Journal of Neurolinguistics 55 (2020) 100911

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

The time course of brain activity in reading identical cognates: An ERP study of Chinese - Japanese bilinguals

Kexin Xiong[a,*], Rinus G. Verdonschot[b], Katsuo Tamaoka[c,d]

[a] *Graduate School of Arts and Letters, Tohoku University, Sendai, Japan*
[b] *Department of Oral and Maxillofacial Radiology, Institute of Biomedical & Health Sciences, Hiroshima University, Hiroshima, Japan*
[c] *Graduate School of Humanities, Nagoya University, Nagoya, Japan*
[d] *College of Foreign Languages, Donghua University, Shanghai, China*

ARTICLE INFO

Keywords:
Bilingual cognate recognition
Language non-selective activation
Chinese - Japanese bilingual
Event-related potential (ERP)
N250
N400

ABSTRACT

Previous studies suggest that bilinguals' lexical access is language non-selective, especially for orthographically identical translation equivalents across languages (i.e., identical cognates). The present study investigated how such words (e.g., 学校 meaning "school" in both Chinese and Japanese) are processed in the (late) Chinese - Japanese bilingual brain. Using an L2-Japanese lexical decision task, both behavioral and electrophysiological data were collected. Reaction times (RTs), as well as the N400 component, showed that cognates are more easily recognized than non-cognates. Additionally, an early component (i.e., the N250), potentially reflecting activation at the word-form level, was also found. Cognates elicited a more positive N250 than non-cognates in the frontal region, indicating that the cognate facilitation effect occurred at an early stage of word formation for languages with logographic scripts.

1. Introduction

The question of how translation equivalents with matching orthography (i.e., cognates) are processed in visual word recognition has been of considerable interest in studies investigating bilingual lexical access (e.g., Cristoffanini, Kirsner, & Milech, 1986; Lemhöfer & Dijkstra, 2004) as it can shed light on whether the bilingual brain considers more than one language at one time or whether language use is restricted to the language-at-hand. Most studies investigating this matter have shown that lexical entries in both languages of a bilingual become activated and can influence reaction times (hereafter, RTs) and accuracy depending on the specific task at hand (e.g., De Groot & Nas, 1991; Dijkstra, Grainger, & van Heuven, 1999). The most prevalent effect known is the so-called cognate facilitation effect (e.g., Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Lemhöfer, Dijkstra, & Michel, 2004), that is, a word such as 'film' (a cognate in Dutch-English-French) typically shows shorter decision latencies compared to a non-cognate word such as 'tree' (i.e., 'arbre' in French, 'boom' in Dutch). This facilitation effect has been widely evidenced in behavioral studies (e.g., Duyck et al., 2007), as well as in studies which used event-related potentials (ERPs; e.g., Midgley, Holcomb, & Grainger, 2011; Peeters, Dijkstra, & Grainger, 2013), which are obtained by measuring the changes in electrical activity generated from participants' brains caused by a specific event (such as experimental stimuli). These and other empirical studies point to an integrated mental lexicon in bilinguals' brains (e.g., Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). As cognate words have various degrees of

---

[*] Corresponding author. Now at Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, 6-3 Aramaki Aoba, Aoba-ku, Sendai city, Miyagi, 980-8578, Japan.

E-mail addresses: xiongkexinkuma@gmail.com (K. Xiong), rinusverdonschot@gmail.com (R.G. Verdonschot), ktamaoka@gc4.so-net.ne.jp (K. Tamaoka).

https://doi.org/10.1016/j.jneuroling.2020.100911
Received 24 June 2019; Received in revised form 19 March 2020; Accepted 28 March 2020

orthographic and/or phonological overlap between languages (e.g., "water" in English and Dutch share the exact same orthography, while "tomato" /təmeɪtoʊ/ is written as "tomaat" /toma:t/ in Dutch), the activation of lexical orthographic representations depends on the degree of orthographical similarity as well as word frequency in each language. Accordingly, the recognition of orthographical identical cognates (e.g., "water" in English and Dutch) is much quicker than that of cognates with similar word forms between languages. This has been reported by Dijkstra et al. (2010) who used an L2-English lexical decision task (LDT) with Dutch-English bilinguals. They reported that the cognate facilitation effects between orthographically similar cognates and identical cognates are not linear. Additionally, Dijkstra et al. (2010) reported that phonological overlap facilitates the visual recognition of identical cognates. This led the authors to propose that non-identical cognates are stored as two orthographic representations, which are connected via an inhibitory link. When a non-identical cognate is presented, lexical information in both languages is activated, thereby facilitating the processing of word meaning. Next, the semantic representation will feed activation back to the orthographic representations. Consequently, the activation of the non-target word would be suppressed by the other lexical representation via an inhibitory link. In contrast, identical cognates are assumed to be stored in one shared orthographic representation, so that this lateral inhibition effect does not suppress the cognate recognition (Dijkstra et al., 2010, p. 299).

To investigate how identical cognates are represented in the bilingual brain, Peeters et al. (2013) performed an L2-English LDT on late French-English bilinguals and manipulated word frequencies in L2-English and L1-French. In this experiment, the phonological distance of each cognate was calculated using an International Phonetic Alphabet (IPA)-based normalized Levenshtein distance, and this phonological distance was controlled between conditions. Along with RTs and accuracy, participants' brain activation was recorded using EEG (electroencephalography). Specific attention was paid to the N400 which is an ERP component with a negative polarity that peaks approximately 400 ms after a stimulus with a semantic violation is presented (e.g., Kutas & Hillyard, 1980), typically reflecting difficulties in the integration of semantic information. Peeters et al. (2013) discussed the representations of cognates based on three theoretical views: (1) shared morphemic representation; (2) form-overlap; (3) and the two-morpheme view. For the first (1)—the shared morphemic representation— this view states that cognates have a *special morphology*, which differs from non-cognates. Under this theoretical view, compared to non-cognates, the input cognates are able to activate the shared morphemic representation, which results in a stronger activation at this level. The second (2) form-overlap view assumes that cognates are represented in a *shared orthographic representation*, *two phonological representations*, and a *shared semantic representation*. Compared to non-cognates, cognates are used in two languages, resulting in stronger lexical activation at both shared orthographic and shared semantic representations. Consequently, the first two views underscore that cognate facilitation is sensitive to *cumulative frequency*. Since participants in Peeters et al. (2013) were late French-English bilinguals, they used L1-French more frequently than L2-English. Therefore, cognates with a high L1-French word frequency but a low L2-English frequency would be processed quicker than cognates with a low L1-French word frequency but a high L2-English word frequency. Lastly and in contrast to the two previous views, the (3) two-morpheme view, which takes a localist connectionist stance, states that whether the L1-French word frequency or the L2-English word frequency have a primary effect on the cognate processing depending on the context. For example, in an L2-*English* LDT, the two-morpheme view predicts a primary effect of L2-English and a secondary effect of L1-French (Peeters et al., p. 318).

Using an L2-English lexical decision task, Peeters et al. (2013) showed that the N400 amplitude elicited by cognates with a high L1-French word frequency but a low L2-English word frequency (e.g., "trait") was larger than the N400 induced by cognates with a low L1-French word frequency but a high L2-English word frequency (e.g., "campus"). In other words, these results indicate that it is difficult to process cognates with a low L2-English frequency even if they are frequently used in L1-French. Based on these results, Peeters et al. (2013) proposed that identical cognates are stored in a single orthographic representation, two phonological and two morphological representations. That is, the activation of one of the representations would be stronger than the other according to the experimental task.

Cognates that have been examined in the abovementioned studies are words with alphabetic scripts (e.g., English and French). On the other hand, in languages using logographic scripts, more specifically, Chinese and Japanese, there are also many translation equivalents that have the same or similar orthographies. In a database for Japanese, Korean, and Chinese languages constructed by Park, Xiong, and Tamaoka (2014a, 2014b, available online at http://kanjigodb.herokuapp.com/, see also Yu & Tamaoka, 2015) which contains 2,058 two-character (kanji in Japanese) compound words (taken from the word lists of the fourth to second levels of the Japanese-Language Proficiency Test; 2007), approximately 1,507 words (or 73.23%) share matching orthographies. Among these words, approximately 1,163 have the same (or closely matching) meaning in Chinese and Japanese (see Xiong, 2018, who added bilingual semantic indexes to this database). Note that, unlike the intrinsic connection between orthography and phonology in languages with alphabetic scripts, the orthographic overlap is not necessarily related to phonological overlap between Chinese and Japanese. For instance, 性格 (meaning "personality", is pronounced as /sei.kaku/ in Japanese and /xìng.gé/ in Chinese; the pronunciation of each character is separated with a dot). Therefore, we define Chinese - Japanese identical cognates (hereafter, cognates) as words that share their *orthography* and *meaning* in two languages but not necessarily their *phonology*. In addition, we define the translation equivalents which do not share orthography in Chinese and Japanese as non-cognates. For example, the word "movie" is written as "映画" (/ei.ga/; lit. "movie picture") in Japanese but "电影" (/diàn.yǐng/; lit. "electric shadow") in Chinese.

There are undoubtedly similarities in cognate processing across languages with alphabetic scripts and logographic scripts; for example, cognates are typically easier to process than non-cognates in certain tasks and conditions (e.g., Nakayama, 2002). There are also considerable differences that can be expected during the visual word recognition due to the characteristics of languages with different writing systems. First, each Japanese kanji (or Chinese character) in cognates has its own meaning which has often been compared to a morpheme in languages with alphabetic scripts. For instance, the cognate "数学" (meaning mathematics, pronounced as /suu.gaku/ in Japanese and /shù.xué/ in Chinese) comprises two kanji characters, in which "数" means "number" and "学" means "knowledge". Additionally, the character "数" can also be used as a verb meaning "to count" in Chinese; for example, "数数"

(/shǔ.shù/) means “to count the numbers”. Accordingly, although the word “数学” is a cognate that shares the same meaning and orthography between languages, it has different morphological representations in Japanese and Chinese. More specifically, cognates are different at the character level in Japanese and Chinese. Second, as we described above, cognates can be pronounced very differently in Japanese and Chinese. In Chinese, characters occasionally contain phonetic radicals that carry some probabilistic information concerning the pronunciation (e.g., “抱”, is pronounced as /bào/ in Chinese, which is similar to the phonetic radical “包” /bāo/). This phenomenon sometimes carries over to the (Chinese derived) ON-reading for Japanese kanji as well (e.g., characters containing the radical 包 in Japanese are typically pronounced as /hou/, e.g., 抱負 /hou.fu/ “ambition”) however phonetic radicals are not informative at all for the Japanese Kun reading (i.e., the Japanese pronunciation of kanji characters; e.g., “抱え” is pronounced as /kaka.e/ in the Kun reading in Japanese; see Verdonschot et al., 2013 for more information). Conversely, visual features of logographs can directly encode semantic information. For example, most characters containing the 虫 part (or radical) on the left side have something to do with insects, such as 蟻 (‘ant’) and 蜂 (‘bee’), although there are occasional exceptions (e.g., 虹 ‘rainbow’). Therefore, during visual word recognition in Chinese and Japanese, extra attention might be paid to the most informative type of information (e.g., such as radicals) during orthographic processing.

1.1. The present study

The present study aims to clarify how Chinese - Japanese cognates are represented in the bilingual mental lexicon, using both RTs and ERPs during an L2-Japanese visual LDT administered to late Chinese - Japanese bilinguals. ERP components that are induced by task events (i.e., word stimuli) can provide valuable information concerning the different stages of visual word recognition (e.g., Zhou, Fong, Minett, Peng, & Wang, 2014). For example, the P200 has been associated with pre-lexical processes such as those pertaining to phonological and orthographical information (e.g., Kong, Zhang, Zhang, & Kang, 2012; Liu, Perfetti, & Hart, 2003). According to Kong et al. (2012), single Chinese characters that are orthographically similar with their primes elicit a reduced P200 compared to the characters that have unrelated orthographical forms with their primes. In addition, a negative deflection that peaks approximately 250 ms after the stimulus presentation has been referred to as the N250 component, which is considered to reflect form-level processing (Holcomb & Grainger, 2006). According to Holcomb and Grainger (2006), the N250 is distributed widely (though frontally prominent), and peaks around 250 ms and is sensitive to the degree of form overlap between the prime and target stimuli. An “amplified” N250 (i.e., more negative) occurs when there is less form overlap between primes and targets. This component has been confirmed by other studies combining a within-language masked priming paradigm with ERP recordings (e.g., Holcomb & Grainger, 2007; Okano, Grainger, & Holcomb, 2013) and has been associated with orthographical and phonological processing at the form-level in the mental lexicon. For cognates, previous evidence using European Portuguese-English bilinguals has suggested that cognates induce a less positive P200 relative to non-cognates (e.g., Comesaña et al., 2012). However, this reduced P200 to cognates was not reported in any other study. This might be as different tasks were employed. Tasks using the masked priming paradigm may invoke early sub-lexical processes, indicating initial discrimination of cognates as a function of their physical properties. Additionally, the N400, an important component when discussing the mechanisms involved in word processing (especially interesting for cognates) is known to reflect semantic processing. Despite a meta-analysis suggesting that activation of word meaning is not explicitly necessary in lexical decision tasks (Murphy, Jogia, & Talcott, 2019), the amplitude of the N400 induced by cognates have generally been shown to be smaller than that induced by non-cognates (Peeters et al., 2013). Therefore, in the present study, we record the brain activity changes of late Chinese - Japanese bilinguals during L2-Japanese logographic word recognition using a simple LDT task (i.e., without priming), focusing on the N400 to clarify what processes are involved in this task using logographs in the bilingual mental lexicon.

We selected cognates with different word frequencies in L1-Chinese and L2-Japanese from a database for Japanese, Korean and Chinese (Park et al., 2014a; 2014b). It is possible to manipulate the word frequencies of cognates in Japanese and Chinese though the written form is identical. For example, the Chinese - Japanese cognate 答案 (/dá.àn/ in Chinese and /too.an/ in Japanese, meaning “answer”) is frequently used in Chinese, but its frequency in Japanese is relatively low. This is because, in Japanese, its synonym 答え (/kota.e/) is more frequently used rather than 答案. In addition to cognates, non-cognates with different word frequencies in L2-Japanese will also be included.

Following the reasoning laid out by Peeters et al. (2013) and Dijkstra et al. (2010), both L2-Japanese and L1-Chinese are non-selectively activated and facilitate cognate processing. Hence, both shorter RTs, as well as a reduced N400, are expected in cognate processing compared to Japanese controls (i.e., a cognate facilitation effect). Furthermore, the size of the cognate facilitation effect is expected to depend on the word frequency, such that a smaller cognate effect is expected for high frequency Japanese words as compared to low frequency words. These two effects will be investigated in the main analyses of the data reported in this study. Additionally, as mentioned earlier, Chinese - Japanese cognates are assumed to be stored in one orthographical shared representation but must contain two-morpheme representations (i.e., at the character level in Japanese and Chinese) due to the characteristics of the logographic writing system. In a Japanese LDT, a primary effect of L2-Japanese knowledge on the cognate processing is predicted, followed by a secondary effect of L1-Chinese information. Therefore, a subsidiary analysis on just the cognates will try to elucidate the effects of word frequencies in Japanese and Chinese on the processing of cognates. Following Peeters et al. (2013), we expect a less negative N400 for cognates with high L2-Japanese word frequencies (hereafter, JWF) and low L1-Chinese word frequencies (hereafter, CWF) compared to those with low L2-JWF but high L1-CWF.

Table 1
Mean (SD) self-ratings using a 7 Likert scale on participants' L1-Chinese and L2-Japanese proficiency (1 = unable to 7 = excellent) as well as the frequency of L1-Chinese and L2-Japanese use (1 = rarely to 7 = very often).

	Listening	Speaking	Reading	Writing
Proficiency				
L1 Chinese	7.00 (0.00)	7.00 (0.00)	7.00 (0.00)	6.96 (0.04)
L2 Japanese	5.42 (0.19)	5.25 (0.19)	5.54 (0.15)	5.00 (0.19)
Frequency				
L1 Chinese	6.54 (0.24)	6.25 (0.33)	6.25 (0.33)	5.17 (0.42)
L2 Japanese	6.25 (0.19)	5.75 (0.21)	5.96 (0.24)	5.75 (0.26)

2. Methods

2.1. Participants

Twenty-four right-handed native Mandarin speakers studying at Nagoya University, Japan (four men, mean age = 25.94, *SD* = 2.63) were recruited. All participants use *simplified Chinese characters* when using Mandarin. They passed the most difficult level (N1) of the Japanese-Language Proficiency Test administered by the joint organization of the Japan Foundation and Japan Educational Exchanges and Services. On average, they studied Japanese for 6 years and 8 months (*SD* = 2.61). A self-rating questionnaire of Chinese and Japanese skills, as well as the frequency of regular use of the two languages, were conducted on all participants using a 7-point Likert scale. As shown in Table 1, our participants were very high-proficiency Japanese L2 speakers. They studied at a Japanese university and used L2-Japanese as frequently as L1-Chinese in their daily life.

2.2. Materials

In the present study, 120 Chinese - Japanese cognates, 120 Japanese controls, and 240 Japanese-like nonwords that are made of two Jōyō kanji (i.e., Chinese characters that are frequently used in Japanese) were administered. Both cognates and controls were selected from a Japanese - Korean - Chinese database (Park et al., 2014a, 2014b; Yu & Tamaoka, 2015). All cognates have no simplified form in Chinese (i.e., 云 [Chinese] = 雲 [Japanese]), meaning that they are visually identical in both languages (i.e., 金 [Chinese] = 金 [Japanese]). All stimuli are nouns both in Japanese and Chinese.

First, we divided the cognates and controls into two groups with high- and low-JWFs (Japanese Word Frequencies). The JWFs of stimuli were collected from an online Japanese lexical database (www.kanjidatabase.com; Tamaoka, Makioka, Sanders, & Verdonschot, 2017). Accordingly, four groups in total were created: (1) cognates with high JWF (HJ), (2) cognates with low JWF (LJ), and (3) controls with HJ and (4) LJ. A 2 (Word Type: cognates vs. controls) × 2 (JWF: HJ vs. LJ) ANOVA was conducted, using natural logarithm transformed JWFs as the dependent variable. The results showed that there was no significant JWF difference between cognates and controls [F (1, 236) = 0.02, p = .90]. Additionally, there was no significant interaction between Word Type and JWF [F (1, 236) = 0.30, p = .58]. The JWF difference was significant between HJ and LJ [F (1, 236) = 218.92, p < .001]. All the above results indicate that these four groups were suitably controlled with respect to the JWFs.

Next, strokes and the number of homophones were collected from the same web-accessible Japanese lexical database (www.kanjidatabase.com; Tamaoka et al., 2017). Strokes of each stimulus were calculated as the total of each kanji. The number of homophones was counted as follows: we first inputted the pronunciation of a certain stimulus, then look up all the two-kanji compound Japanese words that had this pronunciation. The number of results that included the stimulus was then counted as the number of homophones. Two one-way ANOVA was conducted on these four groups, using the strokes and the number of homophones as dependent variables. The results showed that there was no significant stroke difference between any of the four groups [F (3, 236) = 1.29, p = .28]. Similarly, the number of homophones among four groups was suitably controlled [F (3, 236) = 0.46, p = .71]. The JWF, the number of homophones and strokes are summarized in Table 2.

The semantic classification of cognates was determined based on the data provided by the 文化庁 (Agency for Cultural Affairs,

Table 2
JWF, strokes, and the number of homophones of four stimuli groups. Both JWFs and the number of homophones were transformed using a natural logarithm.

Characteristics	Cognates				Controls			
	HJ		LJ		HJ		LJ	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
JWF	9.43	0.86	7.71	0.88	9.38	1.03	7.78	0.66
Strokes	15.22	5.47	15.87	4.41	16.67	4.83	16.83	5.67
The number of homophones	0.50	0.70	0.48	0.64	0.42	0.54	0.38	0.62

Table 3
The means and standard deviations of word frequencies, strokes, and the number of homophones in Japanese and Chinese for the four groups of cognates. The word frequencies and the number of homophones were transformed using a natural logarithm.

Characteristics	HJ				LJ			
	HC		LC		HC		LC	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Word frequency								
Japanese	9.46	0.77	9.39	0.95	7.76	0.90	7.65	0.88
Chinese	11.21	0.58	9.30	1.21	11.14	0.98	9.04	0.48
Strokes								
Japanese	15.00	4.95	15.43	6.03	16.33	3.96	15.40	4.85
Chinese	14.83	4.92	15.30	6.07	16.33	3.99	15.60	5.08
The number of homophones								
Japanese	0.46	0.64	0.54	0.77	0.55	0.64	0.40	0.63
Chinese	0.53	0.60	0.46	0.57	0.87	0.94	0.66	0.75

1978) as well as a Chinese - Japanese dictionary of kanji-compound words (Zhang, 1987). Cognates were orthographically identical two-kanji compound words which had been classified semantically as the same words in either of these two indexes. High- and low frequency cognates (in Japanese) were further divided into those with high and low frequency in Chinese, using the CWF collected from the genre of newspaper in the BCC corpus (http://bcc.blcu.edu.cn/index.php; Xun, Rao, Xiao, & Zang, 2016). The result of a two-way ANOVA using Japanese word frequencies as a dependent variable showed that the main effect of JWF (HJ vs. LJ) was significant [F (1, 116) = 114.73, $p < .001$] but neither the main effect of CWF (HC vs. LC) [F (1, 116) = 0.30, p = .58] nor the interaction of JWF and CWF was significant [F (1, 116) = 0.01, p = .91]. Similarly, the result of a two-way ANOVA using Chinese word frequencies as a dependent variable revealed that the main effect of CWF (HC vs. LC) was significant [F (1, 116) = 161.45, $p < .001$], meanwhile neither the main effect of JWF (HJ vs. LJ) [F (1, 116) = 1.08, p = .30] nor the interaction of JWF and CWF was significant [F (1, 116) = 0.39, p = .53]. To further control the strokes and the number of homophones, four one-way ANOVAs on these four groups of cognates were conducted. The number of homophones in Chinese were collected from an online Chinese word dictionary (https://www.mdbg.net/chinese/dictionary). We confirmed that the strokes in Japanese [F (3, 116) = 0.38, p = .77] and Chinese [F (3, 116) = 0.46, p = .71], the number of homophones in Japanese [F (3, 116) = 0.32, p = .81], and the number of homophones in Chinese [F (3, 116) = 1.76, p = .16] were controlled between the four groups (See Table 3).

In addition to the properties above, we also collected the phonological distances for the cognates between Japanese and Chinese, which were calculated by Yu, Kim, and Tamaoka (2018) using the *cba* package (Buchta & Hahsler, 2016) in *R*. The phonological distances of cognates in the present study ranged from 1 to 12 (see Fig. 1). A larger value indicates the pronunciation in Japanese is

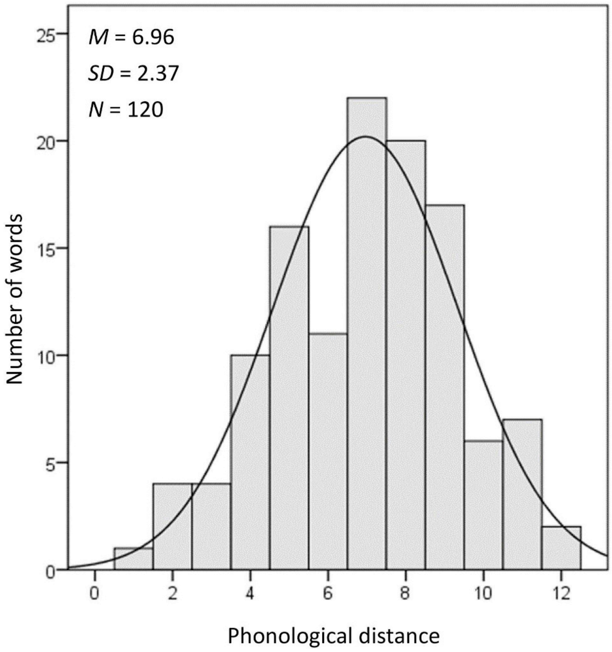

Fig. 1. The distribution of the phonological distances of cognates in the present study.

Table 4
The results of fixed effects in the GLMM model for native speakers.

	Estimate	SE	*t*	*p* (> |t|)
(Intercept)	489.96	9.69	50.58	< .001
Word Type	−4.44	5.60	−0.79	.43
JWF	−13.01	5.75	−2.26	.02
Word Type × JWF	9.78	8.77	1.12	.26

increasingly different from that in Chinese. For instance, the phonological distance of 内科 (“internal medicine”, pronounced as /nai.ka/ in Japanese and /nèi.kē/ in Chinese) is four, while the phonological distance of 食堂 (“cafeteria”, pronounced as /sho-ku.dou/ in Japanese and /shí.táng/ in Chinese) is eleven. Accordingly, the pronunciations in Japanese and Chinese of 内科 (“internal medicine”) sound more similar rather than those of 食堂 (“cafeteria”). A two-way ANOVA showed that there was no significant difference between the four groups in the phonological distance [F (3, 116) = 0.83, p = .48].

To verify whether cognates and controls were equally difficult for Japanese speakers who did not know Chinese, 38 native Japanese speakers (16 men; mean age = 19.23, *SD* = 0.43), participated in a Japanese LDT. In this pilot experiment, native Japanese speakers were asked to decide, as quickly and accurately as possible, whether the two-kanji-compound (target) is a Japanese word or not. All 480 stimuli which were made for Chinese - Japanese bilinguals were presented randomly during the experiment. The mean accuracy of the controls was 94.28% (*SD* = 0.23), which was as high as the accuracy of cognates (*M* = 95.81%, *SD* = 0.20). The accuracy of nonwords was 96.38% (*SD* = 0.19). Next, we analyzed the RTs using correct responses only.

For the analysis of RTs, we adopted Generalized linear mixed-effect models (GLMMs), using R version 3.5.1 (R Development Core Team, 2018). According to Lo and Andrews (2015), raw RTs are preferred to be used using GLMMs, as this avoids potentially undesirable effects on interactions caused by transformations. We analyzed RTs using *Word Type* (cognates = 0.5 vs. controls = −0.5), *JWF* (HJ = 0.5 vs. LJ = −0.5), and the interaction of *Word Type* and *JWF* as fixed effects. In addition to the fixed effects, by random intercepts for both subjects and items were also included as random effects. We compared an LME model, a Gamma Distribution GLMM, and an Inverse Gaussian Distribution GLMM with an identity link function. As a result, both these two GLMM models were found to be well fitted. We adopted the Inverse Gaussian GLMM model for the analysis of pilot experiment data. Data outliers with absolute standardized residuals exceeding 2.5 standard deviations were removed (2.61% of the data). As shown in Table 4, there was no significant difference between cognates (*M* = 457 ms, *SD* = 86) and controls (*M* = 459 ms, *SD* = 87). However, the RT of HJ (*M* = 451 ms, *SD* = 84) was significantly shorter than that of LJ (*M* = 465 ms, *SD* = 88). Additionally, the interaction was not significant, indicating that the facilitation effect of JWF on cognates was equal to the effects on the controls.

2.3. Apparatus and procedure

Participants were tested individually in a soundproof room after giving their informed consent. The stimuli were presented on a 17-inch monitor in black characters on a white background in Courier New font (size 28). At a viewing distance of approximately 70 cm from the screen, the visual angle for each character was estimated to be 1.8°. We used E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA) to present stimuli and to send event markers to the continuous EEG signal. The experimental instructions were presented in Japanese on the screen and orally by the experimenters.

The experimental procedure was set according to Peeters et al. (2013). As shown in Fig. 2, a fixation mark (+) was presented for 200 ms at the center of the screen, then after a blank screen was displayed for 200 ms, two kanji characters were presented at the same location as the fixation mark for 1300 ms. Participants were asked to decide whether the two-kanji-compound (target) was a Japanese word or not within 1300 ms as quickly and accurately as possible (visual lexical decision task; LDT). If the target was a Japanese word, participants had to press the “Yes” button on a response box using their right index finger; if it was not, participants had to press the “No” button using their left index finger. The presentation of stimuli remained on the screen for 1300 ms irrespective of whether participants made judgments or not. During the target presentation, participants were required to minimize eye blinks as

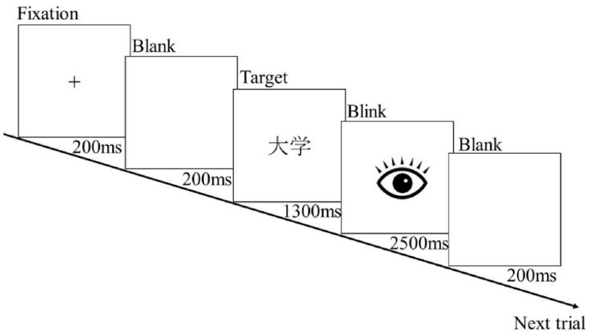

Fig. 2. Procedure for one trial.

well. The duration from the presentation of the target to the participants' decision was recorded as the RT. Next, a symbol of an eye was presented for 2500 ms for participants to relax their eyes. Finally, after a 200 ms blank interval, the next trial started.

All 480 stimuli were randomly distributed within eight blocks. The sequence of blocks and stimuli in each block was unique for each individual participant. No more than two words or nonwords were presented in sequence to prevent repetition effects. The experiment contained eight blocks with a small break between each block. After four blocks, participants could take a more extended break. On average, the experiment lasted approximately 60 min, excluding preparation for EEG recording. Before the real experiment started, we included 16 practice trials to familiarize participants with the task.

2.4. *EEG recording and pre-processing*

Electroencephalogram (EEG) data were recorded continuously using a BioSemi ActiveTwo system (BioSemi B.V., Amsterdam, Netherlands) with sampling at 2048 Hz. We collected the EEG from 32 Ag–AgCl electrodes which were mounted on the scalp according to standard positions within the international 10/20 system. Common Mode Sense (CMS) was used as an online reference (see https://www.biosemi.com/faq/cms&drl.htm). Eye blinks were monitored with two electrodes attached to the sub- and supra-orbital ridges of the left eye, respectively. Eye movements were measured by two electrodes attached to the left and right external canthi of each eye.

The EEG data were processed offline using EEGLAB (Delorme & Makeig, 2004) mostly based on the procedure laid out in Verdonschot, Tokimoto, and Miyaoka (2019). One participant was excluded from further analyses because of extensive noise artifacts. The preprocessing was conducted as follows. (1) The data were downsampled to 250 Hz. (2) A band-pass filter of 1–50 Hz was applied. (3) The channel locations were edited using an MNI coordinate file for the BEM dipfit model. (4) Line noise was removed using the CleanLine plug-in within EEGLAB. (5) The continuous data were cleaned using ASR (Artifact rejection). (6) The data were re-referenced to a common average reference. (7) The data were decomposed using AMICA (Palmer, Kreutz-Delgado, Rao, & Makeig, 2007). (8) The EEG data were further cleaned using SASICA (Chaumon, Bishop, & Busch, 2015). The dipole fit residual variance was set to 15%. In addition, the independent components related to eye blinks and eye movements were also excluded from further analysis. (9) Epochs of 1500 ms were selected from −0.5 to 1 s around the stimulus. ERPs were computed to all the words for the epochs that were followed by the correct responses (5,344 trials), with the baseline set as 100 ms before the onset of the stimuli.

3. Results

3.1. *Cognate facilitation effect*

The average accuracy of words that contained cognates and controls was 96.88% ($SD = 0.17$), while the accuracy of nonwords was 94.34% ($SD = 0.23$). Furthermore, the accuracies of both cognates ($M = 98.54\%$, $SD = 0.12$) and controls ($M = 95.21\%$, $SD = 0.21$) were higher than 95%, which indicated that our bilingual participants were very proficient in L2-Japanese. Next, we analyzed the RTs using correct responses only.

As mentioned above, we analyzed the RTs using GLMMs. To examine the cognate facilitation effect and the word frequency effect during the lexical processing of late Chinese - Japanese bilinguals, we analyzed RTs using *Word Type* (cognates = 0.5 vs. controls = −0.5), *JWF* (HJ = 0.5 vs. LJ = −0.5), and the interaction of *Word Type* and *JWF* as fixed effects in our GLMM model. As in the pilot experiment, we compared an LME model, a Gamma Distribution GLMM, and an Inverse Gaussian Distribution GLMM with an identity link function. It was shown that the Inverse Gaussian Distribution GLMM provided the best-fitted model. Consequently, we opted for this model in subsequent analyses. Data outliers with absolute standardized residuals exceeding 2.5 standard deviations were removed (2.38% of the data). The behavioral results of cognates and controls are summarized in Table 5. These results indicate that the main effects of *Word Type* ($p < .001$) and *JWF* ($p = .01$) were significant. The mean RT of cognates ($M = 516$ ms, $SD = 93$) was 39 ms faster than that of controls ($M = 555$ ms, $SD = 100$). Also, HJ ($M = 529$ ms, $SD = 94$) was processed 13 ms faster than LJ ($M = 542$ ms, $SD = 104$). Additionally, the interaction of *Word Type* with *JWF* was also significant ($p = .004$). The size of the cognate facilitation effect was larger in LJ ($\Delta = 52$ ms) than in HJ ($\Delta = 28$ ms).

3.2. *Japanese/Chinese word frequency effect in cognates*

To further clarify the effects of L1-Chinese knowledge as well as L2-Japanese knowledge, we constructed a GLMM model that contained both *JWF* (HJ = 0.5 vs. LJ = −0.5) and *CWF* (HC = 0.5 vs. LC = −0.5), as well as the interaction terms of these two

Table 5
The results of fixed effects in the model of cognates and controls.

	Estimate	SE	*t*	*p* (> |t|)
(Intercept)	556.65	9.70	57.37	< .001
Word Type	−42.93	5.86	−7.33	< .001
JWF	−15.06	5.91	−2.55	.010
Word Type × JWF	28.56	9.91	2.88	.004

Table 6
The results of fixed effects in the model of JWF and CWF.

	Estimate	SE	*t*	*p* (> |t|)
(Intercept)	527.10	12.12	43.50	< .001
JWF	−0.46	6.13	−0.07	0.94
CWF	−11.05	6.11	−1.81	0.07
JWF × CWF	2.73	11.76	0.23	0.82

factors as fixed effects. In addition to the fixed effects, random intercepts for both subjects and items were also included in the model as random effects. Data outliers (2.26% of the trials) exceeding 2.5 standard deviations were removed. As summarized in Table 6, neither the effect of JWF ($M_{\text{HJ}} = 515$ ms, $SD_{\text{HJ}} = 89$; $M_{\text{LJ}} = 518$ ms, $SD_{\text{LJ}} = 97$) nor the effect of CWF ($M_{\text{HC}} = 511$ ms, $SD_{\text{HC}} = 91$; $M_{\text{LC}} = 522$ ms, $SD_{\text{LC}} = 95$) reached significance.

3.3. Post hoc analysis

To confirm that our cognate effect was indeed due to knowledge of Chinese and not to uncontrolled stimulus characteristics, we further analyzed the RTs for bilinguals and the pilot group of native Japanese speakers. A GLMM model was constructed with *Participant Group* (native Japanese speakers = 0.5 vs. bilinguals = −0.5), *Word Type* (cognates = 0.5 vs. controls = −0.5), and the interaction of these two factors as independent variables, along with two random effects: by random intercepts for subjects and items. Data outliers with absolute standardized residuals exceeding 2.5 standard deviations were removed (2.74% of the data). As shown in Table 7, the interaction of *Participant Group* and *Word Type* was significant ($p < .001$). For the bilingual group, cognates ($M = 516$ ms, $SD = 93$) were processed 39 ms faster than controls ($M = 555$ ms, $SD = 100$). As previously noted, our native Japanese speakers processed cognates ($M = 457$ ms, $SD = 86$) as fast as controls ($M = 459$ ms, $SD = 87$). The results indicated that the cognate facilitation effect found in the present study was caused by the knowledge of the non-target language (here, Chinese) and *not* due to any uncontrolled variables.

3.4. Electrophysiological results

3.4.1. The effects of Word Type and JWF

To test the cognate facilitation effect and the JWF effect, the mean amplitudes of every 50 ms time-window from 50 ms to 700 ms latency were submitted to a series of three-way repeated ANOVAs based on previous studies, using Word Type (cognates vs. controls), JWF (HJ vs. LJ), three regions of interest (ROIs, frontal: AF3/4, F3/4, F7/8, Fz; central: FC1/2, CP1/2, C3/4, Cz; parietal: P3/4, P7/8, PO3/4, Pz), and the interactions of these three factors as independent variables. Greenhouse and Geisser (1959) corrections were applied to all analyses. The results are summarized in Table 8.

The three-way ANOVA showed that the main effect of Word Type was not significant in any of the time windows, while the main effects of JWF were significant in the 400–650 ms time-windows. There were also significant two-way interactions between Word Type and ROI in the 250–300 ms time-window and in the 400–550 ms time-windows. In the early time-window, the amplitudes elicited by controls were more negative than that of cognates in the frontal region (See Fig. 3 and Fig. 4). On the other hand, in the 400–550 ms time-windows, a more negative trend was induced by cognates at the frontal region than controls, while the amplitudes elicited by controls were more negative than cognates at the central-parietal region. Furthermore, regardless of Word Type, LJ elicited a more negative trend relative to HJ in the time-windows of 400–650 ms.

3.5. The effects of JWF and CWF on cognates

To clarify how the L2-Japanese representations and L1-Chinese representations affect the cognate visual recognition, we further analyzed the ERPs using a repeated three-way ANOVA including JWF, CWF, ROI and the interaction of these factors as fixed effects. Consequently, no significant main effects of CWF were found in any of these six time-windows, though a marginally significant main effect of CWF was found in 150–200 ms [$F(1, 22) = 3.90, p = .06$]. In addition, marginally significant interactions of CWF with ROI were found in the 150–200 ms [$F(1, 22) = 2.85, p = .07$] and 450–500 ms [$F(1, 22) = 3.54, p = .06$] time-windows. Regarding the effects of JWF, we found a significant main effect of JWF in 400–450 ms [$F(1, 22) = 6.86, p = .02$], 450–500 ms [$F(1, 22) = 14.08$,

Table 7
The results of fixed effects in the model for the bilingual group and native speakers.

	Estimate	SE	*t*	*p* (> |t|)
(Intercept)	523.30	6.43	81.39	< .001
Participant Group	−68.41	8.14	−8.40	< .001
Word Type	−24.88	5.63	−4.42	< .001
Participant Group × Word Type	36.68	2.54	14.45	< .001

Table 8
The *F*-values of ANOVAs on mean ERP amplitudes for every 50 ms time-windows from 50 to 700 ms latency.

	50+	100+	150+	200+	250+	300+	350+
Word Type							
JWF							
ROI	5.13*		50.20***	22.05***			
Word Type × JWF							
Word Type × ROI					4.52*		
Frontal					8.77**		
Central							
Parietal							
JWF × ROI							
Frontal							
Central							
Parietal							
Word Type × JWF × ROI							
Frontal							
Central							
Parietal							

	400+	*450+*	*500+*	*550+*	*600+*	*650+*
Word Type						
JWF	5.16*	14.82***	11.11**	14.28**	4.68*	
ROI		16.46***	20.11***	11.57**	12.64***	6.95**
Word Type × JWF						
Word Type × ROI	7.35**	14.15***	7.21**			
Frontal	8.20**	17.83***	9.53**			
Central		5.05*				
Parietal	7.76*	13.11**	6.04*			
JWF × ROI						
Frontal						
Central						
Parietal						
Word Type × JWF × ROI						
Frontal						
Central						
Parietal						

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

$p = .001$], and 550–600 ms [$F(1, 22) = 5.75$, $p = .03$] time-windows. Cognates with LJ elicited a more negative trend relative to cognates with HJ; this JWF effect in cognates processing was widely distributed.

4. Discussion

To investigate how cognates are processed in late Chinese - Japanese bilingual's mental lexicon, we conducted an L2-Japanese LDT with cognates and Japanese controls (i.e., non-cognates) following the experimental layout of Peeters et al. (2013). The results of both RTs and ERPs in the present study partially replicated the findings of Peeters et al. (2013), but there were also some differences likely due to the characteristics of the logographic writing system.

First, our findings substantiated the cognate facilitation effect reported in several earlier studies (e.g., Davis et al., 2010), indicating that during the processing of cognates by late Chinese - Japanese bilinguals, L1-Chinese lexical knowledge was non-selectively activated. Additionally, the cognate facilitation effect on LJ (Δ = 52 ms) was much stronger than that on HJ (Δ = 28 ms). It is more difficult for bilinguals to retrieve L2-Japanese lexical information for cognates that are used less frequently in Japanese, therefore bilinguals may have benefitted more from the automatically activated L1-Chinese lexical information, which might have caused a larger cognate effect in words having low Japanese word frequency. However, inconsistent with the results of Peeters et al. (2013), neither the L1-Chinese nor the L2-Japanese word frequency significantly affected the RTs of cognates.

Regarding the ERP results, a more negative waveform elicited by controls relative to cognates in the 250–300 ms time-window (which peaked at approximately 250 ms after stimuli onset) was found in the frontal region. This early component is assumed to be related to the processing at the form-level in the mental lexicon. Accordingly, an amplified N250 was induced by controls relative to cognates, indicating that cognate facilitation effect occurred at the sub-lexical representations during the L2-Japanese two-kanji compound word visual processing. Though we used an experimental design which closely matched that of Peeters et al. (2013), no such component was found in their study. Indeed, as far as we know, this is the first report of this component using a between-language task with no masked priming paradigm. We suppose this component was found in the present study but not previous studies mainly because of the experimental materials we used (i.e., Japanese words composed of two kanji characters). Each character of the Japanese two-kanji compound words could be considered as a morpheme which has its meaning (e.g., such as ‘bedroom’ in English).

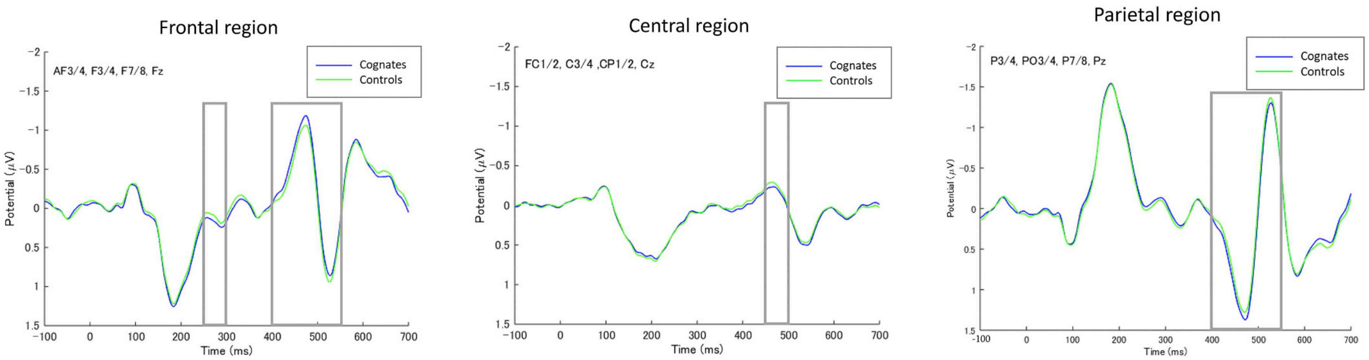

Fig. 3. Averaged ERP waveforms for the effect of Word Type. Negative voltage is plotted upwards.

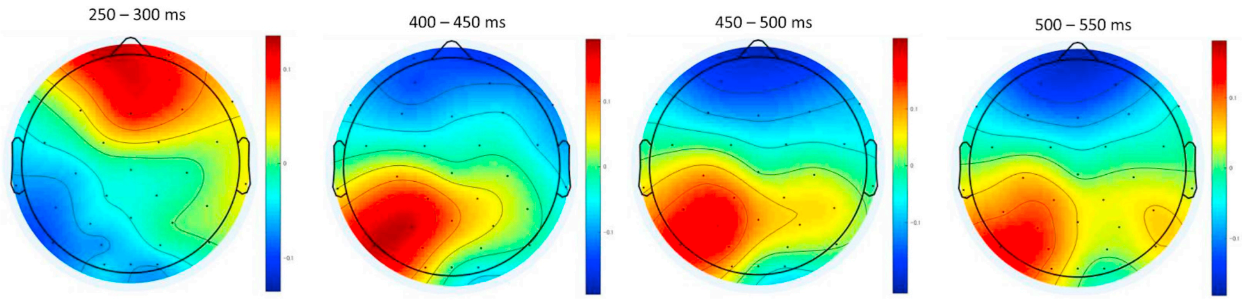

Fig. 4. Topographic maps of the time-windows that showed significant voltage differences between cognates and controls.

Many of the characters can be further decomposed to radicals associated with phonological or semantic information. Unlike cognates in languages with alphabetic scripts, (late) Chinese - Japanese bilinguals might have been able to get more information from sub-lexical representations before they access whole-word representations. Therefore, it is reasonable that Chinese - Japanese bilinguals might have paid more attention to the word forms. Note that we only found the N250 in the frontal region and not in the other two regions. As we mentioned earlier, unlike the languages using alphabetic scripts, many Chinese - Japanese cognates have dissimilar pronunciations (e.g., 大学, "university", /dai.gaku/ in Japanese but /dà.xué/ in Chinese; for this specific word, the phonological distance between Japanese and Chinese is 8, as calculated according to Yu et al., 2018), suggesting that the frontal N250 (hereafter, FN250) found in the present study was elicited by the orthographic information of cognates.

In addition to FN250, controls induced larger N400 than cognates in the central-parietal region: 450–500 ms for central region and 400–550 ms for the parietal region. This N400 effect was consistent with findings of both Midgley et al. (2011) and Peeters et al. (2013), suggesting that cognates are more easily processed than Japanese controls. In Peeters et al. (2013), a P600 at the 600–900 ms was also reported, but this component was not found by Midgley et al. (2011). Peeters et al. (2013) suggested that this late positive component might be associated with an effect caused by the lexical decision task. Nevertheless, though we also conducted a task which requires participants to give lexical decisions, there was no such P600 component in the present study. Instead, there was a peak at around 540 ms, which is possibly caused by the effort the participants made to decide whether the target was a Japanese word or not. According to Falkenstein, Hohnsbein, and Hoormann (1994), the P540 has been linked to the cognitive stage that maps a target word onto a response. Similarly, using a semantic categorization task, Shinoda, Nakagome, Mimura, and Homma (2006) found a negative component which peaked at around 540 ms (N540) localized in the right anterior cingulate gyrus and the left posterior superior temporal gyrus, suggesting that this component may be related to stimulus identification. It is therefore possible that the 540 ms peak is associated with the specific lexical decision task we conducted.

In addition to the FN250, N400, and the P/N540, we also found a more negative waveform induced by cognates compared to controls occurred in the frontal region in 400–550 ms, which has not been reported in any previous study. According to Peeters et al. (2013), a more negative N400 was elicited by cognates with low L1-French word frequency than those with high L1-word frequency in the frontal region during L2-English cognate processing, suggesting the possibility that the frontal N400 in Midgley et al. (2011) was due to the L1 frequency of their cognate stimuli. Contrary to these two previous studies, an amplified frontal N400 was elicited by cognates in comparison to controls in the present study. N400 with a frontal distribution has also been reported in other studies indexing the familiarity and recency of stimuli (e.g., Curran, 2000; Danker et al., 2008). Danker et al. (2008) found that FN400 distinguished old from new items for verbal stimuli and better remembered stimuli elicited more positive-trending FN400s. On the other hand, there is also evidence that FN400 is functionally similar to N400 (Kutas & Federmeier, 2011 for a review). Although whether FN400 is a distinctive component remains controversial, it seems that a more negative FN400 reflects more difficult processing regardless of the type of stimuli or task at hand. Given the evidence mentioned above, the amplified FN400s evoked by cognates in the present study might be construed as representing increased difficulties in cognate visual processing. This difficulty might originate from response competition between Japanese and Chinese since cognates exist in both languages. Conversely, the amplitudes of FN250 and N400 elicited by cognates were smaller than controls, indicating that cognates are more easily processed than controls at both the orthographic level and semantic level. It would be important to also investigate the function of FN400 specific to word type (i.e., cognates or not) in future work.

In addition to the effects of word type, words with low-JWF induced a larger N400 compared to words with high-JWF. In other words, we believe that words with high Japanese word frequencies are more easily processed than words with low frequencies, which replicates the frequency effect reported in previous research. However, inconsistent with Peeters et al. (2013), there was no significant but only marginally significant L1 word frequency effect on cognate processing in the present study in both behavioral and ERP results. There are three possible reasons: (1) each condition only included 30 trials, which may not have constituted enough statistical power to detect Chinese word frequency effects; (2) cognate facilitation effects mainly occurred at pre-lexical representations but not at the semantic representations; (3) cognates in the present study are very frequently used in both Japanese and Chinese, which lead to the absent significant Chinese word frequency effect. First, it is possible that the absent effect of Chinese word frequency is due to the small number of trials for each condition. As mentioned above, we manipulated the JWF and CWF within the cognates, resulting in only 30 trials for each condition which may have not constituted enough statistical power to detect the CWF effect in our data. Second, unlike languages with alphabetic scripts, each character in a Chinese - Japanese cognate contains several sources of information such as semantic information and/or phonological information. It is reasonable to assume that Chinese - Japanese bilinguals would have paid greater attention to the orthographic information, leading to a decreased FN250 for cognates than for controls and absent N400 for the L1-Chinese word frequency effect. Third, in the present study, all the stimuli were selected from a database (Park et al., 2014a; 2014b) composed of words marked as level 4 (the easiest level) to level 2 (medium level) in the Japanese-language proficiency test. This means that though a variety of word frequencies was employed, all of the words are frequently used in both Chinese and Japanese. Given the evidence found in the present study that a smaller cognate facilitation effect occurs in the processing of words with high Japanese frequency than those with low Japanese word frequency and the evidence reported in the previous studies (e.g., Duyck, Vanderelst, Desmet, & Hartsuiker, 2008), the missing frequency effect might have been due to frequent use of cognates in Japanese and Chinese.

5. Theoretical implications

In the mental lexicon of bilinguals using languages with alphabetical scripts (e.g., Dutch-English bilinguals), it is assumed that cognates share *one same orthographic representation* but have two phonological and two morphological representations (e.g., Peeters et al., 2013). This localist connectionist model suggests that when cognates are processed in a bilingual's mental lexicon, both L1 and L2 are automatically activated and one of these two representations would be more highly activated than the other depending on the task at hand. To determine whether cognates in languages with logographic scripts are processed in bilinguals' mental lexicon similarly to those with alphabetic scripts, we examined how cognates are processed in late Chinese - Japanese bilingual's brain using both RTs and ERPs during an L2-Japanese LDT.

ERPs showed that before access to the integrated semantic representation, cognates were quickly activated at the orthographic representations in L1-Chinese and L2-Japanese, resulting in a reduced FN250 of cognates compared to controls. Subsequently, the activation of the orthographic representation further activated the shared lexico-semantic representation, facilitating the visual recognition of cognates. For example, when the cognate “数学” (“mathematics”, which is pronounced as /shù.xué/ in Chinese but /suu.gaku/ in Japanese) was shown, Japanese and Chinese orthographic information was activated immediately at the character level. This was followed by the activation moving from the orthographic representation towards the semantic representation, where the meaning of the cognate “数学” (mathematics) is activated. Since cognates share the same orthographic word form and meaning between L1-Chinese and L2-Japanese, lexical information of both languages facilitated the recognition of the word. Note that in the lexical processing models for Chinese/Japanese monolinguals, the activation of character level plays a vital role during the visual processing of two-character compound words (e.g., Yan, Tian, Bai, & Rayner, 2006 for Chinese; Tamaoka & Hatsuzuka, 1995 for Japanese). In a recent study, Miwa, Dijkstra, Bolger, and Baayen (2014) investigated the processing order of Japanese two-character words using eye tracking. They reported that during the early stage of processing, character effects were larger in magnitude and more robust than radical and whole-word effects, suggesting a character-driven processing model in Japanese monolinguals. This further support our assumptions that the activation at the character level is crucial to cognates visual processing.

Future research is needed to further verify whether the absent Chinese word frequency effect during cognate processing was due to the early effect at the pre-lexical level or not. Similarly, additional studies need to be undertaken to elucidate the cause of the larger FN400 induced by cognates. Finally, as we mentioned above, in languages using logographic scripts, each character is a morpheme which is composed of radicals containing phonological/semantic information, and it remains unclear how L1-Chinese and L2-Japanese pre-lexical information interact with each other during the processing of cognates. In future endeavors it will be critical to explore what is occurring at the character and/or radical level when Chinese - Japanese bilinguals process two-kanji compound words. The present study indicates that cognates in languages with logographic scripts (Chinese and Japanese in the current study) are processed similarly to cognates with alphabetic scripts, which are assumed to be represented in one shared semantic representation, one orthographic representation and two morphological representations in visual word recognition. In addition to the similarities between alphabetic and logographic visual word processing, the cognate facilitation effect occurs at an early stage of processing at orthographic level and seems to be specific to logographic writing systems.

CRediT authorship contribution statement

Kexin Xiong: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration, Funding acquisition. **Rinus G. Verdonschot:** Conceptualization, Software, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition. **Katsuo Tamaoka:** Conceptualization, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration.

Declarations of competing interest

None.

Acknowledgements

This study was funded by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for JSPS Research Fellows granted to

KX (18F18302). RV is supported by a Grant in Aid for Scientific Research-C (17K02748). The authors would like to thank Shingo Tokimoto, Sachiko Kiyama and Shaoyun Yu, for their invaluable advice on the data analyses.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jneuroling.2020.100911.

References

Agency for Cultural Affairs (1978). *Tyuugokugo to taioo-suru kango [Kanji-compound words corresponding to Chinese]*. Tokyo, Japan: Agency for Cultural Affairs in Japan.

Buchta, C., & Hahsler, M. (2016). *Cba: Clustering for business analytics. R package version 0.2-17.* http://cran.r-project.org/package=cba.

Chaumon, M., Bishop, D. V., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods, 250*, 47–63.

Comesaña, M., Sánchez-Casas, R., Soares, A. P., Pinheiro, A. P., Rauber, A., Frade, S., et al. (2012). The interplay of phonology and orthography in visual cognate word recognition: An ERP study. *Neuroscience Letters, 529*, 75–79.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology Section A, 38*, 367–393.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28*(6), 923–938.

Danker, J. F., Hwang, G. M., Gauthier, L., Geller, A., Kahana, M. J., & Sekuler, R. (2008). Characterizing the ERP Old–New effect in a short-term memory task. *Psychophysiology, 45*(5), 784–793.

Davis, C., Sánchez-Casas, R., Garcia-Albea, J. E., Guasch, M., Molero, M., & Ferré, P. (2010). Masked translation priming: Varying language experience and word type with Spanish - English bilinguals. *Bilingualism: Language and Cognition, 13*, 137–155.

De Groot, A. M., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*, 90–123.

Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21.

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*, 496–518.

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*, 284–301.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 663–679.

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review, 15*(4), 850–855.

Falkenstein, M., Hohnsbein, J., & Hoormann, J. (1994). Effects of choice complexity on different subcomponents of the late positive complex of the event-related potential. *Electroencephalography and Clinical Neurophysiology: Evoked Potentials Section, 92*(2), 148–160.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*(2), 95–112.

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience, 18*, 1631–1643.

Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain Research, 1180*, 39–58.

Kong, L., Zhang, B., Zhang, J. X., & Kang, C. (2012). P200 can be modulated by orthography alone in reading Chinese words. *Neuroscience Letters, 529*(2), 161–165.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11*, 99–116.

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

Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language & Cognitive Processes, 19*, 585–611.

Liu, Y., Perfetti, C. A., & Hart, L. (2003). ERP evidence for the time course of graphic, phonological, and semantic information in Chinese meaning and pronunciation decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1231–1247.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*, 1171.

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*, 1634–1647.

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*, 445–463.

Murphy, K. A., Jogia, J., & Talcott, J. B. (2019). On the neural basis of word reading: A meta-analysis of fMRI evidence using activation likelihood estimation. *Journal of Neurolinguistics, 49*, 71–83.

Nakayama, M. S. (2002). The cognate status effect in lexical processing by Chinese - Japanese bilinguals. *Psychologia, 45*, 184–192.

Okano, K., Grainger, J., & Holcomb, P. J. (2013). An ERP investigation of visual word recognition in syllabary scripts. *Cognitive, Affective, & Behavioral Neuroscience, 13*(2), 390–404.

Palmer, J. A., Kreutz-Delgado, K., Rao, B. D., & Makeig, S. (2007). Modeling and estimation of dependent subspaces with non-radially symmetric and skewed densities. In M. E. Davies, C. C. James, S. A. Abdallah, & M. D. Plumbley (Vol. Eds.), *International conference on independent component analysis and signal separation, lecture notes in computer science book series (LNCS: Vol. 4666*, (pp. 97–104). Berlin; Heidelberg, Germany: Springer.

Park, S., Xiong, K., & Tamaoka, K. (2014a). Dookei nizi kanzi go no hinsisei ni kansuru nikkantyuu deeta beesu no gaiyoo [A description of the database of grammatical categories for orthographically-similar two-kanji compound words among the Japanese, Korean and Chinese languages]. *Kotoba no Kagaku [Studia Linguistica], 27*, 3–23.

Park, S., Xiong, K., & Tamaoka, K. (2014b). Dookei nizi kanzi go no hinsisei ni kansuru nikkantyuu deeta beesu [A database of grammatical categories for orthographically-similar two-kanji compound words among Japanese, Korean and Chinese]. *Kotoba no Kagaku [Studia Linguistica], 27*, 53–111.

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*, 315–332.

Shinoda, J., Nakagome, K., Mimura, M., & Homma, I. (2006). Source localization of event-related potentials related to cross-modal semantic interference effect using scalp-skull-brain dipole tracing method. In I. Homma, & S. Shioda (Eds.). *Breathing, feeding, and neuroprotection* (pp. 139–146). Tokyo, Japan: Springer-Verlag Tokyo.

Tamaoka, K., & Hatsuzuka, M. (1995). Kanzi nizi zyukugo no syori niokeru kanzi siyoohindo no eikyoo [The effects of kanji printed-frequency on processing Japanese two-morpheme compound words]. *Dokusyo Kagaku [The Science of Reading], 39*, 121–137.

Tamaoka, K., Makioka, S., Sanders, S., & Verdonschot, R. G. (2017). A new interactive online database for psychological and linguistic research on Japanese kanji and their compound words. *Psychological Research, 81*, 696–708. Www.kanjidatabase.com.

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*, 2023–2038.

Verdonschot, R. G., Tokimoto, S., & Miyaoka, Y. (2019). The fundamental phonological unit of Japanese word production: An EEG study using the picture-word interference paradigm. *Journal of Neurolinguistics, 51*, 184–193.

Xiong, K. (2018). *Cognitive processing of orthographically and semantically similar two-kanji compound words by native Chinese speakers learning Japanese*Unpublished doctoral dissertation. Japan: Nagoya University.

Xun, E., Rao, G., Xiao, X., & Zang, J. (2016). The construction of the BCC Corpus in the age of big data. *Corpus Linguistics, 3*, 93–118. http://bcc.blcu.edu.cn/.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97*(2), 259–268.

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 similarities 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.

Yu, S., & Tamaoka, K. (2015). Nikkantyuu dookei nizi kanzigo no hinsisei web-kensaku enzin [A web-accessible search engine for grammatical category of orthographically-similar two-kanji compound words between Japanese, Korean and Chinese]. *Kotoba no Kagaku [Studia Linguistica], 29*, 43–61. http://kanjigodb.herokuapp.com/.

Zhang, S. (1987). *Tyuu niti kango taihi ziten [A Chinese - Japanese comparison dictionary of Sino-Japanese Words].* Tokyo, Japan: Yumani Shobo.

Zhou, L., Fong, M. C. M., Minett, J. W., Peng, G., & Wang, W. S. (2014). Pre-lexical phonological processing in reading Chinese characters: An ERP study. *Journal of Neurolinguistics, 30*, 14–26.