A Study of the Mutual Phonetic Resemblance Between Japanese and Chinese: Quantification of the Difficulty of Phonetic Cross-Comprehension

Yuji Obataya, University of Geneva, Switzerland

The Asian Conference on Education 2019 Official Conference Proceedings

Abstract

This study investigated the mutual phonetic resemblance of Chinese ideograms between Japanese and Chinese by using a database of 1078 kanji (Chinese ideograms in Japanese) extracted from the two volumes of the Japanese grammar textbook (Second Edition, 2013 and 2015) used in the department of East Asian Studies at University of Geneva. The initial aim of this analysis is to help learners from non-kanji backgrounds to study simultaneously these two languages. Firstly, since most of kanji have multiple readings (on 'yomi [Chinese reading] and kun'yomi [Japanese reading]), the rate of use of on'yomi in each 1078 kanji was calculated :59.72% [1], by taking into account the factor of frequency of all words (total 9233 words) who contain these kanji and are classified in the JLPT word list. Secondly, the basic phonetic resemblance has be figured out at 19.6% [2] according to the result of survey of twelve¹ Chinese native speakers, who teach Chinese to Japanese people, nine of which passed N1 and three of which, N2 level of Japanese-Language Proficiency Test (JLPT). In comparison with the shape resemblance (71%), semantic resemblance (about 90 %) in the same database, analyzed in my former two studies (OBATAYA 2018a, 2018b), this relative low rate of phonetic similarity (12.4 %, judging from these values [1] and [2]) visualize significantly a gap between the "interdependence" of the shape and the meaning aspects and the "independence" of the phonetic aspect, and emphasize the importance of phonetic crosscomprehension for learners of these two East Asian languages.

Keywords: Asian languages, Japanese, Chinese, phonetic, logographic, cross-comprehension, simultaneous learning

iafor

The International Academic Forum www.iafor.org

¹ I should add that when I wrote the abstract, the number of respondents to the survey was 9, but 3 joined after that. In the end, the survey in the current study was completed by 12 respondents.

Introduction

This study investigates the mutual phonetic resemblance in Chinese ideograms between Japanese and Chinese by using a database of 1,078 *kanji* (Chinese ideograms in Japanese) extracted from the two volumes of the Japanese grammar textbook, *Minna no Nihongo* (Second Edition, 2013 and 2015) used in the department of East Asian Studies at University of Geneva.

The initial aim of this analysis was to help the students of Japanese (or learners of Chinese, or both simultaneously) at University of Geneva to learn how to read Chinese characters and pronounce them correctly. It is hoped that this study may be useful not only for the students at University of Geneva but also for simultaneous Japanese and Chinese learners in non-*kanji* areas. For instance, it can be effective for beginner Chinese learners who have prior knowledge of Japanese.

For this purpose, my previous research studies were based on the characters required to pass each language's proficiency test: the Japanese-Language Proficiency Test (JLPT) for Japanese and the *Hanyu Shuiping Kaoshi* (HSK) Chinese Proficiency Test for Chinese.² Figure 1 illustrates the approximate correspondence of the proficiency levels between the JLPT and the HSK.



Figure 1: Correlative table between JLPT and HSK proficiency levels.

The increasing number of candidates for both the JLPT and the HSK reflects the global interest in these two East Asian languages; in 2013, a total of 571,075 examinees took the JLPT, while in 2017, that number rose to 887,380. For the HSK, a total of 189,691 examinees took the test in 2013, while in 2017, 470,807 took it.³

² For the purpose of these studies, I took the HSK exams and passed with a proficiency level of HSK 6—the highest level—in 2014.

³ In comparison, the number of people who participated in the DELF / DALF exams was 389,120 in 2015.





Figure 2: Number of candidates for JLPT and HSK⁴

The increased interest has encouraged approximately 20% of students at department of East Asian Studies at University of Geneva to select them as their two majors required to complete their bachelor's degree, even though a survey⁵ has revealed that learners from non-*kanji* backgrounds find learning Japanese and Chinese simultaneously to be a heavy burden.

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

One of the difficulties that students face in learning these two languages at the same time is the complexity of the Chinese ideograms used in both Japanese and Chinese. For example, the simplified form of Chinese ideograms used in Japanese—called *kanji*—sometimes differ in shape, meaning, and pronunciation from the simplified Chinese characters as used in the People's Republic of China (Yoshida, 2014, p.19).⁶

Contemporary Japanese has three main graphic systems: *kanji*, Chinese ideographs used in Japanese, and two syllabaries (hiragana and katakana).⁷ In order to better understand the relationship between Japanese *kanji* and Chinese characters, it is necessary to review the history of the three graphic systems of Japanese.

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 *manyôgana*, the name of which was derived

⁴ The JLPT's numbers are quoted from the JLPT's homepage (https://www.jlpt.jp/e/statistics/archive.html). The HSK's numbers were taught by the Confucius Institute of Geneva University (2018.07.11).

⁵ Berger C., & Obataya Y (2014), pp.150-152.

⁶ This paragraph is taken from Obataya Y. (2018a), p.2.

⁷ This does not take into account the romaji, or Latin alphabet.

from the title of an anthology of poetry written using this simplified script during the Nara era (710-794), 'Collection of Ten Thousand Leaves', or *Manyôshû*. Manyôgana would give rise to the 9th century hiragana syllabary, as well as to the katakana syllabary, which was also created in the 9th century. Its current form, however, became fixed in the 12th century. Figures 3 and 4 show examples of the three writing systems as well as their proportion of use in Japanese.



Figure 3: Three writing systems in Japan⁸

Chinese Chinese characters 汉字								
Kanji	Hiragana	Katakana						
漢字 30 %	40%	30 %						

Figure 4: Approximate proportion of the use of the three systems in contemporary Japanese

Japanese continued to borrow Chinese characters according to their meaning; however, since the pronunciation differed between Japanese and Chinese, Japanese *kanji* could be read by Japanese speakers in the Japanese way (*'kun'yomi'*, henceforth *kun*-reading) as well as in the (original) Chinese way (*'on'yomi'*, henceforth *on*-reading). Furthermore, a *kanji* sometimes has two, or three or more types of *on*-reading (*go-on*-reading, *kan-on*-reading, and *tô-on*-reading, among others⁹), because *kanji*'s phonetics have evolved in various ways over the centuries in the different eras and regions from China.¹⁰

⁸ Document used for the annual "open campus" session at Geneva University for high school student

⁹ These different *on*-reading are indicated in my database for the analysis of the graphic field (Obataya, 2018a, 9th note).

¹⁰ **'Phonetic value of** *kanji*' (...) *Go-on* is considered to be the oldest *kanji* sound from China. *Go* is the name of the old State of Wu, which is the current area of Jiangsu Province (Nanjing). It broadly includes the Yangtze River area in South China and refers to the sounds that were used in that region. Japan traded in the 5th and 6th centuries, before the Nara era, with the

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

Figure 5: Examples of kanji that have different on-reading types (Yano, 2012, p.42)

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

These different points may cause great confusion for simultaneous learners, especially complete beginners, and are likely to disturb their learning of ideograms. Therefore, this study aims to support such students' effective learning by building a database that identifies the commonalities and differences between *kanji* and Chinese characters. The database comprises all 1,078 *kanji* that appear in the two volumes of the Japanese grammar textbook that the department of East Asian Studies at University of Geneva has adopted for instruction.

South of China and was also in a close relationship with the southern part of the Korean Peninsula. (...)

Kan-on is generally referred to as the sound of the region of Shenyang and Chang'an. From the Nara period to the beginning of the Heian period, exchanges with China became more and more active, and many foreign students were sent with the Japanese missions to Sui China or Tang China. These people learned the Shenyang and Chang'an pronunciation, both the capitals of China at the time, and brought it back to Japan. In the Heian period, *kan-on* was considered as the official pronunciation. (...)

[&]quot;*Tô* (*Tang* in current Chinese pronunciation)" in the word *Tô-on* indicate China (not only the *Tang* dynasty). From the 11th century until the Edo period, especially during the Kamakura and Muromachi period, it was a collective term for *kanji* sounds. These sounds were brought back from the Chinese mainland by Zen Buddhist monks, merchants and private diplomats. (...)" Yano (2012), pp. 40-42, translated by Obataya.

Previous Studies on the Graphic and Semantic Resemblance Between *Kanji* and Chinese Characters

The analysis of the graphic aspect and the semantic aspect of this database has been effectuated in Obataya (2018a) and Obataya (2018b), respectively.



Table 1: The number of *kanji* covered in the textbooks (I and II)¹¹

The first study demonstrated that the textbook not only covers 100% of the *kanji* of the JLPT N5 (the lowest level) and the N4 but also approximately 70% of the Chinese characters of the HSK1 (the lowest level) and HSK2, as well as 60% of the HSK3.

In addition, the results of an analysis of the degree of mutual similarity in the list indicated that 71% of Chinese characters are identical (or have only a slight difference) in the two languages.

After this analysis of graphical resemblance, we effectuated the analysis of semantic resemblance, with the results revealing that 89% of the characters are identical or have only a slight difference.

The current study aims to quantify this final field—namely, phonetic resemblance—by using the same database.

Previous Studies on Phonetic Resemblance

Several studies exist on the phonetic resemblance between Japanese and Chinese words and characters: Kayamoto (1995), Matsushita (2009), and Gi (2017). While the current study highlights 'parallel characters' in terms of phonetic resemblance, Matsushita (2009)¹² and Gi (2017) focused mainly on 'parallel words'.

¹¹ This is the 'Table 1' in Obataya (2018a) that was modified and some images were added.

¹² This study on the phonetic resemblance of Matsushita (2009) is based on the data provided by Kayamoto (1995).

Analytical Approach

In the current study, the scores calculated by the multiplication of two values are considered to be the real phonetic resemblance of Chinese ideograms utilised in both 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 pronunciation in Chinese. In order to calculate the latter value, I applied the method of Kayamoto (1995) by distributing questionnaires concerning the 1,078 *kanji* of my database to Chinese native speakers who have considerable knowledge of the Japanese language and calculating the mean values. In our study, we did not use audio recordings of Japanese or Chinese pronunciation (unlike Gi's study, which used audio recordings of both languages' pronunciation) by pointing to a website that contains the sound sources of Japanese pronunciation.

Calculation of the Frequency of On-Reading Usage

No previous study has taken into consideration the frequency of the *on*-reading usage of a *kanji*, which is one of the most important criteria in this work. A *kanji* is pronounced in various ways, including *on*-reading, *kun*-reading, and other ways of reading, such as an idiomatic way of pronunciation. In this research, the values that signify the usages of *on*-reading in the target *kanji* data are identified.

First, the usage of each *kanji* in words was investigated. Here, 9,233 words in the list of the previous JLPT¹³ were categorised into '*on*-reading', '*kun*-reading', and 'other', according to their pronunciation.

In order to reflect the frequency of *on*-reading usage according to the levels of difficulty designated in JLPT, all five levels were counted differently. For example, a *kanji* in N1, the highest proficiency level in JLPT, was calculated as one point, whereas a *kanji* in N5, the easiest in JLPT level, was calculated as five points. It was considered that the easier a *kanji* is, the more frequently it is used.

¹³ In the previous JLPT (= Old Japanese Language Proficiency Test Levels 1 to 4), before the renewal of the current JLPT (= Levels N1 to N5) in 2010, 'Test Content Specifications' was published for helping the candidates. This book contained a word list, a *kanji* list, and grammar list, respectively classified in four levels. The current JLPT does not publish such a book, so for comparing the new JLPT levels to current work, I reference the Jisho.org website, and Jonathan Waller's JLPT Resources page. According to Waller, New JLPT N1 is equivalent to the old JLPT 1; JLPT N2, to the old JLPT2; JLPT N3, to halfway between the old JLPT 2 and JLPT 3; JLPT N4, to the old JLPT 3; JLPT N5, to the old JLPT4.

[d]	[e]	[f]	[g]	[i]	[j]	[k]
JP Kyokashot ai font	Kun-reading	JLPT Word list	JLPT Word list (lecture)	Calculation of points (JLPT levels)	Proportion de On reading (%)	On-reading
社	やしろ				97	シヤ
	yashiro		0			sha
1	kaisha	会社	かいしゃ	5		
2	shakai	社会	しゃかい	4		
3	shachô	社長	しゃちょう	4		
4	jinja (jin+sha)	神社	じんじゃ	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	shôsha	商社	しょうしゃ	2		
9	nyûsha	入社	にゅうしゃ	2		
10	shakô	社交	しゃこう	1		
11	shataku	社宅	しゃたく	1		
12	shussha	出社	しゅっしゃ	1		
13				32		
14	yashiro	社	やしろ	1		
15				1		

Figure 6: Example of calculation of the on-reading usage

For example, with a *kanji*, \bigstar , there are 12 words that use this *kanji* with the /sha/ pronunciation. All of the scores for these words were counted according to the levels of difficulty and added to be the total score. There is only one word that uses the *kun*-reading: /yashiro/, categorised as N1. Therefore, the score of *kun*-reading for this *kanji* was 1. This means that the frequency of *on*-reading for this *kanji* was designated as 97%.

The *kanji* that tend to be written in hiragana at present were not counted. Instead, they were listed in the section of Figure 7 emphasised in red. For example, a *kanji*, 何, was not included in the JLPT list with the pronunciation of /ka/. Therefore, 'Not on the JLPT list for *on*-reading' was indicated (see Figure 7).



Figure 7: Cases of the kanji that tend to be written in hiragana at present

Questionnaires Concerning Phonetic Resemblance Between Japanese and Chinese

Another important criterion for this study is the value of a rather psychological aspect of the phonetic resemblance of each *kanji*, as evaluated by the questionnaires (see Figure 8).

The survey method introduced by Kayamoto (1995) was applied in this study.¹⁴ However, there are several differences in the data collection methods between Kayamoto (1995) and this research.

Firstly, the number of points for the rating scale was reduced to five in order to allow the respondents to easily answer the questions, whereas a seven-point scale was employed in Kayamoto (1995). Another advantage of the application of a five-point scale was that it enabled us to calculate the rate of phonetic resemblance quickly.

Secondly, most of the survey respondents (a total of 12 people) in the current work had not only acquired N1 of JLPT but had also been teaching Chinese to Japanese students,¹⁵ although in Kayamoto (1995), the 11 respondents were postgraduate students and research students (*kenkyû-sei*) whose mother tongue was Chinese.

Thirdly, the target databases of *kanji* differ. In Kayamoto (1995), the database included 996 *kanji* that had been previously designated for instruction by the Ministry of Education from 1977 to 1989, whereas we analysed 1,078 *kanji* from the two volumes of the Japanese grammar textbook, *Minna no Nihongo*, the first volume of which was printed in 2013 and the second in 2015.

My database contains 72% of the *kanji* analysed in Kayamoto (1995) and 74% of the *kanji* currently designated for instruction since 1989 (a total of 1,006 *kanji*).

¹⁴ Like Kayamoto's research, because the four tones in Chinese are not clearly differentiated in Japanese, respondents were formerly informed not to care too much about them before evaluating each case.

¹⁵ All 12 respondents were Chinese native speakers who are teaching Chinese to Japanese students; 9 respondents have acquired JLPT N1 and 3 have passed N2.

中漢字(約1000字)の		JP	Pinyin	音読み		- 11	٦	xia4	カ		79	女	nv3		
音韻的類似度の調査	1	人	ren2	ジン	-11	41	F	xia4	ゲ		80	家	jia1	カ	
スイス・ジュネーヴ大学)	1	人	ren2	ニン	-11	42	Ĥ	yuan2	エン		80	家	jia1	ケ	
	2	先	xian1	セン	-11	43	百	bai2	ヒャク		81	去	qu4	キョ	
以下の「日中漢字の皆領的類 41歳の過去」にごぬ力を25種い	3	生	sheng1	セイ		55	千	qian1	セン		81	去	au4	1	
いたします。本調査はジュネーブ	3	生	sheng1	ショウ		45	万	wan4	マン		82	年	nian2	ネン	
大学(スイス)が、中国語と日本 語り回時に登している学校の第	4	学	xue2	ガク		45	万	wan4	パン		83	子	zi0/3	2	
零についての理解を深めるために	5	会	hui4	カイ		46	室	shi4	シツ		83	子	zi0/3	ス	
行うものです。	6	社	she4	シャ		47	堂	tang2	ドウ		84	族	zu2	ソク	
たがって、「正しい」答えも「問	7	行	hang2	コウ		18	隶	shi4	ジ		85	次	ci4	ジ	
違った」答えもありません。張玉浩	7	行	hang2	ギョウ		49	受	shou4	ジュ		85	次	ci4	シ	
来は、国際シンボジワムで発表す る予定です。	7	行	xing2	コウ		50	付	fu4	フ		86	뽑	pu3	フ	
どうぞよろしくお願い致し京す。	7	行	xing2	キョウ		51	屭	wu1	オク		87	通	tong1	ツウ	
ご協力ありがとうございます。 市内時以下系子 "中日公本的言葉	8	大	da4	ダイ		52	洗	xi3	セン		88	急	ji2	キュウ	
美观察的课程"。 此次课度是目的表示部(注1)为了。2017	8	大	da4	91		\$3	段	duan4	ダン		89	特	te4	17	
》同时学习学文和目语的学生们对于汉	2	何	he2	カ		54	自	zi4	ジ		90	甲	jia3	コウ	
这不是考试解以既没有"正确" 5	10	来	lai2	ライ		55	自	zi4	シ		90	甲	jia3	カン	
四會13次而"错误"的回答。要们打算 将课至结果在国际研讨会上进行发表。	11	名	ming2	X		55	催	cui1	サイ		91	城	cheng2	ジョウ	
请多多关照。 谢说出约分助1	12	前	qian2	ゼン		56	起	qi3	キ		92	昭	zhao1	ショウ	
コロシインド「泰姓ス」の類似	13	方	fang1	ホウ		57	寝	qin3	シン		93	皇	huong2	コウ	
Ķ ∉A(100%),B(75%),€	16	教	jiao4 /1	キョウ		SB	休	xiu1	キュウ		93	皇	huong2	オウ	
(50%), D (25%), E (0%) で許 (51~) アンボカム	15	医	yi1	1	_11	59	分	fen1	フン	_11	94	吸	xi1	キュウ	
備忠评价一下"续音"与日语诗	16	者	zhe3	シャ	_11	59	分	fen1	ブン	_11	95	肉	rou4	=7	
"音話み"之间的相似性。 从A(100%) 回知(0%)之间进行选择。	17	研	yan2	ケン	_11	59	分	fen1	7	_11	96	野	ye3	ヤ	
	18	究	jiu1	キュウ	-11	60	半	ban4	ハン	-11	97	果	guo3	カ	
	19	病	bing4	ビョウ	_11	61	千	wu3	з	_11	98	物	wu4	ブツ	
55 爱 ai4 71 A	20	初	chu1	ショ	_11	62	朝	zhao1	チョウ	_11	98	物	wu4	モツ	
	21	失	shi1	279	-11	62	朝	chao	チョウ	-11	99	奈	cha2	チャ	
n3 興 xing4/1 キョウ E	22	礼	li3	11	-11	63	昼	zhou4	チュワ	-11	99	杀	cha2	サ	
	23	本	ben3	ホン	-11	64	夜	ye4	*	-11	100	牛	niu2	キュワ	-
」をドデークは http://inho.org/ 25[play audio]機能などを承受に	24	于	shou3	シュ	-11	65	映	ying4	11	-11	101	酒	jiu3	シュ	
TOPEN	25	1	tu3	1	-11	66	<u>月</u>	yuc4	79	-11	102	店	dian4	テン	
式参考http://inhoorg/的[F-29 pudio] 示使获取音频数量。	25	1	tu3	7	-11	67	人	nuo3	73	-11	103	尔	Cuo1	79	-
□全部で、1195ペアあります。	26	央	ying1	1	-11	68	小	snui3	<u>^1</u>	-11	104	510	Juana	7/	-
たの書与は、単複(miler)またに、 欠者(miler1000)もあります。	27	1	r14	ニナ	-11	69	木	mu4	モク	-11	105	米	ca14	71	-
具有1125时。 万侧的号码有重复的	27	1	114	29	-11	69	不	mu4	ホッ	-11	105	オレ	rus	229	
The state of the state of the state of the state	20	-40 (a)	CIZ	2	-11	70	1	jin 1	41	-11	107	10	SU-1	219	
■ 日本總 記力試験 (JLPT N1 たた: N2)の有無を載えてください。	29	JPJ	C19	*	-11	70	亚	Jin 1	127	-11	108	达出	song4	2020	
请告诉我,您是否拿到了日话就力 **********	30	111	ji i	+	-11	n	1次	pian2	E/	-11	109	110	jie4	2000	
	37	111	shid	47		71	风便	biand	EV.		111	小門	ful	7	\square
ILPT NI ILPT N2	37	七	2003	++		71	(人) (伊)	hiand	~		112	る	mu3	#	
※19日(アいス使生ご - 2	33	誠	teno2	<u>۱</u>		72	局	iu2	キョク		113	米	zhuane4	ジョウ	\square
べて用している減子 データ ペースについて:	34	一位	wil	1	-11	73	飛	fan1	12	-11	114	荷	ho2	71	
https://paparetafor.org/submissior2.99907	35	#	jina3	+11	-11	74	문	haod	-117	-11	115	静	iina4	71	
	35	#	jing3	ショウ	-11	75	丘	bing1	21	-11	116	有	vou3	ユウ	
	36	林	lin2	リン	-11	75	丘	hing1	トョウ		116	有	vou3	ウ	
	37	清	ging1	セイ		76	校校	xiao4	コウ		117	元	vuan2	ガン	\square
委責任者:小橋谷友二 OPATAVA Vall	38	合	shi2	ショク		77	た友	vou3	ユウ		117	元	vuan2	ガン	
ジュネーブ大学文学部	39	国	guo2	コク		78	へ彼	bi3	E		118	利	114	ŋ	
東アジア研究学科	40	油	de0	4		79	4	nv3	ジョ		119	1	xiao3	ショウ	
estar/unites/japonus/accav2/	10	地	de0	2		79	4	nv3	ニョウ		120	古	gu3	3	

Figure 8: The first page of the questionnaire (total of six pages)

	JP	Pinyin	音読み	On-reading
1	人	ren2	ジン	
1	人	ren2	=>	ער
2	先	xian1	セン	Т
3	生	sheng1	セイ	1
3	生	sheng1	ショウ	1
4	学	xue2	ガク	1
5	会	hui4	カイ	
6	社	she4	シャ	1
7	行	hang2	コウ	
7	行	hang2	ギョウ	
7	行	xing2	コウ	1
7	行	xing2	ギョウ	ער
8	大	da4	41	1
8	大	da4	91	
9	何	he2	カ	1

Figure 9: Example of the plural on-reading or Chinese pronunciations

Evaluation of 1,095 Pairs of *Kanji* for Phonetic Resemblance

The number of characters in the database is 1,078, but the number evaluated in this questionnaire was 1,095. This is because, in the case of multiple *on*-readings or Chinese pronunciations existing in one *kanji*, respondents evaluated them separately. As shown in Figure 9, a *kanji*, A, 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 $\overline{77}$, which has two ways of *on*-reading, /kou/ and /gyou/, and two Chinese pronunciations, /hang2/ and /xing2/. In this case, the respondents must have assessed the phonetic similarity in four pairs independently.

Findings of the Two Statistical Analyses

(1) Findings from the Frequency of the On-Reading Usage

- The rate of the frequency was 59.7%.
- 294 kanji (27% of the total number of the data) only have on-reading pronunciations.
- 152 *kanji* only have *kun*-reading pronunciations.

(2) Findings from the Survey on Phonetic Resemblance by Means of Questionnaires

• The mean value for the phonetic resemblance from the survey was 19.6%. Compared with the average of the survey outcome from Kayamoto (1995), the result was slightly lower (Kayamoto's mean value was 34%, $2.38/7^{16}$).

• Very few *kanji* had complete correspondence between Japanese and Chinese (only n = 9 *kanji*: 医, 伊, 信, 衣, 敷, 因, 他, 愛, and 膚). The rate of the ones with more than 90% resemblance was only 4% (n = 42)¹⁷.

• Most of the *kanji* with perfect correspondence were first and fourth tones.

• Nine per cent of *kanji* (n = 208) were considered to have no resemblance between Japanese and Chinese. Furthermore, the ones whose resemblance was less than 10% comprised more than half of the data (n = 572).

 $^{^{16}}$ As a result, the average of the overall rating was 2.38 (SD: 1.32) (Gi, 2017, p.63). The average of 2.38 (/7) is 34% when converted to a percentage.

¹⁷ These 42 *kanji* are 医伊信衣敷因他愛膚付利理辛心痢意部離来父富引府民打那符新印 婦負賴夫林苦腐流麻飲太里隠.

Chart 1 shows the mean and standard deviation.



Chart 1: Distribution of the mean value of the questionnaires for the phonetic resemblance of 1078 *kanji* and standard deviation¹⁸

Chart 2 shows the distribution of the number of *kanji* in the range of the mean value, presenting the number of *kanji* on the vertical axis and the range of the mean value on the horizontal axis.



Chart 2: Distribution of the number of averages

As illustrated in this chart, the majority of dissimilarity is distributed on the left and the similarity on the right part of the chart.

¹⁸ The *kanji* at the bottom is representative of each stage.

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

• The score obtained by multiplying the values from the analysis of (1) and (2) was 12.4%.

• Only five *kanji* (医, 伊, 信, 愛, and 膚) have 100% frequency of usage in *on*-reading as well as 100% phonetic similarity.

The distribution graph (Chart 3) indicates a high rate of dissimilarity between Japanese and Chinese.



Chart 4, which clarifies this dissimilarity, shows the distribution of the number of *kanji* in the range of the mean value, presenting the number of *kanji* on the vertical axis and the range of the mean value on the horizontal axis. The number of *kanji* in the range of 0-10% is 697, whereas the number of *kanji* in the range of 0-50% are the majority (95%). This means that most of the pronunciations of *kanji* used in both languages differ drastically.



¹⁹ The *kanji* at the bottom is representative of each stage.



Chart 5: The number of averages from 50% to 100%

Chart 5 shows a zoomed-in view, from 50% to 100%, of the range of the mean values in the Chart 4, showing the minority group of *kanji* that phonetically resemble each other. Figure 10 shows the final results for the three fields.



Figure 10: The final results for the three fields

Conclusion

To conclude, after quantifying the phonetic resemblance of 1,078 *kanji* between Japanese and Chinese, it was found that this resemblance was very low compared to the high resemblance in shape and meaning. Such clarification of the three values of shape, semantic, and phonetic resemblance will allow us to explain the '*kanji* paradox', whose notion was applied in one of our previous works to explain that the simultaneous acquisition of Chinese and Japanese is, in fact, difficult, despite the image of easy cross-comprehension due to the common usage of Chinese ideograms (Berger, C., & Obataya, Y. (2014), pp.162-163).

Despite the findings of this study, it is not necessary to fully deny the attempt of simultaneous learning or cross-comprehension of Japanese and Chinese. In fact, there are considerable advantages and incentives in the resemblance of forms and meanings in the common *kanji* for motivating learners to study two languages simultaneously. By making students aware of the imbalanced rates of resemblance among the three components, sounds, forms, and meanings at the beginner's stage of learning, it may be possible to make their studying process more effective.

This means that it is important for students to develop special learning strategies to study *kanji* for acquiring listening and speaking skills, which should differ from the strategies for developing reading and writing skills in learning Chinese and Japanese simultaneously.

Acknowledgement

I would like to express my special gratitude to Pierre François Souyri, Professor Emeritus at the University of Geneva, 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://ttcn.co.jp/) for their cooperation with the questionnaire, Ayako Shibata for her comments on an earlier version of the manuscript for improvement, Ping Zeller for her corrections to the Chinese part of the questionnaire, and Ikuo Obataya (Quantum Design Japan) for his help with measurements and analysis. Finally, I would like to thank administrators of Jisho.org and Jonathan Waller, creator of the JLPT resources pages for deciding the JLPT levels in my database.

References

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.

The Japan Foundation and Japan Educational Exchanges and Services. (2007). *Japanese-Language Proficiency Test: Test Content Specifications* (Revised Edition). Tokyo, Japan: Bonjinsha.

Matsushita, T. (2009). A Japanese-Chinese Macro Contrastive Study of Common *Kanji* Vocabulary: The Distribution of Similarities and Differences in Frequency, Form and Meaning (*Makuro ni mita jôyô kanji go no nicchû taishô – hindo. keitai, imi no icchi to zure no bunpu*). JSAA-ICJLE 2009 International Conference.

(https://www.researchgate.net/publication/299180565_makuronijiantachangyonghanziyunoriz hongduizhaopinduxingtaiyiweinoyizhitozurenofenbu A Japanese-

Chinese_Macro_Contrastive_Study_of_Common_Kanji_Vocabulary_The_Distribution_of_Si milarities_and_Differenc)

Yano, M. (2012). *Chinese characters culture and Japan (Chûgoku kanji bunka to Nihon)*. Tokyo, Japan: Surugadai shuppansha.

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.

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.

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.

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.

References for data

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.

Contact email: Yuji.Obataya@unige.ch