

A Study of the Mutual Phonetic Resemblance between Japanese Kyōiku Kanji and Chinese

Yuji Obataya, University of Geneva, Switzerland

The IAFOR International Conference on Education – Hawaii 2021
Official Conference Proceedings

Abstract

This study investigated the mutual phonetic resemblance of Chinese ideograms between Japanese and Chinese using the official list (expanded in 2020) of *Kyōiku kanji* (Chinese ideograms in Japanese, taught in elementary school). This analysis aimed to determine methods for helping learners from non-*kanji* backgrounds to quantify the phonetic gap between these two languages. First, since most kanji symbols can be read in multiple ways (*on'yomi* [Chinese reading in Japan] and *kun'yomi* [Japanese reading]), the rate of the use of *on'yomi* for each of the 1,026 *kanji* was calculated at 66.1% by accounting for the factor of the frequency of all 9,292 words that contain these *kanji* and are classified in the Japanese-Language Proficiency Test word list. Second, 12 Chinese native speakers were surveyed using questionnaires (with a rating score of 0–100%) about the phonetic approximation between *on'yomi* in each of these 1,026 *kanji* and Chinese reading in China. In this survey, no noticeable trends were identified between each rating score and the four Chinese tones. However, the statistical analysis of *pinyin* decomposed into vowels and consonants revealed the following: (a) the rating score was high for characters beginning with a vowel; (b) the score was high for characters that contain apical consonants (a mean value of 32.6) or labial consonants (25.8); and (c) almost all characters that contain retroflex consonants (in particular, "zh" [mean 7.5], "r" [7.5], and "ch" [10.6]) were identified as having no phonetic similarity with their Japanese counterparts.

Keywords: Japanese, *Kyōiku Kanji*, Chinese, Phonetic, Logographic, Cross-Comprehension, Simultaneous Learning

iafor

The International Academic Forum
www.iafor.org

1. Introduction

1.1 Background

The number of language learners taking Japanese and Chinese language proficiency tests, JLPT (*Nihongo Nôryoku Shiken*) for Japanese, and HSK¹ (*Hanyu Shuiping Kaoshi*) for Chinese, has been steadily increasing (see Figure 1 below), as has the number of learners who are passing the tests.²

	2012	2013	2014	2015	2016	2017
JLPT	572169	571075	594682	652519	755802	887380
HSK	166313	189691	236403	365409	407479	470807
Total	738482	760766	831085	1017928	1163281	1358187

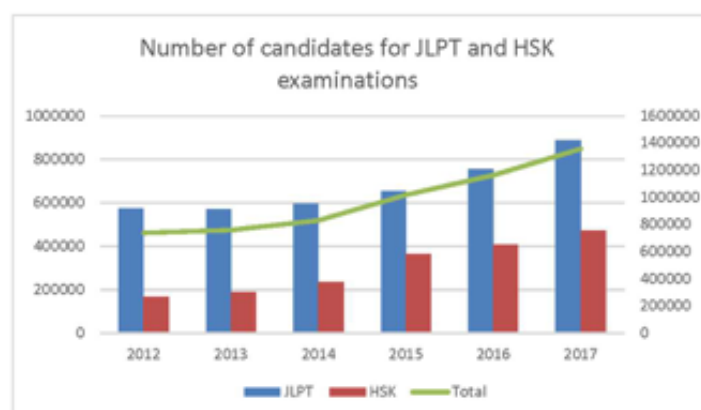


Figure 1. *Number of candidates for JLPT and HSK* (Obataya 2019, pp.1–2)

The increased interest in these two East Asian languages has encouraged students at Geneva University, who must choose two majors for their bachelor's degree, to take two Asian languages as their primary subjects. According to a survey conducted in 2012, students experienced difficulties in simultaneously learning Japanese and Chinese without any prior knowledge of them. However, this has not precluded a number of students from choosing Japanese and Chinese as their main subjects even today³. Between 2010 and 2018, about 20% of the students, on average, chose Japanese and Chinese. While the research previously had focused only on the graphic resemblances between these two languages,⁴ in 2018, a study was conducted of a database based on the characters required to pass each language's proficiency test: the JLPT and HSK.

The main objective of these studies was initially to integrate a systematic contrast of Japanese kanji and simplified Chinese characters into our teaching methods and materials in an effort to minimize mutual interferences.

¹ For the purpose of this series of study, the author took the HSK exams and passed with a proficiency level of HSK 6 – the highest level – in 2014.

² In comparison, the number of people who participated in the DELF/DALF exams was 389,120 in 2015.

³ Berger C., & Obataya Y. (2014).

⁴ Obataya (2018), p. 2.

1.2 Research Aim and Objectives

According to my previous studies on the same database, the phonetic differences between the Japanese and Chinese languages required further study. However, comparative studies on phonetics are still scarce (for example, Kayamoto 1995, Gi 2017, and Obataya 2019). In addition, there is still room in my previous work for a more detailed analysis. Therefore, the study aim is as follows:

- to conduct a more detailed quantification of the resemblance between the two languages.

To analyze the degree of resemblance in more detail, the Chinese *pinyin* were broken down into vowels and consonants. In addition, by using the recently renewed Japanese official list of *kanji*, more universal data had to be collected to fulfill the study aim. Therefore, the study objectives are as follows:

- to modify the database for collecting the recent data
- to examine the new database by breaking down the *pinyin* into vowels and consonants
- to verify whether the phonetic resemblance between Japanese and Chinese has a relationship with the Chinese four tones.

2. The Difficulties of Learning Japanese and Chinese Simultaneously and the Introduction of Chinese Characters to the Japanese Writing System

One of the difficulties students face when learning these two languages simultaneously is the complexity of the Chinese ideograms used in both languages. For example, Chinese ideograms used in Japanese – called “*kanji*” – sometimes differ in shape, meaning, and pronunciation from simplified Chinese characters in the People’s Republic of China (PRC). Since the pronunciation is different between Japanese and Chinese, *kanji* can be read by Japanese speakers in the Japanese way (“*kun’yomi*,” henceforth referred to as *kun*-reading) or in the (original) Chinese way (“*on’yomi*,” henceforth referred to as *on*-reading). Furthermore, as shown in Figure 2, a *kanji* sometimes has two, three, or more *on*-readings (*go-on*-reading, *kan-on*-reading, *tô-on*-reading, and other *on*-readings), because *kanji* phonetics have evolved differently over centuries (depending on epoch and region).⁵

	呉音 <i>Go-on</i>	漢音 <i>Kan-on</i>	唐音 <i>Tô-on</i> (宋音 <i>Sô-on</i>)
京	きょう <i>kyô</i> (東京)	けい <i>kei</i> (京城)	きん <i>kin</i> (南京)
経	きょう <i>kyô</i> (純経)	けい <i>kei</i> (経済)	きん <i>kin</i> (看経)
行	ぎょう <i>gyô</i> (行列)	こう <i>kô</i> (行動)	あん <i>an</i> (行宮)
外	げ <i>ge</i> (外科)	がい <i>gai</i> (外交)	うい <i>ui</i> (外郎)
頭	ず <i>zu</i> (頭痛)	とう <i>tô</i> (没頭)	じゅう <i>jû</i> (饅頭)
明	みょう <i>myô</i> (明年)	めい <i>mei</i> (名月)	みん <i>min</i> (明朝)

Figure 2. Examples of *kanji* with different *on-readings* types (Yano, 2012, 42/Obataya, 2019).

⁵ In addition, a simplification of the Chinese characters in the PRC took place in the 1950s, following the Japanese government’s simplification in 1926. These simplifications resulted in the lists of *Tôyô kanji* in 1946 and *Jôyô kanji* in 1981.

3. The Introduction of Chinese Characters to the Japanese Writing System

Contemporary Japanese has three main graphic systems: *kanji*, Chinese ideographs used in Japanese, and two syllabaries (*hiragana* and *katakana*).⁶ In order to better understand the relationship between Japanese *kanji* and Chinese characters, it is necessary to review the history of these three graphic systems.

The Japanese, having no writing system of their own, imported Chinese characters, or *kanji*, as early as the 4th or 5th century. Some three centuries later, a cursive and simplified form of *kanji* appeared, chosen for its phonetic value: the *manyōgana*, the name of which was derived from an anthology of poetry written using this simplified *kanji* during the *Nara* era (710–794) entitled “Collection of Ten Thousand Leaves,” or *Manyōshū*. *Manyōgana* would give rise to both the *hiragana* and *katakana* syllabary in the 9th century. Its current form, however, became fixed in the 12th century. Figure 3 provides examples of the three writing systems in Japanese.



Figure 3: *Three writing systems in Japan*⁷ (Obataya, 2019).

4. Previous Studies on the Graphic, Semantic and Phonetic Resemblance

The analysis of three aspects – the graphic, semantic, and phonetic – of another database has been undertaken in my three previous works, respectively Obataya (2018a), Obataya (2018b), and Obataya (2019).

An analysis of the degree of mutual graphic similarity indicated that 71% of Chinese characters are identical in both languages. After graphic resemblance analysis, I carried out a semantic resemblance analysis, which revealed that 89% of the characters are identical or only show slight variation. Finally, the result of the analysis of phonetic resemblance was only 12% (see Figure 4). Such clarification of the three values of graphic, semantic, and phonetic resemblance helps to explain the “*kanji* paradox,” a notion applied in one of our previous works. This notion explains that the simultaneous acquisition of Chinese and Japanese is, in fact, difficult despite the assumption of easy cross-comprehension due to the common usage of Chinese ideograms.

⁶ This does not take into account the *romaji*, or Latin alphabet.

⁷ The document used for the annual “open campus” session at Geneva University for high school students.

The database used in my previous research was a commercial textbook for foreign learners of Japanese, namely, *Minna no Nihongo I* (translation, 2nd ed., 2013) and *Minna no Nihongo II* (translation, 2nd ed., 2015).

In the current study, a new database was created of the official *kanji* list for instruction that Japanese students learn in elementary school in order to compare it with its Chinese equivalent. This list is used in my new database not only because it is the official list but also because it was recently modified. The "*Kyōiku kanji* (literally “education *kanji*”)" is an official Japanese list of 1,026 *kanji*, revised in 2017 and implemented in Japanese elementary schools starting in the 2020 academic year (in April in Japan).⁸

Several studies have been conducted on the phonetic similarity between Japanese and Chinese words and characters using the ancient lists of *Kyōiku kanji* (see Obataya, 2019). However, research on this new list has not yet been undertaken.

6. Analytical Approach

In this study, the method of Obataya (2019) was adopted in order to quantify the phonetic resemblance between the current *Kyōiku kanji* and Chinese.

The scores calculated by the multiplication of two values were taken to be the real phonetic resemblance of Chinese ideograms utilized both in Japanese and Chinese. These two values are (1) the frequency of a *kanji* in *on*-reading and (2) the rate of the phonetic resemblance of a *kanji* between *on*-reading and the Chinese pronunciation. To calculate the latter value, I distributed questionnaires concerning 1,026 *Kyōiku kanji* as a database to Chinese native speakers with considerable knowledge of the Japanese language.

Here is the summary of the profile of survey respondents:

- All 12 respondents were Chinese native speakers teaching Chinese to Japanese students.
- 9 respondents have acquired JLPT N1, and three have passed N2.

[c]	[e]	[g]	[h]	[o]	[k]	[f]
<i>Kyokasho-tai</i> font	<i>Kun</i> -reading	JLPT Word list	JLPT Word list (reading)	Calculation of points (JLPT levels)	Proportion of <i>On</i> -reading (%)	<i>On</i> -reading
社	やしろ yashiro				97	シャ sha
1 kaisha		会社	かいしゃ	5		
2 shakai		社会	しゃかい	4		
3 shachō		社長	しゃちょう	4		
4 jinja [jin+sha]		神社	じんじや	4		
5 shinbunsha		新聞社	しんぶんしゃ	4		
6 shakaikagaku		社会科学	しゃかいかがく	2		
7 shasetsu		社説	しゃせつ	2		
8 shōsha		商社	しょうしゃ	2		
9 nyūsha		入社	にゅうしゃ	2		
10 shakō		社交	しゃこう	1		
11 shataku		社宅	しゃたく	1		
12 shussha		出社	しゅつしゃ	1		
13				32		
14 yashiro		社	やしろ	1		
15				1		

Figure 6. Example of a calculation of a *kanji*.⁹

⁸ Japanese elementary schools follow a six-year system.

⁹ Obataya, 2019, 8.

7. The Calculation of the Frequency of *on*-Reading Usage

In this research, the values signifying the usages of *on*-reading in the target *kanji* data were identified. First, the usage of each *kanji* in words was investigated. Next, 9,292 words in the previous JLPT list¹⁰ were separated into three categories according to their pronunciations: *on*-reading, *kun*-reading, and others.

All five levels were counted differently to reflect the frequency of *on*-reading usage according to the levels of difficulty designated by the JLPT. For example, a word that contained a *Kyōiku kanji* in N1, the highest proficiency level in JLPT, was calculated as 1 point, whereas a word in N5, the easiest in the JLPT, was calculated as 5 points. It was assumed that the easier a *kanji* is, the more frequently it is used.

For example, with the *kanji* 社, twelve words use this *kanji* by pronouncing it as /sha/. All the scores of these words were counted according to the difficulty levels and added to the total score. There is only one word categorized in N1 that uses *kun*-reading, /yashiro/. Therefore, the score of *kun*-reading for this *kanji* was 1. The frequency of *on*-reading of this *kanji* was, therefore, found to be 97%.

8. Questionnaires Concerning Phonetic Resemblance between *Kyōiku kanji* and their Chinese Counterparts

Another important criterion for this study was the value of the phonetic resemblance of each *kanji* evaluated by questionnaires. For *kanji* that had already been collected in my previous study (Obataya, 2019), these data were used. For about 23% of *kanji* pairs (247 pairs), a new questionnaire was administered under the same conditions, and the total was calculated.

¹⁰ In the previous JLPT (= Old Japanese Language Proficiency Test Levels 1–4), before the renewal of the current JLPT (= Levels N1 to N5) in 2010, “Test Content Specifications” was published to help the candidates (The Japan Foundation and Japan Educational Exchanges and Services, 2007). This book contained a word list, *kanji* list, and grammar list, respectively classified into four different levels. The current JLPT does not publish such a book; therefore, to determine the new JLPT levels for the current work, I referenced the Jisho.org website and Jonathan Waller’s JLPT Resources page. According to Waller (<http://www.tanos.co.uk/jlpt/aboutjlpt/>), the new JLPT N1 is equivalent to the old JLPT 1, JLPT N2 to the old JLPT N2, JLPT N3 to halfway between the old JLPT 2 and JLPT 3, JLPT N4 to the old JLPT 3, and JLPT N5 to the old JLPT4.

	JP	Pinyin	On-reading
1	人	ren2	ジン
1	人	ren2	ニン
2	先	xian1	セン
3	生	sheng1	セイ
3	生	sheng1	ショウ
4	学	xue2	ガク
5	会	hui4	カイ
6	社	she4	シャ
7	行	hang2	コウ
7	行	hang2	ギョウ
7	行	xing2	コウ
7	行	xing2	ギョウ
8	大	da4	ダイ
8	大	da4	タイ
9	何	he2	カ

Figure 8. Example of the plural on-reading or Chinese pronunciations (Obataya 2019, 10).

9. Evaluation of 1,084 Pairs of *kanji* for Phonetic Resemblance

The number of characters in the renewed *Kyōiku kanji* list is 1,026, but this questionnaire evaluated 1,084 characters. In the case of plural *on*-readings or Chinese pronunciations existing in one *kanji*, respondents evaluated them separately. As shown in Figure 8, a *kanji* 人 has two ways of *on*-reading /jin/ and /nin/ and one Chinese pronunciation /ren2/. In this case, two comparative evaluations between (1) /jin/ and /ren2/ and (2) /nin/ and /ren2/ were conducted. Another example is 行, which has two ways of *on*-reading /kou/ and /gyou/ and two Chinese pronunciations /hang2/ and /xing2/. In this case, the respondents assessed the phonetic similarity in four pairs independently.

10. Findings of the two Statistical Analyses

10.1 Findings from the Frequency of the *on*-Reading Usage

The following represents the findings from the frequency of *on*-reading usage:

- The rate of frequency was 66.0%.
- 297 *kanji* (29% of the total number of the data) only have *on*-reading pronunciations.
- 7 *kanji* only have *kun*-reading pronunciations.

10.2 Findings from the Survey on Phonetic Resemblance by Means of Questionnaires

The following represents the findings from the survey on phonetic resemblance by means of questionnaires:

- The mean value for the phonetic resemblance from the survey was 19.8%.¹¹
- The number of *kanji* with complete correspondence between Japanese and Chinese was minimal (n=6 *kanji*, 医他愛衣信 and 因).
- The rate of *kanji* with more than 90% resemblance was only 3.3% (n=34).¹²

¹¹ Compared with the average of the survey outcome from Kayamoto (1995), the result was slightly lower. (Kayamoto's mean value was 34%, 2.38/7)

¹² These 34 *kanji* are 医他愛衣信因心理意部付利引新父来打負印府富民婦夫林飲苦流在異太里移臨.

- There was no resemblance for 16% of *kanji* (n=165).

Furthermore, the ones with a less than 10% resemblance made up almost half of the data (n=462). Chart 1 shows the mean and standard deviation.

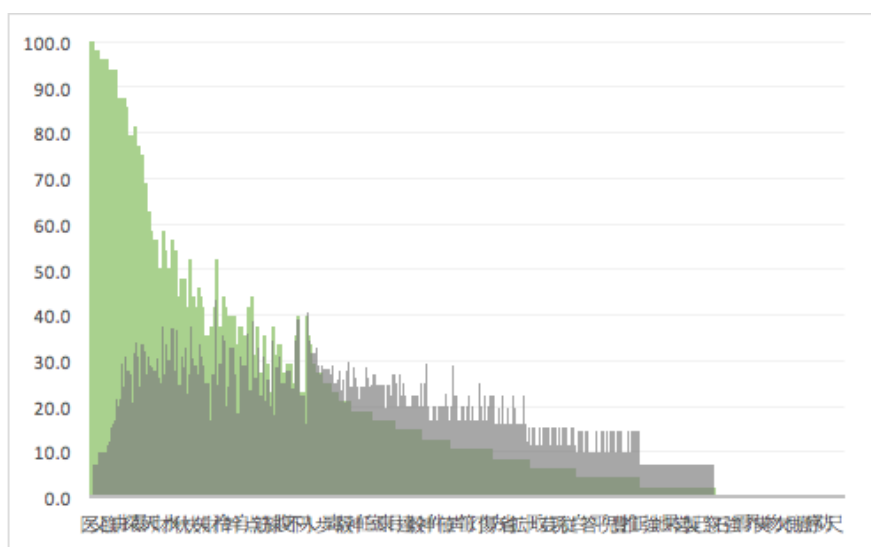


Chart 1. *Distribution of the mean and the standard deviation.*¹³

Chart 2, which clarifies this dissimilarity, illustrates the distribution of the number of *kanji* in the mean value range, showing the number of *kanji* on the vertical axis and the range of the mean value on the horizontal axis. As can be seen, the major dissimilarity is pictured on the left, and the minor similarity is pictured on the right of the chart.

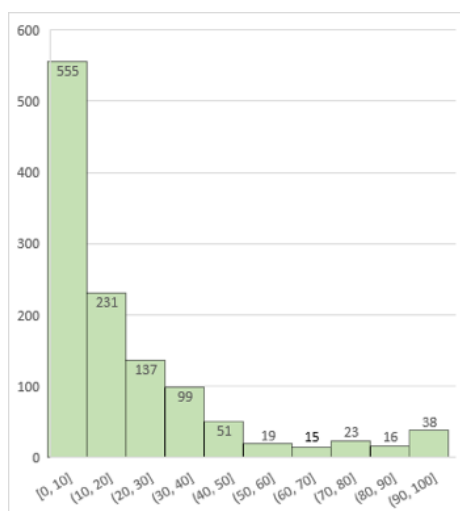


Chart 2. *Distribution of the number of averages.*

The previous survey indicated a possible correlation with the four Chinese tones since “most of the *kanji* with perfect correspondence were first and fourth [Chinese] tones.”¹⁴ Therefore, the potential correlation between the four Chinese tones and resemblances was investigated. However, the means and standard deviations were similar, and there was no noticeable trend.

¹³ The *kanji* on the bottom are representative at each stage.

¹⁴ The result of the current work was almost the same: 医(yi1)、他(ta1)、愛(ai4)、衣(yi1)、信(xin4) and 因(yin1).

The following represent the values for each tone

	Numbers	Mean	Standard deviation
1 st tone:	242	18.54	15.01
2 nd tone:	278	17.90	14.36
3 rd tone:	200	20.45	15.81
4 th tone:	456	20.40	15.69

11. Findings from the Total Scores by Multiplying the Values of (1) and (2)

The following is the summary of the “Findings from the total scores by multiplying the values of (1) and (2)”:

- The score of multiplication of the values from the analysis of (1) and (2) was 13.8%.
- Only three *kanji* (医愛信) have 100% frequency of usage of *on*-reading as well as 100% phonetic similarity.

Chart 3 is the distribution graph. This chart shows a strikingly high rate of dissimilarity between Japanese and Chinese.

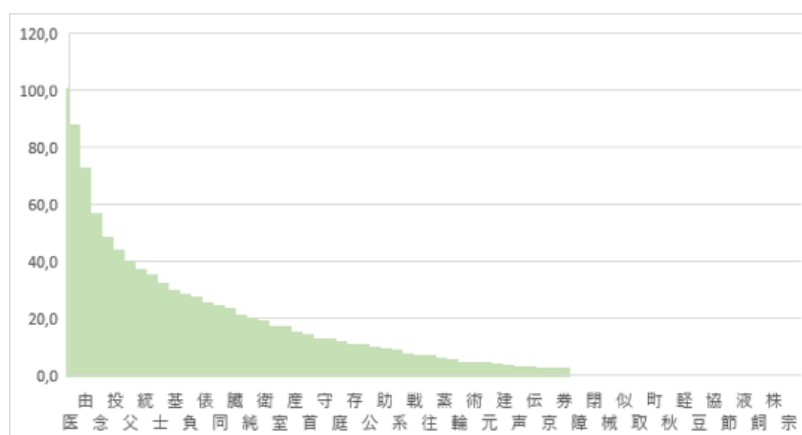


Chart 3. “Real” phonetic resemblance degree: $(1) \times (2)$.¹⁵

12. The Analysis is Broken Down into Vowels and Consonants

Next, the Chinese *pinyin* was broken down into vowels and consonants, and all the Chinese characters in our database were analyzed. For example, the characters with a pronunciation that begins with /zh/ have low phonetic similarity to Japanese (about 7.5 on average).

¹⁵ The *kanji* on the bottom are representative at each stage.

Serial number	JP	Pinyin	Four tones	On'yomi	2	3	4	6	7	8	9	10	11	1	5	12	Means	Standard deviation	Av ± σ	Av - σ	Av + σ
126	側	zhai1	1	ソク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
121	宅	zhai2	2	タク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
140	張	zhang1	1	チャウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
109	長	zhang3	3	チャウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
128	帳	zhang4	4	チャウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
124	着	zhao2	2	チャク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
124	着	zhe0	0	チャク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
115	折	zhe2	2	セツ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
122	爭	zheng1	1	ソウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
113	政	zheng4	4	セイ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
43	正	zheng4	4	ショウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
102	証	zheng4	4	ショウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
109	織	zhi1	1	ショク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
112	直	zhi2	2	ショク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
97	植	zhi2	2	ショク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
110	職	zhi2	2	ショク	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
92	質	zhi4	4	シツ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
94	種	zhong	3	シュ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
94	種	zhong4	4	シュ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
135	仲	zhong4	4	チュウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
113	裝	zhuang1	1	ショウ	0	0	0	0	0	0	0	0	0	0	0	0	0,0	0	0±0	0	0
138	展	zhan3	3	テン	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
92	障	zhang4	4	ショウ	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
105	整	zheng3	3	セイ	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
43	正	zheng4	4	セイ	0	0	0	0	0	25	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
109	織	zhi1	1	シキ	0	0	0	0	0	0	25	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
112	直	zhi2	2	ジキ	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
92	質	zhi4	4	シチ	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
116	製	zhi4	4	セイ	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
64	至	zhi4	4	シ	0	0	0	0	0	0	0	0	0	25	0	0	2,1	7,22	2,1±7,2	0	9,3
82	終	zhong1	1	シュウ	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
138	築	zhu2	2	チュク	0	0	0	0	0	0	25	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
96	祝	zhu4	4	シュク	0	0	0	0	25	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
139	貯	zhu4	4	チョ	0	0	0	0	0	0	0	0	0	25	0	0	2,1	7,22	2,1±7,2	0	9,3
104	專	zhuan1	1	セン	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
135	轉	zhuan3	3	テン	25	0	0	0	0	0	0	0	0	0	0	0	2,1	7,22	2,1±7,2	0	9,3
77	札	zha2	2	サツ	0	0	0	25	0	0	25	0	0	0	0	0	4,2	9,73	4,2±9,7	0	13,9
74	者	zhe3	3	シャ	50	0	0	0	0	0	0	0	0	0	0	0	4,2	14,4	4,2±14,4	0	18,6
94	針	zhen1	1	シン	25	0	0	0	0	0	0	0	0	25	0	0	4,2	9,73	4,2±9,7	0	13,9
59	中	zhong1	1	ジュウ	0	0	0	0	0	25	0	25	0	0	0	0	4,2	9,73	4,2±9,7	0	13,9
59	中	zhong1	1	チュウ	0	0	0	0	0	0	0	0	0	0	0	50	4,2	14,4	4,2±14,4	0	18,6
95	周	zhou1	1	シュウ	25	0	0	0	0	0	0	0	0	25	0	0	4,2	9,73	4,2±9,7	0	13,9
87	諸	zhu1	1	ショ	0	0	0	0	0	0	0	0	0	25	0	25	4,2	9,73	4,2±9,7	0	13,9
130	著	zhu4	4	チョ	0	0	0	0	0	0	0	0	0	0	50	0	4,2	14,4	4,2±14,4	0	18,6
113	裝	zhuang1	1	ソウ	0	0	0	0	0	0	25	0	0	0	25	0	4,2	9,73	4,2±9,7	0	13,9
101	真	zhen1	1	シン	25	0	0	0	0	25	0	0	0	25	0	0	6,3	11,3	6,3±11,3	0	17,6
63	止	zhi1	1	シ	25	0	25	0	0	0	25	0	0	0	0	0	6,3	11,3	6,3±11,3	0	17,6
134	置	zhi4	4	チ	25	0	0	25	0	0	0	0	0	25	0	0	6,3	11,3	6,3±11,3	0	17,6
129	忠	zhong1	1	チュウ	0	25	0	0	0	0	25	0	0	0	25	0	6,3	11,3	6,3±11,3	0	17,6
79	衆	zhong4	4	シュ	0	0	25	0	0	0	25	0	0	0	25	0	6,3	11,3	6,3±11,3	0	17,6
101	招	zhao1	1	ショウ	0	0	0	0	0	0	0	0	25	25	0	50	8,3	16,3	8,3±16,3	0	24,6
111	制	zhi4	4	セイ	0	0	25	0	0	0	0	25	0	0	50	0	8,3	16,3	8,3±16,3	0	24,6
86	重	zhong4	4	チュウ	0	0	25	0	0	0	0	0	0	75	0	0	8,3	22,2	8,3±22,2	0	30,5
58	竹	zhu2	2	チュク	75	0	0	0	0	0	25	0	0	0	0	0	8,3	22,2	8,3±22,2	0	30,5
119	戰	zhan4	4	セン	25	0	0	25	0	0	25	0	0	0	0	50	10,4	16,7	10,4±16,7	0	27,1
111	朝	zhao1	1	チャウ	25	0	0	0	0	25	0	0	0	25	0	50	10,4	16,7	10,4±16,7	0	27,1
111	朝	zhao1	1	チャウ	25	0	0	0	0	25	0	0	0	25	0	50	10,4	16,7	10,4±16,7	0	27,1
103	照	zhao4	4	ショウ	25	0	0	0	0	0	0	25	25	50	0	0	10,4	16,7	10,4±16,7	0	27,1
93	蒸	zheng1	1	ジョウ	50	0	25	0	0	0	0	0	0	0	50	0	10,4	19,8	10,4±19,8	0	30,2
68	紙	zhi3	3	シ	75	0	0	0	0	0	25	0	0	0	25	0	10,4	22,5	10,4±22,5	0	32,9
65	指	zhi3	3	シ	25	0	25	25	0	0	25	0	0	0	25	0	10,4	12,9	10,4±12,9	0	23,3
69	誌	zhi4	4	シ	50	0	0	0	0	25	25	0	0	25	0	0	10,4	16,7	10,4±16,7	0	27,1
79	衆	zhong4	4	シュウ	0	0	25	0	0	0	25	25	0	0	50	0	10,4	16,7	10,4±16,7	0	27,1
80	州	zhou1	1	シュウ	25	0	0	0	0	25	25	0	0	0	0	50	10,4	16,7	10,4±16,7	0	27,1
75	主	zhu3	3	シュ	25	0	0	50	0	0	25	0	0	25	0	0	10,4	16,7	10,4±16,7	0	27,1
106	狀	zhuang4	4	ジョウ	25	0	0	0	0	50	0	25	0	0	25	0	10,4	16,7	10,4±16,7	0	27,1
94	章	zhang1	1	ショウ	25	25	75	0	0	0	25	0	0	0	0	0	12,5	22,6	12,5±22,6	0	35,1
88	治	zhi4	4	チ	25	0	0	50	25	25	0	0	0	25	0	0	12,5	16,9	12,5±16,9	0	29,4
84	志	zhi4	4	シ	50	0	25	25	0	25	0	0	0	25	0	0	12,5	16,9	12,5±16,9	0	29,4
127	值	zhi4	4	チ	50	0	0	0	0	50	25	0	0	25	0	0	12,5	19,9	12,5±19,9	0	32,4
86	重	zhong4	4	ジュウ	0	0	0	0	0	50	0	0	0	0	25	75	12,5	25	12,5±25	0	37,5
128	宙	zhou4	4	チュウ	25	0	0	0	0	25	25	25	0	0	25	0	12,5	13,1	12,5±13,1	0	25,6
90	助	zhu4	4	ジョ	25	0	0	0	0	25	0	0	0	0	75	25	12,5	22,6	12,5±22,6	0	35,1
125	注	zhu4	4	チュウ	25	0	50	0	0	25	0	0	0	75	0	0	14,6	24,9	14,6±24,9	0	39,5
91	昭	zhao1	1	ショウ	25	0	0	0	0	25	0	25	0	50	0	75	16,7	24,6	16,7±24,6	0	41,3
82	支	zhi1	1	シ	25	0	50	25	0	25	50	0	0	0	25	0	16,7	19,5	16,7±19,5	0	36,1
77	遇	zhou1	1	シュウ	25	0	0	0	0	25	50	0	0	50	0	50	16,7	22,2	16,7±22,2	0	38,9
130	追	zhui1	1	ツイ	25	0	25	25	0	25	0	0	0	50	25	25	16,7	16,3	16,7±16,3	0,4	33
85	枝	zhi1	1	シ	25	50	25	25	0	0	25	0	0</								

Serial number	JP	Pinyin	Four tones	On'yomi	Means
126	側	zhai1	1	ソク	0
121	宅	zhai2	2	タク	0
140	張	zhang1	1	チョウ	0
109	長	zhang3	3	チョウ	0
128	帳	zhang4	4	チョウ	0
124	着	zhao2	2	チャク	0
124	着	zhe0	0	チャク	0
115	折	zhe2	2	セツ	0
122	争	zheng1	1	ソウ	0
113	政	zheng4	4	セイ	0

Serial number	JP	Pinyin	Four tones	On'yomi	Means
43	正	zheng4	4	ショウ	0
102	証	zheng4	4	ショウ	0
109	織	zhi1	1	ショク	0
112	直	zhi2	2	チョク	0
97	植	zhi2	2	ショク	0
110	職	zhi2	2	ショク	0
92	質	zhi4	4	シツ	0
94	種	zhong	3	シュ	0
94	種	zhong4	4	シュ	0
135	仲	zhong4	4	チュウ	0

Figure 10. Twenty characters beginning with “zh” that have no similarity at all.

More precisely, Figure 8 shows the characters with Chinese *pinyin* that begins with “zh.” There were 86 characters beginning with “zh.” Of these, 20 *kanji* were judged to have no similarity at all (0% [Figure 10]).

The results of the statistical analysis revealed the following:

- The average score was higher for characters beginning with a vowel.
- The average was higher for characters that contain apical consonants [mean value 32.6] or labial consonants [25.8].
- Almost all characters that contain retroflex consonants (in particular, /zh/ [mean 7.5], /r/ [7.5], and /ch/[10.6]) have no phonetic similarity with their Japanese counterparts.

The following table shows the mean and standard deviation for each of the 21 consonants.

consonnes	子音グループ	平均値	標準偏差
l	舌尖音 apical consonant	38,28627451	19,89579777
k	舌根音 radical consonant	37,17894737	21,56495511
f	唇音 labial (sound)	35,77804878	15,87978845
n	舌尖音 apical consonant	34,70434783	18,64720862
t	舌尖音 apical consonant	33,19767442	20,82330342
s	舌歯音 linguadental	31,34583333	22,93881481
d	舌尖音 apical consonant	24,31594203	17,8003113
m	唇音 labial (sound)	24,20454545	18,8237082
b	唇音 labial (sound)	22,59090909	18,27161936
p	唇音 labial (sound)	20,72777778	21,85060758
z	舌歯音 linguadental	17,60232558	17,84275477
c	舌歯音 linguadental	17,48387097	17,84327541
sh	舌上音 retroflex consonant	17,35483871	18,11704646
x	舌面音 dorsal consonants	14,64722222	10,53723287
g	舌根音 radical consonant	12,052	15,18573507
ch	舌上音 retroflex consonant	10,62325581	12,4953234
q	舌面音 dorsal consonants	10,33913043	12,22004874
j	舌面音 dorsal consonants	8,871428571	10,62687754
r	舌上音 retroflex consonant	7,51	11,23613907
zh	舌上音 retroflex consonant	7,508235294	11,59644148
h	舌根音 radical consonant	7,48627451	9,976812952

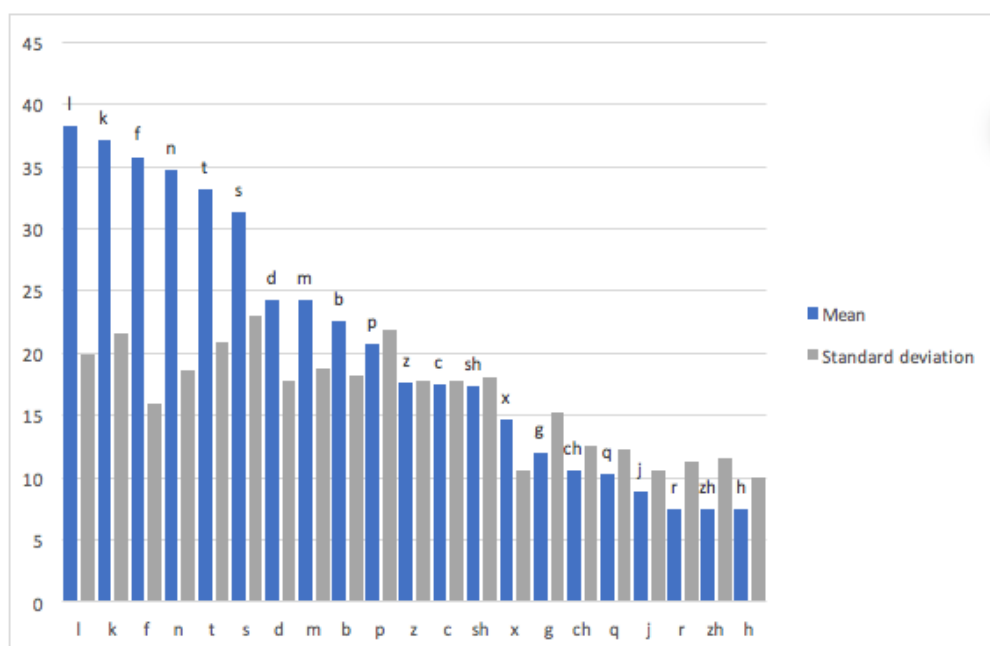


Figure 11. *Phonetic resemblance degree by 21 consonants:*
(l, k, f, n, t, s, d, m, b, p, z, c, sh, x, g, ch, q, j, r, zh, and h) .

As a Japanese person who took and passed HSK6, the highest level of the Chinese test, I knew from experience that it is challenging to pronounce and understand the retroflex consonants. For example, "Japanese people" is pronounced as "*nihonjin*" in Japanese, but in Chinese, it is pronounced as "*ri4 ben3 ren2*," despite the fact that the Chinese characters are the same. Since the retroflex consonant "r" does not exist in Japanese, it is quite difficult to pronounce and understand. However, the way in which the current work quantifies the phonetic resemblance will make it easier to identify the sounds that need to be focused on during language training.

13. The Advantages of *kanji* for Improving Cross-Comprehension of Japanese and Chinese

13.1 Prescription to Prevent Learning Detours or Fixation on Mispronunciation

According to our analysis, the final score of the phonetic resemblance was 13.8%. Despite this low rate, it is unnecessary to reject attempts at simultaneous learning or cross-comprehension of Japanese and Chinese. There are considerable advantages to the resemblance of forms and meanings in the common *kanji* (see Figure 4). By informing students during the early stages of learning about the imbalanced rates of resemblance among the three components – sounds, forms, and meanings – it is possible to improve the efficacy of their study and to prevent learning detours. Moreover, following the analysis of the consonants, Japanese speakers should focus on the practice of retroflex consonants during the earliest stages of learning.

13.2 From “*kanji* Paradox” to Unexpected “*kanji*-Sphere Bonus”

The aforementioned concept of the “*kanji* paradox” entails a somewhat negative view of the languages. However, based on the quantification of my research, this notion might be reframed in a positive way, that is, as “*kanji* advantages” or “*kanji*-sphere bonus.”

Even if Japanese people have never studied modern Chinese, they can infer the meaning to some extent because they know many Chinese ideograms. The reverse holds true as well. This view of languages is useful in disproving the stereotype that we are studying only one language. Of course, in terms of European languages, this is not a new concept. In fact, European languages are interconnected with Latin, Greek, and other languages on many levels, such that many elements in one language can be used to learn other languages.¹⁶ Edgard Pich sums up this notion of languages well: “You are not studying one language. You are learning several languages at the same time.”¹⁷ This view of language is useful in disproving the stereotype that we are studying only one language and can be applied as well to Asian languages, in this case, Japanese and Chinese.

14. Discussion

14.1 General Discussion

This study supports the findings of Obataya (2019, p.15), which showed the importance of developing special learning strategies for students studying *kanji* in order to acquire listening and speaking skills. By using the JLPT levels and the deviation value, the quantification of the current database will be beneficial for all foreign learners of the Japanese and Chinese languages.

14.2 Limitations and Recommendations for Future Research

There are some limitations to this study. First, there is still room for further analysis of this database. The graphic and semantic aspects of the current *Kyōiku kanji* should be analyzed in

¹⁶ For example, the adjectival form of the English word “water” is “aquatic.” The Latin word *aqua*, from which this word originated, is used as the Italian word for water. (The spelling is slightly different.)

¹⁷ When I was a graduate student, I studied French with Professor Emeritus Edgar Pich (1938–) of the University of Lyon II Lumière in France. These phrases were heard during his French lessons.

future studies to address these limitations.¹⁸ Second, as an extension of the *Kyōiku kanji*, it would be worth examining the degree of resemblance to the *Jōyō kanji* (about 2,000 characters) through the same method.

14.3 Theoretical and Practical Implications

The findings have several theoretical and practical implications. As for the theoretical implications, this study confirmed that through the static analysis of *pinyin* decomposed into vowels and consonants, the phonetic resemblance between Japanese and Chinese languages could be quantified in further detail. As to the practical implications, according to our analysis, teachers could advise Japanese speakers who are learning Chinese to focus on the practice of retroflex consonants at the beginner's stage of learning.

15. Conclusions

15.1 Review of the Study Aim and Objectives

The interest in Japanese and in Chinese languages is increasing not only among our students but also across the world. However, few works have analyzed the phonetic aspect of the latest character lists between these two languages. Therefore, it was necessary to develop an effective teaching method for the phonetic cross-comprehension of these two East Asian languages. Thus, the study aimed to conduct a more detailed quantification of the resemblance between the two languages.

After creating the database on the renewed official *kanji* list, the phonetic resemblance between Japanese and Chinese on this database was analyzed. Moreover, by breaking down the Chinese *pinyin* into vowels and consonants, the resemblance degree was examined in more detail.

15.2 Synthesis of the Main Findings

The main findings are divided into two parts: the quantification of the phonetic resemblance by means of a questionnaire and the analysis decomposed into vowels and consonants. First, it was statistically confirmed that each character is pronounced differently regarding the attempt to quantify the phonetic resemblance. Second, the average was higher for characters that contain apical consonants or labial consonants. Almost all characters containing retroflex consonants (in particular, /zh/, /r/, and /ch/) have no phonetic similarity with their Japanese counterparts.

Acknowledgments

I would like to express my special gratitude to Pierre François Souyri, Professor Emeritus at the Geneva University, for allowing me to take charge of all the modules of *kanji*. Without his help, it would have been impossible for me to collect the data for this study. I am also immensely grateful to the teachers of 天天中文 (<https://ttn.co.jp/>) for their cooperation with the questionnaire and Ikuo Obataya (Quantum Design Japan) for his help with measurements

¹⁸ As for form and meaning, we analyzed the previous database but not the current one. However, 70% of the data in the previous and current databases are the same. Thus, we can say that the trend is the same regarding form and meaning.

and analysis. Finally, I would like to thank the administrators of Jisho.org and Jonathan Waller, creator of the JLPT resources pages, for determining the JLPT levels in my database.

References

- 3A Network. (2013). *Minna no Nihongo, Book 1 Translation & Grammatical Notes, French version* (2nd ed.). Tokyo, Japan: 3A Network.
- 3A Network. (2015). *Minna no Nihongo, Book 2 Translation & Grammatical Notes, French version* (2nd ed.). Tokyo, Japan: 3A Network.
- Berger, C., & Obataya, Y. (2012). Simultaneously Learning Chinese and Japanese as Third Languages: Synergy or Futility? *The Teaching of Asian Languages in the 21st Century*. Paper Presentation at the International Symposium on Asian Languages and Literatures. Kayseri, Turkey.
- Berger, C., & Obataya, Y. (2014). Simultaneously Learning Chinese and Japanese as Third Languages: Synergy or Futility? In Ali Küçükler and Hüseyin İçen (Eds.), *The Studies on the Teaching of Asian Languages in the 21st Century*. Cambridge, England: Cambridge Scholars.
- Gi, N. (2017). Similarity between Japanese and Chinese vocabulary in *kanji* - Focusing on Phonological Similarity – (*Nicchû kanji goi no ruijisei ni tsuite – onyinteki ruiji do o chûshin ni*). *Journal of Japanese as a Second Language Kanji Research Group*, 9(1), 62-67.
- Kayamoto, Y. (1995). A Survey on the Similarity between Chinese and Japanese readings in the same *kanji* (*Dôitsu kanji ni okeru Chûgoku go on to Nihon go no on'yomi no ruiji do ni kansuru chôsa*). *Bulletin of the Department of Teaching Japanese as a Second Language*, Hiroshima University, 5, pp.67-75.
- Kume, Isao. *Kanjishido no tebiki Gakushu shidô yōryō jyunkyo* [A Guide to Teaching Kanji - Study Guideline Compliance], (8th ed.). Tokyo, Japan: Kyōiku shuppan, 2017.
- Obataya, Y. (2018a). A Study on the Mutual Similarity between Japanese and Chinese for Simultaneous Learning, in the Asian Conference on Education & International Development (ACEID2018). Kôbe, Japan.
- Obataya, Y. (2018b). “Une étude des similarités graphique et sémantique du japonais et du chinois pour un apprentissage simultané” (Graphic and Semantic Similarities of Japanese and Chinese during Simultaneous Learning), [oral presentation in French]. International Symposium on ADLES 2018. Lausanne, Switzerland.
- Obataya, Y. (2019). A Study of the Mutual Phonetic Resemblance between Japanese and Chinese: Quantification of the Difficulty of Phonetic Cross-comprehension, in the Asian Conference on Education (ACE2019). Tokyo, Japan.
- The Japan Foundation and Japan Educational Exchanges and Services. (2007). *Japanese-Language Proficiency Test: Test Content Specifications* (Revised Edition). Tokyo, Japan: Bonjinsha.
- Yano, M. (2012). *Chinese characters culture and Japan (Chûgoku kanji bunka to Nihon)*. Tokyo, Japan: Surugadai shuppansha.

Yoshida, M. (2014). Comparison of Chinese character teaching between Japan and China – notes in Chinese character teaching (*Nicchû kanji kyôiku hikaku: kanji shidô ni okeru chûiten*). *Journal of Japanese as a Second Language Kanji Research Group*, 6(1), 18-24.

Internet sources

Jōyō kanji [Regular-use Chinese characters]

https://www.bunka.go.jp/kokugo_nihongo/sisaku/joho/joho/kijun/naikaku/kanji/ [cit. 2021-01-17]

Websites for referring to the JLPT levels of a *kanji* or a word: <https://jisho.org/> [cit. 2021-01-17] <http://www.tanos.co.uk/jlpt/> [cit. 2021-01-17]

Contact email: Yuji.Obataya@unige.ch