



LUND
UNIVERSITY

Bachelor's Programme in Japanese

Swedish students and the phonetic component of Kanji

by

Jesper Gustafsson

JAPK11

Bachelor's Thesis (15 credits)

June 2019

Supervisor: Axel Svahn

Examiner: Arthur Holmer

Abstract

This thesis was written to answer the questions *How much do Swedish students of Japanese know about the phonetic component of kanji?* and *Does awareness of the phonetic component help students guess the reading of kanji?*. The phonetic component of kanji gives you the reading of kanji, and can give you an edge in correctly reading unfamiliar kanji. By doing some background research into the curricula of Swedish universities and kanji knowledge of their students, an online questionnaire fit to their presumed knowledge was constructed. The questionnaire measured the participants' ability to guess the reading of unfamiliar kanji before and after being given information regarding the phonetic component of kanji. The results of this questionnaire indicates that the participants in general seem to know of the phonetic component, but are not utilising it very well. There was a significant improvement in most of the participants, and as such, introducing, or focusing more on this topic in school could prove to be beneficial.

Keywords: Japanese language, kanji, phonetic component

Acknowledgements

I would like to thank my supervisor for help and guidance when trying to decide on my research topic. I also highly appreciate the help I received from the universities professors from answering my questions and helping me find a suitable set of kanji for the questionnaire. In addition, I would also like to thank Leonardo Boiko from namakajiri.net for the wonderful introduction to the phonetic component, which sparked my interest in this topic.

Contents

1	Introduction	5
1.1	Research questions	6
1.2	Thesis layout	6
2	Background	7
2.1	The history of kanji	7
2.2	Radicals: The building blocks of kanji	8
2.3	Kanji categories	8
2.4	The phonetic component	9
2.5	How to utilise the phonetic component	10
2.6	Frequency in JLPT & Japanese curriculum	10
2.7	Curricula at Swedish universities	11
2.7.1	Lund University	11
2.7.2	Gothenburg University	12
2.7.3	Dalarna University	13
2.7.4	Stockholm University	14
2.7.5	Summary	15
2.8	Textbooks	15
2.8.1	Genki	15
2.8.2	Tobira	15
2.8.3	Kanji Look and Learn	16
2.8.4	Heisig's Guide to Remembering the Kanji	16
3	Method	18
3.1	The approach	18
3.2	Before the questionnaire	18
3.3	The questionnaire	18
3.4	After the questionnaire	19
3.5	How the kanji were chosen	19
3.6	Questions & Information	20
4	Results	22
4.1	Kanji knowledge	22

4.2	Change in guesses	23
4.3	Individual kanji	24
5	Conclusion	25
5.1	Limitations & Problems	25
5.2	Research answers	26
5.3	Pedagogical implications	27
5.4	Future research	27
	References	28
A	Phonetic Series with KanjiVG kanji	30
B	Genki Kanji	32
C	Potential Phonetic Series	34
D	Data	36
D.1	Comments	43
E	Questionnaire	44

1 Introduction

Japanese is a language with four scripts, hiragana, katakana, kanji and rōmaji. Kanji, literally translated as Chinese characters, are logographic characters imported from China. In Japanese they typically represent lexemes and can be read in multiple ways.

Hiragana and katakana, collectively known as kana, are a Japanese invention and function more similarly to the latin alphabet. Each character represents a mora, the smallest unit of sound in Japanese. Any word in Japanese can be written using only hiragana, but they are most commonly used as grammatical markers or are attached to a kanji and change depending on the inflection of the word.

Rōmaji is the romanization of the Japanese sounds using the Latin alphabet as a base. There are a few systems for romanization. Two of these systems are the Hepburn system and kunrei-shiki.

Out of these four scripts, this thesis will focus on the kanji script. The subject of kanji has been researched extensively, and the specific areas of kanji readings are no exception. This thesis aims to investigate a method of deciphering the reading of kanji and test it on a small scale to see if it can help students figure out the reading of kanji.

The aforementioned deciphering method utilises the kanji component which represents the sound of the kanji, the so called phonetic component. It is not unheard of foreigners assuming that every kanji are purely pictographic or drawings, with no rules of the associated sound. Because of this assumption among foreigners, this thesis will also investigate if this component is known to Swedish students of Japanese.

In reality, if we look at the tōyō kanji set (See glossary), just over a tenth of all kanji are pictographic while nearly two-thirds are of the phono-semantic type (Suzuki, 2007:59). Ivarsson (2016:48) also provides some statistics for kanji classifications. She mentions that in the more modern kanji set, the jōyō kanji set, two-thirds are of the phono-semantic type. In addition, she also mentions as much as 80% of all kanji could be of this type. Phono-semantic kanji consist of a phonetic component, which represents sound and a semantic component, representing meaning. One thing to note is that as previously mentioned, a kanji can have multiple readings. There are two types of readings: kun-yomi (Japanese reading) and on-yomi (Chinese reading). The phonetic component only plays a role when it comes to on-yomi.

1.1 Research questions

This thesis will aim to answer two questions.

1. How much do Swedish students of Japanese know about the phonetic component of kanji?
2. Does awareness of the phonetic component help students guess the reading of unfamiliar kanji?

While research into the phonetic component is not unheard of, there seems to be a lack of examining the usage from a learning perspective. Ivarsson (2016:64) provides a meta-analysis of a few studies which implies that Chinese children seem to develop an awareness of the phonetic component somewhere between ages six and nine years. Japanese children on the other hand, seem to develop this awareness a bit later, still having limited radical awareness and difficulty reading words with on-yomi into the 7th grade (Ivarsson, 2016:79-81). The reason for choosing to investigate these topics is that it would be interesting to know to what degree the phonetic component could help students when reading kanji. If there is a noticeable change it could be of interest to spend some time on this in the classroom as it could make it easier for students to learn kanji both inside and outside of the curriculum.

1.2 Thesis layout

Excluding this introductory chapter the thesis contains five more chapters. The second chapter, *Background*, provides a short history lesson on kanji, their readings and previous research on the subject of kanji reading and the phonetic component. It also goes into the method of teaching kanji that some textbooks use, as well as how Swedish universities teach kanji. The third chapter, *Method*, explains the method of acquiring data to answer the research questions. The fourth chapter, *Results*, presents the results gathered through the method in the form of tables. The fifth chapter, *Analysis*, provides analyses of the results. The sixth chapter, *Conclusion*, provides a summary of the thesis as well as discusses the method used, results and analysis as well as possible practical implications and possible future research.

2 Background

This chapter examines some previous research into kanji, the phonetic component and the curricula of Swedish universities. Firstly a short history lesson on the origin and development of kanji and their readings through history will be given. It will then take a deeper look into kanji, their categories and building blocks, including the phonetic component. In addition, it will also provide some information regarding the curricula at the Swedish universities and common textbooks. Finally, a short summary will be given.

2.1 The history of kanji

Ivarsson (2016:32-36) provides a short description of the history of kanji. As previously mentioned, kanji are characters imported from China. However, Japanese kanji and Chinese hanzi can be quite different. They can be different in two ways, the phonetic aspect and the graphic aspect. In Chinese, the characters generally only have one reading while kanji more often than not carry at least two (Ivarsson, 2016:34). The readings of kanji can be divided into two categories: kun-yomi and on-yomi. Kun-yomi refers to the reading of kanji for native Japanese words while on-yomi refers to the reading of Chinese origin, used in Sino-Japanese words. This study focuses on the on-yomi reading and the kun-yomi is as such sidelined. Apart from the fact that Chinese and Japanese are languages with different sounds there are other reasons for the difference in the on-yomi of kanji and the Chinese reading of its hanzi counterpart. It is in part due to time itself as the pronunciation developed differently in the countries separated by sea. Another reason for the difference is the manner of import. Not only did the process of adoption span several centuries, but also across many regions and thus, the dialect it was imported from also differed significantly. In the beginning of the adoption process, imports came from the southern Chinese dialect via Korea. Between the 7th and 9th century, the import had shifted to the central Chinese dialect, and lastly from the northern dialect of Chinese. Due to this gradual and inconsistent adoption, the on-yomi of kanji does not necessarily correspond to their hanzi counterpart. One example of these kanji is 行 (walk). 行 can be read as *gyō* as in 行列 (*gyō-retsū*, “procession”), based on the southern dialect, *kō* as in 行進 (*kō-shin*, “parade”), based on the central dialect as well as *an* as in 行脚 (*an-gyō*, “walking tour as pilgrimage”), based on the northern dialect. Kanji has also in time been modified, leading to various forms of the same kanji. Suzuki (2007:60) include the examples 学 (learn), simplified from 學 and 国 (country), the simplified form of 國.

2.2 Radicals: The building blocks of kanji

In the first chapters of Seely (2016), the origin and structure of kanji are explained. Kanji can be quite complex characters, ranging from one single stroke to more than twenty-five strokes. They are compounds of different smaller characters sometimes referred to as radicals or components. There are a total of 214 radicals, with some having multiple forms, such as the alternate form of 人 (person), 亻, found in 袋 (bag). There are also variants of radicals, such as 儿 for 人 and 足 for 疋. There is also 月 which carries the meaning of moon, but is at the same time an alternate form of 肉 (meat) and 舟 (boat) (Seely, 2016:34-40). Variants can also be seen in kanji and hanzi. One old example is 魑 and 魅, both meaning “malevolent spirit”. Variants are however nowadays less common, most likely due to more widespread education and official script simplifications in both countries (Seely, 2016:16).

2.3 Kanji categories

Seely (2016:12-14) mentions the four larger out of the six traditional categories of Chinese characters which are classified based on their original formation. The first classification are pictographic characters (象形 *shōkei*). These represent pictures of something concrete, such as 木 (tree), and 山 (mountain). However, an issue with this type of kanji is that it can be quite hard to represent more abstract concepts such as “above” or “below”. These more abstract ideas are represented with kanji classified as diagrammatic characters (指事 *shiji*). These include concepts such as direction and numbers: 上 (above), 下 (below), 一 (one), 二 (two) and 三 (three). Simply adding a horizontal line to the existing pictographic kanji for tree 木 can represent parts of a tree: 本 (root) and 末 (apex). Combining these pictographic and diagrammatic characters leads to another classification, semantic composites (会意 *kaii*). One obvious example is 林 (grove) and 森 (forest) consisting of two and three 木 (tree). Another example is 峠 (mountain pass). The left side represents a mountain, and the right side represents “up” and “down”. In ancient China one would often borrow kanji that had the same reading to represent a different word, eventually leading to a lot of kanji carrying multiple meanings. Let us take 莫 for instance. 莫 originally had the meaning of “sunset”, but was borrowed to also represent “there is none, not any” due to its similar reading. This could lead to confusion and eventually semantic radicals were added to these kanji to differentiate the different words. For our example, this resulted in 暮, with the added semantic radical 日 (sun) carrying the meaning of “sunset”, and 莫 keeping the meaning of “there is none, not any”. These kanji came to be known as phono-semantic composites (形声 *keisei*). The characters consist of

one part representing a meaning (metal, tree, bird) and one part representing the pronunciation. One more example of this is 憬 (yearning), where the left side 忄 is the semantic part meaning "heart" or "mind", and the right part 景 representing the pronunciation of the kanji, *kei*. This phonetic component can be quite useful to be aware of for reasons discussed below.

2.4 The phonetic component

According to Ivarsson (2016:48-49), the majority of kanji actually fall into the phono-semantic composites (PSC) classification. Within the jōyō kanji, around two thirds are considered phono-semantic composites. However, unfortunately not every PSC kanji can be consistently read correctly. Only 57.6% are considered completely consistent, that is to say that they all have the same on-yomi as the phonetic component dictates. Combining these percentages one can accurately guess the pronunciation of a character by the phonetic component 40% of the time. This inconsistency of PSC kanji can partly be explained as previously said, to the modification or change of kanji and pronunciation throughout history. One example of a modified PSC kanji is 仮. 仮 is read as *ka* even though its phonetic component 反 is read as *han*. This inconsistency is due to the simplification of the character. The character originally was written as 假, with its right side representing the pronunciation *ka*. In other words, the simplified character 仮 retained the pronunciation of 假 while replacing it graphically with 反. Another example of this happening is 読 derived from the original character 讀 (Seely, 2016:96). These inconsistent kanji make up about 9.7% of the jōyō kanji. The other 32.7% are considered partially consistent, meaning the pronunciation is similar but not exact. An example of a partially consistent kanji is 海 which is read as *kai* even though the phonetic component 每 is read as *mai*.

Interestingly enough, Ivarsson (2016:64) mentions that while this component plays such a huge role, awareness of the phonetic component seems to develop quite late for both Chinese and Japanese students. This development can be observed even later among Japanese students, likely due to kanji having several readings while hanzi typically only has one reading per character. Due to this difference in kanji and hanzi, Japanese students need to learn more kanji characters before being able to make connections between readings of kanji and phonetic component. Ivarsson (2016:66-67) goes on to cover another study comparing second-language learners and natives which concludes that second-language learners depend more on phonological information compared to native speakers.

2.5 How to utilise the phonetic component

In theory, using the phonetic component could be a great tool when it comes to learning or recalling kanji, but how can one actually use this kanji to determine kanji? Looking at the various phonetic lists of kanji included in the appendix of this thesis, a pattern that appears is that the phonetic component is a large majority of the time on the right side of the kanji when the kanji can be divided vertically, such as 館, with the semantic component 食 (food, eat) on the left and the phonetic component 官 (*kan*) on the right. When kanji can be divided horizontally, the phonetic component is generally on the bottom, as in the case of 芥 with 艹 (grass) on the top and 介 (*kai*) on the bottom. There are quite a few exceptions to these rules, for example 効 and 資, where the phonetic component is to the left and top respectively. Once the phonetic component has been identified, recalling readings of kanji with the same component should lead to a fairly high chance of guessing correctly, or at the very least lead to the reading that has the highest probability of being correct. Let us see an example of this process for the kanji 餃. Step one is identifying the phonetic component, it is highly probable that it is the component on the right side of the kanji, 交. Step two is trying to recall readings of kanji with this component. A very common word, 高校 (high school, *kōkō*) includes 校 (*kō*). It is a quick two-step process, which can further be reduced to a one-step process when getting more familiar with these phonetic components.

2.6 Frequency in JLPT & Japanese curriculum

Suzuki (2007:59-60) provides the two tables 2.1 and 2.2 which include data of kanji category frequency in an example JLPT (Japanese Language Proficiency Test), as well as frequency of kanji categories in the Japanese education system. From these tables, one can see that for both the easiest level of the example JLPT, N4 as well as the first grade kanji mostly consist of pictographic kanji. However, with each level and grade, this amount greatly reduces while other types becoming more frequent. The majority of the kanji for N2, N1, fifth grade and sixth grade are phono-semantic kanji. This makes sense as the phono-semantic and semantic compounds are made out of pictographic and diagrammatic kanji and learning these first would make learning subsequent kanji easier. The information here is based on the jōyō kanji list before it was updated in 2010.

Table 2.1: Table of kanji frequency in an example JLPT test (Suzuki, 2007:59)

type of kanji	N4	N3	N2	N1	jōyō
pictographic	35(42.5%)	32(17.7%)	98(13.0%)	75(8.1%)	216(11.7%)
diagrammatic	9(11.3%)	–	1(0.0%)	–	10(0.5%)
semantic compounds	20(25.0%)	65(35.9%)	218(29.0%)	190(20.5%)	413(22.3%)
phono-semantic	19(12.5%)	82(45.4%)	430(57.2%)	658(71.1%)	1211(65.5%)

Table 2.2: Table of kanji frequency in the Japanese education system's first six grades (Suzuki, 2007:60)

type of kanji	1st grade	2nd grade	3rd grade	4th grade	5th grade	6th grade
pictographic	44(57.9%)	47(32.4%)	24(12.3%)	17(8.7%)	13(6.7%)	25(13.2%)
diagrammatic	5(6.6%)	–	–	–	–	–
semantic compounds	15(19.7%)	38(26.2%)	53(27.2%)	68(34.9%)	43(2.1%)	48(25.3%)
phono-semantic	12(15.8%)	60(41.4%)	116(59.5%)	109(55.9%)	139(71.3%)	117(61.6%)
total	76	145	195	195	195	190

2.7 Curricula at Swedish universities

According to the statistics, one could argue that utilising the phonetic component may be an advantage when it comes to learning kanji, but to what extent do the Swedish universities and their textbooks discuss or utilise the phonetic component? This section of the thesis will look at the kanji-teaching side of Swedish universities. I will base the teachings of Lund University on my own experiences of the courses and the others on comments from professors of the respective universities.

2.7.1 Lund University

Lund University offers a full-time bachelor's program. As many students from Lund University spend their third year as exchange students in Japan with different textbooks and curricula, only the courses for the first two years will be looked at.

In the course plan from the first 20-week course JAPC01 (Japanese Beginners' Course)¹ it says that at the end of the course the students are expected to be able to describe the three writing systems, including the systematic structure of kanji. They are expected to master hiragana and katakana as well as 175 basic kanji whose systematic structure is taught out.

In the course plan for the following 20-week course JAPC02 (Japanese: Level 1)² it says that 250 additional kanji are taught and that the students are expected to be able to identify and write over 400 kanji, and being able to use them in phrases and sentences.

The courses of the second year are JAPC03 (Japanese: Level 2) and JAPK11 (Japanese B.A. Course). As of the publication of this thesis, I am currently attending the JAPK11 course, and have passed the other three. I will base the information for these courses on my own experiences.

In JAPC01 we were introduced to the concepts of hiragana, katakana and kanji. We were also introduced to a few semantic radicals along with learning new kanji. The existence of the phonetic component was however mentioned briefly enough to forget or not mentioned at all. That is not to say that there was no strategy behind the method of teaching us kanji. For example, if a kanji was similar to one we had learned before, this was mentioned. And when components looked like a kana character, this was also mentioned, such as 亼 in 私 and 強 or 力 in 動 and 助. The same strategy was utilised in all courses, but we were no longer expected to learn new semantic radicals past JAPC01. As for learning the reading of kanji, this was generally done by learning vocabulary which included the kanji. As a result, there were quite a few kanji which we only learned one reading of. An example is 私. We had already learned the word kun-yomi reading, *watashi*, but the on-yomi reading *shi* was not taught when we were introduced to the kanji.

2.7.2 Gothenburg University

The following questions were asked to a professor at Gothenburg University.

1. Are the students taught about the different categories of kanji? (*keisei, shokei, kiai, shiji moji*)
2. Are the students taught about the phonetic component in *keisei* kanji?
3. Do you use the phonetic component when you teach new kanji?
4. Do you have the impression that your students actively use the phonetic component when they read/learn new kanji?

¹<http://kursplaner.lu.se/pdf/kurs/sv/JAPC01>, accessed 05-05-2019

²<http://kursplaner.lu.se/pdf/kurs/sv/JAPC02>, accessed 05-05-2019

The answer to all of these questions was "yes". The professor was also asked which kanji the students are expected to know at the end of the first two courses, which was the kun-yomi and on-yomi of all kanji in Genki I and Genki II.

2.7.3 Dalarna University

The following questions were asked to a professor at Dalarna University.

Question 1: "Are the students taught about the different categories of kanji? (keisei, shokei, kiai, shiji moji)"

Answer: "Yes, very briefly and roughly, but yes."

Question 2: "Are the students taught about the phonetic component in keisei kanji?"

Answer: "Same as above."

Question 3: "Do you use the phonetic component when you teach new kanji?"

Answer: "We don't teach kanji much during the lessons because of the lack of teaching hours, so we have kanji worksheets with some instructions as well as one online lecture with several online tests. We mention about the point a little bit both in the online lecture and the worksheet instruction, but not very much."

Question 4: "Do you have the impression that your students actively use the phonetic component when they read/learn new kanji?"

Answer: "They are learning the first book of GENKI right now, so there are not so many kanjis they can apply their knowledge to read new kanji, but as you suggest, we should probably emphasize more on that point."

Question 5: "In the course plans for "Japanska I: Grundläggande språkfärdighet" it says that after the course, the student should be able to "redogöra för kanji-tecknens ursprung och grundläggande struktur". Could you perhaps explain this a bit more? Are they supposed to be able to break down each kanji into their components and/or radicals, should they know the names and/or meaning of the radicals?"

Answer: They need to answer/pass a small online test after watching a lecture about kanji. The questions are about the history of Kanji, different categories, guessing the reading of new kanji from the component."

Dalarna University does mention the phonetic component but does not utilise it very much in their teaching methods. They do have online test(s) to test the students on kanji history, categories and ability to read unfamiliar kanji.

2.7.4 Stockholm University

The following questions were asked to a professor at Stockholm University.

Question 1: Are the students taught about the different categories of kanji? (keisei, shokei, kiai, shiji moji)

Answer: In my kanji lessons (japanska II), I don't teach the categories. I intentionally avoided introducing them because I thought it will be too much information in this level.

Question 2: Are the students taught about the phonetic component in keisei kanji?

Answer: No. Among 指示文字・会意文字・象形文字・形成文字, I feel that 形成文字 is the most complicated one. With this reason, I don't see any reason to introduce the concept of 形成文字 to the students.

Question 3: Do you use the phonetic component when you teach new kanji?

Answer: Do you mean if I teach both kun- and on- reading? If so, yes.

Question 4: Do you have the impression that your students actively use the phonetic component when they read/learn new kanji?

Answer: I don't know. Since I speak Japanese as a mother tongue, I use the phonetic component when I read unknown kanji. However, the students in Stockholm may not necessarily see it in the same way. If they have never seen a kanji, they may simply give up without associating or actively using phonetic component to identify an unknown kanji.

Question 4: In the course plans for Kanji IA, IB, IIA, IIB it says that after the course, the student should be able to "uppvisa kännedom om tecknens struktur och uppbyggnad". Could you perhaps explain this a bit more? Are they suppose to be able to break down each kanji into their components and/or radicals, should they know the names and/or meaning of the radicals?

Answer: Both in Kanji I and II, students are introduced both the names and meaning of radicals. In addition to that, students are explained kanji by emphasizing on 会意文字 which is the combinations of two/three components in one kanji.

The third question was misunderstood, and a follow up question was asked.

Follow up question: From your other answers I would assume that you do not explicitly use the term phonetic component, but maybe you would mention that component? Would you also mention that they share the same reading?

Answer: Now I understand your question, and the answer is no. I don't mention the term phonetic component or sharing the same reading at all.

Stockholm University does not mention or utilise the phonetic component at all in their teachings.

2.7.5 Summary

Gothenburg University seems to be the only university that uses the phonetic component extensively in their teaching method. The other universities only briefly or do not mention these categories at all. But they all seem to at least touch on the subject radicals. All the universities but Lund University mention the different categories.

2.8 Textbooks

2.8.1 Genki

Genki (Banno et al., 2011a,b) is a two-part textbook series with a workbook for each part. *Genki* is used by all the Swedish universities. The first textbook very briefly mentions the phono-semantic kanji (In the book referred to as "phonetic-ideographic compounds") when it lists up classifications of kanji (Banno et al., 2011b:31). The textbooks do however not use this or any other classification in the chapters intended to teach the students to write and read the kanji, and therefore 一 and 時 are not differentiated by classification. What is listed together with the kanji is: stroke order, some vocabulary, some readings and their meaning. *Genki* does in other words not employ the phonetic component when teaching kanji. *Genki* I and *Genki* II include 145 and 172 kanji respectively for a total of 317 characters.

2.8.2 Tobira

Tobira (Oka, 2009) is another popular series, which is used at Lund, Gothenburg and Dalarna for the second year. *Tobira* uses a very similar list of kanji. One key difference is that *Tobira* also includes a space for the radical of kanji and the radical's name.

2.8.3 Kanji Look and Learn

Kanji Look and Learn (Banno and Ikeda, 2009) is a textbook used at Gothenburg University in their second year together with Tobira. Similarly to Genki before it, it shows the on-yomi and kun-yomi readings of the kanji, stroke order as well as some vocabulary. What is unique to this book is that they use pictures to help readers remember how to draw them and what their meanings are.

2.8.4 Heisig's Guide to Remembering the Kanji

Heisig's Guide to Remembering the Kanji (Heisig, 2011, 2012) is a book series that utilises the phonetic component to teach 3000 kanji. Heisig (2011:10) argues that it is inefficient to learn the writing and reading of kanji simultaneously, which may be why he split up his guide into three parts. The first book advises the use of mnemonics and stories to remember how to write kanji. You can often find a series of kanji in a row that have similar components since the guide typically introduces one or a few components and then the following kanji consists of this component. The guide does however only give the meaning and a story as well as the stroke order of a kanji and does not list neither the kun-yomi nor the on-yomi. However, in book 2 Heisig does focus on the reading of kanji and introduce what he calls "signal primitives" which is the same as phonetic components. While book 1 and book 2 consists of the same 2200 kanji, the order is different. Heisig groups the kanji into 10 groups:

1. *The Kana and Their Kanji*: Characters from which kana are derived from.
2. *Pure Groups*: Characters where the signal primitive dictates the reading with no exceptions.
3. *One-Time Chinese Readings*: Characters whose on-yomi readings are unique.
4. *Characters with No Chinese Readings*: Characters with no on-yomi reading.
5. *Semi-Pure Groups*: Characters where the signal primitive dictates the reading with a single exception.
6. *Readings from Everyday Words*: Characters from everyday words.
7. *Mixed Groups*: Characters where the signal primitive dictates the reading of at least two characters and do not belong to the *Pure* or *Semi-Pure* Groups.
8. *Readings from Useful Compounds*: Characters that are frequent in non-conversation. Such as newspapers, billboards and signs.

9. *A Potpourri of Readings*: Characters that do not fit into the previous groups categories but are too common to be left to the last group.
10. *Supplementary Readings*: Characters that are less common and do not fit into the previous groups.

The categories which utilise the signal primitive are the *pure*, *semi-pure* and *mixed* groups and together they contain the majority of all the kanji in the book. There is also the third part in the series which includes an additional 800 kanji, which expands upon the groups found in the second book.

3 Method

3.1 The approach

In order to answer the two research questions (*How much do Swedish students of Japanese know about the phonetic component of kanji?*, *Does awareness of the phonetic component help students guess the reading of kanji?*), an online questionnaire was sent to four Swedish universities: Lund University, Gothenburg University, Stockholm University and Dalarna University. Before constructing the questionnaire, professors of each university were asked to check a list of phonetic series of kanji to make sure that the questionnaire would contain relevant kanji and questions.

3.2 Before the questionnaire

In order for the questionnaire to give something of substance, how the questions were formulated had to be carefully thought over. Students may use the phonetic component of kanji without realising it and because of this, the questions need to be formulated in a manner that can access this knowledge without explicitly asking for it.

3.3 The questionnaire

The questionnaire can be split into three parts.

1. Asking the participant to guess the reading of a set of kanji.
2. Questions and information regarding kanji and the phonetic component.
3. Asking the participant to guess the reading of the same set of kanji again.

The reason for this ordering is because of the second research question (*Does awareness of the phonetic component help students guess the reading of kanji?*). By analysing the proposed readings of each kanji in the first part and comparing them to the third part, one can get an idea of whether the information helped them guess correctly.

3.4 After the questionnaire

To answer the first question, an analysis of the answers from the second part of the questionnaire was made. In order to answer the second question, an analysis of the answers from the first and third part was made. The analyses are explained in more detail in the next chapter. Due to time constraints, all the parts were sent as a single questionnaire. One could argue that it might be better to have a period of time between the second and third part or maybe follow up with a fourth part later on because the results might be skewed if they have just been primed to look for the phonetic component.

3.5 How the kanji were chosen

In order for a kanji to be chosen for this questionnaire it had to pass two criteria and preferably a third criteria.

1. It should not be included in the students' course curricula.
2. It should share a phonetic component with at least one kanji that the student knows.
3. The student should also preferably be confident in the on-yomi of the known kanji.

One kanji that satisfies all of these criteria is 郊. It is not included in any of the students' curriculum. It shares the phonetic component 交 (*kō*) with 校, which is included in the common word 高校 (high school, *kōkō*) which is present in all the first-years' curriculum, which all students should be comfortable with.

To find a set of kanji based on these criteria, two lists were compared. The first list were of kanji that all students should know (See appendix B). As all universities used the Genki books in the curricula for their first year, this list was a list of all kanji from Genki I and Genki II. The second list was based on phonetic series found online in the form of a spreadsheet¹ created by Leonardo Boiko, author of the blog *Namakajiri* (Boiko, 2011). From this spreadsheet, only entries with a reading-accuracy above 50% as well as a size of minimum 5 characters were considered. The two lists were then compared to find entries of phonetic series which contained at least one Genki kanji to make a third list (See appendix A). The entries of this third list contains kanji which satisfy the first two criteria, as all kanji outside the kanji found in Genki are not included in the students' course curricula and all kanji share a phonetic component with the kanji from Genki. To satisfy the third criteria, the professors of the universities were given words using the on-yomi reading of the kanji were searched for in the vocabulary found in Genki² and asked if their

¹https://namakajiri.net/data/kanji/components_phonetic.kanjivg.tsv, accessed 05-05-2019

²http://genki.japantimes.co.jp/resources/saku_tango, accessed 05-05-2019

students should be comfortable in the reading of the word, further reducing the number of potential kanji a bit. From the potential kanji, a final set of 36 kanji were chosen from 18 different phonetic series, with two from each series. This final set can be seen below, with kanji from the same series next to each other. For the questionnaire, the order of appearance was randomized using an online randomizer³ but the randomized order was the same for all participants.

Table 3.1: Chosen kanji

衙	梧	按	鯨	藁	犒
跣	筭	絆	裨	舘	棺
岱	玳	劒	巖	抹	茱
廂	孀	鯨	鳩	撼	緘
鉞	囧	峙	埶	漣	蓮
价	芥	柞	搾	澗	癩

3.6 Questions & Information

To answer the first research question (*How much do Swedish students of Japanese know about the phonetic component of Kanji?*), a number of questions for the questionnaire were produced. They had to be framed and ordered in a way that does not prime the participants to the degree that their answers are influenced. Since the participants may unknowingly make use of the phonetic component while not knowing of its existence or purpose, questions about their experience with reading kanji were put first, while questions explicitly asking about the phonetic component were put afterwards. To see the questions in detail, see appendix E. As introductory questions the students were asked which university they attend, which year and how many kanji they estimate that they know. Next they were asked to guess the reading of 36 different kanji. After this, they were asked questions about their thought process when guessing as well as questions about kanji categories and the phonetic component. They were also given information about how to utilise component and then asked to guess the reading of the kanji again. As an ending question they were asked if they feel that they had an easier time guessing the second time around. The information given in the questionnaire concerning kanji and the phonetic component is given in full below.

Kanji can be distributed into four different categories (pictographic / pictograms, diagrammatic / simple ideograms, semantic / compound ideograms and phono-semantic / phonetic-ideographic). Pictographic kanji are depictions of something concrete such as a tree/mountain (木/山), diagrammatic are more abstract such as up/down (上/下), semantic kanji combines characters from the previous categories to represent words such as forest/mountain pass (森/峠). The fourth category is the phono-semantic category. Phono-semantic kanji can be broken into two components, the phonetic and the semantic component. An example are the kanji for different types of metal. 銅 (bronze), 鉛 (lead), 銑 (pig iron), 鋼 (steel), 鉍 (ore) all share the semantic component 金 (metal) on their left side, but what does the right side mean? The right side (the phonetic component) does not convey any meaning but rather the on-yomi of the kanji. Unfortunately, due

³<https://www.browserling.com/tools/random-letters>, accessed 04-20-2019

to both the writing of some kanji and their reading changing throughout history, it is not 100% sure that the reading of kanji is the same as the phonetic component. 銀 (silver) is read as *gin* and not *gon* like 良. However, using this information, it is possible to correctly guess the reading of kanji that you have never seen before. 広, 𠂔, 𠂔, 𠂔 and 𠂔 are all part of the phonetic-series of 広, and are all read the same as 広, *kō*. Out of the jōyō kanji, around two thirds are phono-semantic characters and out of these, slightly over half are completely consistent, meaning that around one third of the kanji has their reading decided by their phonetic component. The phonetic component is usually to the right or at the bottom of kanji, but they can appear in other positions.

4 Results

This chapter will reference the tables found in Appendix D. These tables contain the answers from all participants and if one wishes to disregard the answers from some participants for reasons discussed below, the numbers might change. In the appendix, each participant has an ID number which are also referred to when specific participants are discussed. The first and third part of the questionnaire where the participants were asked to guess the readings of the kanji are differentiated as "the first kanji test" and "the second kanji test".

A total of 27 students participated in the questionnaire. Out of the 27 participants, 14 of them guessed correctly more often during the second kanji test. Since 16 of the participants are from Stockholm University and only five from Gothenburg University, three from Lund University and two from Dalarna University, there probably is not much to say about the difference of the participants' answers depending on their university, this coupled with only three second year students participating, with the rest being first year students, no difference will be made between participants from different universities or year groups.

One reason why one may want to disregard some answers is that some participants opted not to answer in one or both of the tests, some participants explicitly said that the reason for not answering on the second kanji test was that they thought they would just answer the same. Some of the participants showed an incredible difference between the first and second kanji test, which could possibly have to do with participants searching for the correct answers in between the two tests. Regardless of the reasons behind their answers, their entries were still kept, but should be taken into consideration when answering the second research question, whether awareness of the phonetic component helps students guess the reading of unfamiliar kanji.

4.1 Kanji knowledge

From Table D.2 we can see that most of the participants use the components of kanji in some way or another when trying to guess the reading of the presented kanji. Only one participant thought that it is impossible to guess the reading from just the kanji. As for how many of the jōyō kanji are keisei kanji, 14 of them answered half or less than half (or that they did not know). Seven participants answered within the range of 50% - 59%. Nine participants answered within the 60-70% range (including the participant who just said "Majority"). Only two participants answered above 70% at 80%. Taking the estimation from Ivarsson (2016:48) of 66% in mind, we can see that only 9 participants were correct

within 10 percentage points. Regarding where they have heard of the phonetic component, most participants seem to have some knowledge of it, and classes or the internet are the most common sources. 10 people have heard of it from class and eight people have heard of it on the internet. Nine participants have not heard of the phonetic component at all. However, this does not mean that they do not know what it is, they could simply be unfamiliar with the specific term "phonetic component". This also applies to the terms used for kanji categorization. Six of the participants found the given information new, 18 of the participants found it partly new, while only three participants did not consider any of the information new. The three participants who did not find the information new had a slight percental change at -3%, 0% and 3% in the second test.

If we look at Table D.1, there does not seem to be an obvious correlation between how many kanji they are familiar with and their general knowledge of kanji. The biggest correlation seems to be that the more kanji you can read, the more likely you are to use the phonetic component actively, but general knowledge of kanji seem irregular. We can also see that participants with knowledge of the phono-semantic category have all but one heard of the phonetic component.

From Table D.3 we can also see that most participants had some knowledge of the information given, but to most, at least some of the information was new. Some changes are dramatic, up to 1000%, while others are not as dramatic. This does however not necessarily mean that the information was less helpful in general for the participants with a lower swing. The higher one scores on the first kanji test, the potential swing is lower. The participants with over 10 correct answers in the first kanji test could be very diligent students, confident in the reading of the kanji from Genki and had an easier time seeing patterns with the kanji already in the first kanji test.

4.2 Change in guesses

From Table D.3 we can see that for 14 participants there was a positive shift toward getting more correct answers after they were given the information about the phonetic component. For three of the participants there was a slight decrease in correct answers. For the remaining 10 participants, there was no difference in correct amount of guesses before and after they were given information. 11 of the participants felt that they had an easier time guessing. Whether the participant felt that it was easier to guess correctly during the second kanji test seems to correlate with the change in correct answers, with the exception of three participants. One of these participants thought it easier the second kanji test and yet had no increase in correct answers. The other participants did not find it easier and yet had a increase from three to five and zero to seven respectively.

One thing to note is that only answers which are 100% correct were considered correct guesses. For instance, one participant answered "kyō" instead of *kō* for the kanji 葉, which

one could say is a pretty close guess. There are also some answers which seem to indicate that the participant did use the phonetic component, but for whatever reason decided to spell out every component in the character. For example "igo" for instead of just *go* for the kanji 箇. Finding the phonetic component, but using the kun-yomi of that component is also common, for example "saki" for 跣, a kun-yomi reading for 先, instead of its on-yomi, *sen*. Looking at one participant (ID 16) specifically and their comment, it seems that just the possibility of being able to guess the reading increased their correct answers from 4 to 20, an increase of 400%. This participant did however already say that they used the phonetic component when guessing in the first kanji test, which could indicate that just the confidence boost or getting reminded of this information, can give a substantial increase in being able to guess correctly.

Three participants (ID 10, 24, 27) in particular are interesting. They all had 0 correct guesses in the first kanji test but had 21, 9 and 7 correct guesses in the second kanji test respectively. This may seem like an unlikely increase and is further discussed in the next chapter.

One participant (ID 21) only had one mistake in the first kanji test, but managed to get it correct during the second kanji test. This kanji was 緘 *kan*, which they mistakenly answered "dou" in the first kanji test. This participant said that they knew the reading of 800 kanji, which is on the higher end of the participants. However, it is still amazing that one participant could score this high. This could also be the case of a participant looking up the reading of the kanji during the questionnaire.

4.3 Individual kanji

Looking at Table D.4 there does seem to be a pattern of a participants guessing more correctly if the phonetic component is on the right hand side of the kanji, as is the most common. 劒, 鳩, 藁, 箇 and 囧 were the only kanji which had a negative or neutral swing. In both 劒 and 鳩 the phonetic component is on the left hand side of the kanji, while in 藁 and 囧, the phonetic component is in the middle. When the phonetic component was on the right side, as in the other kanji of the pair, the swing was positive. One thing to keep in mind is that unfortunately one kanji, 廂 was mistakenly included twice, instead of the kanji 孀 in the first kanji test. This could be the reason why two participants (ID 1, 21) had a change of one correct answer and could also partially contribute to the change in several of the participants answers.

5 Conclusion

5.1 Limitations & Problems

As mentioned in the *Methods* chapter, one issue of this thesis was time constraint. Due to time constraints, all the parts were sent as a single questionnaire. One could argue that it might be better to have a period of time between the second and third part or maybe follow up with a fourth part later on because the results might be skewed if they have just been primed to look for the phonetic component. Another issue was the choice of kanji. Which kanji do the students know? Which phonetic series would be good to choose from? How many kanji should be included in the questionnaire?

Luckily, all the universities used the Genki books in their first year courses, and I was able to find the set of kanji that the students should be comfortable with quite easily. The professors of the universities were also very helpful in confirming which kanji they should be able to read. It was quite hard to find a good list of phonetic series for kanji. I did find a Wiktionary article¹ with an extensive listing of Chinese phonetic series, but it could be hard to remove all the hanzi without kanji counterparts. It also did not list the Japanese reading, which would mean I would need to find the reading for a huge number of kanji. Eventually the spreadsheet² from Boiko (2011) was decided upon as it had information regarding predicted reading, size of phonetic series and how consistent the series were. But this spreadsheet is not without faults. The phonetic series are not based on actual phonetic components, but rather components in general. This meant for example that small components such as ” 丶 ” would have their own ”phonetic series”. Luckily these series had a very low consistency (some lower than 1%), as they are not actual phonetic components. One could also argue that whether a component is officially a phonetic component does not matter, as long as it seems to be a consistency of reading between the kanji it is in.

Another concern is the balance of how consistent and how big the series should be. I ultimately decided on a size of at least 5 kanji and a consistency of 50%. 50% may seem low, but the list of the phonetic series included all on-yomi readings of kanji, not just the most common one. For example, the phonetic series of 𠂔 has 9 kanji in its series, all of which can be read as *han*, but three of the characters also have extra readings, which gives the series a potentially misleading 75% reading coverage.

Regarding the three participants (ID 10, 24 and 27) pointed out in the last chapter, there

¹https://en.wiktionary.org/wiki/Wiktionary:About_Chinese/phonetic_series, accessed 05-05-2019

²https://namakajiri.net/data/kanji/components_phonetic.kanjivg.tsv, accessed 05-05-2019

may have been an issue of them looking up answers in the middle of the questionnaire. Participant 10 entered nothing at all in their first round of guessing, and said "I can't guess. They mean nothing to me." when asked what they were thinking about when guessing the reading of the kanji. The information given was partly new, but an increase from 0 to 21 correct guesses is quite unlikely. Perhaps this could be a case of the participant looking up the answers. This participant did think that it was easier the second time around. Participant 24 did also not answer anything the first time around but unfortunately they did not leave a comment regarding how they base their guesses. They also interestingly did not think it was easier the second time around. Participant 27 did answer the first time around, but they also unfortunately did not leave a comment regarding how they based their guesses. This participant did think that it was easier the second time around. While the drastic change for participant 10 is a cause for alarm, the change for these two may not be.

These issues could perhaps have been avoided by requiring the participants to answer every answer as well as trying to make sure that they would not look up the answers in between the two tests. The reasoning behind not requiring an answer for every single question was that some participants may have quit the questionnaire as there were a total of 72 times that they had to guess the reading of kanji. It was unfortunately not mentioned in the questionnaire that the participants should not look up the reading before they finished the questionnaire, and it could have easily been mentioned. For future studies, a physical test rather than an online one would most likely reduce the chance of these issues occurring.

5.2 Research answers

The two research questions for this thesis were *How much do Swedish students of Japanese know about the phonetic component of kanji?* and *Does awareness of the phonetic component help students guess the reading of kanji?*. Did the research help answering these questions? For the first question, there definitely seems to be some knowledge of the phonetic component among most of the participants, but to which degree is very scattered and less than half heard about it during class. The information about the phonetic component was new or at least partly new to most of the participants of the questionnaire. As for the answer to the second question, half of the participants had a noticeable increase in guessing the readings correctly, some benefiting greatly. Only three participants had a slight decrease while most of the stagnant ones did not appear to try especially hard to guess, which could imply that even more participants could benefit from this. These results indicates that yes, awareness of the phonetic component does indeed help students guess the reading of kanji.

5.3 Pedagogical implications

Given the fact that there seems to be a noticeable improvement for most of the participants after being given a short explanation of the phonetic kanji, maybe it would be a good idea to take advantage of this during lectures. Gothenburg University and Stockholm University does mention the component in their curricula, but it is possible that they would see an improvement if they focused a bit more on it. It is also important to note that just because you can read the on-yomi of a kanji, that does not mean that you can understand the word. While being able to read an additional 50% of kanji sounds great, this does not mean that you will increase your vocabulary very much. Most words using the on-yomi reading are also compounds of several kanji, meaning that you will be less likely to be able to read all of the kanji of a word. One more thing to keep in mind is that almost all participants of this questionnaire were first year students. It is possible that there is a higher occurrence of second year students who already know all of this and they may not benefit as much. As an anecdote I could mention that I did not know about the phonetic component before my second year, but I did notice that there seemed to be a orthographic similarity between kanji that shared readings. When I first learned of the phonetic component during the start of my second year while browsing the internet, I definitely seemed to gain more confidence and could guess readings of unfamiliar kanji more often. I would also imagine that if you put your mind to it and try to remember the most consistent series' phonetic components and their readings, you will be able to read many unfamiliar kanji. Given the small investment and high reward, I do not see why this could not be implemented, or at least taken into consideration when designing the courses. Another issue was the number of participants. While I do think that I managed to answer my research questions even with the number of participants I got, more would be even better. Specifically, more from other universities. More than half my the responses came from Stockholm University. Perhaps one could observe a difference based on the participants' universities if there were more from the other universities. I think I could have gotten a lot more responses from Lund University had I went to one of their classes, but unfortunately the questionnaire was not ready before their last class.

5.4 Future research

As always, an experiment should be repeated before making too many assumptions, if one were to repeat this experiment, I would recommend doing it in a classroom setting with paper and pen rather than an online test. This could avoid previously mentioned issues of the participants not even trying to answer the questions or looking up answers in the middle of the questionnaire. I would also recommend meeting the students in person, or at least join them in the classroom via a live video-chat using a program such as Skype

or Zoom. The professor from Stockholm University and I arranged a Zoom-meeting and it is likely the reason for why there was a great number of responses from Stockholm University.

What would probably be the best, most accurate and also fairly simple to perform experiment, could be to hold a short lecture (the information in this questionnaire could easily be held in less than 15 minutes, including potential questions from students) for one class and see if there is a desirable increase in the students' ability to read kanji compared to previous classes.

Another thing to research could be when the best time is to introduce this aspect of kanji. Introducing it at the end of the first year or the start of the second year seems to be a good point, as the majority of the participants of this questionnaire were first year students. It could also be interesting to do a similar test on native Japanese speakers. I have had some (unrecorded) conversations about this topic with native Japanese speakers, albeit a very small set, and interestingly, they do not all seem to know about the phonetic component, and when giving them similar (unrecorded) tests, they scored about the same as the highest-performing participants of this questionnaire. Another thing to pursue could be how much the position of the phonetic component matters, the results of this thesis' questionnaire showed that students with knowledge of the phonetic component in general may think that the phonetic component is always on the right hand side of the kanji.

Bibliography

- Banno, E. and Ikeda, Y. (2009). *Kanji Look and Learn*, Tokyo: Japan Times.
- Banno, E., Ikeda, Y. and Ohno, Y. (2011a). *Genki: An Integrated Course in Elementary Japanese II*, 2nd edn, Tokyo: Japan Times.
- Banno, E., Ikeda, Y. and Ohno, Y. (2011b). *GENKI I: An Integrated Course in Elementary Japanese*, 2nd edn, Tokyo: Japan Times.
- Boiko, L. (2011). *Testing the predictive power of phonetic components in Japanese kanji*, viewed 2 March 2019, <<https://namakajiri.net/nikki/testing-the-power-of-phonetic-components-in-japanese-kanji/>>.
- Heisig, J. W. (2011). *Remembering the Kanji (Book 1)*, 6th edn, Hawaii: University of Hawaii Press.
- Heisig, J. W. (2012). *Remembering the Kanji (Book 2)*, 4th edn, Hawaii: University of Hawaii Press.
- Ivarsson, F. (2016). *A Study of the L2 Kanji Learning Process: Analysis of reading and writing errors of Swedish learners in comparison with level-matched Japanese schoolchildren.*, PhD thesis, Gothenburg University.
- Oka, M. (2009). *Tobira: Gateway to Advanced Japanese Learning Through Content and Multimedia*, Tokyo: Kuroshio Shuppan.
- Seely, C. (2016). *The Complete Guide to Japanese Kanji: Remembering and Understanding the 2,136 Standard Characters*, Tokyo: TUTTLE Publishing.
- Suzuki, Y. (2007). *Kanji Kyōiku Ni Kansuru Oboegaki (A Note on Kanji-Education)*, Departmental Bulletin Paper, Waseda University, viewed 26 December 2018, <<hdl.handle.net/2065/26475>>.

A Phonetic Series with KanjiVG kanji

Phonetic series with at least 50% reading coverage and a series size of at least 5, with at least one kanji from the Genki book series in each series. The kanji set KanjiVG (6300 kanji). The series are sorted by reading coverage and then size. Reading coverage takes all readings of all kanji into account, which is why the percentage is lower than 100% even though every kanji in some series can be read the same.

Component	Kanji in phonetic series	Predicted readings	Size of phonetic series	Readings coverage %	Kanji with extra readings
安	安按晏案鞍鯪	アン	6	100	
介	介价堺界畍疥芥	カイ	7	87.5	芥
広	広拡眩砧紘鉦	コウ	6	85.714	拡
連	漣漣漣蓮連鏈	レン	6	85.714	漣
知	智痴知蜘踟	チ	5	83.333	智
五	五伍吾唔囿寤悟 晤梧梧語語語	ゴ	14	82.353	囿衛
吾	吾唔囿寤悟晤梧 梧語語語語	ゴ	12	80	囿衛
半	伴判半叛拌畔絆 胖絆	ハン	9	75	伴判叛
官	官棺管館菅館	カン	6	75	縮菅
秋	啾愀愁楸湫鞣秋 萩鞣鞣	シュウ	11	73.333	愀湫鞣
代	代垚岱玳袋貸黛	タイ	7	70	代垚袋
化	化囧花訛貨鉞靴	カ	7	70	化囧花
次	咨姿恣次瓷粢茨 諮資	シ	9	69.231	懿次盜茨
高	囑塙敲槁槁稯稿 縞膏蒿藁鎬高	コウ	13	68.421	亭塙嵩藁
先	先洗潛筵跣銑	セン	6	66.667	讚贊鑽
交	交佼倣効咬效校 狡皎絞纈蛟較郊 餃鮫鵄	コウ	17	60.714	佼咬校狡皎較較 餃鮫
相	孀廂想相箱霜	ソウ	6	60	廂想湘相

兪	儉儉劔劔劔劔劔 嶮検検験験険険 験験嶮	ケン	17	58.621	兪匳斂斂激簽臉 験験嶮
間	嫺潤爛癩簡綱間	カン	7	58.333	潤爛簡綱間
奴	努呶奴拏帑努怒 駑	ド	8	57.143	呶拏帑怒拏
咸	咸喊憾憾憾緘轄 鹹鹹	カン	9	56.25	咸喊減箴轄鹹鹹
無	嚙撫撫撫無蕪	ブ	6	54.545	嚙撫撫無蕪
韋	偉圍幃緯葦違韋	イ	7	53.846	幃衛諱韓韜
弟	剃弟悌梯涕睇鶖	テイ	7	53.846	弟悌梯涕鶖
周	凋彫稠蝸調雕鯛	チョウ	7	53.846	周惆稠綢週
父	交佼倣効咬效校 狡皎絞纈蛟較郊 餃餃鵄	コウ	17	53.125	佼咬斧校父爺狡 皎皎較釜餃駁
古	估個倨涸古固姑 居怙故枯糊沽涸 湖瑚痼箇糊罟胡 葫蛄蝴裾詬踞辜 醐鉅鋸錮餽鵠	コ	34	50.746	個倨倣克兢涸剋 嫡尅居据摘敵涸 滴瑚箇糊胡苦裾 謫踞適餽鋸鎬
乍	乍作咋搾昨柞炸 窄窄昨酢	サク	11	50	乍作咋怎炸柞昨 詐鮓
寺	侍峙寺峙恃持時 峙痔蒔	ジ	10	50	侍峙待恃特峙痔 等蒔詩
弔	剃弟悌梯涕睇第 鶖	テイ	8	50	弔弟悌梯涕第鶖
同	同恫桐洞胴銅	ドウ	6	50	恫桐筒銅興
末	抹末沫秣茱韎	マツ	6	50	末沫秣茱韎

B Genki Kanji

L3:	一	二	三	四	五	六	七	八	九	十	百	千	万	円	時	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
L4:	日	本	人	月	火	水	木	金	土	曜	上	下	中	半		
	16	17	18	19	20	21	22	23	24	25	26	27	28	29		
L5:	山	川	元	気	天	私	今	田	女	男	見	行	食	飲		
	30	31	32	33	34	35	36	37	38	39	40	41	42	43		
L6:	東	西	南	北	口	出	右	左	分	先	生	大	学	外	国	
	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	
L7:	京	子	小	会	社	父	母	高	校	毎	語	文	帰	入		
	59	60	61	62	63	64	65	66	67	68	69	70	71	72		
L8:	員	新	聞	作	仕	事	電	車	休	言	読	思	次	何		
	73	74	75	76	77	78	79	80	81	82	83	84	85	86		
L9:	午	後	前	名	白	雨	書	友	間	家	話	少	古	知	来	
	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	
L10:	住	正	年	売	買	町	長	道	雪	立	自	夜	朝	持		
	102	103	104	105	106	107	108	109	110	111	112	113	114	115		
L11:	手	紙	好	近	明	病	院	映	画	歌	市	所	勉	強	有	旅
	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131
L12:	昔	々	神	早	起	牛	使	働	連	別	度	赤	青	色		
	132	133	134	135	136	137	138	139	140	141	142	143	144	145		
L13:	物	鳥	料	理	特	安	飯	肉	悪	体	空	港	着	同	海	昼
	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161
L14:	彼	代	留	族	親	切	英	店	去	急	乗	当	音	楽	医	者
	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177
L15:	死	意	味	注	夏	魚	寺	広	転	借	走	建	地	場	足	通
	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193
L16:	供	世	界	全	部	始	週	以	考	開	屋	方	運	動	教	室
	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209
L17:	歳	習	主	結	婚	集	発	表	品	字	活	写	真	步	野	
	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	
L18:	目	的	力	洋	服	堂	授	業	試	験	貸	図	館	終	宿	題
	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
	春	秋	冬	花	様	不	姉	兄	漢	卒	工	研	究	質	問	多

Table B.1 continued from previous page

L19:	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256
L20:	皿	声	茶	止	枚	両	無	払	心	笑	絶	対	痛	最	続	
	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	
L21:	信	経	台	風	犬	重	初	若	送	幸	計	遅	配	弟	妹	
	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	
L22:	記	銀	回	夕	黒	用	守	末	待	残	番	駅	説	案	内	忘
	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302
L23:	顔	情	怒	変	相	横	比	化	違	悲	調	査	果	感	答	
	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	

C Potential Phonetic Series

Component	Predicted reading	Kanji from Genki (chapter)	Word from Genki (chapter)	Phonetic series
安	アン	安 (13)	不安 (13)	安 按 晏 案 鞍 鰻
介	カイ	界 (16)	世界 (13)	介 价 堺 界 畛 疥 芥
広	コウ	広 (15)	広告 (13)	広 拡 眩 砵 紘 鉦
連	レン	連 (12)	連絡 (21)	噠 漣 漣 蓮 連 鏈
五	ゴ	語 (7)	日本語 (1)	五 伍 吾 唔 圀 寤 悟 晤 梧 牾 瑯 衙 語 齧
半	ハン	半 (4)	半 (1)	伴 判 半 叛 拌 畔 絆 胖 拌
官	カン	館 (18)	映画館 (15)	官 棺 管 館 菅 館
代	タイ	代 (14)	時代 (14)	代 岱 岱 玳 袋 貸 黛
化	カ	化 (23)	文化 (11)	化 囧 花 訛 貨 鉹 靴
高	コウ	高 (7)	高校 (1)	嚙 塙 敲 槁 犒 藁 稿 縞 膏 蒿 藁 鎬 高
先	セン	先 (6)	先生 (1)	先 洗 潛 筭 跣 銑
交	コウ	交 (7)	高校 (1)	佼 倣 効 咬 效 校 狡 皎 絞 纈 蛟 較 郊 餃 蛟 鵄
相	ソウ	相 (23)	相談 (14)	孀 廂 想 相 箱 霜

兪	ケン	験 (18)	試験 (9)	兪 儉 劔 劔 劔劔劔嶮検 検験臉険険 験験嶮
間	カン	間 (9)	時間 (4)	嫺 涸 爛 癩 簡綱間
咸	カン	感 (23)	感動 (13)	咸喊感憾撼 緘 輓 鹹 鹹
乍	サク	作 (8)	作文 (9)	乍 作 咋 搾 昨 柞 炸 窄 笮 酢 酢
寺	ジ	時 (3)	時間 (4)	侍 埒 寺 峙 恃 持 時 峙 痔 蒔
末	マツ	末 (22)	週末 (13)	抹 末 沫 秣 茱 𪗇

D Data

The results are divided into four tables. The first table shows the responses to some of the short answers. The second table shows paraphrased longer answers. The third table shows correct answers for the two kanji tests, as well as the difference in percentage. It also includes answers to question 10 and 13, which are the most relevant to the potential change in reading accuracy. The fourth table shows the collective correct answers for each individual kanji. For each table, each row represents a single participant. The same participant is represented by the same row in each table. In other words, the participant represented by row 5 in table 1 is also represented in row 5 in the other tables. Each participant has an ID to make it easier to traverse the different tables. The participants are ordered by university and then by year. As the questions are sometimes quite long, they cannot fit neatly into the table and for that reason, all the columns that have the header *Q1*, *Q2*, *Q3*, etc. corresponds to *Question 1*, *Question 2*, *Question 3*, etc. The relevant questions can be seen below each table. Comments from participants are also included at the end of this appendix.

Table D.1: Short answers

ID	University	Year	Q1	Q2	Q4	Q5	Q6	Q7	Q11	Q12
1	Dalarna	1	A few hundreds	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2	Dalarna	1	2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3	Göteborg	1	300	Yes	No	Yes	No	Yes	No	No
4	Göteborg	1	200	Yes	Yes	Yes	No	Yes	No	No
5	Göteborg	1	100-150	Yes	No	Yes	Yes	Yes	No	No
6	Göteborg	2	200	Yes	Yes	No	No	Yes	Yes	No
7	Göteborg	2	700	Yes	Yes	No	No	Yes	Yes	Yes
8	Lund	1	30	Yes	No	No	No	No	No	No
9	Lund	1	About 340-350	Yes	No	No	Yes	Yes	Yes	No
10	Lund	1	~200	No	No	Yes	No	No	No	No
11	Stockholm	1	1000	Yes	Yes	No	No	Yes	Yes	Yes
12	Stockholm	1	100	No	No	Yes	No	Yes	No	Yes
13	Stockholm	1	400	No	No	No	No	No	No	No
14	Stockholm	1	100	Yes	Yes	No	Yes	Yes	Yes	No
15	Stockholm	1	350	Yes	No	Yes	No	No	No	No
16	Stockholm	1	350	Yes	No	No	No	Yes	Yes	Yes
17	Stockholm	1	400	No	Yes	Yes	Yes	Yes	Yes	Yes
18	Stockholm	1	150	Yes	No	Yes	No	No	No	Yes
19	Stockholm	1	150	Yes	No	Yes	Yes	Yes	No	No
20	Stockholm	1	approximately 400	Yes	No	Yes	Yes	Yes	Yes	No
21	Stockholm	1	800	Yes	Yes	Yes	Yes	Yes	Yes	Yes
22	Stockholm	1	150	Yes	No	Yes	Yes	Yes	No	No
23	Stockholm	1	100	Yes	No	Yes	No	No	No	No
24	Stockholm	1	200	No	No	No	Yes	No	No	No
25	Stockholm	1	200	Yes	Yes	Yes	Yes	Yes	Yes	Yes
26	Stockholm	2	Like over 100	No	No	Yes	No	No	No	No
27	Stockholm	2	40	Yes	No	Yes	Yes	Yes	No	No

Q1: How many kanji do you think that you can read? (on-yomi) **Q2:** Do you practice writing or reading Japanese outside of the class-material? **Q4:** Do you find that you can sometimes correctly guess the reading of kanji that you have never seen before? **Q5:** Did you know that kanji can be grouped into different categories? **Q6:** Have you heard about the phono-semantic/phonetic-ideographic category of kanji? **Q7:** Have you ever heard of the phonetic component of kanji? **Q11:** Do you consciously use the phonetic component when trying to guess the reading of kanji? **Q12:** Do you consciously use the phonetic component when trying to remember the reading of kanji?

Table D.2: Longer answers

ID	Q3	Q8	Q9
1	2nd part of a kanji	On the Internet	20%
2	Radicals	On the Internet	20%
3	Nothing	On the Internet	Minority
4	Trees and fish etc	The Kanjibook (genki)	50%
5	Phonetic component	In class	50% - 60%
6		Nowhere	Majority
7	Radicals	In class	70%
8	Particles	I have not heard about it	No idea
9	Radicles	On the Internet	60%
10	I can't guess. They mean nothing to me.	I have not heard about it	I don't know.
11	Parts, "feeling"	In class	50%
12	Nothing, impossible to guess	In class	80%
13	Parts I know from other kanji	I have not heard about it	30%
14	Components, wild guesses	In class	60%
15		I have not heard about it	No idea
16	Radicals, common on-yomi	From classmates, On the Internet	30%
17	Parts I know	From classmates	60%
18	Radicals, wild guesses	I have not heard about it	70%
19	Elements, "general feel"	In class, On the Internet	15%
20	Phonetic component	In class, On the Internet	70%
21	Phonetic component	On the Internet	65%
22	Similar kanji, wild guesses	In class, From classmates	50%
23	Parts	I have not heard about it	60%
24		I have not heard about it	50%
25	Right hand side of kanji	In class, From classmates	80%
26	Instinct and radicals	I have not heard about it	33%
27		In class	60%

Q3: When guessing the readings of the kanji, what did you think about? **Q8:** Where have you heard about the phonetic component? **Q9:** How many percentage of jouyou kanji do you think are phono-semantic/phonetic-ideographic?

The full, un-edited answers for Question 3 can be found on the next page.

1. "I have learned that the 2nd part of a kanji is often the on reading, while the first denote meaning"
2. "radical"
3. "Nothing, it's very hard to think of a lead that would make me come up with a possible reading. I've heard radicals can help you but I have no idea in what way."
4. "Trees and fish etc, I really don't get all the coincidental multi-meaningbearing short sounds yet. Haven't figured out how you" would talk about a spec kanji without writing it down and showing it. But guessing you express it WITH the trail of hiragana but still, then its also several different words.. so wierd haha."
5. "Phonetic component"
6. ""
7. "I look at the radicals and if it has a radical I know from another kanji I often assume it has the same reading. It can be difficult however if it has 2 different radicals that I believe is the radical containing the reading."
8. "The what are they called, particles they consist of"
9. "The radicals (just a little bit though, as I don't know how, when or where in the kanji they ever effect the pronunciation or if radicals that by themselves also are kanji have a separate pronunciation as radicals). For some I simply used "feeling"."
10. "I can't guess. They mean nothing to me."
11. "On reading of the parts they consist of (or in the case of some of them I already new the on reading)"
12. "I could not think of anything, in my point it is impossible to figure out the onyomi by just randomly guessing."
13. "Parts that I knew how to read/pronounce"
14. "Some are completely wild guesses, others I guessed based on components that I know the on-yomi of."
15. ""
16. "What common on-yomi exist, and also the radicals as they sometimes contribute to readings."
17. "Identifying one part of the kanji with an on-yomi reading I know, and guessing that reading."

18. "I try to look at the radicals of the Kanji, or take a wild guess. "
19. "Their elements and general feel"
20. "I thought about which kanji can be found as components, and guess the reading based on what I think is the phonetic component. "
21. "I tried seeing if there were any obvious phonetic component(s) in the kanji, or parts that were similar to other kanji"
22. "1.If there was anything i recognised from other kanji i knew, maybe some with similar symbols have the same reading." 2. Straight up guessing, but with the ""rules"" of kanji in mind. I have so far not seen a single kanji with more than 3 (maybe 4) mora/moura (don't know how that is spelled). So I made my guesses with that ""rule"" in mind. I have also seen A LOT of しゅ、しゃ、ぎょ、じゅ and so on, so I made sure to include some of those in there."
23. "Guessed basing on kanjis that they had parts of"
24. ""
25. "I mostly guessed the reading by comparing the right side of the kanji to kanji that I know."
26. "Instinct and radicals"
27. ""

Table D.3: Before and After

ID	Correct answers before explanation	Correct answers after explanation	% change	Q10	Q13
1	29	28	-3%	No	No
2	6	6	0%	Yes	Partly
3	1	11	+1000%	Yes	Partly
4	3	3	0%	No	Partly
5	21	18	-14%	No	Partly
6	15	19	+27%	Yes	Partly
7	26	26	0%	No	Partly
8	0	0	0%	No	Yes
9	3	17	+467%	Yes	Partly
10	0	21	∞ %	Yes	Partly
11	22	20	-9%	No	Partly
12	0	0	0%	No	Yes
13	3	5	+67%	No	Yes
14	14	21	+50%	Yes	Partly
15	0	0	0%	No	Yes
16	4	20	+400%	Yes	Partly
17	21	23	+10%	Yes	Partly
18	1	5	+400%	Yes	Yes
19	0	0	0%	No	Partly
20	13	19	+46%	Yes	Partly
21	35	36	+3%	Yes	No
22	1	11	+1000%	Yes	Partly
23	0	0	0%	No	Yes
24	0	9	∞ %	Yes	Partly
25	25	25	0%	No	No
26	1	1	0%	No	Partly
27	0	7	∞ %	No	Partly

Q10: Was the information about keisei kanji and the phonetic component new to you?

Q13: Do you feel that you had an easier time guessing the reading of kanji the second time around?

Table D.4: Change for individual kanji

衙	梧	按	鯨	藁	犒
8 / 8	7 / 11	8 / 14	6 / 14	8 / 7	9 / 16
(0%)	(57,1%)	(75%)	(133,3%)	(-12,5%)	(77,8%)
跣	筧	絆	裨	館	棺
11 / 16	8 / 16	10 / 14	9 / 13	11 / 16	11 / 15
(45,5%)	(100%)	(40%)	(44,4%)	(45,5%)	(36,4%)
岱	玳	劒	鯨	抹	茱
6 / 7	11 / 15	6 / 4	7 / 10	5 / 10	7 / 9
(16,7%)	(36,4%)	(-33,3%)	(42,9%)	(100%)	(28,6%)
廂	孀 *	鯨	鵠	撼	緘
1 / 2	? / 2	7 / 11	5 / 4	11 / 15	8 / 12
(100%)	(? %)	(57,1%)	(-20%)	(36,4%)	(50%)
銚	囧	峙	埶	漣	蓮
4 / 6	5 / 5	4 / 7	4 / 7	5 / 7	6 / 7
(50%)	(0%)	(75%)	(75%)	(40%)	(16,7%)
价	芥	柞	搾	澗	癩
5 / 9	5 / 8	6 / 8	5 / 8	9 / 13	5 / 6
(80%)	(60%)	(33,3%)	(60%)	(44,4%)	(20%)

This table shows the correct amount of guesses for each kanji in the first kanji test and the second kanji test. It also shows the swing in percentage points. Unfortunately, in the first kanji test, the kanji 廂 was mistakenly included twice, instead of the kanji 孀, which means that there was no data for it during the first test.

D.1 Comments

4. "I did the best I could the first time, sry. "
7. "I'm sorry I just skipped it the second time because I feel that I would have put in the exact same answers since I pretty much used this method already. This was fun though and I learned some new stuff, good luck on your report!"
13. "This is really good information to know, but I felt that I couldn't really use it on many of the kanji because I don't know/ remember the onyoumi readings for many of the kanji. "
16. "I feel like without knowing about the phonetic components, you feel a lot less confident in guessing the readings. The other time around, despite not really knowing any more, just knowing that I had a certain chance of being correct was better than just guessing blindly, so it gives you more confidence to guess."

E Questionnaire

Reading of kanji

This questionnaire aims to measure your ability to guess the on-yomi (Chinese reading) reading of kanji that you have not learned.

**Obligatorisk*

1. Which University do you study at? *

Markera endast en oval.

- ☐ Gothenburg University
- ☐ Lund University
- ☐ Stockholm University
- ☐ Dalarna University
- ☐ Övrigt: _____

2. Which year are you in? *

Markera endast en oval.

- ☐ 1st year
- ☐ 2nd year
- ☐ Higher

3. How many kanji do you think that you can read (on-yomi)? *

4. Do you practice writing or reading Japanese outside of the class-material? *

Markera endast en oval.

- ☐ Yes
- ☐ No

Guess the reading! - Part 1

Keep in mind that I assume that you have not yet learned these 20 kanji. Even if you don't have a clue on how to read the kanji, just try to guess the on-yomi. Please use either hiragana (あ), katakana (ア) or romaji (a/A) when filling in.

5. 衛

6. 按

Reading of kanji

<https://docs.google.com/forms/d/1xRUA-HajK6FtR9HrHcljfNlhqCYd...>

7. 潤

8. 藁

9. 跳

10. 絆

11. 梧

12. 舘

13. 岱

14. 劔

15. 玳

16. 棺

17. 抹

18. 裨

19. 廂

20. 絞

21. 撼

22. 銚

23. 絨

24. 齧

25. 峙

26. 漣

27. 价

28. 囧

29. 柞

30. 痢

31. 蓮

32. 擗

33. 笄

34. 犒

35. 廂

36. 鳩

37. 耒

38. 塹

39. 芥

40. 鯨

Questions about kanji

41. When guessing the readings of the kanji, what did you think about?

42. Do you find that you can sometimes correctly guess the reading of kanji that you have never seen before? *

Markera endast en oval.

- ☐ Yes
☐ No

43. Did you know that kanji can be grouped into different categories? *

Markera endast en oval.

- ☐ Yes
☐ No

44. Have you heard about the phono-semantic/phonetic-ideographic category of kanji? *

Markera endast en oval.

- ☐ Yes
☐ No

45. Have you ever heard of the phonetic component of kanji? *

Markera endast en oval.

- ☐ Yes
☐ No

The phonetic component & jouyou kanji

The phonetic component is one of the components that can make up a kanji. There are several types of kanji, the ones making use of this component are the phono-semantic/phonetic-ideographic (keisei) kanji. The jouyou kanji are the 2136 kanji that are taught in Japanese schools from primary until high-school and are the ones that allowed to be used in official governmental documents.

46. Where have you heard about the phonetic component? *

Markera alla som gäller.

- ☐ In class
☐ From classmates
☐ On the Internet
☐ I have not heard about it
☐ Övrigt: _____

47. How many percentage of jouyou kanji do you think are phono-semantic/phonetic-ideographic? *

The Phonetic Component & Keisei kanji

Kanji can be distributed into four different categories (pictographic/pictograms, diagrammatic/simple

ideograms, semantic/compound ideograms and phono-semantic/phonetic-ideographic). Pictographic kanji are depictions of something concrete such as a tree/mountain (木/山), diagrammatic are more abstract such as up/down (上/下), semantic kanji combines characters from the previous categories to represent words such as forest/mountain pass (森/峠). The fourth category is the phono-semantic category. Phono-semantic kanji can be broken into two components, the phonetic and the semantic component. An example are the kanji for different types of metal. 銅 (bronze), 鉛 (lead), 鉄 (pig iron), 鋼 (steel), 鉱 (ore) all share the semantic component 金 (metal) on their left side, but what does the right side mean? The right side (the phonetic component) does not convey any meaning but rather the on-yomi of the kanji. Unfortunately, due to both the writing of some kanji and their reading changing throughout history, it is not 100% sure that the reading of kanji is the same as the phonetic component. 銀 (silver) is pronounced gin and not gon like 良. However, using this information, it is possible to correctly guess the reading of kanji that you've never seen before. 広, 𠂔, 𠂔, 𠂔 and 𠂔 are all part of the phonetic-series of 広, and are all pronounced as 広 (kou). Out of the jouyou kanji, around two thirds are phono-semantic characters and out of these, slightly over half are completely consistent, meaning that around one third of the kanji has their reading decided by their phonetic component. The phonetic component is usually to the right or at the bottom of kanji, but they can appear in other positions.

48. Is this information new to you? *

Markera endast en oval.

- ☐ Yes
☐ No
☐ Partly

49. Do you consciously use the phonetic component when trying to guess the reading of kanji? *

Markera endast en oval.

- ☐ Yes
☐ No

50. Do you consciously use the phonetic component when trying to remember the reading of kanji? *

Markera endast en oval.

- ☐ Yes
☐ No

Guess the reading! - Part 1

Keep in mind that I assume that you have not yet learned these 20 kanji. Even if you don't have a clue on how to read the kanji, just try to guess the on-yomi. Please use either hiragana (あ), katakana (ア) or roomaji (a/A) when filling in.

51. 衡

52. 按

53. 潤

54. 藁

55. 跳

56. 絆

57. 梧

58. 舘

59. 岱

60. 劔

61. 玳

62. 棺

63. 抹

64. 裨

65. 嬬

66. 鯨

67. 撼

68. 銚

69. 絨

70. 齟

71. 峙

72. 漣

73. 价

74. 囧

75. 柞

76. 痢

77. 蓮

78. 擗

79. 笄

80. 犒

81. 廂

82. 鳩

83. 秉

84. 塙

85. 芥

86. 鯨

End of Questionnaire questions

Thank you so much for participating in this questionnaire.

87. Do you feel that you had an easier time guessing the reading of kanji the second time around? *

Markera endast en oval.


☐ Yes

☐ No

Reading of kanji

<https://docs.google.com/forms/d/1xRUA-HajK6FtR9HrHcljfNlhqCYd...>

88. Any other comments?

Tillhandahålls av
 Google Forms

