HARMONIZING SOUNDS: THE NAVIGATION OF PHONOLOGICAL PROCESSING IN PINYIN AND HANZI BY EARLY JAPANESE CFL (CHINESE AS A FOREIGN LANGUAGE) LEARNERS

Jingyi ZHANG[1] and Katsuo TAMAOKA[2), 3)]

[1] *University of Miyazaki, Japan*
[2] *Shanghai University, China*
[3] *Nagoya University, Japan*

This study investigated the unique challenges encountered by Japanese speakers learning Chinese as a foreign language (CFL). Through a paper-based test, it was revealed that Japanese CFL learners, leveraging their first language (L1) knowledge in Chinese characters (*kanji*), demonstrated advanced understanding of word meanings in *hanzi* compared to *pinyin*. Regardless of their Chinese ability, words in pinyin consistently exhibited higher naming accuracy than their hanzi counterparts. However, naming latencies for pinyin words consistently appeared longer, which can be attributed to the application of regular pinyin-to-sound conversion rules. Conversely, with increasing Chinese ability, accuracy for hanzi words improved, albeit accompanied by extended naming latencies, indicating a 'speed-and-accuracy tradeoff.' As Japanese CFL learners advanced in their Chinese ability, they would learn to suppress Japanese pronunciations for words in hanzi and use Chinese equivalents instead. With increased ability, they became more adjusted to potential mispronunciations due to similar sounds, leading to slower but more accurate naming. This study highlighted the complexity of phonological processing among Japanese CFL learners, shaped by various interacting factors.

Key words: Chinese as a foreign language (CFL), Japanese CFL learners, *hanzi*, *pinyin*, *kanji*, Chinese ability, cognate, phonological similarity, speed-and-accuracy tradeoff

Introduction

When acquiring a foreign language, the linguistic resemblance between the learner's first language (L1) and the target language plays a significant role. *Pinyin*, a phonetic system, is employed to facilitate the pronunciation of Mandarin Chinese. Native English speakers learning Chinese as a foreign language (CFL), particularly in countries like America, Australia or Canada, adopt the alphabetic script of pinyin to grasp the pronunciation and meanings of Chinese words. Given their prior knowledge of the

This research was supported by the Grant-in-Aid for Japan Society for the Promotion of Science (#21K13055, PI: J.Z., University of Miyazaki, Japan and #23K00628, PI: K.T., Nagoya University, Japan). We would like to express our gratitude to the reviewers and all those who contributed to this research.

Correspondence concerning this article should be addressed to Katsuo Tamaoka, Professor, School of Foreign Languages, Shanghai University, 99 Shangda Road, Baoshan District, Shanghai 200444, China (email: ktamaoka@gc4.so-net.ne.jp).

English alphabet (Everson, 1998; McGinnis, 1997), learning Chinese words through pinyin appears more accessible than attempting a direct acquisition from Chinese characters of *hanzi*. Consequently, a common and effective approach involves initially learning Chinese words in pinyin and subsequently progressing to hanzi. This method has demonstrated success among English CFL learners in hanzi acquisition (e.g., Dai & Lu, 1985; Huang & Hanley, 1997; Jordan, 1971). However, it is essential to recognize that despite the prevalence of pinyin learning, Chinese text is fundamentally composed in hanzi.

Contrastingly, the Japanese language incorporates Chinese characters as one of its scripts, and many Chinese hanzi share similarities with Japanese *kanji*. Consequently, native Japanese speakers, without formal Chinese study, can roughly comprehend text when encountering Chinese hanzi in daily life, such as on street signs. However, since the Japanese language seldom utilizes the alphabet as a script, the use of pinyin is unfamiliar to native Japanese speakers. Consequently, when native Japanese speakers learn Chinese, their reliance on pinyin, particularly to access the meaning of Chinese words, may be limited. Given that Japanese CFL learners already possess knowledge of the majority of Chinese hanzi from their L1 Japanese script, connecting Chinese hanzi to meanings, and possibly to pronunciations, might be more straightforward. Therefore, this study investigated the phonological processing mechanism used by CFL learners, specifically native Japanese speakers with a unique kanji background.

Similarities Between Chinese Hanzi and Japanese Kanji

In modern Japanese, three distinct scripts are utilized: *hiragana* and *katakana*, two phonetic scripts, and kanji, a morphemic script. Kanji is commonly employed for writing lexical morphemes, such as nouns and the roots of verbs and adjectives, while hiragana is used for grammatical morphemes and katakana is reserved for alphabetic loanwords (Miller, 1967; Tamaoka, 2014, 2015; Verdonschot et al., 2013). The Japanese kanji script finds its origins in Chinese hanzi. According to Chen (2002), approximately 54.5% of Japanese kanji compounds in mainland China and 55.1% in Taiwan share the same character shapes and meanings. Hishinuma (1983, 1984) further noted that, even accounting for minor orthographic differences, about 98% of the 1,945 items in the previous commonly-used kanji list (*Kyû Zyôyô Kanzi Hyô*) established in 1981 by the Japanese government are common with Chinese hanzi.

Due to the similarity between kanji and hanzi, native Chinese speakers learning Japanese as a foreign language (JFL) exhibit quicker and more accurate lexical decision-making with kanji-presented words compared to JFL learners without Chinese characters in their L1 scripts (Matsumoto, 2013; Tamaoka, 1997, 2000). When contrasted with Korean JFL learners, Chinese JFL learners comprehend words in kanji more rapidly than those in hiragana and katakana (Yamato & Tamaoka, 2013). Nakayama (2002) also noted that Chinese JFL learners made faster lexical decisions for orthographically identical Chinese-Japanese cognates than for non-cognates. The advantage of script similarity

(Djojomihardjo et al., 1994; Tamaoka, 2000, 2014, 2015, 2022) for Chinese JFL learners, specifically those with a background in Japanese kanji, is noteworthy when compared to learners from non-kanji backgrounds. Although the situation for Chinese JFL learners (L1 Chinese to Japanese) represents a processing direction opposite to that of Japanese CFL learners (L1 Japanese to Chinese), it is reasonable to anticipate that the shared use of the hanzi/kanji script in both Chinese and Japanese languages is advantageous for Japanese CFL learners in understanding Chinese words written in hanzi.

Can the knowledge of L1 kanji among Japanese CFL learners effectively contribute to the phonological processing of Chinese words in hanzi? Japanese kanji typically have multiple pronunciations, categorized into *on*-readings and *kun*-readings (Tamaoka & Makioka, 2004; Tamaoka & Taft, 2010; Verdonschot et al., 2013). On-readings are primarily derived from Chinese pronunciations, while kun-readings usually stem from traditional Japanese pronunciations. For instance, the on-reading of 人 'person' is /nin/ or /zin/, and the kun-reading is /hito/. Native Japanese speakers possess a good ability to distinguish between on-readings and kun-readings (Tamaoka & Taft, 2010). Japanese CFL learners may roughly predict Chinese hanzi pronunciations from Japanese kanji on-readings. For example, the three-hanzi word 图书馆 'library' is written with three-kanji 図書館 in Japanese. Although hanzi has been simplified, the word 'library' is essentially written the same way in hanzi and kanji. This three-hanzi word is pronounced /tú shū guǎn/ in Chinese and is pronounced similarly, with on-reading as /to syo kan/ (described in the Kunrei-style romanization) in Japanese. However, this subtle difference in pronunciation between the two languages could lead to mispronunciations.

There are conflicting arguments regarding the effectiveness of the relationship between the pronunciations of Chinese hanzi and Japanese kanji in the phonological processing of both languages. Previous studies on Chinese JFL learners (Matsumi et al., 2012, 2016; Tome et al., 2012) reported positive results, demonstrating that both orthographic and phonological similarity influenced the processing of Japanese kanji cognates. Matsumi et al. (2012, 2016) highlighted facilitatory effects on both orthographic and phonological similarity in the processing of Japanese kanji words by Chinese JFL learners.

On the contrary, negative results have been observed. Fei (2013, 2015) and Fei et al. (2022) investigated the impact of Chinese-Japanese orthographic and phonological similarities on lexical processing. They suggested a facilitatory effect of orthographical similarity when processing Chinese words for Chinese JFL learners but identified an inhibitory effect of phonological similarity. Other studies (Chiu, 2003; Hong, 2004; Ishida, 1986) also reported that phonological similarities between Chinese and Japanese could inhibit the pronunciation of Chinese words for Chinese JFL learners. The coexistence of two different pronunciations for the same word in both languages may interfere with the phonological processing of Japanese words in kanji. Although these studies (Chiu, 2003; Fei, 2013, 2015; Fei et al., 2022; Hong, 2004; Ishida, 1986) focused on Japanese words by Chinese JFL learners, their findings may also provide insights for Japanese CFL learners. Japanese CFL learners may experience adverse effects on pronunciation when attempting to pronounce equivalent Chinese words due to the influence of L1 Japanese words.

The question of whether or not the effects of orthographically similar Chinese-Japanese morphemes work for or against phonological processing remains unresolved. It appears that both orthographical and phonological similarities between hanzi and kanji could influence the lexical processing of Chinese words for Japanese CFL learners. Thus, the present study examined the influence of Japanese CFL learners' L1 kanji knowledge, specifically focusing on the lexical processing of Chinese words presented in pinyin and hanzi.

Use of Pinyin for L1 Chinese Speakers and CFL Learners

A total of 407 syllables consisting of 21 consonants and 36 vowels are used in Chinese pinyin (J. Zhang et al., 2021). Pinyin is an alphabetic phonetic coding system that transcribes hanzi into phonemes with tones to enable hanzi pronunciation. According to Anderson et al. (2013), Chinese children need to learn 3,500 Chinese hanzi corresponding to approximately 1,200 tonal syllables. The majority of pinyin-to-sound mappings exhibit a basic one-to-one correspondence. Therefore, the relationship between spelling and pronunciation in pinyin is much more regular than with the English alphabet. Pinyin is utilized as a phonetic tool to represent the pronunciation of Chinese hanzi. Native Chinese children are usually taught pinyin before they start learning hanzi. School textbooks in China are typically written in hanzi accompanied by pinyin (Ding, 1988; Zhou, 1959).

Previous research (e.g., Ho & Bryant, 1997; Ma et al., 2020; Shu & Liu, 1994; Shu et al., 1993; Wang et al., 2014) has consistently shown that reading comprehension test scores for L1 Chinese students instructed using pinyin were significantly higher than those relying solely on hanzi. Moreover, Zhang et al. (2020) identified a robust positive correlation between pinyin proficiency and hanzi recognition among 159 kindergarten children in China. Similarly, Lee and Kalyuga (2011) conducted a study involving 240 L1 Chinese students in Hong Kong who were also proficient in English. Their research indicated that when learning Chinese, students taught with both pinyin and hanzi outperformed those taught with either only pinyin or only hanzi. Consequently, pinyin appears to play a substantial role in the acquisition of Chinese language among L1 Chinese students.

Research on the use of pinyin has been conducted on CFL learners whose native languages utilize the alphabetic script. Previous studies (Chung, 2003; Xiao et al., 2020; H. Zhang et al., 2021) have suggested that effective utilization of pinyin-to-hanzi significantly contributes to Chinese proficiency among native Indonesian and English CFL learners. Everson (1998) highlighted that knowledge of Chinese lexical pronunciations exhibited a very strong correlation ($r = .96, p < .001$) with participants' ability to explain the meaning of Chinese words. This was observed among adult CFL learners, specifically among university students studying Chinese over two consecutive years, whose languages are written in alphabet (excluding native Japanese and Korean speakers, and those of Chinese background). These students seem to undergo education with a primary focus on Chinese listening and speaking skills, predominantly using pinyin.

Regarding the relationship between pinyin and hanzi, Chung (2003) examined the role of pinyin in hanzi learning with eight English L1 secondary school students in Sydney, Australia. The study suggested that Chinese words presented verbally with pinyin were better remembered than those only presented verbally. In a larger-scale study, Xiao et al. (2020) investigated the use of pinyin for Chinese learning among Indonesian CFL learners. Since the alphabet is also used to phonetically describe the Indonesian language, like pinyin for the Chinese language, their path analysis indicated that pinyin dictation (writing down orally presented sentences in pinyin) and pinyin tagging (writing out corresponding pinyin syllables according to hanzi) influenced reading comprehension through the depth of vocabulary knowledge. Their findings suggest that pinyin plays a crucial role not only in hanzi lexical understanding but also in Chinese reading comprehension.

In contrast to previous studies on CFL learners who use alphabetic scripts in their L1 language, the Japanese language uniquely incorporates Chinese hanzi as its script, specifically in the form of kanji. It is anticipated that Japanese CFL learners' kanji knowledge would positively influence their understanding of the meaning of Chinese words in hanzi. However, the impact of pinyin on the conceptual understanding and phonological processing of Chinese words for Japanese CFL learners remains unclear.

To investigate the phonological processing mechanism of Chinese words among Japanese CFL learners, the present study employed two methodologies: (1) a paper test, and (2) a naming experiment. Initially, a paper test was conducted to assess the ease of accessing meanings for cognate (i.e., words sharing the same orthographic hanzi/kanji characters with the same or similar meanings) and non-cognate words presented in both pinyin and hanzi. The objective was to determine whether pinyin or hanzi was more effective in accessing lexical concepts among Japanese CFL learners. We anticipated that the paper test would reveal a hanzi advantage over pinyin for accessing lexical meanings among Japanese CFL learners.

Subsequently, a naming experiment was conducted, encompassing cognate and non-cognate Chinese compound words presented in both pinyin and hanzi. It was hypothesized that naming accuracies and latencies for Chinese words written in hanzi would be strongly correlated with overall Chinese proficiency. This assumption stems from the notion that Japanese CFL learners access the conceptual meaning of Chinese words before delving into lexical pronunciations in hanzi, but not in pinyin. Conversely, pinyin may assist Japanese CFL learners in accurately pronouncing Chinese words using the regular pinyin-to-sound conversion rules.

Background Factors Influencing the Pronunciation of Chinese Words

Japanese CFL learners possess a distinctive advantage in knowing Chinese characters, due to the historical integration of these characters into the Japanese language. This historical connection has led to a substantial overlap of lexical items shared between the two languages, along with familiarity with the hanzi script. Additionally, the borrowed Chinese words have adopted phonetic similarities with Japanese, resulting in a degree of resemblance between the two languages. Given this background, four factors can be considered in the phonological processing of Chinese words by Japanese CFL learners.

The first factor is the cognate relationship between Japanese and Chinese languages. Due to substantial similarity in written forms and meanings of characters and words between the two languages, Japanese CFL learners are inclined to access the meaning of a Chinese word from its hanzi before deciding its pronunciation. Although shared meanings facilitate accurate and rapid access with the pronunciation of Chinese words, the inherent difference in pronunciation between Japanese kanji and Chinese hanzi may result in errors, particularly among learners at lower levels of Chinese ability who have not yet mastered accurate Chinese pronunciation.

The second factor involves phonological similarities in words between Japanese and Chinese. The adaptation of Chinese hanzi into Japanese was also accompanied by their hanzi pronunciations. The pronunciations adapted during various dynasties in China have led to multiple pronunciations for a single kanji. While there is a certain level of similarity in the pronunciations of Chinese characters between the two languages, the extent to which this similarity influences the pronunciation of Chinese words by Japanese CFL learners remains to be explored. Related to the first factor regarding the cognate relationship, Japanese CFL learners with lower Chinese ability may directly apply Japanese kanji pronunciations to Chinese hanzi.

The third factor under consideration is Chinese language ability. Given that native Japanese speakers are already highly familiar with Chinese hanzi, their Chinese ability may not significantly impact their ability to understand the meanings conveyed by hanzi. However, it is anticipated that Chinese ability may influence the pronunciation of Chinese words among Japanese CFL learners. The interaction between Chinese ability and the aforementioned factors of cognate status and phonological similarity remains unclear. It is possible that lower proficiency in Chinese could result in inaccuracies in recalling Japanese pronunciations for Chinese-Japanese cognate words, while higher proficiency in Chinese may facilitate more accurate pronunciation, particularly for cognates with easily recognizable meanings via orthographic similarities of kanji and hanzi.

The fourth factor involves the distinction in script types, namely pinyin and hanzi. Given the significant overlap of characters between Japanese and Chinese, hanzi serves as a familiar writing system for Japanese CFL learners. The factors of cognate relationships and phonological similarities via hanzi are expected to contribute to the accurate and prompt pronunciation of Chinese words. Although Japanese CFL learners have some familiarity with the alphabetical script of pinyin due to their exposure to English, the differences in pronunciation between pinyin and English spelling demand careful consideration. Nonetheless, the regular relationship between pinyin and pronunciation enables Japanese CFL learners to achieve accuracy, albeit potentially at a slower pace. The script difference between pinyin and hanzi is anticipated to exert a strong influence on the pronunciation of Chinese words by Japanese CFL learners.

The present study investigated how these four factors of cognate relationships, phonological similarities, Chinese language ability, and script type affect the phonological processing of Chinese words by Japanese CFL learners.

Paper Test

A paper test was conducted with Japanese CFL learners to assess their understanding of the meaning of words presented in pinyin and hanzi.

Participants

Participants in the paper test were undergraduate Japanese CFL learners from a university in Japan (N = 40, 28 females, 12 males, mean age = 19 years 9 months, SD = 11 months), studying various disciplines including education, engineering, agriculture, and regional innovation (management). Upon entering the university, they had chosen to study CFL for 11 months. Participants dedicated an average of 139 minutes per week (SD = 52 mins) to studying Chinese, determined by a simple questionnaire before the experiment. None of them had prior experience learning Chinese or had visited China. Their Chinese abilities were at a beginner's level. The present study involving human participants was reviewed and approved by the Research Ethics Committee of Nagoya University, Japan. Participants signed informed consent forms before the paper test, and at the end of the experiment they received a token payment.

Stimulus Words

The paper test consisted of 28 stimuli words in total, 14 words in hanzi and 14 words in pinyin. All stimulus words were selected from the textbook (J. Zhang et al., 2021) which participants had been using during their 11 months of studying Chinese. The 28 words in pinyin and in hanzi were selected from the textbook with the same level of difficulty. The words included 13 nouns, 14 verbs and 1 pronoun and were controlled to consist of 14 cognates and 14 non-cognates. The complete list of stimulus words is shown in Appendix 1. Lexical characteristics for stimuli in the paper test are shown in Table 1. The average stroke numbers for the 28 words were 16.00 strokes (SD = 4.58 strokes). As shown in Table 1, these 28 words were divided into the different script types of pinyin (n = 14) and hanzi (n = 14), each of which was further classified into cognates and non-cognates. Lesson numbers in the textbook (J. Zhang et al., 2021) were used as the index of familiarity for participants. The average of lesson numbers (familiarity) for 28 words was 8.71 (SD = 1.98). Stroke numbers as an index of visual complexity, $t(26) = 1.163$, $p = .255$, Cohen's $d = .44$, and lessons as the familiarity index, $t(26) = –.758$, $p = .455$, Cohen's $d = –.29$, showed no difference between words presented in pinyin and hanzi.

The present study computed phonological similarities for 2-hanzi compound words using the formula developed by Zhang (2018). Three types of phonological similarities were calculated: (1) first hanzi, (2) second hanzi, and (3) an average of both hanzi.

Table 1. Characteristics of Words in *Pinyin* and *Hanzi* for the Paper Test

Stimulus type (*N* = 28)		Stroke numbers		Familiarity (Lesson no.)		Parts of speech		
		M	*SD*	*M*	*SD*	Noun	Verb	Pronoun
Pinyin (*N* = 14)	Cognate (*n* = 7)	15.43	4.16	9.00	2.16	4	3	0
	Non-cognate (*n* = 7)	14.57	5.62	9.00	2.16	5	2	0
Hanzi (*N* = 14)	Cognate (*n* = 7)	16.00	2.71	7.86	1.57	3	4	0
	Non-cognate (*n* = 7)	18.00	5.54	9.00	2.16	1	5	1
	Total	16.00	4.58	8.71	1.98	13	14	1

Table 2. Phonological Similarities in Stimulus Words for the Paper Test

Lexical similarity		First *hanzi*		Second *hanzi*		Whole words	
		M	*SD*	*M*	*SD*	*M*	*SD*
Pinyin (*N* = 14)	Cognate (*n* = 7)	.56	.30	.34	.17	.45	.17
	Non-cognate (*n* = 7)	.37	.36	.47	.18	.42	.16
Hanzi (*N* = 14)	Cognate (*n* = 7)	.40	.12	.40	.32	.40	.15
	Non-cognate (*n* = 7)	.71	.27	.54	.38	.62	.27
	Total	.51	.29	.44	.27	.47	.20

Chinese hanzi were transcribed in pinyin, while Japanese kanji were transcribed in the Kunrei-style Roman script (*Roma-ji*). The ratio of similar phonemes was determined based on pinyin in Chinese and Roma-ji in Japanese.

For instance, consider the Chinese 2-hanzi word 现在, meaning 'now,' pronounced as /xiàn zài/. Its Japanese cognate word is 現在, pronounced as /gen zai/. Only 1 phoneme /n/ between the 4 phonemes of /xiàn/ and the 3 phonemes of /gen/ was shared. Therefore, the first hanzi-and-kanji similarity was calculated as (1 + 1)/(4 + 3) = 2/7 = .29. Similarly, the second hanzi is pronounced with 3 phonemes of /zài/, which is the same as its cognate kanji in Japanese. Thus, the second hanzi-and-kanji similarity was calculated as (3 + 3)/(3 + 3) = 6/6 = 1.00. The overall lexical phonological similarity of 现在 was calculated by averaging both the first and second hanzi-and-kanji similarities, resulting in (.29 + 1.00)/2 = .64.

As indicated in Table 2, the average phonological similarity for 28 two-hanzi words was .47 (*SD* = .20). No significant differences were observed in phonological similarities among the first hanzi, $t(26) = .760$, $p = .454$, Cohen's $d = .29$, the second hanzi, $t(26) = .581$, $p = .566$, Cohen's $d = .22$, and 2-hanzi words, $t(26) = .915$, $p = .368$, Cohen's $d = .35$, between words presented in pinyin and hanzi.

Table 3. Accuracies of Meanings of Words Presented in *Pinyin* and *Hanzi*

Type of words	Number of words	*M*	*SD*
		95% CI	
Words presented in *pinyin*	14	11.60 [10.83, 12.30]	2.43 [1.85, 2.86]
Cognate	7	5.80 [5.38, 6.20]	1.34 [1.05, 1.56]
Non-cognate	7	5.80 [5.40, 6.18]	1.29 [0.96, 1.54]
Words presented in *hanzi*	14	13.13 [12.80, 13.40]	0.94 [0.72, 1.11]
Cognate	7	6.95 [6.88, 7.00]	0.22 [0.00, 0.34]
Non-cognate	7	6.18 [5.90, 6.43]	0.87 [0.67, 1.03]

Note. 95% CI = 95% confidence interval according to 1,000 bootstrapping iterations.

Hence, the stimulus words presented in both pinyin and hanzi for the paper test were controlled for various characteristics.

Procedure

A paper test was conducted with 40 Japanese CFL learners in a quiet classroom setting. Participants were tasked to provide the Japanese meanings for 14 Chinese words in hanzi and 14 words in pinyin. The test had a time limit of 20 minutes for completion.

Data Analysis of Test Scores

A 2 (script type: pinyin and hanzi) × 2 (lexical similarity: cognate and non-cognate) analysis of variance (ANOVA) with repeated measures was performed on the test scores of 40 Japanese CFL learners. Means and standard deviations, along with a 95% confidence interval (CI), were computed through 1,000 bootstrapping repetitions, as detailed in Table 3.

Results indicated a significant main effect of script type (pinyin and hanzi), $F(39, 1) = 22.97$, $p < .001$, $\eta_p^2 = .37$. The comprehension of words presented in hanzi ($M = 13.13$, $SD = .94$, 95% CI [12.80, 13.40]) was significantly superior to that of words presented in pinyin ($M = 11.60$, $SD = 2.43$, 95% CI [10.83, 12.30]). Moreover, the interaction between script type and lexical similarity was significant, $F(39, 1) = 12.50$, $p < .001$, $\eta_p^2 = .24$. A simple contrast revealed no difference between cognates ($M = 5.80$, $SD = 1.34$, 95% CI [5.38, 6.20]) and non-cognates ($M = 5.80$, $SD = 1.29$, 95% CI [5.40, 6.18]) for words in pinyin. In contrast, for words in hanzi, the meanings of cognates ($M = 6.95$, $SD = .22$, 95% CI [6.88, 7.00]) were significantly better understood than those of non-cognates ($M = 6.18$, $SD = .87$, 95% CI [5.90, 6.43]), $F(39, 1) = 12.50$, $p < .001$, $\eta_p^2 = .24$.

Cognates in hanzi demonstrated high accuracy, scoring 6.95 points, or 99.29%, while non-cognates in hanzi also achieved a notable score of 6.18 points, equivalent to 88.29% accuracy, out of a maximum of 7 points. Additionally, the main effect of Chinese-and-Japanese lexical similarity (cognate and non-cognate) was found to be significant, $F(39, 1) = 14.88$, $p < .001$, $\eta_p^2 = .28$. Chinese-and-Japanese cognates ($M = 12.75$, $SD = 1.43$, 95% CI [12.30, 13.15]) were comprehended more accurately compared to the meanings of Chinese-and-Japanese non-cognates ($M = 11.98$, $SD = 1.87$, 95% CI [11.40, 12.55]).

Discussion

The inherent closeness between Japanese kanji and Chinese hanzi significantly enhanced the semantic comprehension of Chinese words for Japanese CFL learners. The shared writing system between these two scripts imparted a distinct advantage to Japanese CFL learners, aiding in the recognition and memorization of Chinese words presented in hanzi. This finding aligned with research on Chinese JFL learners, which consistently shows the facilitative effect of orthographic similarities between hanzi and kanji on accessing the meanings of Chinese words (Fei, 2013, 2015; Matsumi et al., 2012, 2016; Nakayama, 2002; Tome et al., 2012).

Naming Experiment

A naming experiment was conducted on Japanese CFL learners, tasking them with pronouncing words in pinyin and hanzi as quickly and accurately as possible.

Participants

Forty-four undergraduate Japanese CFL learners (25 females, 19 males, mean age = 19 years 0 months, $SD = 13$ months) from a university in Japan participated in this experiment. Participants pursued various disciplines, including education, engineering, agriculture, medicine, nursing, and regional innovation (management). All participants belonged to the same university as those who took the paper test, ensuring similar student characteristics. These participants had opted for CFL studies for 6 months after entering university, dedicating an average of 116 minutes ($SD = 31$ mins) weekly to Chinese language study. None of the participants had previous experience learning Chinese before university, indicating a beginner's level of proficiency. Like the paper test, the naming experiment involving human participants was reviewed and approved by the Research Ethics Committee of Nagoya University. Participants provided signed informed consent forms before the paper test, and at the end of the experiment, they received a token payment.

Stimulus Words

For the naming task, 28 Chinese 2-hanzi words (10 cognates and 18 non-cognates) were selected. All stimulus words were sourced from the textbook (J. Zhang et al., 2021) used in the participants' classroom during their 6 months of studying Chinese. Words

Table 4. Phonological Similarities in Chinese Words With Equivalent Japanese Words

Lexical similarity	First *hanzi*		Second *hanzi*		Whole words	
	M	*SD*	*M*	*SD*	*M*	*SD*
Cognate (n = 10)	.43	.28	.41	.31	.42	.14
Non-cognate (n = 18)	.35	.29	.30	.28	.33	.19
Total (N = 28)	.38	.29	.34	.29	.36	.18

were presented in both pinyin and hanzi scripts (refer to Appendix 2 for the complete list of stimulus words).

The ratio of phonological similarities between Chinese hanzi and Japanese kanji was calculated using the formula developed by Zhang (2018). Table 4 presents the means and standard deviations. Independent samples t-tests were conducted on phonological similarities to compare cognates and non-cognates between Chinese words and Japanese words. No significant differences were found in the phonological similarities for the 3 indexes between cognates and non-cognates: (1) the first hanzi, $t(26) = -.775$, $p = .445$, Cohen's $d = -.31$, (2) the second hanzi, $t(26) = -.895$, $p = .379$, Cohen's $d = -.35$, and (3) the entire words consisting of 2 hanzi, $t(26) = -1.364$, $p = .184$, Cohen's $d = -.54$. Consequently, phonological similarities remained consistent between cognates and non-cognates. As phonological similarities and cognates could potentially influence the speed and accuracy of the naming task, the index of phonological similarities for the entire word and cognates were included in the analysis of task performance.

Stroke numbers represent the calculated strokes required to write 2 hanzi. The average stroke number for the 28 words was 14.89 strokes (SD = 4.32 strokes). Regarding the index of frequencies, the Beijing Language and Culture University Corpus Center (BCC) Corpus, containing approximately 9.5 billion characters, is commonly used. However, since this extensive corpus is designed for native Chinese speakers, it may not be suitable for Japanese CFL learners at the beginner level. Instead, lesson numbers from the textbook (J. Zhang et al., 2021) were utilized as the index of lexical familiarity. Words appearing in early lessons (lessons 1 and 2) were considered more familiar than those in later lessons (lessons 4 and 5). The average lesson number for the 28 words in the textbook was 3.00 (SD = 1.52).

For the naming experiment, a counterbalancing method (Latin-square design) was employed to control the word conditions of hanzi and pinyin equivalently. It is important to note that 巴士 'bus' is a loanword presented in a 2-hanzi combination, and 是的 'yes' functions as an adverb. These words are commonly used in Chinese conversation, so they were included in the list of stimuli. None of these lexical characteristics of words in hanzi and pinyin reached significance: stroke numbers, $t(26) = .649$, $p = .522$, Cohen's $d = .25$, and lessons, $t(26) = 1.819$, $p = .080$, Cohen's $d = .69$. This indicates that these stimulus words were adequately controlled for use in the naming task.

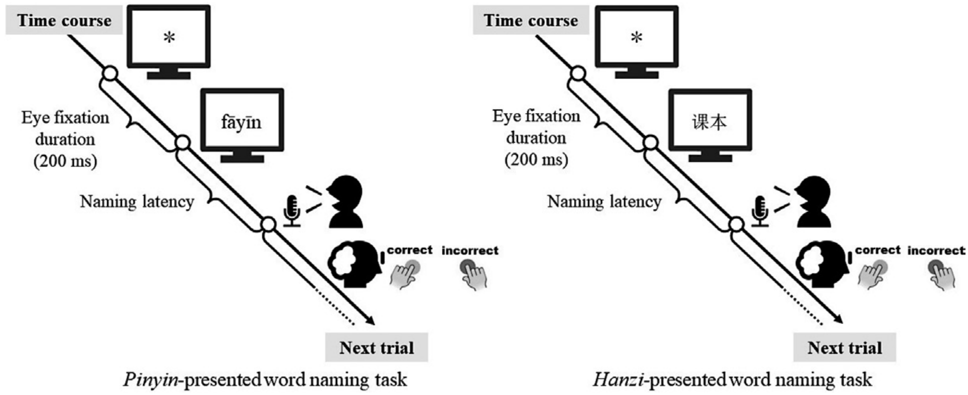

Fig. 1. A Single Trial of the *Pinyin*-Presented and *Hanzi*-Presented Word Naming

Procedure

The experiment was conducted using E-Prime 3.0. As illustrated in Fig. 1, an asterisk served as an eye fixation point, initially displayed at the center of a computer screen for 200 milliseconds (ms). Following this, the fixation point was replaced by a target word. Participants were instructed to pronounce the presented word, whether in hanzi or pinyin, as quickly and accurately as possible. To prevent the same word in a different script from being assigned to a single participant, all stimulus words were divided into 2 lists for distribution to 2 participant groups, employing a counterbalanced design. Each list contained 14 words in pinyin and 14 words in hanzi. Additionally, the experimental order of hanzi and pinyin naming was reversed for half of the participants (22 participants each). Naming latency, from visual presentation to the initiation of word naming, was automatically recorded by a computer using a voice key. Accuracy was assessed by an experimenter who was a native Chinese speaker with extensive experience in teaching Chinese to native Japanese speakers at the university level. Subsequent trials appeared every 600 ms, with all stimulus words presented randomly to each participant. Prior to the actual experiment, each participant received 5 practice items.

Measuring Chinese Ability

The Chinese ability of 44 native Japanese speakers, following a 6-month learning period, was assessed in a face-to-face classroom setting. The proficiency test covered various aspects of Chinese language skills. Most test questions were adapted from Tamaoka and Zhang (2022) and comprised 31 questions, distributed across the following categories.

(1) Understanding Quantifiers (10 points): Participants selected the correct quantifier to complete noun phrases, as in “五（　　）狗 (5 dogs).”

(2) Sentence Comprehension (6 points): Participants chose expressions to semantically complete sentences, as in “汉语语法（　　）难 (Chinese grammar is (　　) difficult).”

(3) Grammatical Understanding (10 points): Participants corrected grammatically

incorrect sentences, as in “医院旁边在车站 (The hospital is next to the station).”

(4) Japanese to Chinese Translation (5 points): Participants translated Japanese sentences into Chinese, as in “机の上に携帯がありますか (Is there a smartphone on the desk?).”

The total score ranged from 0 to 31 points, providing an overall measure of Chinese ability. Scores among the 44 Japanese participants varied from 7 to 31 points. The mean Chinese ability score was 23.64 points (SD = 5.37 points). The Cronbach’s reliability coefficient (α) for this Chinese ability test (N = 44) was notably high, α = .872.

Data Analysis for Naming Accuracies

Binomial accuracy data collected from the naming task underwent analysis using the linear mixed-effects (LME) model (Baayen et al., 2008) with the *lme4* package (Bates et al., 2014) in R (R Development Core Team, 2014). The *glmer* function, employing a logit link function, calculated the z distribution via maximum likelihood (Laplace approximation). A total of 1,232 responses (44 participants × 28 items) were subject to analysis. Fixed factors comprised Chinese ability, script (pinyin or hanzi), cognate, phonological similarity, and trial. Trial was standardized into z-values, denoted as *trial.z*. The 2 random factors were participants and stimulus words. Model comparisons, utilizing Akaike information criterion (AIC; Anderson et al., 2000), determined that the best-fit LME model was *glmer* (acc ~ script*chiability*cognate*phosimilarity + trial.z + (1|participant) + (1|stim), data, family = binomial). Here, ‘chiability’ refers to Chinese ability, and ‘phosimilarity’ refers to phonological similarity between Chinese and Japanese. Satterthwaite’s approximations (Satterthwaite, 1946) were employed via the *lmerTest* package to generate p values for each model (Kuznetsova et al., 2014), utilizing restricted maximum likelihoods (Harville, 1977).

Table 5 presents the results of the LME model. A significant main effect was observed for script type, $\beta = 4.938$, $z = 2.944$, $p = .003$, indicating that pinyin-presented words (M = 73.70%, SD = .44) were pronounced with greater accuracy compared to hanzi-presented words (M = 43.67%, SD = .50). Additionally, there was a significant main effect of Chinese ability, $\beta = .153$, $z = 2.913$, $p = .004$. However, neither the factor of cognate, $\beta = -2.134$, $z = -.646$, $p = .518$, nor phonological similarity, $\beta = 3.525$, $z = 1.105$, $p = .269$, reached significance. Furthermore, the LME analysis did not reveal a significant main effect for trial, $\beta = -.047$, $z = -.683$, $p = .494$.

A significant interaction between script type and Chinese ability regarding naming accuracy was observed, $\beta = -.143$, $z = -2.085$, $p = .037$. As depicted in the blue regression line in Fig. 2, the naming accuracy of words in pinyin exhibited a consistent trend, suggesting that Japanese CFL learners could pronounce words in pinyin with relatively higher accuracy. In contrast, the green regression line representing Chinese ability and naming accuracies for words written in hanzi displayed an upward slope. This suggested that Japanese CFL learners with higher Chinese ability who could leverage their L1 kanji knowledge to access the corresponding Chinese hanzi, exhibited improved naming accuracy for words in hanzi. Furthermore, a significant 3-way interaction among script type, cognate, and phonological similarity was observed, $\beta = -19.138$, $z = -2.031$, p

Table 5. Results of the LME Model Analysis for Accuracies

Variables	Estimate	*SE*	*z* value	Pr (> |z|)	*p*
(Intercept)	–3.978	1.331	–2.988	.003	**
script type	4.938	1.677	2.944	.003	**
Chinese ability	0.153	0.053	2.913	.004	**
cognate	–2.134	3.302	–0.646	.518	
phonological similarity	3.525	3.189	1.105	.269	
trial.z	–0.047	0.068	–0.683	.494	
script type*Chinese ability	–0.143	0.069	–2.085	.037	*
script type*cognate	7.990	4.197	1.904	.057	
Chinese ability*cognate	–0.031	0.129	–0.242	.809	
script type*phonological similarity	–3.306	4.475	–0.739	.460	
Chinese ability*phonological similarity	–0.103	0.126	–0.814	.416	
cognate*phonological similarity	3.079	7.072	0.435	.663	
script*Chinese ability*cognate	–0.250	0.171	–1.463	.144	
script*Chinese ability*phonological similarity	0.112	0.185	0.607	.544	
script*cognate*phonological similarity	–19.138	9.423	–2.031	.042	*
Chinese ability*cognate*phonological similarity	0.052	0.278	0.189	.850	
script*Chinese ability*cognate*phonological similarity	0.669	0.390	1.716	.086	

Note. Participants = 44. Items = 28. Total Observations = 1,232. *glmer* (acc ~ script*chiability*cognate*phosimilarity + trial.z + (1|participant) + (1|stim), data, family=binomial) where 'chiability' refers to Chinese ability and 'phosimilarity' refers to phonological similarity between Chinese and Japanese. LME = Linear Mixed-Effects.
**p* < .05, ***p* < .01.

= .042.

The main effects were observed for both script type and Chinese ability in terms of accuracy, with these factors displaying a significant 2-way interaction. As depicted in Fig. 2, it appeared that accuracy for word naming in pinyin remained consistent, while accuracy for word naming in hanzi increased with higher levels of Chinese ability. However, this trend appeared to be more intricate. Notably, there was a significant 3-way interaction involving Chinese ability, cognate status, and phonological similarity. Fig. 3 illustrates the relationships among these 3 factors within the context of script type, distinguishing between pinyin and hanzi.

As shown in Fig. 3, when the phonological similarity was 0, it became evident that the higher the Chinese ability, the better the accuracy in naming words in hanzi, regardless of whether they were cognates or non-cognates. In contrast, for the non-cognate condition, the accuracy of naming words in pinyin remained constant, but for

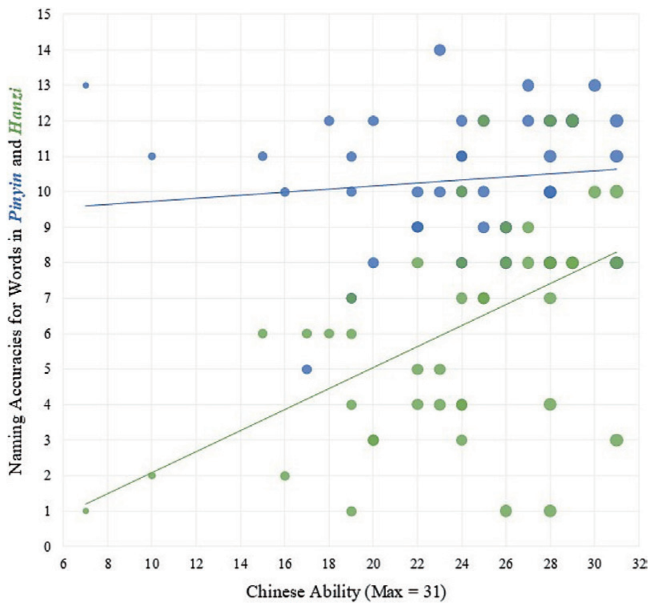

Fig. 2. Plotting of Naming Accuracies for Words in *Pinyin* and *Hanzi* With Chinese Ability
Note. Pinyin = blue line. *Hanzi* = green line.

cognates, the accuracy of words in pinyin decreased with increasing Chinese ability. This trend persisted even when phonological similarity values were at .2 and .3. However, a shift in this pattern occurred when phonological similarity reached .5. In the case of non-cognates, the accuracy for the naming of words in pinyin remained constant, regardless of Chinese ability. For words in hanzi, accuracy increased as Chinese ability improved. In cognates, both words in pinyin and hanzi saw improvement as Chinese ability increased. This trend held true even when phonological similarity reached .7.

Examining Fig. 3 comprehensively, it became evident that as Chinese ability improved, the accuracy of naming words in hanzi also improved, regardless of the specific relationship between phonological similarity and cognate status. In contrast, words in pinyin appeared to be strongly influenced by the condition of being cognates or non-cognates. When words were non-cognates, Japanese CFL learners maintained a consistently higher level of accuracy, irrespective of their Chinese ability. However, for cognate words, phonological similarity exerted a significant influence on naming accuracy, leading to a notable shift in the pattern. When phonological similarity was low, the accuracy of naming words in pinyin decreased as Chinese language ability improved.

This trend reversed when phonological similarity was higher, much like words in hanzi, the accuracy of words in pinyin also improved as Chinese ability increased. However, this shifting trend was mostly neutralized when considering overall accuracy, and consistently displayed higher naming accuracy for words in pinyin. This phenomenon might be attributed to the fact that during the phonetic conversion from pinyin to sound, Japanese CFL learners recognized words that existed in Japanese, enabling them to

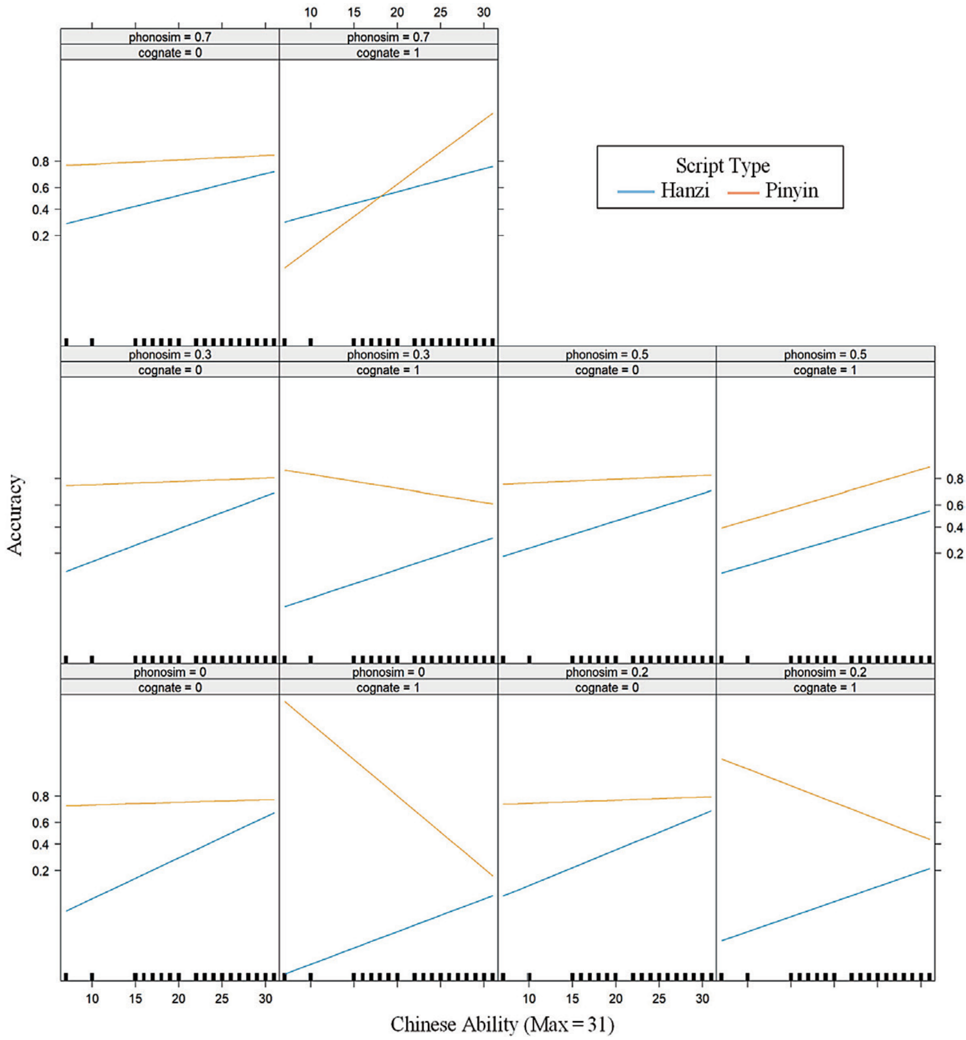

Fig. 3. The Interplay of Script, Cognate, Phonological Similarity and Chinese Ability
Note. The term 'phonosim' refers to phonological similarity. '0' in cognate refers to non-cognate, while '1' in cognate refers to cognate.

pronounce cognate words in pinyin more accurately as their Chinese ability improved. In contrast, accuracy for words in hanzi consistently improved as Chinese ability increased. As Japanese CFL learners' Chinese ability improves, they may adeptly utilize a hanzi-syllable conversion approach for word naming, leveraging the intricate interactions among phonological, orthographic, and semantic similarities between Japanese kanji and Chinese hanzi.

Data Analysis for Overall Naming Latency

After eliminating 509 incorrectly named items from a total of 1,232 items (responses),

Table 6. Results of the LME Model Analysis for Naming Latencies

Variables	Estimate	*SE*	*df*	*t* value	Pr (> |t|)	*p*
(Intercept)	6.511	.154	61.307	42.313	.000	***
script type	0.402	.082	647.730	4.896	.000	***
Chinese ability	0.005	.006	57.965	0.760	.450	
phonological similarity	0.223	.070	658.291	3.193	.001	**
trial.z	–0.015	.007	638.210	–2.208	.028	*
script type*Chinese ability	–0.010	.003	645.806	–3.177	.002	*
script type*phonological similarity	–0.282	.081	648.152	–3.459	.001	***
Chinese ability*phonological similarity	–0.007	.003	643.606	–2.687	.007	**
script*Chinese ability*phonological similarity	0.011	.003	646.617	3.533	.000	***

Note. Participants = 44. Items = 28. Total Observations = 705. *lmer* (log(rt) ~ script*chiability*phosimilarity + trial.z + (1|participant) + (1|stim), data) where 'chiability' refers to Chinese ability and 'phosimilarity' refers to phonological similarity between Chinese and Japanese. LME = Linear Mixed-Effects.
$*p < .05$, $**p < .01$, $***p < .001$.

the remaining 723 correctly answered items were analyzed for naming latencies. Employing the Box-Cox power transformation technique (Box & Cox, 1964; Venables & Ripley, 2002), a natural log transformation was applied to naming latencies to attenuate skewness in their distribution. Naming latencies were subjected to analysis using the *lmer* function with restricted maximum likelihood (Harville, 1977). Fixed factors, random factors, variable centering, and model comparisons were consistent with this accuracy analysis. Following AIC-based model selection, an optimal LME model was determined as *lmer* (log(rt) ~ script*chiability*phosimilarity + trial.z + (1|participant) + (1|stim), data). The fixed factor of cognate was excluded from the formula as it did not significantly contribute to naming speed. Based on this best-fit LME model, potentially influential outliers with absolute standardized residuals exceeding a standard deviation of 2.5 were removed (Kirk, 2013). In this operation, 18 responses were accordingly removed. The final result of the LME model analysis for 705 responses is reported in Table 6.

As shown in Table 6, the results of the LME analysis indicated that trial was a significant factor, $\beta = -.015$, $t(638.210) = -2.208$, $p = .028$, suggesting that naming latency shortened as the experiment progressed. The script type was also a significant factor, $\beta = .402$, $t(647.730) = 4.896$, $p < .001$, indicating that pinyin-presented words (M = 903 ms, SD = 261 ms) took longer to pronounce than hanzi-presented words (M = 745 ms, SD = 202 ms). Furthermore, phonological similarity was a significant factor, $\beta = .223$, $t(658.291) = 3.193$, $p = .001$, with a positive estimate, suggesting that phonological similarity slowed down naming speed. However, the main effect of Chinese ability was not significant, $\beta = .005$, $t(57.965) = .760$, $p = .450$.

An important finding was the significant interaction between script type and Chinese

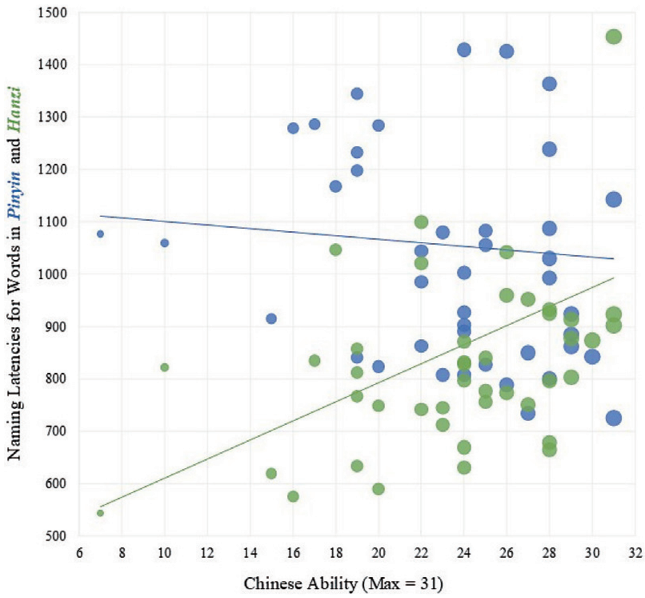

Fig. 4. Plotting of Naming Latencies for Words in *Pinyin* and *Hanzi* With Chinese Ability
Note. Pinyin = blue line. *Hanzi* = green line.

ability, β = –.01, *t*(645.806) = –3.177, *p* = .002, indicating that Chinese ability had a differential impact on pinyin and hanzi scripts. To further understand the effect of this interaction, naming latencies for words written in pinyin and hanzi were plotted against Japanese CFL learners' Chinese ability, as illustrated in Fig. 4. The blue regression line for pinyin displayed a flat slope, signifying a consistent decoding speed for words (naming latency) in pinyin. Since pinyin could be pronounced based on regular pinyin-to-sound conversion rules, the naming latency remained constant, although longer than the whole syllable unit conversion of hanzi. In contrast, recalling pronunciation from hanzi was more challenging, resulting in slower naming latency for words in hanzi as Chinese ability increased, as depicted by the green regression line.

Combining the trends observed in Fig. 2 for accuracy and Fig. 4 for naming latency, it became evident that the accuracy of words in hanzi increased with higher Chinese ability, but the naming latency also extended. This phenomenon might be attributed to Japanese CFL learners pronouncing hanzi more cautiously to avoid errors. In contrast, for words in pinyin, both accuracy and naming latency remained consistent regardless of Chinese ability. This stability was likely because Japanese CFL learners could correctly pronounce words in pinyin when they took their time and applied the pinyin-to-sound conversion rules accurately. These results suggest a pattern among hanzi-presented words, indicating more accurate word pronunciation, yet slower processing speed with increased Chinese ability, depicting an 'accuracy-and-speed trade-off' relationship.

The interaction between script type and phonological similarity was also found to be significant, β = –.282, *t*(648.152) = –3.459, *p* < .001. Additionally, the interaction

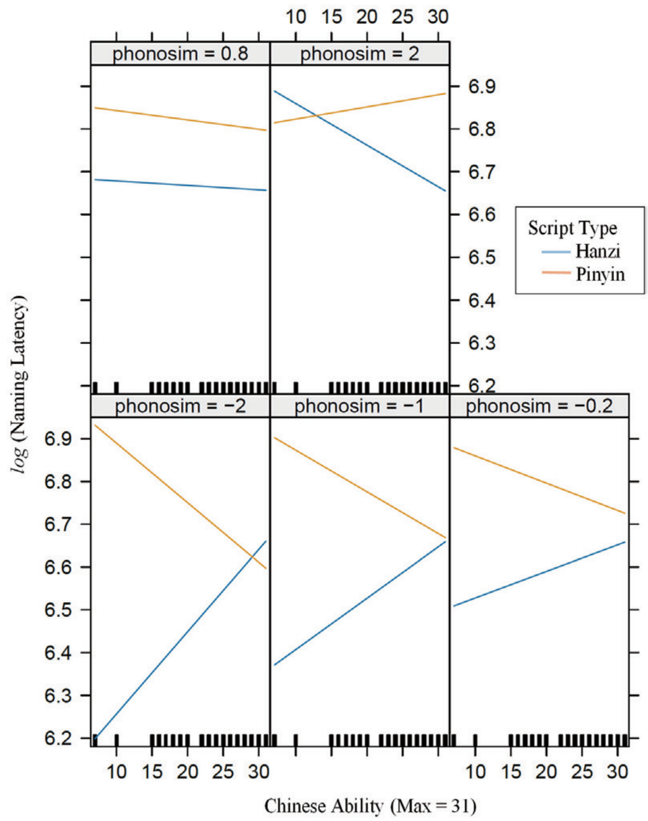

Fig. 5. The Three-Way Interaction of Script, Phonological Similarity and Chinese Ability

Note. The term 'phonosim' refers to phonological similarity. Values of phonological similarity were standardized (*z*-score), ranging from –2 to 2. Naming latency is presented as a natural log-converted value.

between Chinese ability and phonological similarity was also significant, $\beta = -.007$, $t(643.606) = -2.687$, $p = .007$. Furthermore, the 3-way interaction of script type, Chinese ability, and phonological similarity was highly significant, $\beta = .011$, $t(646.617) = 3.533$, $p < .001$. Therefore, naming latencies were illustrated separately for hanzi-presented words and pinyin-presented words under different conditions of Chinese ability and phonological similarity in Fig. 5.

In Fig. 5, phonological similarity was standardized to *z*-values ranging from –2 to 2, and the script type of pinyin and hanzi was examined alongside phonological similarity with increasing Chinese proficiency. At low phonological similarity values of –2, –1, and –.2, naming latencies for words in hanzi slowed down with increasing Chinese ability, while the opposite was true for words in pinyin, where naming speed increased with higher Chinese ability. When the phonological similarity reached .8, both words in hanzi and pinyin remained relatively constant, regardless of Chinese ability. However, when the phonological similarity increased to 2, the naming latency lengthened for words in pinyin as Chinese ability improved, whereas for words in hanzi, the naming latency shortened as Chinese ability improved.

In summary, as the phonological similarity of Chinese and Japanese increased, the time to reach pronunciation became progressively faster as Chinese ability improved. However, for words in pinyin, the opposite was observed; as phonological similarity increased, the time to reach pronunciation gradually slowed down with improved Chinese ability. Overall, the naming latency for words in hanzi appeared to lengthen as Chinese ability improved, but shorten for words in pinyin. However, the effect of phonological similarity played a significant role in these patterns.

Discussion

Accuracy and latency in the naming experiment yielded interesting results. Words presented in pinyin were pronounced with higher accuracy compared to words in hanzi. Increased Chinese ability resulted in improved accuracy. Neither cognate status nor phonological similarity had a significant impact on accuracy. An interaction effect revealed that word accuracy for pinyin remained consistent regardless of Chinese ability, while word accuracy for hanzi improved with higher Chinese ability. For pinyin, accuracy depended on both cognate status and phonological similarity, leading to a noticeable shift in this pattern. Regarding naming latencies, words in pinyin took longer to pronounce than those in hanzi. Higher phonological similarity slowed down naming speed, but Chinese ability alone did not significantly affect naming latency. An interaction effect showed that naming latency of words in pinyin remained consistent. In contrast, naming latency of words in hanzi increased with higher Chinese ability. In the case of hanzi, there appeared to be a 'trade-off' between accuracy and latency; accuracy improved, but latency lengthened with increased Chinese ability. These findings underscore the intricate relationship between script type (pinyin versus hanzi), Chinese proficiency, cognate status, and phonological similarity in the naming latency and accuracy of Japanese CFL learners.

General Discussion

Chinese text is typically written only in hanzi without pinyin. Since Japanese CFL learners have already acquired Chinese characters in their L1 Japanese, they possess advanced knowledge of orthographically-equivalent morphosyllabic units of Chinese hanzi (Chen, 2002; Hishinuma, 1983, 1984). Due to this advantage in script similarity (Djojomihardjo et al., 1994; Tamaoka, 2000, 2014, 2015, 2022), Japanese CFL learners would be able to readily utilize their Japanese kanji knowledge to access the meanings of Chinese words in hanzi, similar to how Chinese JFL learners benefit from their advantage of having learned Japanese kanji (Matsumoto, 2013; Nakayama, 2002; Tamaoka, 1997, 2000; Yamato & Tamaoka, 2013). However, the interplay of the Japanese kanji advantage for lexical processing of Chinese words in hanzi with other factors, including the involvement of pinyin, Chinese ability, cognate status and phonological similarity, has not been clearly indicated. In the present study, both a paper test and an experimental study were conducted to explore these relationships.

The paper test revealed a greater comprehension of words presented in hanzi compared to those in pinyin. Furthermore, cognates were understood more accurately than non-cognates, irrespective of script type. Notably, a significant interaction between script type and lexical similarity (cognate) was observed. In pinyin, there was no substantial difference in comprehension between cognates and non-cognates. However, in hanzi, cognates were markedly better understood than non-cognates. These findings underscored the advantage that Japanese CFL learners have in comprehending Chinese words presented in hanzi, attributable to the similarity between hanzi and Japanese kanji. This observation was consistent with previous research on Chinese JFL learners, highlighting the role of orthographic and semantic similarities between hanzi and kanji in enhancing word comprehension (Fei, 2013, 2015; Matsumi et al., 2012, 2016; Nakayama, 2002; Tamaoka, 1997, 2000; Tome et al., 2012; Yamato & Tamaoka, 2013). Overall, the paper test offered valuable insights into the influence of script type and lexical similarity (cognate) on the semantic comprehension strategies employed by Japanese CFL learners in learning Chinese vocabulary.

The complexity of the relationship between hanzi and sound can significantly impact phonological processing. The "orthographic depth hypothesis" (Ellis et al., 2004; Katz & Feldman, 1983; Katz & Frost, 1992; Koda, 1988, 1990) categorizes hanzi as a deep orthography and pinyin as a shallow orthography. According to this hypothesis, phonological processing for the deep orthography of hanzi is expected to take longer than for the shallow orthography of pinyin. Phonological conversions from hanzi to syllabic sounds are complex, lacking regular script-to-sound rules, whereas pinyin can be consistently converted to their corresponding sounds. Thus, it was predicted that the ease of script-to-sound consistency would enable more accurate pronunciations of Chinese words written in pinyin compared to those in hanzi. Although Japanese CFL learners acquire knowledge of hanzi from their L1 kanji, they must carefully convert syllabic sounds of hanzi to correctly pronounce Chinese words.

With the inclusion of Chinese ability as a factor, the naming experiment yielded an unexpected outcome. Despite the orthographic depth hypothesis predicting a significant influence of script type, with higher accuracy and shorter naming latency anticipated for words presented in pinyin compared to hanzi, the findings diverged from these expectations. Contrary to predictions, naming latency indicated that Japanese CFL learners took longer to pronounce words in pinyin than those in hanzi although words in pinyin showed constantly higher accuracy than those in hanzi. This suggested that words in pinyin required more time for accurate decoding and pronunciation. Furthermore, irrespective of the Chinese ability of Japanese CFL learners, words in pinyin consistently exhibited higher accuracies than those in hanzi, despite their longer naming latencies. Conversely, as Chinese ability increased, word accuracy for hanzi also increased. However, this improvement in accuracy was accompanied by longer naming latencies, indicating a 'speed-and-accuracy tradeoff' phenomenon for words presented in hanzi.

The script type also interacted with factors such as cognates, phonological similarity, and Chinese ability, contributing to variations in the naming performance of Japanese

CFL learners. In the naming experiment involving Japanese CFL learners, cognate words notably influenced pronunciation accuracy, particularly within the hanzi script. These cognate words, sharing similar meanings in both Japanese and Chinese, exhibited a significant impact on pronunciation accuracy. Surprisingly, cognate words were pronounced less accurately compared to non-cognate words, irrespective of script type (pinyin or hanzi). Notably, an interaction between script type (pinyin and hanzi) and cognate words was observed, leading to a significantly lower accuracy rate for cognate words presented in hanzi compared to those in pinyin. This finding suggests that cognate words in hanzi may impede accurate pronunciation, potentially influenced by their L1 Japanese pronunciation. Conversely, cognate words in pinyin demonstrated relatively high accuracy. This consistency in pronunciation accuracy for cognates in pinyin was attributed to the rule-based phonetic conversion inherent in pinyin, enabling learners to pronounce these words more accurately. Cognate words exerted a distinct influence on pronunciation accuracy in the naming experiment, with their impact differing between the pinyin and hanzi scripts. While cognates in hanzi displayed lower accuracy, cognates in pinyin were pronounced more accurately.

Cognates between Japanese and Chinese encompass semantic, orthographic, and phonological similarities. For instance, the word 'material' is a cognate, written as 资料 /zī liào/ in Chinese and 資料 /si ryô/ in Japanese. While cognates exhibit strong orthographic resemblance, allowing Japanese CFL learners to easily recognize their lexical meaning, they also present a challenge in pronunciation. The Japanese pronunciation of /si ryô/ is quickly activated, leading to potential mispronunciation when attempting to pronounce the cognate in Chinese. Consequently, Japanese CFL learners must suppress the Japanese pronunciation /si ryô/, and activate the Chinese pronunciation of /zī liào/. As Chinese ability increases, Japanese CFL learners become more aware of the mispronunciation process due to multiple resemblances.

The phonological similarity between Chinese hanzi and Japanese kanji emerged as a significant factor influencing naming accuracy and latency among Japanese CFL learners. Phonological similarity positively impacted naming accuracy, with words sharing higher similarity between Chinese hanzi and Japanese kanji being more accurately pronounced by Japanese CFL learners. Furthermore, a significant interaction was observed between script type (pinyin and hanzi) and phonological similarity, indicating a stronger influence of phonological similarity on pronunciation accuracy for hanzi words. Interestingly, while higher phonological similarity facilitated accuracy, it also affected naming latency. Hanzi words with greater phonological similarity were associated with slightly shorter naming latencies. Conversely, pinyin-presented words showed a slight increase in naming latency with a higher phonological similarity. Thus, phonological similarity between Chinese hanzi and Japanese kanji exerted a dual impact on naming accuracy and latency, enhancing accuracy, particularly for hanzi words, while also influencing naming speed. These findings underscored the nuanced role of phonological similarity in shaping the naming strategies adopted by Japanese CFL learners, impacting both accuracy and processing speed.

Chinese ability, as assessed among Japanese CFL learners, significantly influenced naming performance when encountering words in both pinyin and hanzi scripts. Notably, Chinese ability had a substantial impact on naming accuracy, especially for words presented in hanzi. The results revealed a positive relation between Chinese ability and the accuracy of naming hanzi words, indicating that Japanese CFL learners with higher Chinese proficiency were more adept at pronouncing hanzi accurately. Moreover, a significant three-way interaction was observed among script type, Chinese ability, and phonological similarity. This interaction suggested that Chinese ability generally contributed to enhanced accuracy, particularly for hanzi words, throughout varying levels of phonological similarity. These findings underscore the pivotal role of Chinese ability in shaping the pronunciation and processing of Chinese characters among Japanese CFL learners.

The significance of this study lies in the early acquisition of Chinese, which not only enhances Chinese language ability but also facilitates the recall of similar pronunciations in both languages through the medium of hanzi. This enables Japanese CFL learners to achieve greater accuracy in pronunciation, albeit at the expense of speed. For Japanese CFL learners, the phonological similarities activated through hanzi introduce a tradeoff between the speed of recall and the accuracy of pronunciation. Considering these findings, Chinese language education for native Japanese speakers should place emphasis on leveraging the phonological similarities facilitated by hanzi in both languages. CFL instructors can incorporate specific strategies and techniques aimed at enhancing students' awareness of these similarities and developing their ability to navigate the intricacies of pronunciation. By doing so, Japanese CFL learners can achieve a balance between speed and accuracy in their language acquisition, ultimately leading to more effective and efficient learning outcomes.

AUTHOR'S CONTRIBUTION

J.Z. conducted the experiments, administered the Chinese comprehension test, collated the data for statistical analyses, performed the statistical analyses, and drafted the manuscript. K.T. assisted in constructing the theoretical framework, conducted statistical analyses, and, as the corresponding author, drafted responses to reviewers. Both authors contributed to the article and approved the submitted version.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

Anderson, D. R., Burnham, K. P., & Thompson, W. L. (2000). Null hypothesis testing: Problems, prevalence, and an alternative. *Journal of Wildlife Management*, *64*(4), 912–923. https://doi.org/10.2307/3803199

Anderson, R. C., Ku, Y.-M., Li, W., Chen, X., Wu, X., & Shu, H. (2013). Learning to see the patterns in Chinese characters. *Scientific Studies of Reading*, *17*(1), 41–56. https://doi.org/10.1080/10888438.2012.689789

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-6*.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252. https://www.jstor.org/stable/2984418

Chen, Y. (2002). Nihongo nizi kanzi goi to sore ni taiô suru tyûgokugo nizi kanzi goi wa onazi ka: Taiwan oyobi tyûgoku no tyûgokugo tono hikaku [Are the Japanese two-kanji words and their corresponding Chinese two-hanzi words the same?: Comparison with Chinese in Taiwan and China]. *Gengo Bunka to Nihongo Kyôiku*, *24*, 40–53.

Chiu, H. C. (2003). Taiwanzin nihongo gakusyûsya no nihongo kanzi zyukugo no on'in syori ni tuite: Tango taipu, tango no syûtoku nenrei, syûzyukudo no kanten kara no kentô [The phonological processing of Japanese kanji words by Taiwanese learners of Japanese: An analysis of word type, age of acquisition, and fluency level]. *Nihongo Kyôiku*, *116*, 89–98.

Chung, K. K. H. (2003). Effects of pinyin and first language words in learning of Chinese characters as a second language. *Journal of Behavioral Education*, *12*(3), 207–223. https://doi.org/10.1023/A:1025560327860

Dai, B.-Y., & Lu, J.-P. (1985). Reading reform in Chinese primary schools. *Prospects*, *15*(1), 103–110. https://doi.org/10.1007/BF02195707

Ding, Y. (1988). Liyong Hanyu *pinyin* keyi tigao Hanyu de jiaoxuelv [Using Hanyu *pinyin* can improve the teaching rate of Chinese language]. In *Hanyu pinyin lunwenxuan* (pp. 107–109). Language Press.

Djojomihardjo, M., Koda, K., & Moates, D. (1994). Development of L2 word recognition. In Q. Jing, H. Zhang, & D. Peng (Eds.), *Information processing of Chinese language* (pp. 153–161). Beijing Normal University Publishing.

Ellis, N. C., Natsume, M., Stavropoulou, K., Hoxhallari, L., van Daal, V. H. P., Polyzoe, N., Tsipa, M.-L., & Petalas, M. (2004). The effects of orthographic depth on learning to read alphabetic, syllabic, and logographic scripts. *Reading Research Quarterly*, *39*(4), 438–468. https://doi.org/10.1598/RRQ.39.4.5

Everson, M. E. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *The Modern Language Journal*, *82*(2), 194–204. https://doi.org/10.2307/329208

Fei, X. (2013). The auditory recognition of Japanese Kanji-words in proficient Chinese learners of Japanese studying in Japan: Effects of the orthographic and phonological similarities between Chinese and Japanese. *Journal of International Student Education*, *18*, 35–43.

Fei, X. (2015). The process of learning Japanese Kanji (Chinese character) words in Chinese-native learners of the Japanese language: Effects of orthographical and phonological similarities between the Chinese and the Japanese languages. *Theory and Research for Developing Learning Systems*, *1*, 57–69. https://ir.lib.hiroshima-u.ac.jp/00039417

Fei, X., Zhao, S., & Liu, J. (2022). Auditory recognition of Chinese-Japanese cognates and homographs by Chinese JFL learners. *Psychologia*, *64*(1), 1–22. https://doi.org/10.2117/psysoc.2021-A144

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, *72*(358), 320–338. https://doi.org/10.2307/2286796

Hishinuma, T. (1983). Nihongo to tyûgokugo no zyôyô zii [Commonly-used kanji in Japanese and Chinese]. *Tyûgoku Kenkyû Geppô*, *428*, 1–20.

Hishinuma, T. (1984). Tyûgokugo no hyôzyun zitai to nihongo no tûyô zitai [The standard orthography in Chinese and the common orthography in Japanese]. *Nihongogaku*, *3*, 32–40.

Ho, C. S.-H., & Bryant, P. (1997). Phonological skills are important in learning to read Chinese. *Developmental Psychology*, *33*(6), 946–951. https://doi.org/10.1037/0012-1649.33.6.946

Hong, S. J. (2004). Kikitori no purosesu ni okeru kanzi no eikyô: Taiwanzin nihongo gakusyûsya (dikutêsyon bunseki) o tyûsin ni site [The influence of kanji in the listening process: Focus on Taiwanese Japanese learner (dictation analysis)]. *Tukuba Daigaku Tiiki Kenkyû*, *22*, 155–176.

Huang, H. S., & Hanley, J. R. (1997). A longitudinal study of phonological awareness, visual skills, and Chinese reading acquisition among first-graders in Taiwan. *International Journal of Behavioral Development*, *20*(2), 249–268. https://doi.org/10.1080/016502597385324

Ishida, T. (1986). Eigo, tyûgokugo, kankokugo ken betu nihongo gakuryoku no bunseki [Analysis of Japanese proficiency among areas of English, Chinese and Korean speakers]. *Nihongo Kyôiku*, *58*, 162–194.

Jordan, D. (1971). *Guide to the romanization of Chinese*. Mei Ya Publications.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 157–166. https://doi.org/10.1037/0278-7393.9.1.157

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Advances in psychology* (Vol. 94, pp. 67–84). Elsevier. https://doi.org/10.1016/S0166-4115(08)62789-2

Kirk, R. E. (2013). *Experimental design: Procedures for the behavioral sciences* (4th ed.). Sage Publications.

Koda, K. (1988). Cognitive process in second language reading: Transfer of L1 reading skills and strategies. *Second Language Research*, *4*(2), 133–155. https://doi.org/10.1177/026765838800400203

Koda, K. (1990). The use of L1 reading strategies in L2 reading: Effects of L1 orthographic structures on L2 phonological recoding strategies. *Studies in Second Language Acquisition*, *12*(4), 393–410. https://doi.org/10.1017/S0272263100009499

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package) (R package version 2.0-33)*. https://cran.r-project.org/web/packages/lmerTest/

Lee, C. H., & Kalyuga, S. (2011). Effectiveness of on-screen pinyin in learning Chinese: An expertise reversal for multimedia redundancy effect. *Computers in Human Behavior*, *27*(1), 11–15. https://doi.org/10.1016/j.chb.2010.05.005

Ma, S., Zhang, X., Hatfield, H., & Wei, W.-H. (2020). Pinyin is an effective proxy for early screening for Mandarin-speaking children at risk of reading disorders. *Frontiers in Psychology*, *11*, Article 327. https://doi.org/10.3389/fpsyg.2020.00327

Matsumi, N., Fei, X., & Cai, F. (2012). Nihongo kanzi tango no syori katei [The lexical processing of Japanese Kanji words]. In K. Hatasa, Y. Hatasa, M. Kudara, & T. Shimizu (Eds.), *Dainigengo syûtoku kenkyû to gengo kyôiku* (pp. 43–67). Kuroshio Publishers.

Matsumi, N., Fei, X., & Zhu, G. (2016). Processing of oral translation from Chinese Kanji words to Japanese words among intermediate Chinese learners of Japanese: Effect of orthographic and phonological similarities between Chinese and Japanese on auditory identification. *Bulletin of the Department of Teaching Japanese as a Second Language, Hiroshima University*, *26*, 29–34. https://doi.org/10.15027/39740

Matsumoto, K. (2013). Kanji recognition by second language learners: Exploring effects of first language writing systems and second language exposure. *The Modern Language Journal*, *97*(1), 161–177. https://doi.org/10.1111/j.1540-4781.2013.01426.x

McGinnis, S. (1997). Tonal spelling versus diacritics for teaching pronunciation of Mandarin Chinese. *The Modern Language Journal*, *81*(2), 228–236. https://doi.org/10.2307/328789

Miller, R. A. (1967). *The Japanese language*. The University of Chicago Press.

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

R Development Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*(6), 110–114. https://doi.org/10.2307/3002019

Shu, H., & Liu, B. (1994). Pinyin zai xiaoxue dinianji ertong zaoqi yuedu zhong zuoyong de yanjiu [The role of pinyin in early reading for Chinese junior primary school children]. *Psychological Development and Education*, *3*, 11–15.

Shu, H., Zeng, H., & Chen, W. (1993). Xiaoxue dinianji ertong liyong pinyin xuexi shengzici de shiyan yanjiu [Research on using pinyin in learning vocabulary by junior primary school children]. *Psychological Development and Education*, *1*, 18–22.

Tamaoka, K. (1997). The processing strategy of words presented in kanji and kana by Chinese and English speakers learning Japanese. *Gengo Bunka Kenkyû*, *17*(1), 65–77.

Tamaoka, K. (2000). The effects of mother-tongue scripts on the phonological processing of the Japanese language by Chinese and English speakers. *Dokusyo Kagaku*, *44*(3), 83–94.

Tamaoka, K. (2014). The Japanese writing system and lexical understanding. *Japanese Language and Literature*, *48*(2), 431–471. https://www.jstor.org/stable/24394417

Tamaoka, K. (2015). Processing of the Japanese language by native Chinese speakers. In M. Nakayama (Ed.), *Handbook of Japanese psycholinguistics* (pp. 583–632). De Gruyter Mouton.

Tamaoka, K. (2022). Nihongo gakusyûsya no kioku mekanizumu to sinteki zisyo no kôzô [The mechanism of memory in relation to the structure of the mental lexicon among learners of Japanese as a second language]. *Dai-2-Gengo to site no Nihongo no Syûtoku Kenkyû*, *25*, 57–83.

Tamaoka, K., & Makioka, S. (2004). New figures for a Web-accessible database of the 1,945 basic Japanese kanji, fourth edition. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 548–558. https://doi.org/10.3758/BF03195601

Tamaoka, K., & Taft, M. (2010). The sensitivity of native Japanese speakers to On and Kun kanji readings. *Reading and Writing*, *23*(8), 957–968. https://doi.org/10.1007/s11145-009-9184-8

Tamaoka, K., & Zhang, J. (2022). The effect of Chinese proficiency on determining temporal adverb position by native Japanese speakers learning Chinese. *Frontiers in Psychology*, *12*, Article 783366. https://doi.org/10.3389/fpsyg.2021.783366

Tome, M., Fei, X., & Matsumi, N. (2012). A survey of phonological similarity of Japanese kanji-words to Chinese-words: Focused on cognates, interlingual homographs, and non-cognates. *Bulletin of the Department of Teaching Japanese as a Second Language, Hiroshima University*, *22*, 41–48.

Venables, W. N., & Ripley, B. D. (2002). Random and mixed effects. In *Modern applied statistics with S* (pp. 271–300). Springer. https://doi.org/10.1007/978-0-387-21706-2_10

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

Wang, Y., Lam, S. S.-Y., Mo, J., & McBride-Chang, C. (2014). Pinyin knowledge as a potentially important marker of early literacy. In K. K. H. Chung, K. C. P. Yuen, & D. M. McInerney (Eds.), *Understanding developmental disorders of auditory processing, language and literacy across languages: International perspectives* (pp. 189–206). Information Age Publishing.

Xiao, H., Xu, C., & Rusamy, H. (2020). *Pinyin* spelling promotes reading abilities of adolescents learning Chinese as a foreign language: Evidence from mediation models. *Frontiers in Psychology*, *11*, Article 596680. https://doi.org/10.3389/fpsyg.2020.596680

Yamato, Y., & Tamaoka, K. (2013). Effects of linguistic similarity on Japanese text processing by native Chinese and Korean speakers. *Koide Kinen Nihongo Kyôiku Kenkyûkai Ronbunsyû*, *21*, 61–73.

Zhang, H., Zhang, X., Li, M., & Zhang, Y. (2021). Morphological awareness in L2 Chinese reading comprehension: Testing of mediating routes. *Frontiers in Psychology*, *12*, Article 736933. https://doi.org/10.3389/fpsyg.2021.736933

Zhang, J. (2018). Background factors for judgements of phonological similarities by native Japanese speakers on Chinese-Japanese cognates. *Studia Linguisica*, *32*, 61–78. https://doi.org/10.18999/stul.32.61

Zhang, J., Tamaoka, K., & Wang, L. (2021). *Netto zidai no tyûgokugo* [The Chinese language in the internet era]. Hituzi Shobo.

Zhang, S.-Z., Georgiou, G. K., Inoue, T., Zhong, W.-W., & Shu, H. (2020). Do pinyin and character recognition help each other grow? *Early Childhood Research Quarterly*, *53*, 476–483. https://doi.org/10.1016/j.ecresq.2020.06.004

Zhou, Y. (1959). Hanyu pinyin jiaoyu he hanzi gaige yundong [Hanyu pinyin education and the Chinese character reform movement]. *Journal of Peking University (Philosophy and Social Sciences)*, *4*(2), 113–121.

(Manuscript received 31 July, 2023; Revision accepted 26 March, 2024; Released online in J-STAGE as advance publication 22 October, 2024)

APPENDIX 1

Chinese Words Used for the Paper Test

#	Stimuli	Script type	Stroke numbers	Familiarity (Lesson no.)	Stimuli type (Cognate)	Parts of speech	Meaning
1	衣服	*pinyin*	14	Lesson 6	Cognate	Noun	cloth
2	学习	*pinyin*	11	Lesson 7	Cognate	Verb	study
3	资料	*pinyin*	20	Lesson 8	Cognate	Noun	information
4	小说	*pinyin*	12	Lesson 9	Cognate	Noun	fiction
5	参加	*pinyin*	13	Lesson 10	Cognate	Verb	participate
6	流行	*pinyin*	16	Lesson 11	Cognate	Verb	popular
7	新闻	*pinyin*	22	Lesson 12	Cognate	Noun	news
8	巴士	*pinyin*	7	Lesson 6	Non-cognate	Noun	bus
9	上课	*pinyin*	13	Lesson 7	Non-cognate	Verb	take a class
10	衬衫	*pinyin*	16	Lesson 8	Non-cognate	Noun	shirts
11	钢琴	*pinyin*	21	Lesson 9	Non-cognate	Noun	piano
12	感冒	*pinyin*	22	Lesson 10	Non-cognate	Verb	flu
13	飞机	*pinyin*	9	Lesson 11	Non-cognate	Noun	airplane
14	报纸	*pinyin*	14	Lesson 12	Non-cognate	Noun	newspaper
15	商店	*hanzi*	19	Lesson 6	Cognate	Noun	shop
16	起床	*hanzi*	17	Lesson 7	Cognate	Verb	get up
17	现在	*hanzi*	14	Lesson 8	Cognate	Noun	now
18	回答	*hanzi*	18	Lesson 9	Cognate	Verb	answer
19	休息	*hanzi*	16	Lesson 10	Cognate	Verb	rest
20	练习	*hanzi*	11	Lesson 9	Cognate	Verb	practice
21	时间	*hanzi*	17	Lesson 6	Cognate	Noun	time
22	睡觉	*hanzi*	22	Lesson 6	Non-cognate	Verb	sleep
23	那里	*hanzi*	13	Lesson 7	Non-cognate	Pronoun	there
24	打扫	*hanzi*	11	Lesson 8	Non-cognate	Verb	clean up
25	跳舞	*hanzi*	27	Lesson 9	Non-cognate	Verb	dance
26	足球	*hanzi*	18	Lesson 10	Non-cognate	Noun	football
27	寄信	*hanzi*	20	Lesson 11	Non-cognate	Verb	send a letter
28	收拾	*hanzi*	15	Lesson 12	Non-cognate	Verb	tidy up

APPENDIX 2

Chinese Words Used for the Naming Task

#	*Hanzi*	*Pinyin*	Phonological similarities			Stroke numbers	Familiarity (Lesson no.)	Stimuli type	Parts of speech	Meaning
			1st *hanzi*	2nd *hanzi*	Two-*hanzi* words					
1	名字	míngzi	0.25	1.00	0.63	12	Lesson 1	Cognate	Noun	name
2	早上	zǎoshang	0.33	0.22	0.28	9	Lesson 1	Non-cognate	Noun	morning
3	下午	xiàwǔ	0.40	0.00	0.20	7	Lesson 1	Non-cognate	Noun	afternoon
4	蛋糕	dàngāo	0.67	0.33	0.50	27	Lesson 2	Non-cognate	Noun	cake
5	课本	kèběn	0.50	0.33	0.42	15	Lesson 2	Non-cognate	Noun	textbook
6	发音	fāyīn	0.33	0.40	0.37	14	Lesson 3	Cognate	Noun	pronunciation
7	水果	shuǐguǒ	0.86	0.00	0.43	12	Lesson 3	Non-cognate	Noun	fruit
8	朋友	péngyou	0.00	0.67	0.33	12	Lesson 3	Cognate	Noun	friend
9	作业	zuòyè	0.00	0.33	0.17	12	Lesson 3	Non-cognate	Noun	homework
10	词典	cídiǎn	0.50	0.29	0.39	15	Lesson 4	Non-cognate	Noun	dictionary
11	邮票	yóupiào	0.67	0.25	0.46	18	Lesson 4	Non-cognate	Noun	stamp
12	银行	yínháng	0.67	0.00	0.33	17	Lesson 5	Cognate	Noun	bank
13	超市	chāoshì	0.50	0.80	0.65	17	Lesson 5	Non-cognate	Noun	supermarket
14	邮局	yóujú	0.67	0.29	0.48	14	Lesson 5	Non-cognate	Noun	post office
15	什么	shénme	0.25	0.00	0.13	7	Lesson 1	Non-cognate	Pronoun	what
16	我们	wǒmen	0.00	0.00	0.00	12	Lesson 1	Non-cognate	Pronoun	we
17	大家	dàjiā	0.00	0.40	0.20	13	Lesson 1	Non-cognate	Pronoun	everybody
18	铅笔	qiānbǐ	0.33	0.33	0.33	20	Lesson 2	Cognate	Noun	pencil
19	哪里	nǎlǐ	0.00	1.00	0.50	16	Lesson 2	Non-cognate	Pronoun	where
20	是的	shìde	0.00	0.33	0.17	17	Lesson 2	Non-cognate	Adverb	yes
21	电脑	diànnǎo	0.57	0.67	0.62	15	Lesson 2	Non-cognate	Noun	computer
22	最近	zuìjìn	0.29	0.67	0.48	19	Lesson 3	Cognate	Noun	recently
23	苹果	píngguǒ	0.29	0.00	0.14	16	Lesson 4	Non-cognate	Noun	apple
24	杂志	zázhì	1.00	0.33	0.67	13	Lesson 4	Cognate	Noun	magazine
25	冰箱	bīngxiāng	0.00	0.22	0.11	21	Lesson 5	Non-cognate	Noun	refrigerator
26	公园	gōngyuán	0.29	0.33	0.31	11	Lesson 5	Cognate	Noun	park
27	医院	yīyuàn	0.67	0.33	0.50	16	Lesson 5	Cognate	Noun	hospital
28	食堂	shítáng	0.50	0.00	0.25	20	Lesson 5	Cognate	Noun	dining hall