.050
Subjective phonological similarity x L2 proficiency	0.00	0.04	0.11	.916	.000

Note. ***$p < .001$. *marginal* $R^2 = .05$. *conditional* $R^2 = .35$.

Finally, our third research question was about how Japanese word frequency, Cantonese word frequency, and L2 proficiency modulate cognate lexical decision. Table 8 shows the results of the final model. It showed statistically significant effects of Cantonese word frequency, $f^2 = .014$, $p < .001$, Japanese word frequency, $f^2 = .018$, $p < .001$, and L2 proficiency, $f^2 = .058$, $p < .001$. These results indicate that these variables were negatively associated with RT. In other words, cognate RTs were shorter when Cantonese word frequency, Japanese word frequency, and L2 proficiency were higher. None of the interactions were statistically significant in this model (all $p > .143$).

As discussed above, we used Japanese reading accuracy to infer Japanese proficiency. However, we also analyzed the data using participants' self-ratings averaged across reading, writing, listening, and speaking as planned in the pre-registration. The results are reported in Supplementary Material V (see Tables S6-S9). All the results were similar, except that the interaction between word type and L2 proficiency for our first research question was absent. Hence, this interaction may require cautious interpretation.

Unlike the analyses for the second research question, our analyses for the first and third research questions did not exclude trials with words that participants could not read. However, because there might be potential phonological effects that could affect the data pattern, we conducted additional exploratory analyses by excluding these trials. The results were very similar (see Supplementary Material VI), showing that the results were robust and again indicating the minimal effects of phonology on cognate lexical decision.

Table 8

The Results of Fixed Effects in the Model of Inverse RT, Cantonese Word Frequency, Japanese Word Frequency, and L2 Proficiency Measured by Japanese Word Reading Accuracy

	Estimate	*SE*	*t*	*p* (> |*t*|)	f^2
(Intercept)	-1.45	0.02	-58.84	< .001	–
Cantonese word frequency	-0.02***	0.00	-5.33	< .001	.014
Japanese word frequency	-0.04***	0.01	-6.20	< .001	.018
L2 proficiency	-0.83***	0.22	-3.83	< .001	.058
Cantonese word frequency x Japanese word frequency	0.00	0.00	1.47	.144	.000
Cantonese word frequency x L2 proficiency	-0.00	0.02	-0.08	.933	.000
Japanese word frequency x L2 proficiency	-0.02	0.03	-0.55	.585	.000
Cantonese word frequency x Japanese word frequency x L2 proficiency	-0.00	0.02	-0.24	.811	.000

Note. ***$p < .001$. *marginal* $R^2 = .08$. *conditional* $R^2 = .35$.

Discussion

The present study examined the influences of phonological similarity, word frequency, and L2 proficiency on cognate lexical decision in bilinguals of two logographic scripts (L1 Cantonese and L2 Japanese). Additionally, the study investigated whether L2 word frequency and proficiency could modulate the cognate facilitation effect. The results showed that the influence of phonological similarity on cognate processing is minimal regardless of whether it is measured objectively or subjectively. On the other hand, as expected, participants responded faster to cognates than to non-cognates, and faster responses were shown for words with higher Japanese word frequency. The cognate facilitation effect was found to be larger for participants with higher L2 proficiency. In addition, cognate lexical decision was facilitated by higher Cantonese word frequency, Japanese word frequency, and L2 proficiency. Each of these results will be discussed below.

Regarding our first research question, we found a cognate facilitation effect, suggesting that Chinese representations were non-selectively activated during a Japanese lexical decision task. This finding corroborates existing evidence from different languages, including those with logographic scripts (Hsieh et al., 2017; Xiong et al., 2020). Therefore, non-selective access is likely to be universal in bilinguals, regardless of the writing system. Importantly, our results suggest that the cognate facilitation effect in logographic languages may not be primarily due to phonological similarity, but mainly to overlapping semantics and orthography. In addition, in line with the findings from Xiong et al.'s (2020) study with Mandarin-Japanese bilinguals, lexical decision was faster for words with higher Japanese word frequency.

However, in contrast to their findings, the interaction between word type (cognate vs. non-cognate) and Japanese word frequency was not statistically significant in the present study. This nonsignificant interaction may be partly due to the difference in Japanese

proficiency of the participants between the two studies. While all the participants in Xiong et al. (2020) had passed N1 of the JLPT and used Japanese more frequently, the participants in the present study included both N1 and N2 holders. When seeing cognates with low Japanese frequency, participants with relatively lower Japanese proficiency may be more likely to treat them as Chinese words instead of relying on Chinese lexical information to make a correct judgment. However, further studies with participants with a wider range of L2 proficiency are needed before a firm conclusion can be drawn.

Regarding the role of proficiency in relation to the cognate facilitation effect, we found a larger cognate facilitation effect in bilinguals who were more proficient in Japanese. This finding was contrary to our hypothesis and previous findings (Andras et al., 2022; Bultena et al., 2014; Casaponsa et al., 2015; Krogh, 2022; Rosselli et al., 2014).

One possible explanation is that the stimuli used in the present study may have been difficult for our participants. Indeed, Casaponsa et al. (2015) found similar results showing that the cognate facilitation effect was positively related to reading performance when readers had low proficiency, but the relationship became negative as proficiency increased. The former result suggests a positive relationship between the degree of reliance on cross-language similarity and proficiency, which can be applied to the present findings. Although the participants in the present study had relatively high proficiency in Japanese, some words may have been unfamiliar to them. In turn, they may have had to rely on connections with the L1 translations and the corresponding semantic representations during lexical decision tasks, leading to a larger cognate facilitation effect in more proficient participants. We should note, however, that this finding should be interpreted with caution, as this effect was only found when the L2 proficiency was measured by Japanese word reading accuracy. This finding was not replicated when we used participants' ratings as an indicator of L2 proficiency (see Table S8).

For our main research question, contrary to our expectation, the results did not provide evidence for the overall effect of phonological similarity on cognate processing. The absence of a phonological effect is inconsistent with previous findings from studies that included at least one alphabetic orthography. To recap, these studies showed a facilitatory (i.e., shorter RT) or inhibitory effect (i.e., longer RT) of phonological similarity on cognate processing (e.g., facilitatory: Allen & Conklin, 2013; Allen et al., 2021; Carrasco-Ortiz et al., 2021; Dijkstra et al., 2010; Haigh & Jared, 2007; Lemhöfer & Dijkstra, 2004; inhibitory: Dijkstra et al., 1999; Frances et al., 2021). In contrast, no such effects were found in our study. Moreover, the present finding does not support the postulate that a facilitatory effect would be found when two languages have similar orthographic depth either, as the orthography-phonology mappings in both Japanese and Cantonese are opaque (Frances et al., 2021).

Instead, the present finding is more consistent with a recent masked priming study by Liu et al. (2023), in which they compared the priming effects of phonologically similar and dissimilar cognates of simplified Chinese and Japanese Kanji characters using a word naming task. Although they found faster naming for phonologically similar targets than for phonologically dissimilar targets, they found equivalent priming effects between the two cognate types. These results showed that the priming effects were not affected by phonological similarity (Experiment 1). The same results were also found when suprasegmental similarity (i.e., Chinese tones and Japanese pitch accent) was taken into account (Experiment 3) and when Chinese-Japanese cross-script cognates (Chinese characters and Japanese katakana) were used (Liu et al., 2022).

Similar results were also found in a masked translation priming study with Chinese-English bilinguals by Zhang et al. (2019). They found no advantageous priming effect (i.e., shorter RT in translation) for cognates over non-cognates when Chinese words were used as

primes and English words were targets. However, the effect was found when the primed and target languages were reversed. These findings suggest that phonological activation may be hindered by the opaque mapping from orthography to phonology in Chinese. Thus, despite different research designs, our present findings corroborate those of Liu et al. (2022, 2023) and Zhang et al. (2019) and suggest that phonological similarity may only play a limited role in cognate processing in logographic scripts.

There are at least two possible explanations for the null effects of phonological similarity on cognate processing in the present study. First, the graphic complexity (Chang et al., 2018; Miton & Morin, 2021) and orthographic composition of logographic scripts may require readers to pay more attention to the orthographic information of the characters than to the phonological information of written words. Since most characters consist of multiple graphic components, including radicals (Shu & Anderson, 1997; Taylor & Taylor, 2014), this information may take priority for readers in logographic scripts. Indeed, the relative importance of orthographic information in logographic scripts has been demonstrated in a previous study with Mandarin-Japanese bilinguals (Xiong et al., 2020), but not in studies with alphabetic-script bilinguals (e.g., Peeters et al., 2013).

Moreover, although previous studies have shown that phonology plays a role in reading Chinese and Japanese Kanji characters (e.g., Hino et al., 2013; Perfetti & Zhang, 1995; Wydell et al., 1993), its role in cognate lexical decision may be limited due to the opaque relationship between print and sound in logographic scripts. This finding is consistent with the orthographic depth hypothesis, which stipulates that transparent orthographies facilitate lexical decision involving phonology. Conversely, opaque orthographies encourage readers to process words by accessing the lexicon and meaning through the word's orthography (Katz & Frost, 1992). This hypothesis also explains the discrepancies between

the present and previous findings from studies involving at least one alphabetic language (e.g., Allen & Conklin, 2013; Carrasco-Ortiz et al., 2021; Miwa et al., 2014).

Another possible explanation for the null effects of phonological similarity is that there may be competing (facilitatory and inhibitory) effects at different stages during lexical decision tasks, which may have counteracted each other. In a previous study using an eye-tracking technique with Japanese-English bilinguals, Miwa et al. (2014) showed that phonological similarity first triggered an inhibitory effect, followed by a facilitatory effect within the same trial. They concluded that the facilitatory and inhibitory effects exist at different stages of lexical decision, leading to the mixed findings in the previous lexical decision literature. In the present behavioral study, we cannot exclude the possibility that the null effects were due to a mixture of facilitatory and inhibitory effects occurring at different stages. Future studies are needed to clarify this issue in logographic script reading, possibly by using an eye-tracking technique.

Overall, our findings, together with similar results from previous studies (Liu et al., 2023; Zhang et al., 2019), did not converge with the premise of theoretical models of bilingual word recognition, including the BIA+ model (Dijkstra & Van Heuven, 2002) and the Multilink model (Dijkstra et al., 2019), which assume that the phonological similarity between words in two languages should benefit cognate lexical decision in bilinguals. Therefore, these models may need to be refined to explain the lack of phonological similarity effects on word recognition in logographic scripts.

One possibility may be to consider the potential moderating effect of script on the strength of the connection from orthography to phonology, as proposed by Zhang et al. (2019). Specifically, the connection between the two may be weaker in logographic scripts than in alphabetic scripts, as the mapping between orthography and phonology is more opaque in the former.

Another possible refinement is related to the relationship between L1 and L2 phonological representations. Liu et al. (2023) proposed an extension of the BIA+ model and suggested that L1 and L2 phonological representations are less tightly linked in logographic-script bilinguals compared to alphabetic-script bilinguals. Instead, the complete orthographic and semantic overlap in logographic scripts may lead to a high degree of connectivity between orthographic and semantic representations. Thus, the weak phonological activation, if there is any, may be overshadowed by the dominant orthographic and semantic coactivation during bilingual visual word recognition (see Figure 1 of Liu et al., 2023). The present findings provide empirical support for these ideas, which may extend the application of the BIA+ and Multilink models to bilinguals of languages that use logographic scripts.

Notably, the absence of a phonological similarity effect on cognate processing appears to be a bilingual phenomenon. This finding does not negate the importance of phonology in processing Japanese or Chinese words, as shown by previous studies. For example, Zhang et al. (2019) reported an advantageous priming effect using English primes with Chinese targets. Similarly, Liu et al. (2023) found that Chinese-Japanese bilinguals named phonologically-similar cognates faster than their dissimilar counterparts in a word naming task (Experiment 1). This finding appears to contradict the results of our current study, where we observed no reduced RTs for cognates with greater phonological similarity. However, the discrepancy may arise because phonological activation is inherently more crucial in word naming tasks than in lexical decision tasks, as the former requires participants to actively retrieve and articulate the target words. Although phonological activation may be subtler in lexical decision tasks, studies indicate that phonological representations are indeed activated in the minds of Chinese and Japanese monolinguals during these tasks (e.g., Cheng, 1992; Morita & Tamaoka, 2002). Therefore, while phonology is activated during the

processing of logographic languages, its effect may be too nuanced to significantly alter cognate processing outcomes.

Finally, our third research question was about whether Japanese word frequency and Cantonese word frequency contributed to cognate processing. We predicted that Japanese and Cantonese cognates with higher frequency should result in shorter RTs. This hypothesis was supported, and it corroborated with previous findings in alphabetic scripts (e.g., Peeters et al., 2013). Thus, the facilitatory effect of word frequency on cognate processing may be similar between logographic and alphabetic scripts. The statistically nonsignificant effect found in the study by Xiong et al. (2020) may, as they explained, be due to insufficient statistical power and the narrow range of word frequency in their stimuli. The results of the present study support these explanations by using a larger sample size (and thus greater statistical power) and stimuli with a wider range of frequency in the experiment.

Despite the contributions of the present study, several limitations need to be noted. First, a relatively large number of trials had to be removed from the analyses, especially for the second research question, as only the trials corresponding to correct naming in the Japanese reading task remained. The deletion of trials may have resulted in a potential loss of statistical power. Therefore, the interaction results may need to be interpreted with caution. Nevertheless, we believe that our remaining observations are sufficient to detect the main effects of phonological similarity on cognate lexical decision, if they exist, which was the main research question of the present study. In fact, our remaining observations are more than most previous studies examining the relationship between phonological similarity and lexical decision (e.g., Allen & Conklin, 2013; Allen et al., 2021; Miwa et al., 2014). These studies recruited approximately 20 participants with similar or fewer trial numbers, and a facilitatory phonological similarity effect was found in Japanese-English bilinguals.

Therefore, our conclusion that phonological similarity has a minimal effect on cognate lexical decision is tenable.

Second, there is a possibility that participants might have relied on their L1 knowledge to make lexical decisions because there were no non-cognate Cantonese words in our stimuli. In other words, when participants saw a Cantonese word, it had to be a real Japanese word. Consequently, they might have adopted a strategy based on the presence of Cantonese words rather than Japanese words, potentially utilizing their L1 phonology to make lexical decisions. This strategy might have obscured a potential phonological similarity effect. While we cannot exclude this possibility completely, we think that it was unlikely to affect our findings due to the following reasons.[3]

One might expect the cognate facilitation effect to be more pronounced for listeners with low L2 proficiency, as they would be more likely to adopt this strategy. However, participants with low proficiency did not exhibit a cognate facilitation effect, as illustrated in Figure 1. Additionally, if this strategy was prevalent, the impact of Japanese word frequency would diminish, while the influence of Cantonese word frequency would increase, particularly for cognates with high Cantonese word frequency. Yet, Japanese frequency emerged as a statistically significant predictor of RT, showing a slightly higher effect size than Cantonese frequency (Japanese frequency: $f^2 = 0.18$; Cantonese frequency: $f^2 = 0.14$; see Table 8). Furthermore, the interaction between Japanese and Cantonese frequencies was not statistically significant and demonstrated a weak effect size.

To further investigate this issue, we analyzed subsets containing only low L2 proficiency or high Cantonese frequency data, based on their medians. Crucially, the patterns reported in Table 8 remained consistent, with Japanese word frequency, Cantonese word frequency, and L2 proficiency being the only statistically significant variables. Therefore, it is unlikely that participants widely adopted this strategy to a degree that would influence our

findings. Nevertheless, future research investigating relevant issues may consider including L1-specific words without cognate equivalents in L2 (e.g., Cantonese words not present in Japanese). This would prevent participants from using their L1 knowledge and allow researchers to examine the effects of stimulus lists on the relationship between phonological similarity and cognate processing.

The third limitation of the present study is that, although participants' ratings may be a better measure of phonological similarity than Levenshtein distance (as discussed in the Introduction), it may not be a perfect measure. One advantage of the participants' ratings is that the suprasegmental information (e.g., Cantonese lexical tone and Japanese pitch accent) can be taken into account. However, the ratings may still not directly reflect participants' phonological representation during the lexical decision task because participants listened to the correct Japanese and Cantonese pronunciation before rating the phonological similarity, but they did not do so in the lexical decision task. Therefore, they might have different representations of Japanese pitch accent in their minds during the lexical decision task. Also, in previous studies, word pairs were mainly presented visually rather than auditorily to tap on participants' own phonological representations for their ratings (e.g., Dijkstra et al., 1999; Liu et al., 2023). In the present study, however, it was essential to present the word pairs auditorily to ensure that participants rated the correct pronunciation. To conclude, future studies may need to pay more attention to the effects of suprasegmental information on cognate processing.

In addition, the present study used only a lexical decision task, which has been reported to be more sensitive to semantic processing and less sensitive to phonological processing (Haigh & Jared, 2007; Kim & Davis, 2003). However, previous studies using the same task have shown a facilitation effect of phonological similarity in different script bilinguals (e.g., Allen & Conklin, 2013; Allen et al., 2021; Miwa et al., 2014). Therefore, we

believe that participants' phonological representations were activated when making lexical decision, albeit to a lesser extent in logographic scripts (e.g., Cheng, 1992; Morita & Tamaoka, 2002). Importantly, Liu et al. (2023) found no differences in the phonological priming effect in a word naming task between cognates with high and low phonological similarity in Mandarin-Japanese bilinguals. The present study showed comparable results and conclusions in a similar group of participants (Cantonese-Japanese bilinguals using traditional Chinese script).

Another limitation is related to the characteristics of our participants. For most of our participants, Japanese was their third/fourth language, as children in Hong Kong generally learn English from kindergarten and Mandarin from primary school. Therefore, it is unknown whether there were any influences from other languages they knew, especially Mandarin, which is also a language with a logographic orthography. However, we believe that the likelihood is low because the majority of our sample (N = 44 out of 55) reported very little time using Mandarin in their daily lives (M = 4.9 out of 100, SD = 5.9). In addition, the remaining 11 participants did not report Mandarin as a language they knew. Therefore, the influences of participants' English and Mandarin proficiency on our results, if there were any, were minimal.

Finally, we only included 120 cognates in the present study, and most of them were Japanese On-reading words. Although a similar number of cognates has been used in previous studies (e.g., Xiong et al., 2020), it is uncertain whether the results can be generalized to other cognates. Further studies are needed to examine the relationship between phonology and cognate processing with more cognates, especially more Kun-reading words.

To conclude, the present study examined the influence of phonological similarity on cognate lexical decision in bilinguals of two logographic scripts (L1 Cantonese and L2 Japanese) with words that are completely identical in orthography and meaning. Together

with some other previous studies (e.g., Allen & Conklin, 2013; Allen et al., 2021; Miwa et al., 2014), the present study highlights the importance of distinguishing the effects of orthography and phonology when examining how each lexical representation may contribute to cognate processing in bilinguals. Our results show that while word frequency of both languages and L2 proficiency may influence the lexical decision speed for cognates, phonological similarity has a limited overall effect. These findings suggest that current theoretical frameworks of word recognition in bilinguals (e.g., Dijkstra & Van Heuven, 2002; Dijkstra et al., 2019) may need to be refined to explain the limited overall effect of phonology on cognate processing in logographic scripts. To consolidate comprehensive models of bilingual representations and improve our understanding of word processing in bilinguals, further studies should examine the mechanisms of cognate processing across bilinguals of diverse sets of languages.

Acknowledgements

We would like to thank Mr. Wong Kwun Pok Oscar and Dr. Shuting Huo for assisting with part of the data collection and for providing statistical advice. We would also like to thank Prof. Urs Maurer for offering his laboratory for data collection. In addition, we would like to thank the reviewers for providing constructive feedback on previous versions of the manuscript.

Declaration of conflicting interests

The Authors declare that there is no conflict of interest

Funding

This work was supported by the General Research Fund [grant number 14617721] from the Hong Kong Special Administrative Region Research Grants Council and the Start-up Fund from The Chinese University of Hong Kong [grant number 4930965] to Tomohiro Inoue.

Footnote

[1] The possibility that cognate facilitation effects mainly occurred at pre-lexical representations, but not at the semantic representations, was also suggested as one of the reasons for the absence of the frequency effects on cognate processing. However, as this postulation needs to be addressed by EEG, it is not discussed in the present behavioral study.

[2] The categorizations of HJLC and LJHC are possible because, although the concepts remain consistent across both languages, the frequencies of the corresponding words can vary between the two. For example, “相撲” (sumo) is a word with high frequency in Japanese but low frequency in Cantonese because it is a more popular sport in Japan. Additionally, the same concepts can be conveyed using different expressions. For instance, “答案” (answer) is a word with high frequency in Cantonese but low frequency in Japanese, where “答え” (answer; a combination of kanji and hiragana) is more commonly used.

[3] We thank an anonymous reviewer for providing useful insights regarding this issue.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Sage.

Allen, D. B., & Conklin, K. (2013). Cross-linguistic similarity and task demands in Japanese-English bilingual processing. *PLoS ONE*, *8*(8), Article e72631. https://doi.org/10.1371/journal.pone.0072631

Allen, D., Conklin, K., & Miwa, K. (2021). Cross-linguistic lexical effects in different-script bilingual reading are modulated by task. *International Journal of Bilingualism*, *25*(1), 168–188. https://doi.org/10.1177/1367006920943974

Andras, F., Rivera, M., Bajo, T., Dussias, P. E., & Paolieri, D. (2022). Cognate facilitation effect during auditory comprehension of a second language: A visual world eye-tracking study. *International Journal of Bilingualism, 26*(4), 405–425. https://doi.org/10.1177/13670069211033359

Audacity Team (2012). Audacity(R) Version 2.0.0.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barton, K. (2019). MuMIn: Multi-Model Inference. R package version 1.47.1. https://CRAN.R-project.org/package=MuMIn.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review, 102*(1), 146–184. https://doi.org/10.1037/0033-295X.102.1.146

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23*(3), 389–411. https://doi.org/10.1037/met0000159

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 137–148. https://doi.org/10.1037/0096-1523.25.1.137

Bultena, S., Dijkstra, T., & Van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *Quarterly Journal of Experimental Psychology, 67*(6), 1214–1241. https://doi.org/10.1080/17470218.2013.853090

Carrasco-Ortiz, H., Amengual, M., & Gries, S. T. (2021). Cross-language effects of phonological and orthographic similarity in cognate recognition: the role of language dominance. *Linguistic Approaches to Bilingualism, 11*(3)*,* 389–417. https://doi.org/10.1075/lab.18095.car

Casaponsa, A., Antón, E., Pérez, A., & Duñabeitia, J. A. (2015). Foreign language comprehension achievement: insights from the cognate facilitation effect. *Frontiers in Psychology*, *6*, Article 588. https://doi.org/10.3389/fpsyg.2015.00588

Chang, L.-Y., Chen, Y.-C., & Perfetti, C. A. (2018). GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages. *Behavior Research Methods, 50*(1), 427–449. https://doi.org/10.3758/s13428-017-0881-y

Cheng, C.-M. (1992). Lexical access in Chinese: Evidence from automatic activation of phonological information. *Advances in Psychology*, *90*, 67–91. https://doi.org/10.1016/s0166-4115(08)61887-7

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Comesaña, M., Sánchez-Casas, R., Soares, A. P., Pinheiro, A. P., Rauber, A., Frade, S., & Fraga, I. (2012). The interplay of phonology and orthography in visual cognate word recognition: An ERP study. *Neuroscience Letters, 529*(1), 75–79. https://doi.org/10.1016/j.neulet.2012.09.010

de Groot, A. M. B., Delmaar, P., & Lupker, S. J. (1999). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *Quarterly Journal of Experimental Psychology*, *53*(2), 397–428. https://doi.org/10.1080/713755891

Dijkstra, T., Grainger, J., & Van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and language, 41*(4), 496–518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284–301. https://doi.org/10.1016/j.jml.2009.12.003

Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition, 22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Dong, J., Li, A., Chen, C., Qu, J., Jiang, N., Sun, Y., Hu, L., & Mei, L. (2021). Language distance in orthographic transparency affects cross-language pattern similarity

between native and non-native languages. *Human Brain Mapping*, *42*(4), 893–907. https://doi.org/10.1002/hbm.25266

Frances, C., Navarra-Barindelli, E., & Martin, C. D. (2021). Inhibitory and facilitatory effects of phonological and orthographic similarity on L2 word recognition across modalities in bilinguals. *Scientific Reports, 11,* Article 12812. https://doi.org/10.1038/s41598-021-92259-z

Frost, R. (2005). Orthographic systems and skilled word recognition processes in reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 272–295). Blackwell. https://doi.org/10.1002/9780470757642.ch15

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(5), 1122–1139. https://doi.org/10.1037/0278-7393.23.5.1122

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 623–644. https://doi.org/10.1037/0278-7393.33.4.623

Hayakawa, K., Yu, S., Chu, X., & Tamaoka, K. (2017). Nitityuu niji Kanjigo niokeru kyakkanteki onin ruijisei sihyoo: syukanteki onin ruijisei sihyoo to no hikaku [Objective indexes of phonological similarity between Japanese and Chinese two-Kanji compound words: A comparative study on objective and subjective indexes]. *Kwanseigakuindaigaku nihongo kyooiku senta kiyoo [Bulletin of the Center for*

Japanese Language Education Kwansei Gakuin University], 6, 21–34. http://hdl.handle.net/10236/00025910

Hino, Y., Kusunose, Y., Lupker, S. J., & Jared, D. (2013). The processing advantage and disadvantage for homophones in lexical decision tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(2), 529–551. https://doi.org/10.1037/a0029122

Hsieh, M.-C., Jeong, H., Kawata, K. H. D. S., Sasaki, Y., Lee, H.-C., Yokoyama, S., Sugiura, M., & Kawashima, R. (2017). Neural correlates of bilingual language control during interlingual homograph processing in a logogram writing system. *Brain and Language, 174,* 72–85. https://doi.org/10.1016/j.bandl.2017.06.006

Hsieh, M.-C., Jeong, H., Sugiura, M., & Kawashima, R. (2021). Neural evidence of language membership control in bilingual word recognition: An fMRI study of cognate processing in Chinese–Japanese bilinguals. *Frontiers in Psychology, 12,* Article 643211. https://doi.org/10.3389/fpsyg.2021.643211

Inoue, T., Zheng, M., Lui, K. F. H., McBride, C., & Ho, C. S.-H. (2023). Early literacy acquisition in logographic orthographies: Evidence from Chinese and Japanese. *Early Childhood Research Quarterly*, *63*, 73–84. https://doi.org/10.1016/j.ecresq.2022.11.009

Japan Foundation & Japan Educational Exchanges and Services. (2009). *The new Japanese-language proficiency guidebook*. http://www.jlpt.jp/e/reference/pdf/guidebook1.pdf

Kass, R. E., & Rafferty, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology,*

morphology, and meaning (pp. 67–84). North-Holland. https://doi.org/10.1016/S0166-4115(08)62789-2

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language, 49*(4), 484–499. https://doi.org/10.1016/s0749-596x(03)00093-7

Koda, K. (2017). Learning to read Japanese. In L. Verhoeven & C. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 31–56). Cambridge University Press. https://doi.org/10.1017/9781316155752.003

Krogh, S. M. (2022). Danish-English bilinguals' cognate processing in L1 and L2 visual lexical decision tasks. *Languages, 7*(3), 228. https://doi.org/10.3390/languages7030228

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(1), 1–26. https://doi.org/10.18637/jss.v082.i13

Lai, R., & Winterstein, G. (2020, May). Cifu: A frequency lexicon of Hong Kong Cantonese. In Proceedings of the 12th Language Resources and Evaluation Conference (pp. 3069–3077).

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition*, *32*(4), 533–550. https://doi.org/10.3758/bf03195845

Liu, C., Lupker, S. J., & Nakayama, M. (2022). Masked cognate priming effects with Chinese–Japanese bilinguals: Still no phonological facilitation for cognates with different scripts. *Studies in Language Sciences*, *20*(1), 1–8. https://doi.org/10.34609/sls.20.1_1

Liu, C., Wanner-Kawahara, J., Yoshihara, M., Lupker, S. J., & Nakayama, M. (2023). Cognate translation priming with Chinese–Japanese bilinguals: No effect of interlingual phonological similarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0001240

Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education, 6*(1), Article 8. https://doi.org/10.1186/s40536-018-0061-2

Lupker, S. J., Nakayama, M., & Perea, M. (2015). Is there phonologically based priming in the same−different task? Evidence from Japanese−English bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 41*(5), 1281–1299. https://doi.org/10.1037/xhp0000087

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94,* 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McBride-Chang, C., Shu, H., Zhou, A., Wat, C. P., & Wagner, R. K. (2003). Morphological awareness uniquely predicts young children's Chinese character recognition. *Journal of Educational Psychology, 95*(4), 743–751. https://doi.org/10.1037/0022-0663.95.4.743

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of Cognitive Neuroscience, 23*(7), 1634–1647. https://doi.org/10.1162/jocn.2010.21463

Miton, H., & Morin, O. (2021). Graphic complexity in writing systems. *Cognition, 214,* Article 104771. https://doi.org/10.1016/j.cognition.2021.104771

Miwa, K., Dijkstra, T., Bolger, P., & Baayen, R. H. (2014). Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals. *Bilingualism: Language and Cognition*, *17*(3), 445–463. https://doi.org/10.1017/s1366728913000576

Morita, A., & Tamaoka, K. (2002). Phonological involvement in the processing of Japanese at the lexical and sentence levels. *Reading and Writing*, *15*(7–8), 633–651. https://doi.org/10.1023/a:1020912217035

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nakayama, M. (2002). The cognate status effect in lexical processing by Chinese-Japanese bilinguals. *Psychologia, 45*(3), 184–192. https://doi.org/10.2117/psysoc.2002.184

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2012). Cross-script phonological priming for Japanese-English bilinguals: Evidence for integrated phonological representations. *Language and Cognitive Processes, 27*(10), 1563–1583. https://doi.org/10.1080/01690965.2011.606669

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2013). Masked translation priming with Japanese–English bilinguals: Interactions between cognate status, target frequency and L2 proficiency. *Journal of Cognitive Psychology, 25*(8), 949–981. https://doi.org/10.1080/20445911.2013.839560

Nakayama, M., Verdonschot, R. G., Sears, C. R., & Lupker, S. J. (2014). The masked cognate translation priming effect for different-script bilinguals is modulated by the phonological similarity of cognate words: Further support for the phonological account. *Journal of Cognitive Psychology, 26*(7), 714–724. https://doi.org/10.1080/20445911.2014.953167

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language, 68*(4), 315–332. https://doi.org/10.1016/j.jml.2012.12.003

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30*(4), 473–485. https://doi.org/10.1016/0749-596X(91)90017-E

Perfetti, C. A., & Dunlap, S. (2008). Learning to read: General principles and writing system variations. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Crosslinguistic relationships in first- and second-language literacy development* (pp. 13–38). Lawrence Erlbaum.

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24–33. https://doi.org/10.1037/0278-7393.21.1.24

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 139–156. https://doi.org/10.1037//0278-7393.27.1.139

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from https://www.support.pstnet.com/.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rosselli, M., Ardila, A., Jurado, M. B., & Salvatierra, J. L. (2014). Cognate facilitation effect in balanced and non-balanced Spanish–English bilinguals using the Boston Naming Test. *International Journal of Bilingualism*, *18*(6), 649–662. https://doi.org/10.1177/1367006912466313

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition, 15*(1), 157–166. https://doi.org/10.1017/S1366728910000623

Shu, H., & Anderson, R. C. (1997). Role of radical awareness in the character and word acquisition of Chinese children. *Reading Research Quarterly, 32*(1), 78–89. https://doi.org/10.1598/RRQ.32.1.5

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the P value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/jgme-d-12-00156.1

Tamaoka, K., Makioka, S., Sanders, S., & Verdonschot, R. G. (2017). www.Kanjidatabase.com: a new interactive online database for psychological and linguistic research on Japanese Kanji and their compound words. *Psychological Research, 81*(3), 696–708. https://doi.org/10.1007/s00426-016-0764-3

Taylor, I., & Taylor, M. M. (2014). *Writing and literacy in Chinese, Korean and Japanese: Revised edition* (Vol. 14). John Benjamins Publishing Company.

Timmer, K., Ganushchak, L. Y., Mitlina, Y., & Schiller, N. O. (2014). Trial by trial: selecting first or second language phonology of a visually masked word. *Language, Cognition and Neuroscience, 29*(9), 1059–1069. https://doi.org/10.1080/01690965.2013.824994

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*(4), 780–789. https://doi.org/10.3758/BF03196335

Verdonschot, R. G., La Heij, W., Tamaoka, K., Kiyama, S., You, W. P., & Schiller, N. O. (2013). The multiple pronunciations of Japanese Kanji: A masked priming investigation. *Quarterly Journal of Experimental Psychology, 66*(10), 2023–2038. https://doi.org/10.1080/17470218.2013.773050

Verhoeven, L., & Perfetti, C. (2022). Universals in learning to read across languages and writing systems. *Scientific Studies of Reading*, *26*(2), 150–164. https://doi.org/10.1080/10888438.2021.1938575

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition, 35*(5), 938–952. https://doi.org/10.3758/bf03193467

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p -values: Context, process, and purpose. *The American Statistician*, *70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Wieling, M., Margaretha, E., & Nerbonne, J. (2012). Inducing a measure of phonetic similarity from pronunciation variation. *Journal of Phonetics, 40*(2), 307–314. https://doi.org/10.1016/j.wocn.2011.12.004

Wydell, T. N., & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition*, *70*(3), 273–305. https://doi.org/10.1016/s0010-0277(99)00016-5

Wydell, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to meaning for Kanji: Is a rows still a rose in Japanese Kanji?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 491. https://doi.org/10.1037/0278-7393.19.3.491

Xiong, K., Verdonschot, R. G., & Tamaoka, K. (2020). The time course of brain activity in reading identical cognates: An ERP study of Chinese-Japanese bilinguals. *Journal of Neurolinguistics, 55,* Article 100911. https://doi.org/10.1016/j.jneuroling.2020.100911

Yu, S., Kim, J., & Tamaoka, K. (2018, May). Nizi kanzigo no onin ruizisei oninteki kyori ni kansuru nikkantyuuetu deetabeesu no onrain kensaku enzin no kootiku [Constructing a web-accessible search engine for the database of phonological similarity and phonological distances for orthographically – similar two-Kanji compound words among Japanese, Korean, Chinese and Vietnamese]. Poster session presentation at the meeting of the Nihongo Kyooiku Gakkai Syunki Taikai [the Conference in Spring 2018 on Teaching Japanese as a Foreign Language]. Tokyo, Japan.

Zhang, J., Wu, C., Zhou, T., & Meng, Y. (2019). Cognate facilitation priming effect is modulated by writing system: Evidence from Chinese-English bilinguals. *International Journal of Bilingualism*, *23*(2), 553–566. https://doi.org/10.1177/1367006917749062

Zhou, H., Chen, B., Yang, M., & Dunlap, S. (2010). Language nonselective access to phonological representations: Evidence from Chinese–English bilinguals. *Quarterly Journal of Experimental Psychology*, *63*(10), 2051–2066. https://doi.org/10.1080/17470211003718705