Following the rules of *rendaku* in Japanese

Native speakers' production of novel compound words

Fabian Sturk

LUND UNIVERSITY

Bachelor's Thesis

Japanese Studies

Spring semester 2018

Lund University

Centre for Languages and Literature

Supervisor: Shinichiro Ishihara

Abstract

The goal of this thesis is to gain knowledge on the psychological reality of some known rules, conditions, and categorizations that have been observed in the phenomenon of Japanese "rendaku" (also known as sequential voicing). The results presented in this thesis are based on a quantitative study that was aimed at native speakers of Japanese. The conducted experiment for this study was designed as a forced-choice test, where the respondents had to select their preferred reading of novel compound words. Eight groups of words, each consisting of five compound words, were tested. Four groups tested the psychological reality of different lexical *rendaku* personalities proposed by Rosen (2001), and one of these groups consisted of words, found by Irwin (2009), which seemingly break Rosen's Rule, a rule based on a prosodic size requirement. The four remaining groups of words tested four different known rules or conditions that tend to systematically either trigger or block *rendaku*. The results show clear differences in Rosen's categorization of words, which suggests that native speakers of Japanese possess an internal categorization as well. However, it is also evident from the results that the internal and lexical categorizations differ considerably from each other. The results further show that lexically rule breaking words translate into irregularities in native speakers' production of novel compound words as well. The results from the four groups of words that tested known rules and conditions only reinforce their positions as rules and conditions.

Keywords: rendaku, sequential voicing, Japanese, Rosen's Rule, prosodic size

Table of Contents

Conventions and abbreviations

General conventions

When a section number is preceded by a §, a section within the thesis is being referenced, e.g. "§2.1" would reference section 2.1 in this thesis.

Romanization

With the exception of place names and/or proper nouns, which will be written as they are commonly written, all Japanese will be romanized using the *kunrei* system in this thesis.

Abbreviations

FJ	Foreign Japanese
NJ	Native Japanese
OCP	Obligatory Contour Principle
PSR	Prosodic Size Requirement
RDB	Rate of Devoice Blocking
RoD	Rate of Devoicing
RoO	Rate of Optionality
RRB	Rate of Rendaku Blocking
RRO	Rate of Rendaku Occurrence
SJ	Sino-Japanese
SLI	Specific Language Impairment

List of Figures

List of Tables

1 Introduction

1.1 Research topic

When compound words (referred to only as compounds hereafter) are created in Japanese, an initial voiceless obstruent of a non-initial element can become voiced. In literature written in English, this phenomenon of sequential voicing in Japanese is commonly referred to as either "Japanese rendaku", or only "rendaku". From this point on in this thesis, the term rendaku will be used without italicization.

When you create a compound from the the words *hana* 'nose' and *koe* 'voice' you get the compound *hana**g**oe* 'nasal voice', as illustrated in (1), where we can see the change from the voiceless obstruent ***k*** to its voiced counterpart ***g***.

(1) [*hana* + *koe*] → *hana**g**oe*

There are several voiceless obstruents in Japanese that can become voiced, listed below in Table 1.

Table 1: Obstruents in Japanese

	Voiceless obstruents		Voiced obstruents
a.	[ɸ]	→	[b]
b.	[h][ç]	→	[b]
c.	[t]	→	[d]
d.	[ts]	→	[dz]
e.	[tɕ]	→	[dʑ]
f.	[s]	→	[dz]
g.	[ɕ]	→	[dʑ]
h.	[k]	→	[g]([ŋ])

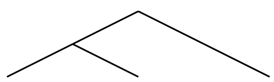
There are some conditions that usually need to be met for rendaku to occur in a compound. There are also constraints that prevent rendaku from occurring. Words that are neither pure nouns[1] nor in the etymological stratum native Japanese (NJ)

[1]Pure nouns are nouns that are neither derived from verbs (such as *yomi* 'reading'), nor adjectives (such as *kuro* 'black').

tend to not undergo rendaku. There are two other main strata of words in Japanese that are not considered native Japanese; Sino-Japanese (SJ) and foreign Japanese (FJ). Although some SJ and FJ words do exhibit NJ-like behavior, they are few in number. Words outside of the NJ stratum that undergo rendaku tend to be older, and might have become "nativized" or "Japan-ized" due to its long and regular usage in the language (see discussion in Kawahara 2016: 39).

Moreover, rendaku is only allowed in the right branches of a compound tree, whereas a left-branching element never voices. This is known as the Right-Branch Condition, proposed by Otsu (1980: 219) and refined later by Itô & Mester (1986). This condition can easily be illustrated with three-element compound trees, which show the two steps in which the compounds are created. If both the second and third element contain a voiceless initial obstruent we can test the effects of the Right-Branch Condition. Using the words *nise* 'fake', *tanuki* 'raccoon', and *siru* 'soup', we can create compounds with two semantically different interpretations. In the first compound, [['fake'+'raccoon']+'soup'], rendaku occurs on both the second and third compound element, as can be seen in (2a). However, in the second compound, ['fake'+['raccoon'+'soup']], the unvoiced *tanuki*, instead of the voiced _*__d__anuki_, appears in a left-branching position. Thus, as can be seen in (2b), this element does not undergo rendaku.

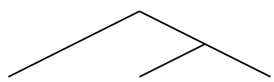
(2) a.



nise ***d****anuki* *ziru*

[['fake'+ 'raccoon'] +'soup']

b.



nise *tanuki* *ziru*

['fake'+['raccoon'+ 'soup']]

In other words, it is possible to distinguish two different semantic interpretations of a three-element compound only by looking at rendaku occurrence.

Furthermore, there is a constraint discovered by Motoori (1822) and Lyman (1894), most likely independent of each other, which is either called Motoori-Lyman's Law, or more commonly Lyman's Law. This constraint states that if there already is a voiced obstruent present in an element, an initial voiceless obstruent of that element cannot undergo rendaku. It is generally understood that Lyman's Law blocks rendaku at a rate of almost 100% (see Itô & Mester 1986, etc.). The word *tabi* 'trip' in the compound *hunatabi* 'trip by boat' therefore remains unvoiced, instead of becoming *hunadabi.

(3) [*huna* + *tabi*] → *hunatabi*

There are, however, exceptions to Lyman's Law, albeit of a small number, similarly to the exceptions of rendaku in non-NJ words (see discussion in Vance 2016: 8). One of the words that violate Lyman's Law is the word *hasigo* 'ladder'. This is a word that already contains the voiced obstruent ***g***, but can nonetheless be found with rendaku in the compound *nawa**b**asigo* 'rope ladder'.

Yet another group of words that tend to block rendaku are *dvandva* or copulative compounds (Irwin 2016b: 85). Although they are included in this group, noun-noun reduplicative words where one word is repeated to create a two-element compound, have been discovered to trigger rendaku in most compounds (see discussion in Rosen 2001: 42; Irwin 2009: 181, 2016b: 83; Vance 2017, etc.). Irwin (2016b: 85) calls these compounds "systematic forcers" of rendaku. Reduplicative compounds even have the ability to trigger rendaku in words that are otherwise considered immune to rendaku. One example of this can be seen in the reduplicative compound *sakizaki* 'distant future' which comes from *saki* 'tip, previous', a word that both Rosen (2001: 256) and Irwin (2016b: 105) claim to be immune to rendaku.

Honorific prefixes, such as *o-*, *go-*, and *mi-*, heavily block rendaku in compounds according to Nakagawa (1966: 314), and Nakagawa further argues that numerical prefixes also heavily dampen rendaku. However, the inclusion of numerical prefixes, as a general condition that heavily dampens rendaku, needs some refining according to a study by Irwin (2012). In his study, Irwin still suggests that even though numerically-related prefixes either heavily dampen or sometimes block rendaku entirely, they should be in the NJ stratum and should not include full numerals (2012: 34). One of the numerical prefixes that always block rendaku is *hito* 'one' (Irwin 2012: 32), and it can be found in the compound *hitotoori* '(in) general' among others.

Despite the seemingly large amount of proposed conditions and constraints surrounding rendaku, there are still irregularities. Linguists have tried, and are still trying, to make sense of rendaku, but regardless of the amount of knowledge that is gained on the subject, new problematic behavior continue to surface.

1.2 Research question and hypothesis

This thesis is based on a couple of research questions that arose from a place of curiosity while reading literature on rendaku. Despite the large number of studies on rendaku, few seem to involve native speakers' production of rendaku in real words. Due to this, the research questions behind this thesis are as follows:

(i) Are the known rules, conditions, and categorizations of rendaku actually realized when native speakers produce novel compounds?

(ii) Do native speakers produce novel compounds with the same irregularities that can be found in existing compounds?

This study was made under the hypothesis that one should be able to observe general tendencies in accordance to known rules, conditions, and categorizations when native speakers produce novel compounds. However, rule breaking behavior is probably more common than studies based on lexical data might suggest.

2 Previous research

Rendaku has been an established topic in Japanese linguistics since the late 20th century (see Suzuki 2004: 19), and extensive research and tests have been conducted on the subject. In this section, previous research that is relevant to the study in this thesis will be discussed. Unanswered and/or remaining questions on rendaku will also be discussed briefly at the end of this section.

2.1 Categorization

Since different groups of words behave differently when it comes to rendaku, it is important to be able to distinguish between these groups in discussion. There are groups that tend to inhibit rendaku, groups that trigger rendaku, and then there are groups that seem to exhibit inherently unpredictable behavior. This section will discuss some groups of words relevant to the study in this thesis, and how they have been categorized by others.

There are different lexical strata present in the modern Japanese language. The three major strata, all of which are etymological, are as follows:

(i) native Japanese (NJ) words, *wago* in Japanese, which are native to the language,

(ii) Sino-Japanese (SJ) words, *kango* in Japanese, which originate from Chinese, and

(iii) foreign Japanese (FJ) words, *gairaigo* in Japanese, which are loanwords from other languages besides Chinese.

Out of the three main strata in Japanese the NJ stratum has the highest percentage of words that can undergo rendaku. Vance (1996: 25) finds that 87% of all NJ words can undergo rendaku, and Irwin (2016b: 93) finds that NJ pure nouns undergo rendaku around 80% of the time. In contrast to this, Vance (1996: 25) finds that only 10% of SJ words can undergo rendaku. In another study on the rendaku rate based on lexical strata, Sano (2014: 184) finds that NJ words undergo rendaku in approximately 66% of compounds in spontaneous speech, while SJ words undergo rendaku in 44%. Rendaku occurrence in the FJ stratum, however, is rare and there are only a couple of known examples (see Ohno 2000: 154; Irwin 2005: 132-133, 2009: 181, 2016b: 85, among others). A fourth stratum (mimetic Japanese, e.g. *onomatopoeia*) is sometimes added to this list of strata. This stratum, however, is neither of etymological nature nor is it relevant to this study and will therefore not be included.

Rendaku eligible Japanese words are divided into five parts of speech. Firstly, there are three different types of nouns; pure nouns, deverbal nouns, and deadjectival nouns. Secondly, there are verbs, and finally adjectives. Analyses on rendaku behavior by parts of speech can be found in Irwin (2016b), where he uses a database called *The Rendaku Database*[2] (Irwin & Miyashita 2013). Using this database, Irwin finds that 58% of every second element are pure nouns and another 29% are deverbal nouns, together making up 88% of every compound in the database (2016b: 89). Next are verbs at 10%, and finally adjectives and deadjectival nouns both at 1% (Irwin 2016b: 90). In a later part of his analyses of the database, Irwin looks at the rate of rendaku occurrence[3] (hereafter RRO) for every part of speech. Irwin finds that the part of speech with the highest RRO is deadjectival nouns at .823 (Irwin 2016b: 93). Following are pure nouns, which have an RRO of .806. Deadjectival nouns and pure nouns the only parts of speech above the average RRO in the database, which is .708. Deverbal nouns have an RRO of .701, adjectives have an RRO of .623, and verbs have the lowest RRO at .184 (Irwin 2016: 93).

Another way of categorizing words is entirely based on their RRO. Rosen (2001: 40)[4] defines three categories based on this criteria;

[2]Irwin uses Version 2.0 of *The Rendaku Database*, which is comprised of 34,359 compounds in its unaltered state. However, when the database is mentioned in this thesis, a modified version called "DATABASE" in Irwin (2016b) is the one referred to. This DATABASE has problematic compounds excised from it and contains 27,856 compounds (Irwin 2016b: 85).

[3]The rate of rendaku occurrence is shown on a scale from 1.000 where 100% of the compounds undergo rendaku, to .000 where 0% of the compounds undergo rendaku.

[4]Rosen's definitions are based on compounds gathered from three sources; the 1999 edition of the NHK Accent and Pronunciation Dictionary, the Kodansha Japanese-English Dictionary, and Version 1.31 of Stephen Chung's Freeware "JWP" Japanese Word Processor (Rosen 2001: 28).

(i) "Immune to Rendaku" which undergo rendaku 0% of the time, i.e. they have an RRO of .000,

(ii) "Rendaku Haters" which have an RRO below .333, and

(iii) "Rendaku Lovers" which have an RRO higher than .666.

Rosen states that it is extremely rare for words to be in the range between .333 and .666, with the only compound being *kawa* 'skin' that voices in two out of six compounds found by Rosen, an RRO of .333 (2001: 40). Rosen does not give a name for words that lie in the middle of rendaku haters and rendaku lovers. Irwin, however, proposes an expanded categorization of words based on their RRO (Irwin 2016b). Irwin expands Rosen's categories from three to five. Irwin decides to give the words in between rendaku haters and rendaku lovers the name "rendaku waverers", which Irwin finds are not as rare as Rosen asserts (Irwin 2016b: 103).

Figure 1: Number of rendaku rate personalities

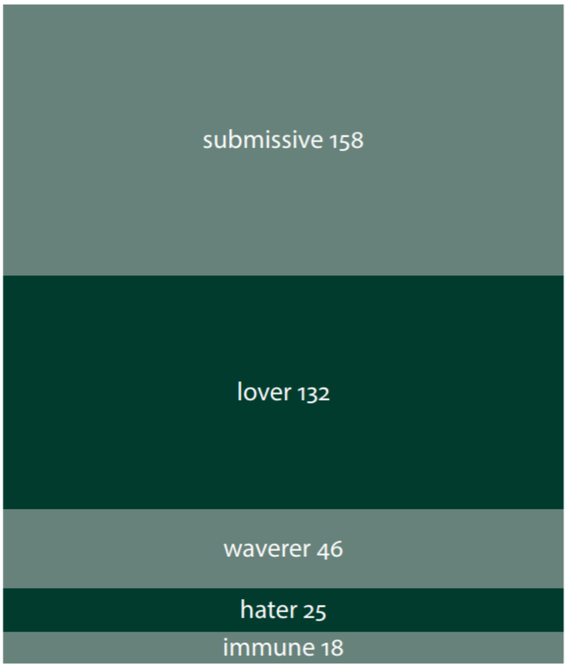


Figure from Irwin (2016b: 103).

This category of waverers comprises 12% of *The Rendaku Database* (Irwin & Miyashita 2013), which is only the third smallest category above haters at 7% and immune words at 5% (Irwin 2016b: 103). The largest category Irwin calls "rendaku submissive" which contains words that always undergo rendaku, i.e. an RRO of 1.000 (Irwin 2016b: 102-103). Rosen puts words with an RRO of 1.000 in the rendaku lover category. When Irwin separates these categories into two, submissive words exist in a larger quantity than those that are lovers, and they comprise 42% and 35% of the database respectively (see Figure 1 with actual numbers for each category).

2.2 Prosodic size and Rosen's Rule

Until the start of this century there had not been many studies on how prosodic size[5] affects rendaku. This, however, changed when Rosen (2001, 2003) proposed a rule which he claimed could predict rendaku behavior when specific requirements on prosodic size are met. This proposed rule has later become known as Rosen's Rule (see Kawahara & Sano 2014c; Vance 2015; Irwin 2016a, etc.), and has been tested both statistically and psychologically. The rule and its definitions, as well as studies on this rule, will be discussed below. Rosen (2001, 2003) proposes the idea of a prosodic size rule that can predict rendaku in a compound if the second member is not rendaku immune. Rosen states that every rendaku eligible non-immune word will undergo rendaku if this prosodic size requirement (hereafter PSR) is met. The specifics of the rule are, however, somewhat unclear (see discussion in Irwin 2016a). This has caused researchers to have different interpretations of Rosen's Rule. Different interpretations of the rule and its requirements has in turn lead to results that are sometimes difficult to compare and relate to results from other studies. Irwin defines two requirements for Rosen's Rule based on Rosen's (2001, 2003) definitions:

(i) **Narrow Rule PSR:** both elements must be native Japanese pure nouns, both elements must be at least 2µ, one element at least 3µ

[5]Prosodic size refers to the moraic length of a word or compound. Moraic length is written as the number of moras in a word followed by the µ symbol. E.g. the prosodic size of the two-element compound *aburagami* 'oil paper', [*abura* + *kami*], is written as 3µ2µ.

(ii) **Broad Rule PSR:** both elements must be native Japanese pure nouns, one element must be at least 3μ Irwin (2016a: 111)

In this thesis, Irwin's definition of the requirements for Rosen's Rule (listed above) will be used. The decision of choosing this definition will result in some criticism of two conducted studies on the psychological reality of Rosen's Rule (Kawahara & Sano 2014c; Sano 2014), which will be discussed later in this section. When Rosen (2001) tests the viability of his proposed rule, he uses a corpus of rendaku eligible compounds (see footnote 4 in §2.1). Rosen finds that 100% of compounds that pass the PSR confine to the proposed rule, and states as a result that one can predict rendaku in these compounds. Rosen restricts compound constructions which he considers rendaku eligible from the following criteria:

(i) Both elements of a two-element compound are NJ pure nouns

(ii) The compound is non-*dvandva*

(iii) The compound does not violate Lyman's Law

Irwin (2009) sets out to test Rosen's Rule, and to show, with the help of an empirical analysis, that further restrictions are needed when deciding if a compound should be eligible for Rosen's Rule or not. Furthermore, Irwin also wants to show that even if one redefines these restrictions, rule breakers do exist. Irwin (2009) uses the Narrow Rule definition when testing Rosen's Rule. Irwin (2009: 191-192) expands Rosen's eligibility criteria by adding the following conditions:

(i) Elements cannot be names, nor should they be obsolete

(ii) All *dvandva* must be disallowed, including reduplicative compounds

With these modifications to Rosen's Rule, Irwin only finds five words that do not behave predictably. Four of these rule breaking words have expected rendaku immunity, but are nonetheless found voiced in some compounds. These four words are *hotaru* 'firefly', *hutokoro* 'bosom'[6], *kanna* 'plane', and *kata* 'person' (Irwin 2009:

[6] The irregular rendaku behavior of *hutokoro* 'bosom' has been noted in other work as well (see Irwin 2005: 130; Vance 2015: 412, etc.).

190-191). The final word that Irwin finds to break his modified version of Rosen's Rule is *koori* 'ice', a word with expected rendaku lover behavior that can be found unvoiced in at least one compound (2009: 191). Irwin, however, points out that while *koori* itself is not obsolete, the compound in which it can be found unvoiced, *kanakoori* 'extremely cold object' ('icicle' in some modern dialects), could be considered "infrequent and probably non-standard" (Irwin 2009: 191). Despite the rule breakers, Irwin's results support the idea of a rule based on Rosen's work. Vance (2015) comes to the same conclusion as Irwin, though by using the Broad Rule definition, that the effects of Rosen's Rule is visible in a majority of PSR pass compounds.

Irwin (2016a) tests Rosen's Rule once again, however, this time using data from *The Rendaku Database* (see footnote 2 in §2.1). With the restrictions set up by Rosen (2001, 2003), further added to by Irwin (2009: 191-192), and finally defined as two versions by Irwin (2016a: 111), 1,865 PSR pass compounds from the database were left to be tested. Irwin finds that both the Narrow Rule and the Broad Rule has an accuracy of 89% (2016a: 112). These results support the viability of Rosen's Rule in a similar fashion to the results from both Irwin (2009) and Rosen's original study (2001, 2003). However, this level of accuracy of Rosen's Rule can only be reached when you put heavy restrictions on compound eligibility, as can be seen in further analysis by Irwin (2016a: 114), where he, step by step, loosens the restrictions. The following analysis by Irwin only tests the Broad Rule. Already in the first step where he allows any NJ noun type for both the first and second element, the accuracy drops to 80%. In the final step, Irwin allows any word as the first element, regardless of etymology and part of speech, and any NJ word as the second element. The accuracy of Rosen's Rule at this level of loosened restriction is only 65%.

One study on Rosen's Rule by Kawahara & Sano (2014c) sets out to look at the psychological effects of this proposed rule. The study is based on an experiment where native Japanese speakers had to select one alternative in a two-way forced "wug" test; either an unvoiced, or a voiced second element alternative. 42 compounds were constructed using 14 real NJ nouns, with either two or three moras as the first element (seven words each), and three NJ-like nonce words, with three moras, as the second element (Kawahara & Sano 2014c: 113-114). The analysis of the results focused on the difference between rendaku application rate based on first element moraic length, i.e. the difference between 2µ3µ- and 3µ3µ-compounds (Kawahara & Sano 2014c: 114-115). As stated earlier in this section, Irwin defines two versions of Rosen's Rule (one Broad Rule and one Narrow Rule), neither of which seem to be adopted by Kawahara & Sano (2014c) as both 2µ3µ- and 3µ3µ-compounds already conform to both versions of Rosen's Rule. They forget to take the moraic length of the second element into account when looking at the PSR. This fact is also predictably reflected by the results from the study which show random (non-)application of rendaku (around 50%), and no significant psychological difference between 2µ3µ-compounds and 3µ3µ-compounds (RRO were .480 and .500 respectively) (Kawahara & Sano 2014c: 114-115). It could be argued that there is another viable criticism to Kawahara & Sano's study in the use of nonce words for the second element, since nonce words might not be interpreted as NJ words by native speakers (see discussion in Kawahara 2016: 38). Moreover, words outside the NJ stratum tend to undergo rendaku less often (see §2.1), and words need to be in the NJ stratum to pass the PSR (see earlier in this section). The use of nonce words when testing Rosen's Rule might not be bad practice in itself. However, it would, in my opinion, be inadvisable to conduct tests on Rosen's Rule using nonce words without taking the above points into consideration.

Another test, although not testing Rosen's Rule per se but similar in design to the one in Kawahara & Sano (2014c), was conducted by Tamaoka et al. (2009). In this test, the goal was to find if there is any correlation between the moraic length of the first element in a two-element compound, and the rate of rendaku occurrence in native speakers' minds. The stimuli consisted of nine different real NJ nouns with moraic lengths ranging from 1µ to 3µ as first elements, and two NJ-like nonce words with a moraic length of 3µ as second elements (Tamaoka et al. 2009: 21-22). The results from this particular test showed more notable differences than the results from Kawahara & Sano (2014c: 114-115). The results in Tamaoka et al. (2009:

23-24) showed that a shorter moraic length of the first element increased the rate of rendaku occurrence (1µ3µ-compounds had an RRO of .813, 2µ3µ-compounds had an RRO of .707, and 3µ3µ-compounds had an RRO of .630). One could try to put these results in comparison to Rosen's Rule, however, this test faces similar problems to the test by Kawahara & Sano (2014c); every compound already conforms to the broad version of Rosen's Rule, and the second elements are nonce words. However, if one were to only adopt the Narrow Rule as definition of Rosen's Rule, and ignoring the etymological problems, these results would suggest that Rosen's Rule is not realized in native speakers' minds. Rosen (2001, 2003) predicts that a PSR pass compound will always undergo rendaku, except when it is rendaku immune, which would imply that 2µ3µ- and 3µ3µ-compounds, to which the Narrow Rule would apply, should in theory have higher RROs than 1µ3µ-compounds.

In another way of testing the psychological reality of Rosen's Rule, Sano (2014) uses a database of speech samples[7] to find out if it is real in native speakers' minds. With the same interpretation of Rosen's Rule as the one adopted in Kawahara & Sano (2014c) (in which it is only the moraic length of the first element that affects the rendaku rate), Sano (2014) looks at rendaku occurrence in compounds depending on the moraic length of the first member, without taking the second element into account. Sano's (2014) adopted interpretation of Rosen's Rule states that a compound has higher chance of undergoing rendaku if the first element exceeds two moras, and the results reflect this idea. Compounds where the first element is either one or two moras show an RRO below .500, and compounds with first elements which exceed two moras have an RRO of over .600 (4µXµ-compounds with the highest RRO: approximately .800) (Sano 2014: 186). These results are quite contradictory to the findings from Tamaoka et al. (2009) on the effects on rendaku based on first element moraic length. Sano's results may support the interpretation of Rosen's Rule that is adopted by both himself and Kawahara & Sano (2014c), but none of the restrictions on either PSR or rendaku eligibility (word excision etc.) suggested by Rosen (2001, 2003) and Irwin (2009, 2016a) are employed when

[7]The *Corpus of Spontaneous Japanese - Relational Database*, Kokuritsu Kokugo Kenkyuujo 2012 (Sano 2014: 181).

testing the compounds. The results from this test and the results from Kawahara & Sano (2014c) are notably different, probably due to different methods of testing. However, if one were to employ consistent and appropriate restrictions, the results could stand next to not only each other, but to results of other studies as well (e.g. Irwin (2016a)). The reality of Rosen's Rule in native speakers' minds needs further, and more consistent, testing.

2.3 Lexicalized pattern or phonological system

Rendaku has long been viewed as a phonological rule based system (see discussion and citations in Kawahara 2016: 36), and has as a result been used as phonological argumentation in several cases (see discussion in Kawahara 2015a: 4). However, it is at the same time widely known that rendaku exhibits lexical irregularity. There are studies that suggest that rendaku might be of lexical nature at its core, where rendaku behavior is decided from lexicalized patterns instead of being phonologically productive (see Ohno 2000; Sano 2015, among others). This section will discuss some research done on both theories, what the implications of either theory being exclusively correct would be, and also the probable dual nature of rendaku.

2.3.1 Lexical argumentation

In a study on the lexical nature of rendaku, Ohno (2000) tests the hypothesis that the (non-)application of rendaku on novel compounds is based on semantic or phonetic analogy, i.e. lexical processes. Ohno created novel compounds with real NJ nouns for both elements. The novel compounds were designed to resemble existing compounds, either semantically or phonetically, where only the first elements were changed from the existing compounds. The selected existing compounds were, for the semantic test, extreme rendaku cases, where the second element behaves oppositely to its general rendaku-pattern. Ohno's idea was that native speakers would use analogy for their (non-)application of rendaku for these novel compounds and select the "incorrect" (non-)application of the second element, instead of following the general rendaku-pattern of the second element. For the phonetic test, Ohno uses second elements which can be found in both rendaku- and non-rendaku-compounds. The results support Ohno's hypothesis for both semantic, and phonetic analogy. The two novel compounds that tested semantic analogy conformed at 87.10% and 86.05% respectively (Ohno 2000: 160). The two novel compounds that tested phonetic analogy conformed at 71.43% each (Ohno 2000: 161). These results indicate that there at least is a lexical component to rendaku in native speakers' minds. Ohno defines the rendaku selection process as follows:

(4) **Rendaku selection** (second version)

Native speakers refer to existing compounds and refer to a semantically and/or phonetically parallel form when they determine the rendaku/non-rendaku form of a novel compound. Ohno (2000: 161)

Ohno further tests moraic length of the first element in novel compounds with the second element *hon* 'book', a word which only undergoes rendaku if the first element is three moras or longer. According to Ohno (2000: 161), moraic length is also part of phonetic analogy. The results of the novel compounds with *hon* as the second element, one 2μ2μ-compound and one 4μ2μ-compound, support Ohno's hypothesis again (the RRO of the 2μ2μ-compound was .400, and the RRO of the 4μ2μ-compound was .886) (2000: 161). However, the test involving moraic length might not be lexical evidence, where moraic length is used as analogy, but rather evidence for the reality of Rosen's Rule[8] in native speakers' minds (see discussion on Rosen's Rule in §2.2).

Sano (2015) tries to build upon Ohno's (2000) model to explain the contradictory rendaku behavior of *kami* 'hair' and *ti* 'blood' in the novel compounds *sirokami* 'gray hair' and *mimizi* 'earbleed' from Ohno's (2000: 160) test. Sano states that the (non-)application of rendaku in infrequent or novel compounds is based on the frequency of variants with or without rendaku in one's lexicon. However, it is the distribution in the actual language use that is the deciding factor of whether one applies rendaku or not, i.e. "token frequency" (Sano 2015: 338-339). Sano explains

[8]At the time of Ohno (2000), Rosen's Rule had not yet been discovered.

the exceptional behavior of Ohno's (2000) novel compounds, by finding that these compounds actually conform to the token frequency (2015: 338). The problem with Sano's (2015) study on rendaku (non-)application in infrequent or novel compounds is that only two items are tested. Ohno's (2000) test contains several novel compounds that are seemingly ignored by Sano. The question is why Sano decided to not mention any of the other novel compounds in his study. One reason might be that at least two of them would oppose his theory (in these two compounds, which also include *kami* and *ti* as second elements, a majority of respondents selected the non-exceptional (non-)application of rendaku, thus opposing Sano's theory). Even though Sano's theory removes the exceptionality in two novel compounds, it is the only thing he decides to show. Until further similar studies are conducted, I will remain skeptical of Sano's (2015) conclusion.

2.3.2 Phonological argumentation

Lyman's Law, which generally blocks rendaku in compounds where the second element already contains a voiced obstruent in a non-initial position, can be seen as a phonological constraint. Lyman's Law is believed to be a result of a cross-linguistic phenomenon, namely OCP(voice).[9] Kawahara (2012) finds that Lyman's Law is active in both loanwords and nonce words, and respondents tended to avoid rendaku when a compound violated Lyman's Law, which argues against it being a lexicalized pattern.

Another effect of OCP in rendaku is the Identity Avoidance constraint, i.e. the avoidance of identical moras next to each other in a word. Adjacent identical moras rarely occur in NJ words, but Identity Avoidance is however possible in compound formation. In an experiment, which used real NJ words as first elements, and nonce words as second elements, it was found that Identity Avoidance can both trigger and block rendaku in compounds (Kawahara & Sano 2014b). Compounds without rendaku application that would violate Identity Avoidance like [*i**ka*** + ***ka**niro*]

[9] See discussion on the relation between Lyman's Law and OCP, together with citations to the original theory on OCP in Itô and Mester 1986: 71-72.

showed higher rates of rendaku than compounds that would not, like [***ika*** + ***ta**niro*] (Kawahara and Sano 2014b: 4-5). This indicates that native speakers apply rendaku to a word to avoid two adjacent identical moras. Similarly, compounds that would violate Identity Avoidance if rendaku was applied, like [***iga*** + ***ka**niro*], which would become ***igaga**niro* with rendaku application, showed lower rates of rendaku than compounds that would not, like [***ida*** + ***ka**niro*], which would become ***idaga**niro* after rendaku application (Kawahara and Sano 2014b: 5-6). This indicates that native speakers decide to not apply rendaku on a compound if the resulting compound violates Identity Avoidance. Kawahara & Sano (2014a) found that Identity Avoidance not only applies to adjacent identical moras, but also identical consonants in adjacent moras. However, they found that Identity Avoidance on consonants was not as strong as for moras (Kawahara and Sano 2014a: 5-6).

In an experiment by Fukuda & Fukuda (1999), children with specific language impairment (SLI), and one control group, were tested on their rendaku production. Children with SLI are known to have difficulties learning linguistic processes, but can still learn lexical information (Paradis & Gopnik 1997, cited in Kawahara 2016: 36). The results showed that the SLI group applied rendaku much less often on infrequent or novel compounds than the control group. This indicates that rendaku is indeed a phonological process. However, there were no major differences between the groups when it came to rendaku (non-)application on frequent compounds, which indicates the existence of a lexical component in rendaku as well (Fukuda & Fukuda 1999, cited in Kawahara 2016: 36-37).

The effect on rendaku based on the first element in compounds was tested by Tamaoka et al. (2009). They found that first elements in both NJ and SJ compounds ending in the moraic nasal *n* had a higher rate of rendaku than compounds ending with other moras (Tamaoka et al. 2009: 31). These results are somewhat supported by Irwin (2016b: 96-97), where he finds that compounds with a first element ending with the moraic nasal *n* have the highest mean rendaku rate. However, the results from Irwin (2016b) do not show as notable differences as the results in Tamaoka et al. (2009). Nevertheless, the results from Tamaoka et al. suggest that there might be a phonological process based on the first element of a compound.

Another phonological process, which is affected by both the first and second element in a compound, is Rosen's Rule (see discussion on Rosen's Rule in §2.2). Even though the psychological and phonological reality of Rosen's Rule is yet to be definitively proven, the fact that we lexically can observe the effects of this prosodic size rule suggests a phonological process at work in the background.

2.3.3 The dual nature of rendaku

It is difficult to imagine that rendaku is either a completely phonological process or completely based on lexicalized patterns. The existence of phonological rules and constraints, such as Rosen's Rule, Lyman's Law, and Identity Avoidance among others, implies that rendaku is at least partly phonological (see discussion in §2.3.2). At the same time, there is evidence of a lexical component to rendaku in analogy when native speakers create novel compounds (see discussion on Ohno's study in §2.3.1). If rendaku was purely lexical at its core, phonological processes like Lyman's Law would be unnecessary. If rendaku was only phonological, we would not be able to explain some exceptional rendaku behavior in where native speakers choose the "incorrect" (non-)application on a word. To say that rendaku is entirely either lexicalized or phonological would be unwise in my opinion. Accepting the dual nature of rendaku would, as a result, allow us to test and to further discover which parts of rendaku is phonological and which parts are lexical.

An extensive discussion on the topic of phonology and lexicality, with further examples, can be seen in Kawahara (2015a).

2.4 Remaining questions

The reality of Rosen's Rule in native speakers' minds is yet to be tested properly, despite the attempts by Kawahara & Sano (2014c) and Sano (2014). It is interesting to note that the test conducted by Kawahara & Sano (2014c) and the one by Tamaoka et al. (2009) produced surprisingly different results, despite the similarity in design. Further research is needed, but testing the psychological reality of

Rosen's Rule could reasonably be done by putting, for example, 2µ2µ-compounds and 3µ2µ-compounds that conform to the eligibility criteria discussed in §2.2 against each other. A test of this kind would probably show the effects of Rosen's Rule more clearly than both Kawahara & Sano's (2014c) test on the psychological effects of the rule and the one on prosodic size based on the first element by Tamaoka et al. (2009).

There is still uncertainty to whether rendaku is a phonological or lexical based system, and further tests need to be conducted to find out which parts of rendaku belong to which system.

3 Study

This section will cover the study conducted for this thesis. The first part will cover the design of the study, with an in-depth explanation of the methodology and the contents of the study itself. The second part will cover the results of the study. In the final part, the study, and its results, will be discussed on a category level, where results from different categories will be compared to one another. On top of this, possible effects of other factors on the results will be discussed.

3.1 Methodology

The data that is discussed in this thesis comes from a study based on an online questionnaire that was sent out to native Japanese speakers. The questionnaire was designed to gather information on how native Japanese speakers produce rendaku when faced with both a voiced and unvoiced alternative. With the goal of choosing what reading felt the most natural, the respondents were given two-element noun compounds written in *kanji*, together with three different alternatives;

(i) a version written in *hiragana* containing an unaltered reading of the second element,

(ii) a version written in *hiragana* where the reading of the second element had an altered voicing, and

(iii) an option which said that both readings felt natural.

The respondents were forced to choose one alternative on every compound in the questionnaire, in a so called forced-choice "wug" test (Berko 1958). One of the stimuli, where the first element *hana* 'nose' and the second element *kuse* 'habit' were combined into a novel compound, had the alternatives "*hanakuse*", "*hana**g**use*", and "both readings are natural". The respondents also provided some basic personal information about themselves. This personal information included age, gender, and place of birth.

The compounds used in the questionnaire were created using actual Japanese words, but they were put together to create novel compounds that would not be found in dictionaries. Since commonly used rendaku eligible compounds have a fixed reading, either with a voiced or unvoiced second element, novel compounds were used to get data on the second element specifically and not the compound in its entirety. Every compound consisted of elements that are pure nouns, i.e. there were no deverbal or deadjectival nouns present.

The questionnaire consisted of 40 compounds which were presented in a random order for every respondent. The questionnaire tested eight categories of words with five compounds per category. Four categories tested the readings of compounds with rendaku immune words, rendaku haters, rendaku lovers, and words that seemingly break Rosen's Rule (Irwin 2009: 191-193). These four categories were designed to test Rosen's categorization of rendaku personalities (Rosen 2001, 2003). The remaining four categories contained compounds which were meant to test some established rules or conditions. These four categories tested the readings of compounds that violate Lyman's Law, Sino-Japanese compounds, compounds in which the first consonant of the second element is already voiced, and lastly, reduplicative compounds.

Every compound was evaluated by at least one native Japanese speaker before the questionnaire was sent out. During the process of selecting words and creating new compounds, there were some words and compounds that either felt unnatural or contained an uncommon word. In these cases, the compounds were either modified or replaced in their entirety.

3.1.1 Stimuli

The novel compounds created for this study were designed to either follow or break some rules or tendencies that have been found in rendaku. As briefly described above, the compounds are divided into two larger categories; compounds that contain words categorized by Rosen, and compounds that have established rules or conditions tied to them. This section will detail the contents of both groups.

There are 20 compounds in the first group, which contains compounds with words categorized by Rosen. All of these compounds are created based on Rosen's three main categories of rendaku personalities; rendaku immune words, rendaku haters, and rendaku lovers. Table 2 shows five of these 20 compounds. These compounds were created using five different pure NJ nouns that Rosen claims never undergo rendaku (an RRO of .000), i.e. rendaku immune words.

Table 2: Compounds created using rendaku immune words

First element	Second element	Alternative 1	Alternative 2
isi 'stone'	***h**ama* 'beach'	*isi**h**ama*	*isi**b**ama*
hukuro 'bag'	***h**ime* 'princess'	*hukuro**h**ime*	*hukuro**b**ime*
kuni 'country'	***h**asi* 'edge'	*kuni**h**asi*	*kuni**b**asi*
yuka 'floor'	***k**atati* 'shape'	*yuka**k**atati*	*yuka**g**atati*
umi 'sea'	***k**emuri* 'smoke'	*umi**k**emuri*	*umi**g**emuri*

Table 3 shows five compounds that were created using five different NJ pure nouns that Rosen claims resist rendaku to some extent in compounds, i.e. rendaku haters which have an RRO below .333.

Table 3: Compounds created using rendaku haters

First element	Second element	Alternative 1	Alternative 2
hone 'bone'	***k**usa* 'grass'	*hone**k**usa*	*hone**g**usa*
hana 'nose'	***k**use* 'habit'	*hana**k**use*	*hana**g**use*
nasi 'pear'	***h**ara* 'field'	*nasi**h**ara*	*nasi**b**ara*
asi 'leg'	***k**i* 'tree'	*asi**k**i*	*asi**g**i*
wata 'cotton plant'	***t**e* 'hand'	*wata**t**e*	*wata**d**e*

Table 4 shows five compounds that were created using five different NJ pure nouns that Rosen claims always undergo rendaku in compounds with an RRO of 1.000, i.e. included in Rosen's category of rendaku lovers which should have an RRO higher than .666.

Table 4: Compounds created using rendaku lovers

First element	Second element	Alternative 1	Alternative 2
sora 'sky'	***k**uti* 'mouth'	*sora**k**uti*	*sora**g**uti*
itigo 'strawberry'	***h**ako* 'box'	*itigo**h**ako*	*itigo**b**ako*
kuti 'mouth'	***k**ami* 'paper'	*kuti**k**ami*	*kuti**g**ami*
sara 'plate'	***s**oko* 'bottom'	*sara**s**oko*	*sara**z**oko*
tori 'bird'	***h**ue* 'flute'	*tori**h**ue*	*tori**b**ue*

Table 5 shows the last five compounds which were created using problematic NJ pure nouns discovered by Irwin (2009) that seemingly break Rosen's Rule in one way or another, i.e. Rosen's Rule breakers. Four of these Rosen's Rule breakers should in reality behave like rendaku immune words, but can in fact be found voiced in some PSR pass compounds (Irwin 2009: 190). The final Rosen's Rule breaker *koori* 'ice' is a rendaku lover, though it appears unvoiced in a PSR pass compound; *kanakoori* 'extremely cold object' ('icicle' in some modern dialects) (2009: 191).

Table 5: Compounds created using Rosen's Rule breakers

First element	Second element	Alternative 1	Alternative 2
kawa 'river'	***h**utokoro* 'bosom'	*kawa**h**utokoro*	*kawa**b**utokoro*
yoru 'evening'	***h**otaru* 'firefly'	*yoru**h**otaru*	*yoru**b**otaru*
yoko 'horizontal'	***k**anna* 'plane'	*yoko**k**anna*	*yoko**g**anna*
tonari 'neighbor'	***k**ata* 'person'	*tonari**k**ata*	*tonari**g**ata*
kumo 'cloud'	***k**oori* 'ice'	*kumo**k**oori*	*kumo**g**oori*

Every first and second element in these 20 compounds are NJ pure nouns, to avoid the answers from being influenced by other potential parameters that might impact the results.

Similarly to the first group that tested compounds with words categorized by Rosen, the second group is made up of 20 compounds divided into four categories with five compounds in each. Table 6 shows the first category which tested the readings of compounds that violate Lyman's Law. These compounds should in general never voice (see discussion in §1.1).

Table 6: Compounds created using words that violate Lyman's Law

First element	Second element	Alternative 1	Alternative 2
mizu 'water'	***k**arada* 'body'	*mizu**k**arada*	*mizu**g**arada*
tuki 'moon'	***h**ige* 'beard'	*tuki**h**ige*	*tuki**b**ige*
kusari 'chain'	***k**agi* 'key'	*kusari**k**agi*	*kusari**g**agi*
maru 'circle'	***t**amago* 'egg'	*maru**t**amago*	*maru**d**amago*
hana 'flower'	***s**ugata* 'figure'	*hana**s**ugata*	*hana**z**ugata*

Table 7 shows the next category which tested the readings of compounds with both elements being SJ words. In this case both the first and second element of every compound is a SJ pure noun. Similarly to compounds that violate Lyman's Law, compounds containing SJ words tend to block rendaku (see §2.1).

Table 7: Compounds created using SJ words

First element	Second element	Alternative 1	Alternative 2
tyuu 'pillar'	***k**otu* 'bone'	*tyuu**k**otu*	*tyuu**g**otu*
heki 'wall'	***t**an* 'edge'	*heki**t**an*	*heki**d**an*
gan 'face'	***t**ei* 'bottom'	*gan**t**ei*	*gan**d**ei*
kyoo 'border'	***k**oo* 'mouth'	*kyoo**k**oo*	*kyoo**g**oo*
gen 'present'	***s**oo* 'contend'	*gen**s**oo*	*gen**z**oo*

Table 8 shows the third category, which tested the readings of compounds where the second element already begins with a voiced obstruent. The alternative given to the original voiced reading was a devoiced reading, that in theory should not occur. Rendaku changes an unvoiced obstruent to its voiced counterpart, and an opposite scenario where this change goes from voiced to unvoiced should not be possible. Table 9 shows the final category which consists of reduplicative words, which tend to undergo rendaku in most compounds (see §1.1). Irwin (2009: 190, 192) states that reduplicative words, which are technically *dvandva*, should be treated as being outside of expected rendaku behavior.

Table 8: Compounds created with devoicable second elements

First element	Second element	Alternative 1	Alternative 2
yoru 'evening'	***b**aba* 'old woman'	*yoru**b**aba*	*yoru**h**aba*
kami 'god'	***g**ake* 'cliff'	*kami**g**ake*	*kami**k**ake*
koori 'ice'	***b**uta* 'pig'	*koori**b**uta*	*koori**h**uta*
ai 'love'	***b**eni* 'crimson'	*ai**b**eni*	*ai**h**eni*
sake 'alcohol'	***d**eko* 'forehead'	*sake**d**eko*	*sake**t**eko*

Table 9: Compounds created using reduplicative words

First element	Second element	Alternative 1	Alternative 2
***s**aka* 'slope'	(reduplication)	*saka**s**aka*	*saka**z**aka*
***s**iru* 'soup'	(reduplication)	*siru**s**iru*	*siru**z**iru*
***k**use* 'habit'	(reduplication)	*kuse**k**use*	*kuse**g**use*
***h**imo* 'string'	(reduplication)	*himo**h**imo*	*himo**b**imo*
***t**aki* 'waterfall'	(reduplication)	*taki**t**aki*	*taki**d**aki*

3.1.2 Respondents

55 native Japanese speakers in total participated in this study. 29 of the respondents were men, and 26 of the respondents were women. The median age was 21 years old, with an average age of 21.6 years old, and they were born between the years 1984 and 1998. A large majority of all respondents, 76.4%, gave an area within the Kanto region as their place of birth. The remaining respondents' places of birth were spread out over six other main regions of Japan. The distribution of the respondents' places of birth can be seen in Table 10.

Table 10: Distribution of place of birth among respondents

Hokkaido	Tohoku	Kanto	Chubu	Kansai	Shikoku	Kyushu
1	1	42	5	3	1	2

3.2 Results

In this section, the results from the questionnaire will be presented in detail. Every tested compound had three alternatives; a reading in *hiragana* where the second element was unvoiced, a reading in *hiragana* where the second element was voiced, and finally an alternative which stated that both readings felt natural.[10] The results will be shown mainly by RRO, as well as a rate of rendaku blocking (hereafter RRB), and a rate of optionality (hereafter RoO). RRO is here calculated by combining the percentage of respondents who selected the voiced alternative, and the percentage of respondents who selected the 'both natural' alternative. For example, if the voiced alternative was 65% and 20% were 'both natural', the RRO would become .850. Similarly, the RRB is calculated by combining the unvoiced percentage and the 'both natural' percentage. For example, if the unvoiced alternative was 15% and 20% were 'both natural', the RRB would become .350. RoO is calculated from the 'both natural' percentage, e.g. if 20% were 'both natural', the RoO would become .200.

3.2.1 Rosen's categorization

The results presented in this section are from 20 compounds with second elements that Rosen (2001, 2003) has categorized. Rosen puts these words into three categories; words that are immune to rendaku, rendaku haters, and rendaku lovers. Five of these words exhibit unpredictable behavior, where all seemingly break Rosen's Rule (see §2.2). All tested compounds, and detailed results can be seen in Table 11 (the number of individual votes each answer received can be seen for all compounds in Appendix A).

Rosen claims that rendaku immune words never voice in compounds. Only one of these tested compounds, [*hukuro* + *hime*], shows an RRB of 1.000, together with a very low RRO of .073. Two compounds, [*isi* + *hama*] and [*umi* + *kemuri*], are not far off with RRBs of .909 and .873 respectively. However, these two compounds

[10] One group of compounds with a devoicable second element had different alternatives; one original voiced reading, and one altered devoiced alternative. The third alternative was however the same.

Table 11: Results from compounds with words categorized by Rosen

Immune to rendaku				**Rendaku haters**			
Compound	RRO	RRB	RoO	Compound	RRO	RRB	RoO
[*isi* + *hama*]	.382	.909	.291	[*hone* + *kusa*]	.745	.545	.291
[*hukuro* + *hime*]	.073	1.000	.073	[*hana* + *kuse*]	.909	.309	.218
[*kuni* + *hasi*]	.473	.782	.255	[*nasi* + *hara*]	.727	.655	.382
[*yuka* + *katati*]	.636	.455	.091	[*asi* + *ki*]	.873	.418	.291
[*umi* + *kemuri*]	.418	.873	.291	[*wata* + *te*]	.927	.273	.200
Mean	.396	.804	.200	**Mean**	.836	.440	.276
SD	.184	.188	.098	**SD**	.084	.143	.065

Rendaku lovers				**Rosen's Rule breakers**			
Compound	RRO	RRB	RoO	Compound	RRO	RRB	RoO
[*sora* + *kuti*]	.764	.564	.327	[*kawa* + *hutokoro*]	.527	.764	.236
[*itigo* + *hako*]	1.000	.073	.073	[*yoru* + *hotaru*]	.818	.600	.418
[*kuti* + *kami*]	.782	.509	.291	[*yoko* + *kanna*]	.600	.618	.218
[*sara* + *soko*]	1.000	.109	.109	[*tonari* + *kata*]	.873	.309	.182
[*tori* + *hue*]	.982	.146	.127	**Mean**	.705	.573	.264
Mean	.906	.280	.186	**SD**	.145	.165	.091
SD	.109	.211	.103	[*kumo* + *koori*][a]	.982	.146	.127

[a]The compound [*kumo* + *koori*] does not have the same expected behavior as the other four Rosen's Rule breakers, and should therefore be analyzed independently from the rest.

show considerably higher RROs at .382 and .418 respectively. The compound [*kuni* + *hasi*] has an RRB of .782 and also shows a similar RRO, although slightly higher, to the previous two compounds; .473. The last compound, [*yuka* + *katati*], has a low RRB of .455. This last compound is the only one with a higher RRO (.636) than RRB, which arguably is opposite behavior from what is expected. Another interesting thing to note about this compound, is that despite its relatively even RRO-RRB-split, it has the second lowest RoO at .091 ([*hukuro* + *hime*] at .073) in its category. The three remaining compounds all have RoOs above .250.

Rendaku haters are words that have an RRO below .333, which is something none of the words that Rosen claims are rendaku haters shows in those tested compounds. The compound among rendaku haters with the lowest RRO is [*nasi* + *hara*] at .727. This compound also shows the highest RRB in its category at .655. The RoO of this compound, .382, is considerably higher than the mean RoO of the entire cateogry at .276. The other compounds in this category show gradually increasing, respectively decreasing, values of RRO and RRB. The highest RRO in this category can be found in the compound [*wata* + *te*] at .927. This compound also has the lowest RRB in its category; .273.

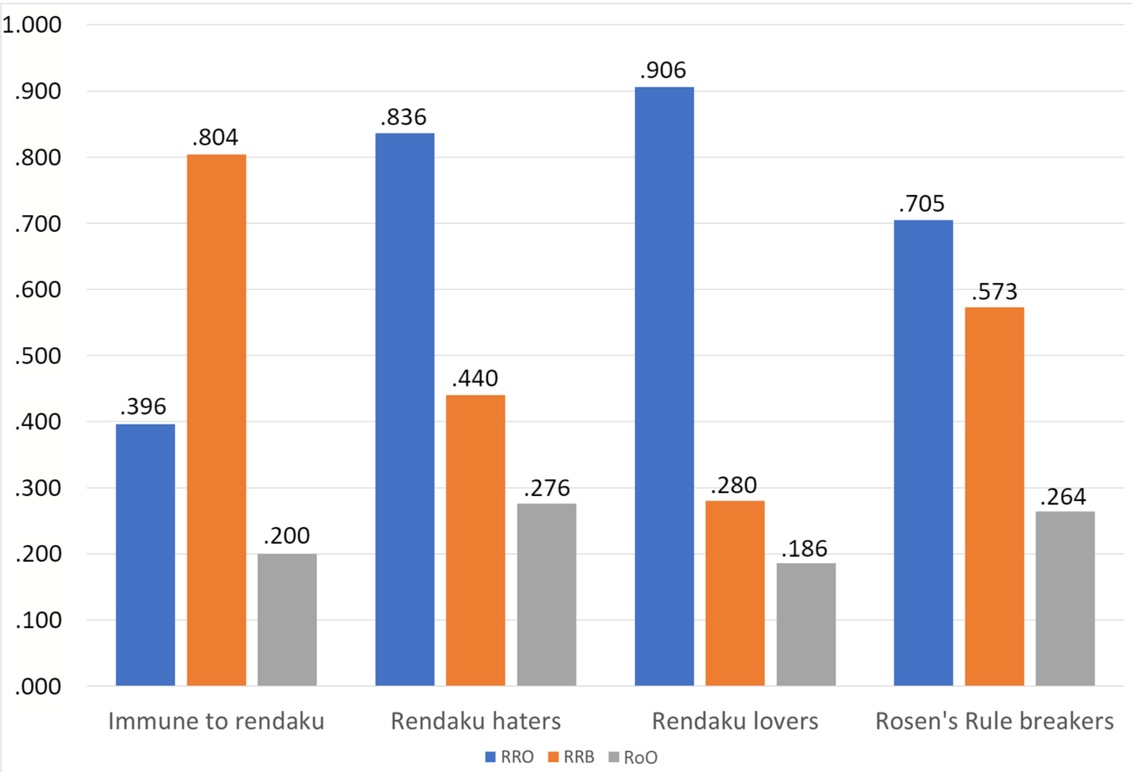
Rosen claims that the words tested for being rendaku lovers in this experiment have an RRO of 1.000. A word needs to have an RRO above .666 in order to classify as a rendaku lover according to Rosen. Out of the words categorized by Rosen, this group of tested rendaku lovers is the only category within the specified RRO. This category had a mean RRO of .906, which is also relatively close to Rosen's 1.000 claim for these words. Two compounds in this tested category, [*itigo* + *hako*] and [*sara* + *soko*], show an RRO of 1.000 and one compound, [*tori* + *hue*], shows an RRO of .982. These three compounds all have relatively low RRBs, ranging from .073 to .146, as well as low RoOs, ranging from .073 to .127. The two remaining compounds, [*sora* + *kuti*] and [*kuti* + *kami*], share similar results to one another. They have RROs between .764 and .782, RRBs between .509 and .564, and RoOs between .291 and .327. The RRBs shown in these compounds are interestingly high, compared to the RRBs shown in the other three compounds. The same thing can be noted in the RoOs, although these are not as high.

The last five compounds contain words that should behave predictably in PSR pass compounds. However, as pointed out by Irwin (2009: 190-191), these words have been found to behave oppositely in some compounds. Four of these words should behave like rendaku immune words in PSR pass compounds, and one should always voice in PSR pass compounds. Among the four compounds which should behave like rendaku immune words, there is little consistency. Two words, [*yoru* + *hotaru*] and [*tonari* + *kata*], show a higher RRO than RRB, whereas the two other compounds, [*kawa* + *hutokoro*] and [*yoko* + *kanna*], show a higher RRB than RRO. The compound [*yoko* + *kanna*] has the most even RRO-RRB-split of any tested compound in this experiment; .600 and .618 respectively. Furthermore, the compound [*yoru* + *hotaru*] has the highest RoO of any tested compound in the experiment at .418. The fifth Rosen's Rule breaker, [*kumo* + *koori*], which should voice in PSR pass compounds, shows expected behavior with an RRO of .982. Interesting to note is that the results for this compound is identical to the results for the rendaku lover [*tori* + *hue*].

Every compound in these four tested groups has an RoO below .300, except for three compounds. These compounds are, in an order of decreasing value, [*yoru* + *hotaru*] (a Rosen's Rule breaker) at .418, [*nasi* + *hara*] (a rendaku hater) at .382, and [*sora* + *kuti*] (a rendaku lover) at .327.

In Figure 2, the mean value of RRO, RRB, and RoO for all four groups of tested compounds based on Rosen's categorization can be seen in comparison to each other.

Figure 2: Comparison of results from Rosen's categorization



Note: The compound [*kumo* + *koori*], with a Rosen's Rule breaker as the second element, is excluded from this comparison. The reason for this exclusion is discussed in §2.2.

The three groups that tested immune to rendaku words, rendaku haters, and rendaku lovers show increasingly high RROs; .396, .836, and .906 respectively. Similarly, they show decreasingly low RRBs; .804, .440, and .280 respectively. The mean value of RoO is similar between compounds with expected rendaku immunity and those with expected rendaku submissiveness (i.e. rendaku lovers), at .200 and .186 respectively. The mean RoO one can observe in rendaku haters is slightly higher at .276. The fourth group of compounds that tested words that seemingly break Rosen's Rule has the closest RRO-RRB-split of any tested group of words in this experiment; an RRO of .705 and an RRB of .573. This group also has a relatively high mean RoO at .264, which is only slightly lower than the mean RoO of the group with rendaku haters.

3.2.2 Known rules and conditions

This section will present the results from 20 compounds that in some way or another test a known rule or condition. The compounds are divided into four groups; compounds that violate Lyman's Law, compounds with SJ words, compounds with devoicable second elements, and reduplicative compounds. All 20 compounds, and results, can be seen in Table 12 (the number of individual votes each answer received can be seen for all compounds in Appendix A).

The first group contains compounds that would violate Lyman's Law after rendaku application, which is something that rarely ever happens (see discussion in §1.1). The results reflect this generally known constraint on rendaku. Three compounds have an RRB of 1.000, and the remaining two compounds have RRBs of .982 and .946 respectively. The latter of these remaining two compounds, [*mizu* + *karada*], shows the highest RRO in this group; .091. This compound also shows the highest RoO at .036. The mean RRB value for this group is .986, which is the highest of any tested category in this experiment. The mean RRO, as well as the mean RoO are below .030 (.025 and .011 respectively), and both are the lowest values of any tested category in this experiment.

The second group contains Sino-Japanese compounds, which heavily dampen rendaku (see discussion in §2.1). Every compound in this group reflects this, where they all show RRBs higher than .900. One compound, [*tyuu* + *kotu*], shows an RRB

Table 12: Results from known rules and conditions

Lyman's Law violaters				**Sino-Japanese compounds**			
Compound	RRO	RRB	RoO	Compound	RRO	RRB	RoO
[*mizu* + *karada*]	.091	.946	.036	[*tyuu* + *kotu*]	.036	1.000	.036
[*tuki* + *hige*]	.000	1.000	.000	[*heki* + *tan*]	.109	.945	.055
[*kusari* + *kagi*]	.018	.982	.000	[*gan* + *tei*]	.073	.982	.055
[*maru* + *tamago*]	.018	1.000	.018	[*kyoo* + *koo*]	.167	.946	.109
[*hana* + *sugata*]	.000	1.000	.000	[*gen* + *soo*]	.036	.982	.018
Mean	.025	.986	.011	**Mean**	.084	.971	.055
SD	.034	.021	.014	**SD**	.049	.022	.030
Devoicing compounds				**Reduplicative compounds**			
Compound	RoD	RDB	RoO	Compound	RRO	RRB	RoO
[*yoru* + *baba*]	.000	1.000	.000	[*saka* + *saka*]	.967	.127	.091
[*kami* + *gake*]	.036	.982	.018	[*siru* + *siru*]	.927	.145	.073
[*koori* + *buta*]	.000	1.000	.000	[*kuse* + *kuse*]	1.000	.055	.055
[*ai* + *beni*]	.000	1.000	.000	[*himo* + *himo*]	.491	.582	.073
[*sake* + *deko*]	.145	.927	.073	[*taki* + *taki*]	.982	.109	.091
Mean	.036	.982	.018	**Mean**	.873	.204	.077
SD	.056	.028	.028	**SD**	.193	.192	.013

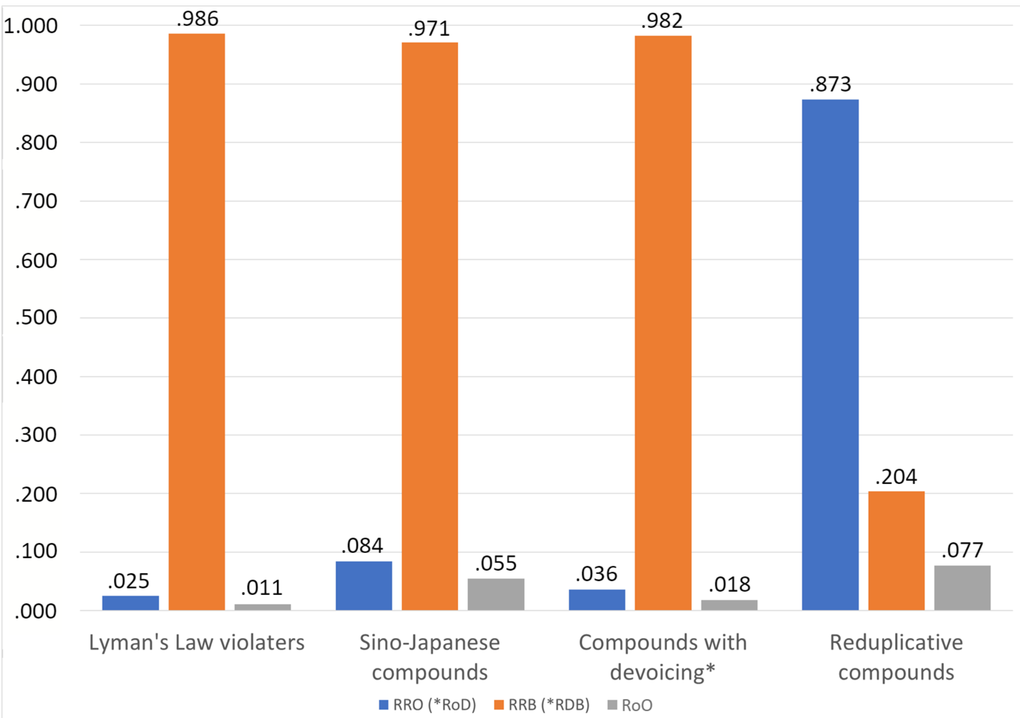
of 1.000. Furthermore, only one compound in this group, [*kyoo* + *koo*], has an RoO above .100 at .109, and this compound also has the highest RRO at .167.

The third group contains compounds where the second element already has a voiced obstruent in its initial position. Seeing that rendaku generally happens when an unvoiced non-initial element becomes voiced, this devoicing should not occur. Since the altered alternative to these compounds are devoiced, RoD, which is the rate of devoicing, is used instead of RRO. RoD is calculated by combining the devoiced percentage and the 'both natural' percentage. Similarly, RDB is the rate of devoice blocking, which will be used instead of RRB. It is calculated by combining the blocking percentage and the 'both natural' percentage. Three of these compounds have an RDB of 1.000, and the remaining two compounds have RDBs of .982 and .927 respectively. Only one compound [*sake* + *deko*] shows an RoD above .100 at .145, whereas three compounds have RoDs of .000, and the final compound has an RoD of .036.

The final group tests reduplicative compounds, which trigger rendaku in most compounds (see discussion in §1.1). Four of these compounds reflect this, and show RROs above .900, one of which has an RRO of 1.000 ([*kuse* + *kuse*]). However, one compound, [*himo* + *himo*], only has an RRO of .491. This compound is the only one in this group which has a higher RRB than RRO. The two values are also relatively close in proximity (RRO: .491, RRB: .582). Despite the even RRO-RRB-split, this compound has a surprisingly low RoO at .073.

A comparison of mean values from all four groups that test some known rules or conditions can be seen in Figure 3. These groups all have mean RoOs below .100. From a perspective of conforming to the expected behavior, compounds that violate Lyman's Law have the highest mean conformity (an RRB of .986). The second most conforming group is the one which contains compounds with devoicable second elements (an RDB of .982). The third most conforming group is the one which contains compounds with SJ elements (an RRB of .971). The fourth most conforming group is the one which contains reduplicative compounds (an RRO of .873).

Figure 3: Comparison of results from known rules and conditions



3.3 Discussion

3.3.1 Method

Contrary to many other psycholinguistic experiments on rendaku, the current experiment was designed as a forced three-way "wug" test, instead of a forced two-way "wug" test. This was done to obtain another factor which two-way tests misses, i.e. the optionality factor. Some native speakers do not see many compounds as either black or white when it comes to rendaku (non-)application. The option of selecting an alternative which corresponds to this feeling gives the results more depth. There have been experiments on rendaku using naturalness tests, where respondents evaluated compound readings on a Likert scale. However, Kawahara (2015b) found that forced-choice tests give better results than naturalness tests do.

Furthermore, one should note that it was important for this study that the compounds were made using real words, since the goal was to test specific categories of words that simply could not have been done using nonce words. The use of nonce words would not have been able to give an insight to the behavior of specific word categorizations. Another reason to use real words when creating novel compounds, instead of using nonce words, is that native Japanese speakers might treat nonce words like they were loanwords. This would be problematic since loanwords tend to either dampen or completely block rendaku (see Irwin 2005: 132-133, 2016b: 85, etc.). This fact, together with the fact that one category of tested compounds in this study contained SJ nouns, would have made it impractical at the very least to use nonce words.

3.3.2 Results from Rosen's categorization

Data from Version 3.2 of *The Rendaku Database* (Irwin et al. 2017) will be used in the discussion of results from Rosen's categorization. This data will be based on combined entries from both *Kōjien* (Shinmura 2008) and *Kenkyusha's new Japanese-English dictionary* (Watanabe et al. 2008). In case of conflict between the two dictionaries, the entry from *Kōjien* will be used. The compounds in *The Rendaku*

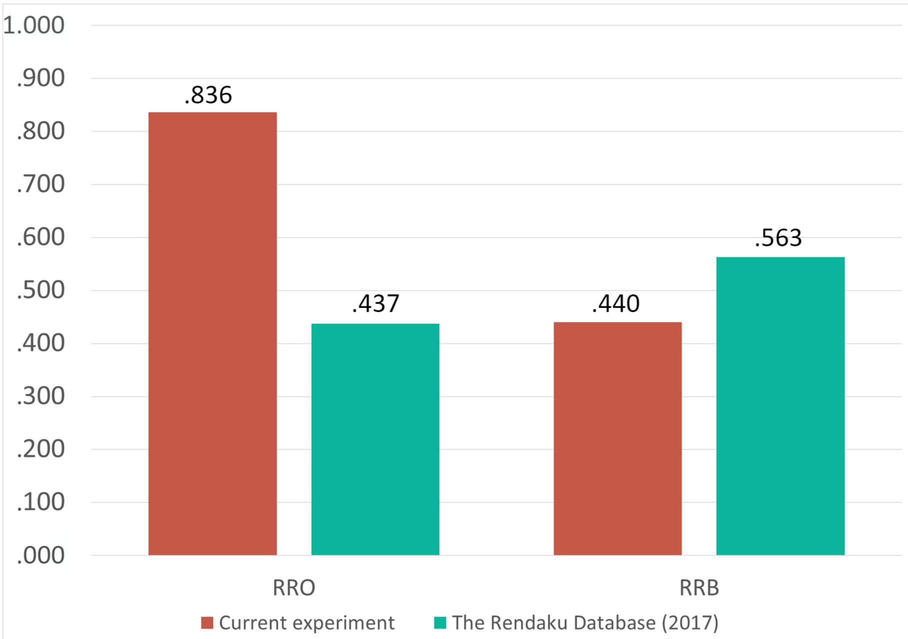
Database that are used for reference will conform to the eligibility criteria from Rosen (2001) and Irwin (2009) (see all criteria in §2.2).

The first group of compounds consists of second elements that Rosen claims are immune to rendaku. This claim is also supported by data from *The Rendaku Database*, where every second element used in this category has an RRO of .000. Despite this, one can clearly observe that there is only a single compound in this experiment that behaves immune-like; [*hukuro* + *hime*] with an RRO of .073 and an RRB of 1.000. Only one other compound in this category, [*isi* + *hama*], has an RRB above .900 at .909. However, this compound does not have a particularly low RRO (.382), which we could see in [*hukuro* + *hime*]. It is apparent that despite a word being lexically immune to rendaku, this immunity might not be psychologically real in native speakers' minds. This is even more evident in the compound [*yuka* + *katati*] with expected rendaku immunity, which shows exceptional rendaku behavior. In contrast to the other compounds in this category where the RRB is always higher than RRO, [*yuka* + *katati*] has an RRO of .636 and an RRB of .455. The exceptional behavior of this compound could perhaps be explained by the rendaku behavior of *kata* 'shape' (shortened form of *katati* 'shape') which almost always voices in compounds (an RRO of .855 in *The Rendaku Database* (Irwin et al. 2017)). However, despite the relatively even RRO-RRB-split in [*yuka* + *katati*], the RoO is only .091, which makes it one of only three compounds with an RoO below .100 of all compounds that test Rosen's categorization. This suggests that the respondents feel certain about their (non-)application preference, despite the large divide in answers. The mean values for the immune to rendaku category show a general tendency for rendaku blocking with an RRB of .804. However, by observing the RRO of .396 and the RoO of .200, one can see that native speakers are able to go against the lexical rendaku behavior. A comparison of results from the test on the immune to rendaku category and the results from *The Rendaku Database* can be seen in Figure 4.

Figure 4: Rendaku immune words in comparison to The Rendaku Database

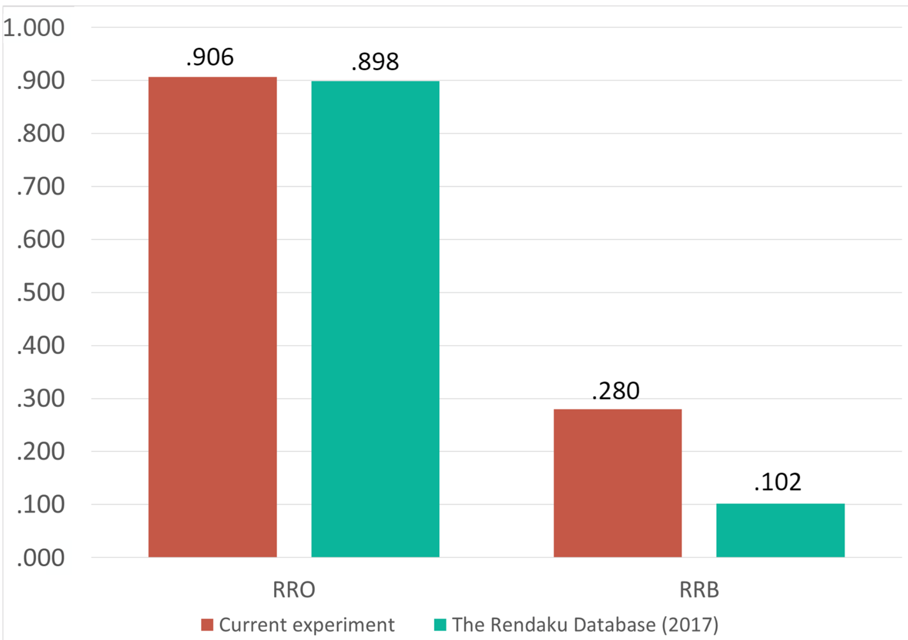
The second group of compounds consists of second elements that Rosen claims are rendaku haters. This claim is, however, not supported by data from *The Rendaku Database*. Only two of the tested second elements in the rendaku haters category have RROs of .333 or below in *The Rendaku Database*; *kuse* 'habit' which has an RRO of .333, and *hara* 'field' which has an RRO of .150. The remaining three compounds all have RROs between .334 and .666, which is a different category of words that Irwin calls rendaku waverers (see discussion in §2.1). One interesting thing to note about the compound in which *hara* is tested in, [*nasi* + *hara*], is that it has the second highest RoO of any tested compound in this experiment at .382. Despite the fact that *hara* 'field' has the lowest RRO in *The Rendaku Database* of any compound in this category, it showed a very high optionality rate among the respondents in the current study. This is another thing that points to the existence of a difference between lexical rendaku rates found in the database and psychological rendaku rates in native speakers. In the comparison between the results from the current experiment and the results from *The Rendaku Database* for these five words, which can be seen in Figure 5, the mean RRO in *The Rendaku Database* is .437, while the RRO from the test results is .836. The results from this category show an even greater disparity between lexical and psychological rendaku behavior than the results from the rendaku immune category do. There is no tendency towards rendaku hating behavior to be seen from the results of this category. Nor is there any tendency to follow lexical rendaku behavior. The fact that the results are completely different and almost opposite from what is lexically expected from these words, together with the high rates of optionality, suggests that there is some particular rendaku behavior for this category. The psychological and lexical differences are interesting since they were not expected, and this category might be worth exploring further in future research.

Figure 5: Rendaku haters in comparison to The Rendaku Database



The third group of compounds consists of second elements that Rosen claims undergo rendaku 100% of the time. These claims are partly supported by *The Rendaku Database*. Two of the words, *hako* 'box' and *hue* 'flute', have an RRO of 1.000 in *The Rendaku Database*, and another word, *kami* 'paper', has an RRO of .983. The word *kuti* 'mouth' has an RRO of .880, and while not being as close to 100% submissiveness to rendaku as the previous three, it is within the range of rendaku lovers (above .666). The final word *soko* 'bottom', however, has an RRO of .625 in *The Rendaku Database*, which would qualify it as a rendaku waverer. It is important to note that this word only appears in eight compounds in *The Rendaku Database*. Despite this, the compound it is tested in here, [*sara* + *soko*], shows an RRO of 1.000, which conforms to Rosen's claim of submissiveness. The results from the compound [*itigo* + *hako*] conforms to both Rosen's claim, and *The Rendaku Database* with an RRO of 1.000. The results from [*tori* + *hue*] almost conforms with an RRO of .982. Both [*sora* + *kuti*] and [*kuti* + *kami*] show lower RROs of .764 and .782 respectively. These two compounds also show considerably higher RRBs than the rest at .564 and .509 respectively. This, together with RoOs around .300, suggests that these two words are not inherently rendaku submissive in native speakers' minds. The other three compounds all have RRBs and RoOs under .150. However, if one looks at the mean values, this category of rendaku lovers shows that there is a tendency towards rendaku submissiveness. The mean RRO from the test, .906, very closely resembles the mean RRO from *The Rendaku Database* at .898. A comparison of results from the test on the rendaku lover category and the results from *The Rendaku Database* can be seen in Figure 6.

Figure 6: Rendaku lovers in comparison to The Rendaku Database

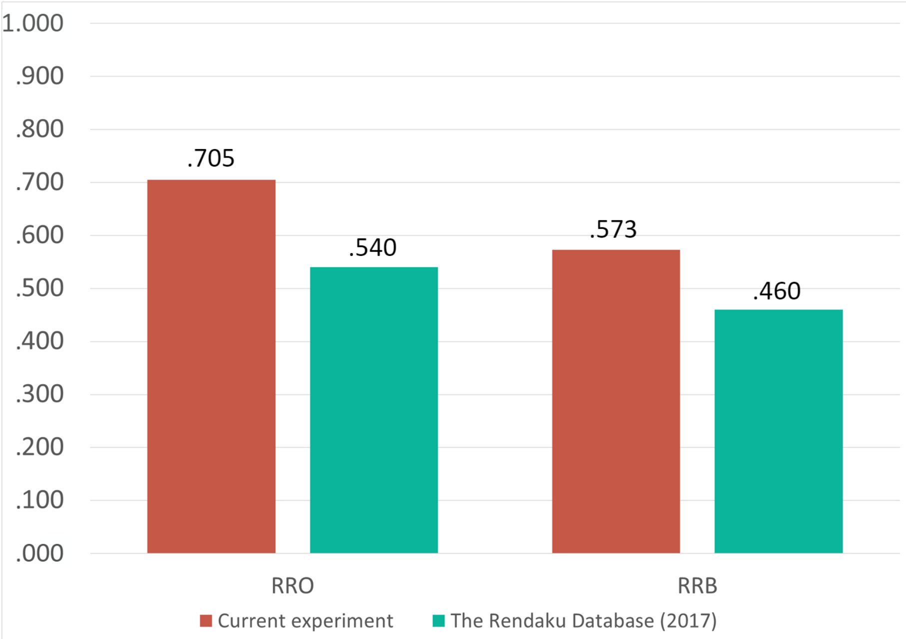


The fourth group of compounds consists of seconds elements that Irwin finds break Rosen's Rule (see discussion in §2.2). Irwin's claims are supported by *The*

Rendaku Database for all five compounds. These words can be found in both unvoiced and voiced PSR pass compounds, which, according to Rosen, should not happen. Four of these words, Irwin claims should behave like rendaku immune words in PSR pass compounds. The compounds in which they are part of in the current experiment all display different rendaku behavior. Two compounds, [*kawa* + *hutokoro*] and [*yoko* + *kanna*], show the most similar results in this group, and both of them have higher RRBs than RROs, contrary to the other two compounds in the group. [*yoko* + *kanna*] has the closest RRO-RRB-split of any compound tested in this experiment at .600 and .618 respectively, which puts it in the range of rendaku waverers. [*kawa* + *hutokoro*] shows the highest RRB in this group at .764, which makes it the most conforming compound of these four. The two compounds [*yoru* + *hotaru*] and [*tonari* + *kata*] both show higher RROs than RRBs, oppositely to the expected immune behavior. The latter shows results of a rendaku lover, and a relatively submissive one, with an RRO of .873 and an RRB of .309. The other compound, [*yoru* + *hotaru*], also shows a relatively high RRO at .818. However, its RoO is the highest of any compound tested in this experiment at .418, which suggests that some native speakers might not have one specific rendaku preference for this word. The mean RRO for these four compounds, which is .705, together with the mean RRB of .573, show no tendency towards the expected rendaku immune behavior. On the contrary, these numbers show no convincing tendency at all. This group shows the most indecisiveness of any group of tested compounds in this experiment. Furthermore, it is evident from the results from these four compounds that words that break Rosen's Rule behave rather inconsistently, almost randomly, when tested in novel compounds on native speakers. However, the mean results are surprisingly close to the results from these words in *The Rendaku Database*. A comparison of results from the test on Rosen's Rule breakers with expected immune to rendaku behavior and the results from *The Rendaku Database* can be seen in Figure 7.

The final compound that breaks Rosen's Rule, [*kumo* + *koori*], has an expected behavior of a rendaku lover. The results from this test suggest that despite the

Figure 7: Rosen's Rule breakers in comparison to The Rendaku Database



Note: The compound [*kumo* + *koori*] is not included in this comparison.

existence of at least one unvoiced PSR pass compound with the second element *koori* 'ice', native speakers still view it as a rendaku lover. The tested compound has an RRO of .982 and an RRB of .165, which is close to rendaku submissive behavior. In contrast to the four other compounds with second elements that break Rosen's Rule, this compound behaves expectedly. However, as pointed out in the discussion about this particular word, *koori*, in §2.2, the compound in which it breaks Rosen's Rule might be considered "infrequent and probably non-standard". This might be an explanation for the results from the compound [*kumo* + *koori*]. It is probably not worth considering this word a Rosen's Rule breaker in future discussions.

Looking at all four groups, and the mean values from their results, we can clearly see the existence of a pattern. In the first three groups, which test Rosen's three categories of rendaku personalities, there is a clear difference to be seen between each group. The RRO increases with each category, which was expected from their lexical counterparts, and the RRB decreases in a similar manner. These results suggest that native speakers possess an internal categorization similar to the lexical existing categorization. However, the internal and lexical categorization evidently differ considerably from each other, and native speakers are often willing to go against the lexically preferred alternative. The words that are lexically immune to rendaku are still accepted with rendaku by almost 40% of respondents in the current study. The rendaku haters show almost opposite behavior from the lexical data, where over 80% of the respondents in the current study accepted these words with rendaku. The words that had expected rendaku loving behavior were more in line with the results from the current study, but even these words were accepted without rendaku by almost 30% of the respondents. Furthermore, one could expect that words that are lexically considered either black or white, i.e. immune to rendaku or rendaku submissive, should have a lower rate of optionality when tested on native speakers. This expected behavior can also be observed in the results where the mean RoO from the groups with immune to rendaku words and rendaku lovers are similarly low at .200 and .186 respectively. The mean RoO from the group with rendaku haters is expectedly higher at .276. The group with words that break Rosen's Rule ([*kumo* + *koori*] excluded) also shows a similarly high RoO of .264, which should not be seen as unexpected due to the rendaku inconsistency of these words.

3.3.3 Results from known rules and conditions

The Rendaku Database (Irwin et al. 2017) will not be used as extensively during the discussion of the four groups of words that test known rules and conditions. There is only relevant data from the database for the group with reduplicative compounds, which will be compared to the results from this experiment.

The first group of tested compounds consists of compounds that would violate Lyman's Law. This is a known constraint that rarely fails to block rendaku in compounds (see discussion in §1.1). The results from this experiment reflect the expected behavior, almost without fail. Three compounds show an RRB of 1.000, and the remaining two compounds show RRBs of .982 and .946. The mean RRO for this group is also the lowest of any tested group of words in this experiment at .025. The same is true for the mean RoO which is also the lowest of any tested group of words in this experiment at .011, which is an extremely low optionality rate. It is evident from these results that native speakers show awareness to Lyman's Law when producing novel compounds using real elements.

The second group of tested compounds have two Sino-Japanese elements, which only undergo rendaku in about 10% of compounds according to Vance (1996: 25) (see discussion in §2.1). The results from this experiment show a similar tendency when observing the mean RRO for the group, which is at .084 (close to the 10% from Vance (1996)). The mean RRB is however very high at .971, with one compound, [*tyuu* + *kotu*], at 1.000. It is difficult to draw any general conclusions about the rendaku behavior in SJ compounds purely from these results. However, at least for the five compounds that were tested, one can observe a high blocking rate, a close to expected rate of rendaku, and a notably lower rate of optionality compared to NJ compounds at .055.

The third group of tested compounds have second elements with an already initial voiced obstruent, which were presented to the respondents together with a devoiced alternative. This "sequential devoicing" is the opposite from regular rendaku behavior, and does in general not occur in compounds. The results reflect the expected behavior in a similar fashion to the results from the Lyman's Law test discussed earlier. This group, with an unvoiced alternative is the second most conforming of any tested groups in this experiment. The rate of devoice blocking (RDB) is .982 and the RoO is only .018, which suggests that native speakers are aware that the only acceptable direction for voicing alteration on initial obstruents in second elements is the direction from unvoiced to voiced.

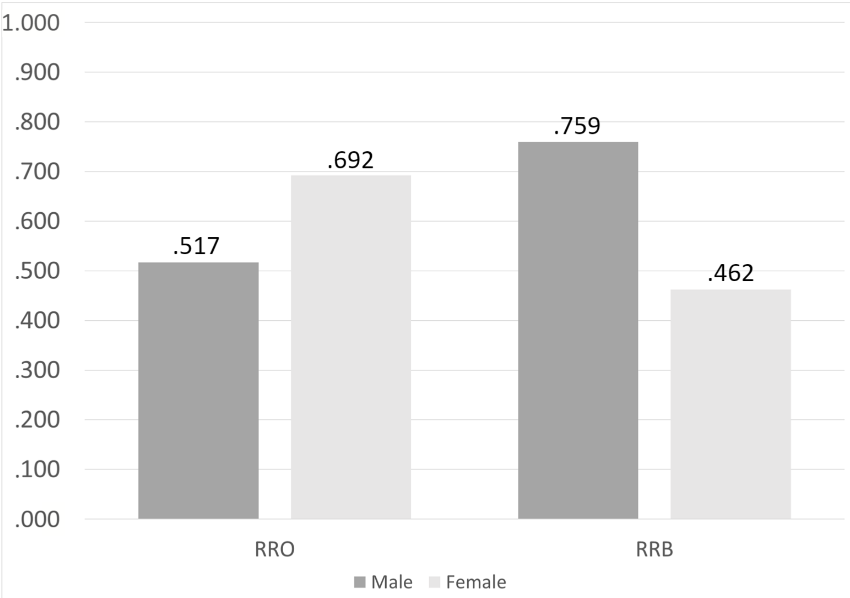
The fourth, and final, group of tested compounds are reduplicative compounds, where one word appears in both elements of a two-element compound. These compounds tend to undergo rendaku in most cases according to several sources (see discussion in §1.1). These claims are somewhat supported by *The Rendaku Database*, which shows an RRO of .736 for reduplicative compounds. Irwin (2016b: 84), however, points out that some of the reduplicative compounds that do not show rendaku are "baby-talk", and some have "strong onomatopoeic overtones and behave just as mimetics, in that rendaku fails to be triggered". This suggests that the actual RRO in *The Rendaku Database* for reduplicative compounds is higher than .736. The results from this experiment suggests that the systematic forcing of rendaku is real in native speakers' minds. For the reduplicative compounds in the current experiment, five words were used that according to *The Rendaku Database* are either rendaku submissive (*taki* 'waterfall' with an RRO of 1.000), rendaku lovers (*saka* 'slope' with an RRO of .750, and *siru* 'soup' with an RRO of .682), a rendaku hater (*kuse* 'habit' with an RRO of .333), or immune to rendaku (*himo* 'string' with an RRO of .000). Every reduplicative compound in this experiment but one shows an RRO above .900. The one compound that does not behave like the others consists of the immune to rendaku word *himo* 'string'. This compound has an RRO of .491, and an RRB of .582, which makes it the only compound in this group with a higher blocking rate than rendaku rate. Despite this, the RoO for this compound, .073, is close to the mean RoO for the entire group (.077). This suggests that despite there being a relatively even divide among respondents on whether it should be voiced or not, most respondents have a clear preference for one. There exists a tendency in many languages, including Japanese, where two adjacent homorganic consonants (like ***b*** and ***m*** in the rendaku version of *himo* 'string') are avoided (see discussion in Kawahara 2015a: 7). This might explain why the results from this compound is different from the other four reduplicative compounds in the current study. Almost half of the respondents accepted the voiced alternative of a word that never appears voiced in any rendaku eligible compound, and goes against a cross-linguistic tendency. This reinforces the idea of reduplicative compounds being systematic forcers of rendaku. The mean RRO for this group of compounds is .873, which can be compared to the considerably lower mean RRO from compounds with these words as second elements in *The Rendaku Database* which is only .553.

3.3.4 Differences between gender

There were no considerable differences on the results from a gender perspective. Both male and female respondents answered similarly for every group of words.

There were however some notable differences between men and women on one compound; the Rosen's Rule breaker [*yoko* + *kanna*]. This compound shows the most even RRO-RRB-split of any tested compound in this experiment, with an RRO of .600 and an RRB of .618. It is the only compound where men and women show close to opposite rendaku preference. A comparison between men and women's rendaku preference for the compound [*yoko* + *kanna*] can be seen in Figure 8.

Figure 8: Men and women's rendaku preference for [*yoko* + *kanna*]



Despite this difference in preference, it is difficult to draw any conclusions from these results. There can be several reasons as to why the results are the way they are for this compound, and the most probable in my opinion would be pure chance. The word *kanna* 'plane' as a second element of a compound might be unfamiliar to many respondents, and could be a reason for these differences in preference. Another possible reason could be that men and women might possess different lexicon, which would be an argument for rendaku being mostly based on lexicalized patterns. More studies on the differences between men and women's rendaku preference would however be needed to make any claim on the matter.

4 Conclusion

The study conducted for this thesis had the goal of figuring out at what degree native speakers of Japanese follow different known rules, conditions, and categorizations when it comes to rendaku in novel compounds. Further, it had the second goal of seeing if native speakers of Japanese produce these novel compounds with the same irregularities that can be found in existing compounds. The results from Rosen's categorization, i.e. immune to rendaku, rendaku haters, and rendaku lovers, show that one can observe distinct differences between the categories in a logical order of rendaku occurrence (immune to rendaku with the lowest rate of rendaku, and rendaku lovers with the highest rate of rendaku), even in real compound production. These results suggest that an internal categorization of rendaku personalities exists in native speakers' minds. However, the results are not within the lexically corresponding ranges of rendaku occurrence, which indicates that the internal and lexical categorizations are very different from each other. All three categories further show relatively high optionality rates. The category with rendaku haters, especially, shows a rendaku occurrence far from its lexical expectation, and an even higher optionality rate than the two other categories. When looking at the group of compounds with words that seemingly break Rosen's Rule with expected rendaku immune behavior, we can observe that native speakers of Japanese show irregularity in their production of novel compounds with these words. This visible irregularity in native Japanese speakers' production is very similar to the lexical irregularities these words exhibit. The final Rosen's Rule breaker, however, shows expected rendaku lover behavior, and should perhaps not be considered a Rosen's Rule breaker in future discussions.

The results from known rules and conditions only reinforce the positions they already hold in literature and research on rendaku. However, I believe that reduplicative compounds should be investigated further due to their peculiar behavior of systematically forcing rendaku and even allowing immune to rendaku words to undergo rendaku.

One final important thing to note is that the analysis of the results from the current study was done without taking lexicalized analogy into account, but focused on the phonological nature of rendaku instead. However, it is possible that one could analyze some of the results with Ohno's (2000) analogy-based model.

For future work on rendaku, I would suggest using Irwin's (2016b) more extensive categorization of rendaku personalities, instead of the somewhat restricted categorization proposed by Rosen (2001). I would also like to recommend using *The Rendaku Database* (Irwin et al. 2017) for any kind of work on rendaku, both as a tool and reference.

References

Berko, Jean (1958). "The childs learning of English morphology". In: *Word* 14, pp. 150–177.

Fukuda, Suzy E. and Fukuda, Shinji (1999). "The operation of rendaku in the Japanese specifically language-impaired". In: *Folia Phoniatrica et Logopaedica* 51, pp. 36–54.

Irwin, Mark (2005). "Rendaku-Based Lexical Hierarchies in Japanese: The Behaviour of Sino-Japanese Mononoms in Hybrid Noun Compounds". In: *Journal of East Asian Linguistics* 14.2, pp. 121–153.

Irwin, Mark (2009). "Prosodic size and rendaku immunity". In: *Journal Of East Asian Linguistics* 18.3, pp. 179–196.

Irwin, Mark (2012). "Rendaku Dampening and Prefixes". In: *NINJAL Research Papers* 4, pp. 27–36.

Irwin, Mark (2016a). "Rosen's Rule". In: *Sequential Voicing in Japanese*. Ed. by Timothy Vance and Mark Irwin. Amsterdam: Benjamins, MLA International Bibliography, pp. 79–106.

Irwin, Mark (2016b). "The Rendaku Database". In: *Sequential Voicing in Japanese*. Ed. by Timothy Vance and Mark Irwin. Amsterdam: Benjamins, MLA International Bibliography, pp. 107–117.

Irwin, Mark and Miyashita, Mizuki (2013). *The Rendaku Database*. Version 2.0. URL: `http://www-h.yamagata-u.ac.jp/~irwin/site/Rendaku_Database.html`.

Irwin, Mark, Miyashita, Mizuki, and Russell, Kerri (2017). *The Rendaku Database*. Version 3.2. URL: `http://www-h.yamagata-u.ac.jp/~irwin/site/Rendaku_Database.html`.

Itô, Junko and Mester, Armin (1986). "The Phonology of Voicing in Japanese: Theoretical Consequences for Morphological Accessibility". In: *Linguistic Inquiry* 17.1, pp. 49–73.

Kawahara, Shigeto (2012). "Lyman's Law is active in loanwords and nonce words: Evidence from naturalness judgment studies". In: *Lingua* 122, pp. 1193–1206.

Kawahara, Shigeto (2015a). “Can we use rendaku for phonological argumentation?” In: *Linguistic Vanguard* 1, pp. 3–14.

Kawahara, Shigeto (2015b). “Comparing a forced-choice wug test and a naturalness rating test: an exploration using rendaku”. In: *Language Sciences* 48, pp. 42–47.

Kawahara, Shigeto (2016). “Psycholinguistic studies of rendaku”. In: *Sequential Voicing in Japanese*. Ed. by Timothy Vance and Mark Irwin. Amsterdam: Benjamins, MLA International Bibliography, pp. 35–45.

Kawahara, Shigeto and Sano, Shin-ichiro (2014a). *Granularity of Identity Avoidance: Consonantal Identity, moraic Identity, and rendaku*. Ms. Keio University.

Kawahara, Shigeto and Sano, Shin-ichiro (2014b). “Identity avoidance and rendaku”. In: *Proceedings of Phonology 2013*.

Kawahara, Shigeto and Sano, Shin-ichiro (2014c). “Testing Rosen’s Rule and Strong Lyman’s Law”. In: *NINJAL Research Papers* 7, pp. 111–120.

Lyman, Benjamin Smith (1894). *The Change from Surd to Sonant in Japanese Compounds*. Oriental studies. Philadelphia PA: Oriental Club of Philadelphia.

Motoori, Norinaga (1822). *Kojiki den* [Commentary on the *Kojiki*]. Edo: Maegawa Rokuzaemon.

Nakagawa, Yoshio (1966). “Rendaku, rensei (kashoo) no keifu [Compounds with Rendaku and Compounds without Rendaku]”. In: *Kokugokokubun* 35, pp. 302–314.

Ohno, Kazutoshi (2000). “The lexical nature of rendaku in Japanese”. In: *Japanese/Korean linguistics* 9. Ed. by Mineharu Nakayama and Charles J. Quinn. Stanford CA: CSLI, pp. 151–164.

Otsu, Yukio (1980). “Some Aspects of Rendaku in Japanese and Related Problems”. In: *Theoretical Issues in Japanese Linguistics* [MIT Working Papers in Linguistics 2]. Ed. by Ann Farmer and Yukio Otsu. Department of Linguistics and Philosophy, Cambridge, Massachusetts, pp. 207–228.

Paradis, Michel and Gopnik, Myrna (1997). “Compensatory strategies in genetic dysphasia: Declarative memory”. In: *Journal of Neurolinguistics* 10, pp. 173–185.

Rosen, Eric (2001). "Phonological Processes Interacting with the Lexicon: Variable and Non-regular Effects in Japanese Phonology". PhD thesis. University of British Columbia.

Rosen, Eric (2003). "Systematic Irregularity in Japanese Rendaku: How the grammar mediates patterned lexical exceptions". In: *The Canadian Journal Of Linguistics La Revue Canadienne De Linguistique* 48, pp. 1–37.

Sano, Shin-ichiro (2014). "Examining Lexical and Phonological Factors on Rendaku in Spontaneous Speech". In: *MIT Working Papers in Linguistics* 73. Ed. by Shigeto Kawahara and Mika Igarashi. Cambridge: MIT Press, pp. 179–190.

Sano, Shin-ichiro (2015). "The Role of Exemplars and Lexical Frequency in Rendaku". In: *Open Linguistics* 1, pp. 329–344.

Shinmura, Izuru, ed. (2008). *Kōjien*. 6th edn. Tokyo: Iwanami Shoten.

Suzuki, Yutaka (2004). "Rendaku no koshoo ga kakuritsu suru made [Establishing the term rendaku]". In: *Kokubungaku Kenkyuu* 142, pp. 11–21.

Tamaoka, Katsuo, Ihara, Mutsuku, Murata, Tadao, and Lim, Hyunjung (2009). "Effects of first-element phonological-length and etymological-type features on sequential voicing (*rendaku*) of second elements." In: *Journal of Japanese Linguistics* 25, pp. 17–38.

Vance, Timothy J. (1996). "Sequential voicing in Sino-Japanese". In: *Journal of the Association of Teachers of Japanese* 30, pp. 22–43.

Vance, Timothy J. (2015). "Rendaku". In: *Handbook of Japanese Phonetics and Phonology*. Ed. by Haruo Kubozono. Berlin: De Gruyter Mouton, pp. 397–441.

Vance, Timothy J. (2016). "Introduction". In: *Sequential Voicing in Japanese*. Ed. by Timothy Vance and Mark Irwin. Amsterdam: Benjamins, MLA International Bibliography, pp. 1–12.

Vance, Timothy J. (2017). "Rendaku or Sequential Voicing in Japanese Phonology". In: *Oxford Research Encyclopedia of Linguistics*. Date of access 14 Mar. 2018. URL: http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-280.

Watanabe, Toshirō, Skrzypczak, Edmund R., and Snowden, Paul, eds. (2008). *Kenkyusha's new Japanese-English dictionary*. Digital supplemented edn. Tokyo: Kenkyusha.

Appendix A: Results

In this appendix, the number of answers for every compound in the questionnaire will be presented. The answers for the original and unaltered readings (UA in the table), the altered readings (AL in the table), and the 'both are natural' answers (BN in the table) are listed in the table in this appendix. All 55 respondents are represented in the table.

Immune to rendaku				**Rendaku haters**			
Compound	UA	AL	BN	Compound	UA	AL	BN
[*isi* + *hama*]	34	5	16	[*hone* + *kusa*]	14	25	16
[*hukuro* + *hime*]	51	0	4	[*hana* + *kuse*]	5	38	12
[*kuni* + *hasi*]	29	12	14	[*nasi* + *hara*]	15	19	21
[*yuka* + *katati*]	20	30	5	[*asi* + *ki*]	7	32	16
[*umi* + *kemuri*]	32	7	16	[*wata* + *te*]	4	40	11
Rendaku lovers				**Rosen's Rule breakers**			
Compound	UA	AL	BN	Compound	UA	AL	BN
[*sora* + *kuti*]	13	24	18	[*kawa* + *hutokoro*]	26	16	13
[*itigo* + *hako*]	0	51	4	[*yoru* + *hotaru*]	10	22	23
[*kuti* + *kami*]	12	27	16	[*yoko* + *kanna*]	22	21	12
[*sara* + *soko*]	0	49	6	[*tonari* + *kata*]	7	38	10
[*tori* + *hue*]	1	47	7	[*kumo* + *koori*]	1	47	7
Lyman's Law violaters				**Sino-Japanese compounds**			
Compound	UA	AL	BN	Compound	UA	AL	BN
[*mizu* + *karada*]	50	3	2	[*tyuu* + *kotu*]	53	0	2
[*tuki* + *hige*]	55	0	0	[*heki* + *tan*]	49	3	3
[*kusari* + *kagi*]	54	1	0	[*gan* + *tei*]	51	1	3
[*maru* + *tamago*]	54	0	1	[*kyoo* + *koo*]	46	3	6
[*hana* + *sugata*]	55	0	0	[*gen* + *soo*]	53	1	1
Devoicing compounds				**Reduplicative compounds**			
Compound	UA	AL	BN	Compound	UA	AL	BN
[*yoru* + *baba*]	55	0	0	[*saka* + *saka*]	2	48	5
[*kami* + *gake*]	53	1	1	[*siru* + *siru*]	4	47	4
[*koori* + *buta*]	55	0	0	[*kuse* + *kuse*]	0	52	3
[*ai* + *beni*]	55	0	0	[*himo* + *himo*]	28	23	4
[*sake* + *deko*]	47	4	4	[*taki* + *taki*]	1	49	5

Appendix B: Questionnaire

In this appendix, the contents of the questionnaire will be presented. The questionnaire was divided into two sections. The first section started with a title, and then a short introductory message which explained the purpose of the questionnaire. Following this message were four questions regarding gender, age, and place of birth. The second section of the questionnaire started with a brief explanation of what kind of task the respondents were expected to complete. Following this explanation, 40 compounds with three alternatives were presented. All fields and all tasks needed to be filled in or completed by the respondents before submitting their answers.

Section 1

Title
アンケート・複合語読み方評価

Introductory text
今、私は複合語（２つの単語を新しい言葉にすることです。例えば「顔」と「形」が「顔形」になります。）について卒業論文を書いています。論文のために、このアンケートを使います。

あなたの個人情報は匿名で、私の研究だけのために使います。

５～１０分ぐらいかかります。

よろしくお願いします。

性別*
- 男性
- 女性

生年月日（年月日）*
Short text answer

出身県*
Short text answer

区や市または群*
Short text answer

Section 2

Task instructions
時々、複合語の二つ目の単語の読み方が変わることがあります。「鼻」と「声」が「はなこえ」じゃなくて、「はなごえ」になります。このアンケートはそのような現象を調べます。

アンケートには色々な複合語が入っていて、読み方の３つの選択肢も入っています。

実際によく使われている単語で作ったのに、アンケートの中には辞書に載っていない複合語しかありません。辞書や携帯などを見ないで、複合語の読み方を評価して、３つの選択肢の中から、最も自然なものを１つ選んで下さい。

Compounds

石浜*
- いしはま
- いしばま
- 両方とも自然

海煙*
- うみけむり
- うみげむり
- 両方とも自然

足木*
- あしき
- あしぎ
- 両方とも自然

口紙*
- くちかみ
- くちがみ
- 両方とも自然

袋姫*
- ふくろひめ
- ふくろびめ
- 両方とも自然

骨草*
- ほねくさ
- ほねぐさ
- 両方とも自然

綿手*
- わたて
- わたで
- 両方とも自然

皿底*
- さらそこ
- さらぞこ
- 両方とも自然

国端*
- くにはし
- くにばし
- 両方とも自然

鼻癖*
- はなくせ
- はなぐせ
- 両方とも自然

空口*
- そらくち
- そらぐち
- 両方とも自然

鳥笛*
- とりふえ
- とりぶえ
- 両方とも自然

床形*
- ゆかかたち
- ゆかがたち
- 両方とも自然

梨原*
- なしはら
- なしばら
- 両方とも自然

苺箱*
- いちごはこ
- いちごばこ
- 両方とも自然

川懐*
- かわふところ
- かわぶところ
- 両方とも自然

夜蛍*
- よるほたる
- よるぼたる
- 両方とも自然

雲氷*
- くもこおり
- くもごおり
- 両方とも自然

横鉋*
- よこかんな
- よこがんな
- 両方とも自然

隣方*
- となりかた
- となりがた
- 両方とも自然

夜婆*
- よるばば
- よるはば
- 両方とも自然

神崖*
- かみがけ
- かみかけ
- 両方とも自然

氷豚*
- こおりぶた
- こおりふた
- 両方とも自然

愛紅*
- あいべに
- あいへに
- 両方とも自然

酒凸*
- さけでこ
- さけてこ
- 両方とも自然

水体*
- みずからだ
- みずがらだ
- 両方とも自然

月髭*
- つきひげ
- つきびげ
- 両方とも自然

鎖鍵*
- くさりかぎ
- くさりがぎ
- 両方とも自然

丸卵*
- まるたまご
- まるだまご
- 両方とも自然

花姿*
- はなすがた
- はなずがた
- 両方とも自然

柱骨*
- ちゅうこつ
- ちゅうごつ
- 両方とも自然

壁端*
- へきたん
- へきだん
- 両方とも自然

顔底*
- がんてい
- がんでい
- 両方とも自然

境口*
- きょうこう
- きょうごう
- 両方とも自然

現争*
- げんそう
- げんぞう
- 両方とも自然

坂々*
- さかさか
- さかざか
- 両方とも自然

汁々*
- しるしる
- しるじる
- 両方とも自然

癖々*
- くせくせ
- くせぐせ
- 両方とも自然

紐々*
- ひもひも
- ひもびも
- 両方とも自然

滝々*
- たきたき
- たきだき
- 両方とも自然