Rendaku in Japanese place names

Michelle Hermina van Bokhorst

1348884

MA East Asian Studies

Leiden University

MA thesis

1 July 2018

Dr. W. Uegaki

13113 words

Content

1. Introduction

During a trip to Hiroshima, a Japanese friend and I took a train in the direction of 糸崎. When I asked my friend if it was pronounced as *Itozaki* or *Itosaki*, she had to think for a while, eventually telling me that it probably was *Itosaki*. However, when the conductor announced the final station of the train five minutes later, it turned out to be *Itozaki* station. Japanese names can often be read differently depending on whether the first consonant of the second kanji of a name is voiced. This phenomenon is called rendaku and can cause a lot of confusion regarding the pronunciation of names and common nouns, even amongst native speakers of Japanese.

This research will focus on the geographical distribution of rendaku in Japanese place names, by focusing on morphemes of which it is known that they have a tendency to undergo rendaku, but not always. Based on a database of the Japan Post, 39 place names and their pronunciations over the entire country will be investigated to see if the differences in rendaku occurrence is related to the region of the place name. By investigating this, I try to answer the following two research questions: "How is rendaku in place names distributed across Japan?" and "Is there a connection in rendaku occurrence between names ending with the same morpheme?" The results showed that the rendaku rate differs for regions, but that morphemes do not really seem to have an influence, except for *kawa*.

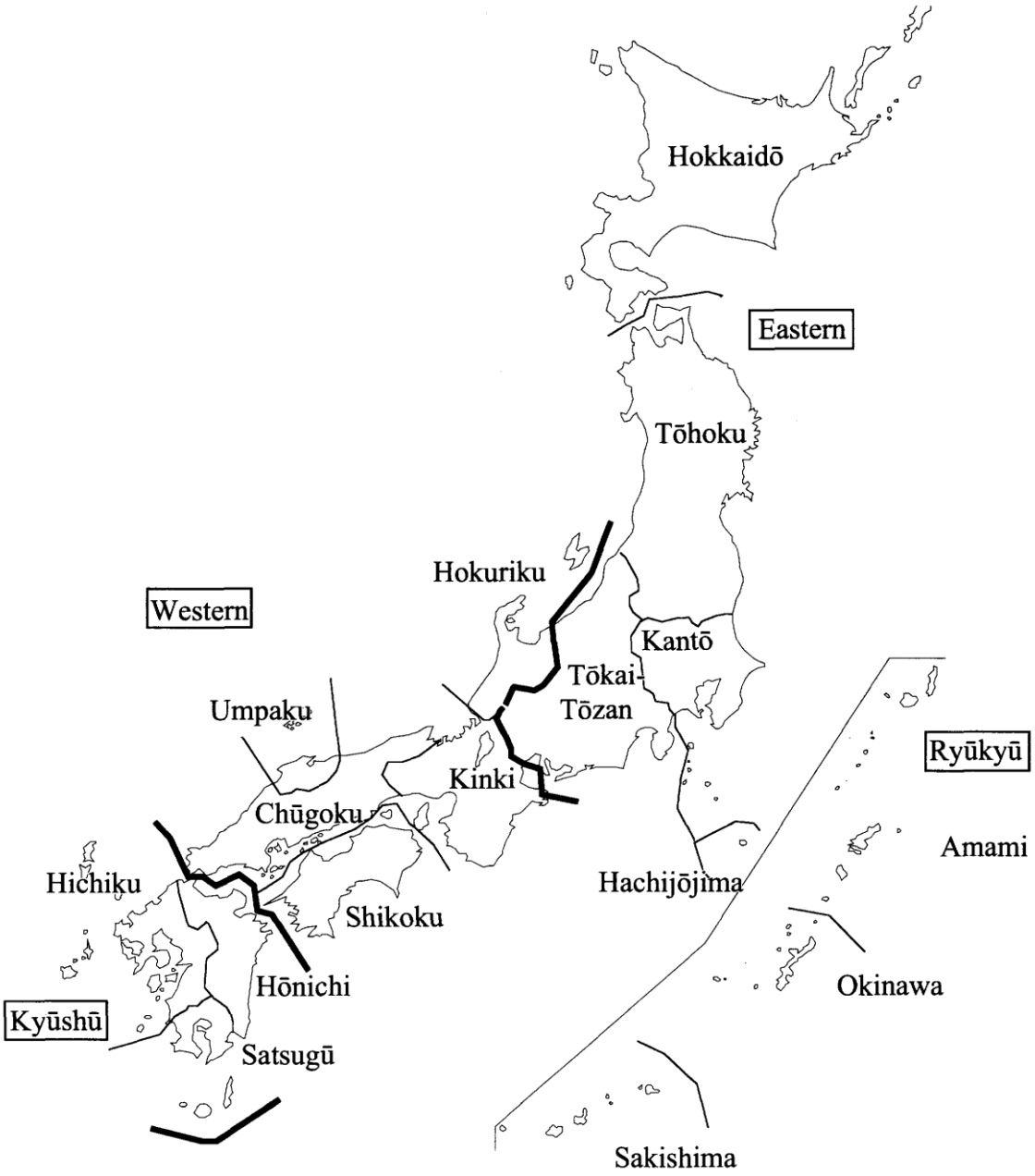
Before addressing the recent discussion on rendaku in names, the paper will first discuss classifications of the Japanese dialects in chapter 2. The third and fourth chapter will provide an overview of rendaku in common nouns and names. Chapter 5 will contain the research question and hypothesis, chapter 6 the method and chapter 7 the results. The discussion will be presented in chapter 8 and lastly the conclusion in chapter 9.

2. Divisions within Japanese dialects

According to Kadooka (2007) and Shibatani (1990), there have been many Japanese dialect divisions proposed over the years, all focusing on different aspects of Japanese. For instance, there are divisions based on lexical items and grammatical patterns, but also on accentual patterns.

Even though these accounts all divide the country in a different manner, there are a few major divisions as well. One of the dialect divisions is presented in figure 1.

Figure 1: Division of Japanese dialects



From Kadooka (2007, 161), based on Kato (1989)

The first major division is the one between the Ryukyuan dialects as spoken on the Ryukyu islands and the mainland dialects of Japan (Kato 1989; Kadooka 2007; Shibatani 1990). The

dialects on the Japanese mainland can be further divided in an Eastern dialect group and a Western dialect group, although the distinction Eastern Japan, Western Japan and Kyushu is not uncommon as well. Shibatani (1990) argues that the division between East and West lies in the middle of the Chubu region, of which Niigata prefecture and Shizuoka prefecture still belong to the Eastern dialect group, while Toyama prefecture and Aichi prefecture belong to the Western dialect group. According to Shibatani (1990), this division of the Chubu region can be attributed not only to linguistic aspects, but also to other social and cultural aspects. Furthermore, there is a clear geographical distinction where the Japanese Alps separate East and West Japan. However, it should be noted that within the three main dialect groups (East, West and Kyushu), there are smaller dialect groups as well, as can be seen in figure 1.

Lastly, Shibatani (1990) proposes a third division, namely the one between central areas and peripheral areas. Since Eastern aspects sometimes occur in the Western dialect group or in geographically separated areas, scholars often argue that the Eastern-type language was first spoken all over Japan, while later Western traits started to spread from Nara and Kyoto (the old capitals of Japan) to the outward areas. This hypothesis is based on the work of Yanagita (1930), who proposed that the central area of Kyoto contained newer words for 'snail', while the outward areas retained older forms.

There are many linguistic phenomena, phonological, morphological and syntactical, which can be explained through the core periphery theory. One of the examples Shibatani (1990) proposes is the palatalization of the fricative /s/. In the regions central in Japan, such as Tokyo, Yokohama, Osaka and Kyoto, palatalization of /s/ occurs before /i/, resulting in [ʃi] instead of [si]. However, until the 15th century palatalization of /s/ also occurred in these dialects in front of /e/, leading to [ʃe]. According to Shibatani (1990), the palatalization of /s/ in front of /e/ is still visible in Kyushu, Tohoku, Hokuriku, parts of Shikoku and even in some parts of Kinki, which can all be considered peripheral areas, while it has disappeared in the more central dialects.

On the other hand, there are also many phenomena where historical change seems to have occurred parallel in different regions and thus provide evidence against Yanagita's core periphery model. For instance, the high vowels /i/ and /u/ can be devoiced in dialects in the Tohoku region,

but also in a few other regions. Shibatani (1990) argues that this phenomenon has simultaneously developed in the San-in and Kyushu areas. Since these regions are far removed from each other, he argues that it is an example of parallel development instead of spread from a core to the peripheral areas.

In this chapter we saw that there are various manners to look at Japanese dialects. The divisions that are most relevant for later discussion are the ones between East, West, Kyushu and the subdivisions within figure 1, the division between East and West Chubu and the division between center and peripheral areas. In the next two chapters, the nature of rendaku in Japanese common nouns and names will be discussed.

3. Rendaku in common nouns

3.1 General overview

Within Japanese compounds, it happens often that the voiceless obstruent of the second morpheme becomes voiced when combined with another morpheme. This phenomenon is called *rendaku*, or *sequential voicing* (Vance 1987). Examples are provided in (1).

(1) a. ao “blue” + sora “sky” aozora “blue sky”

b. te “hand” + kami “paper” tegami “letter”

One of the problems with this phenomenon is that there are many exceptions. The literature on rendaku in Japanese common noun is extensive, with many scholars aiming to explain these variations or restrictions on rendaku. One of the most famous restriction was discovered by Benjamin Lyman (Vance 1987). According to Lyman, rendaku is inhibited in compounds when the second morpheme of the word already contains a voiced obstruent, such as in (2).

(2) Hitori “alone” + tabi “travel” hitoritabi/*hitoridabi “solo travel”

This condition is known as *Lyman's Law*. Since the word *tabi* in (2) already contains a voiced /b/, rendaku cannot occur anymore. There are instances where a voiced obstruent in the preceding morpheme can also have an influence, this is called *Strong Lyman's Law* (Kubozono 2005). This law will be further discussed in paragraph 4.2.

Besides Lyman's Law, there are other aspects that can influence the occurrence of rendaku. According to Rosen (2003), featural specification of the morpheme as well as the length of the compound can determine whether rendaku occurs. He argues that the many irregularities scholars found regarding rendaku actually seem to be very systematical, since they are connected to certain morphemes. Rosen (2003) argues that it is possible to classify these morphemes into four groups. Firstly, there are rendaku-immune nouns that never undergo rendaku, not even in long compounds. Examples are for instance *kita* 'north', *hasi* 'edge' and *hime* 'princess'. Secondly, there are nouns called 'rendaku resisters', which sparsely undergo rendaku in compounds where both nouns consist of two mora, but always do in longer compounds. An example is the noun *kusa* 'grass', which mainly occurs without rendaku (3a) and only seldom with rendaku in short compounds (3b). In long compounds, it always undergoes rendaku (3c).

(3) a. *aki* 'autumn' + *kusa* = *akikusa* 'autumn plants'
sita 'under' + *kusa* = *sitakusa* 'undergrowth'
b. *no* 'field' + *kusa* = *nogusa* 'wild grasses'
c. *hotaru* 'firefly' + *kusa* = *hotarugusa* 'firefly grass'
matu 'wait' + *yoi* 'evening' + *kusa* = *matuyoigusa* 'evening primrose'
Data from Rosen (2003, 10)

Thirdly, there are also nouns that voice in every compound. For instance, *hara* 'belly', *hana* 'flower', *kuti* 'mouth' and *soko* 'bottom' are seldom seen without rendaku. Lastly, there are nouns that voice in most compounds, but not in all of them. A few examples are *tori* 'bird', *kumo* 'cloud', *kawa* 'hide'

and *tama* 'ball, jewel'. According to Rosen (2003), the nouns that block rendaku are lexically specified to do so.

3.2 Dialectal variation

As Irwin & Vance (2015) point out, there has been a substantial amount of research regarding rendaku in Standard Japanese, but very little research regarding Japanese dialects. There are a few instances where scholars focus on a certain occurrence of rendaku within a particular dialect, although comparative studies are especially scarce. In this paragraph, I will discuss two studies where rendaku was compared across Japanese dialects.

The first research was conducted by Tamaoka & Ikeda (2010). They investigated the occurrences of rendaku in different *shoochuu* names across six regions, namely Kagoshima, Oita, Fukuoka, Yamaguchi, Hiroshima and Shizuoka. They distributed a survey to 405 undergraduate students from the six different regions, in which the students were asked for the pronunciation of five different types of *shoochuu*, namely *imo-joochuu*, *soba-joochuu*, *mugi-shoochuu*, *kome-shoochuu* and *kokutoo-shoochuu*. They expected that rendaku would occur more often in the Kagoshima prefecture for *imo-joochuu*, since this is the most frequent type of *shoochuu* in the region. Similarly, for the Oita prefecture they expected *mugi-shoochuu* to have a higher rendaku rate, since the people are more used to this type of *shoochuu*. The results showed that for instance *imo-joochuu* had a higher rendaku rate than the other types in every region, although the dialect region does not seem to have an influence on rendaku (Tamaoka & Ikeda 2010). However, except for the Shizuoka prefecture, the regions were located very close to each other (the Kyushu and Chugoku regions), which may contribute to the lack of variation in rendaku rate amongst the used dialects in words containing *shoochuu*.

Irwin & Vance (2015) investigated rendaku in common nouns amongst four dialects. They chose four cities that each represented a large dialect group based on Shibatani (1990). In their experiment, they included the Eastern Kyushu subdivision of the Kyushu dialect group as spoken in Miyazaki City (Miyazaki prefecture), the Tohoku subdivision of the Eastern Japanese dialect group as spoken in Kahoku Town (Yamagata prefecture), the Shikoku subdivision of the Western

Japanese dialect group as spoken in Matsuyama City and Ozu Town (Ehime prefecture) and lastly, the Kinki subdivision of the Western Japanese dialect group, as spoken in Kobe City (Hyogo prefecture). From each city or town, they asked participants to read 27 sample sentences out loud, which contained 32 rendaku targets in total. These rendaku targets were common nouns; chosen to include as many different types of nouns as possible. The results illustrated that there were a few rendaku targets that showed variation, although there was no significant difference in rendaku rate among the four groups.

Based on the two studies mentioned above, rendaku does not seem to vary that much across dialects for common nouns. However, as the scholars concluded themselves as well, more word types and dialects should be investigated in order to draw adequate conclusions. The following chapter will discuss the research regarding rendaku in Japanese names.

4. Rendaku in names

A subfield of rendaku in which the literature is less extensive, is rendaku in Japanese names. Sugito (1967) first made the observation that surnames behave differently from common nouns, although it took decades before other scholars started to investigating rendaku in Japanese names. Lately, research on rendaku in Japanese names has been extensively increasing, for instance by Kubozono (2005), Zamma (2005), Tanaka (2017), Shirooka (2014) and more. Each of these scholars provides different explanations for the same phenomenon, which I will briefly discuss in this chapter.

4.1 Accents and morphemes

Research regarding Japanese names started with scholars who investigated the relationship between rendaku and accent. Japanese is a pitch language, which means that it has tones, but contrary to tone languages such as Chinese not every mora needs to bear a tone. Standard Japanese is said to make a distinction between high and low tones. The place in a word where the high tone falls is said to be the place where the accent comes (Shibatani 1990).

Sugito (1967) conducted a study in which she investigated Japanese surnames ending with the morpheme 田 *ta*. She asked speakers of Standard Japanese and the Osaka region to pronounce 362 surnames consisting of three mora which ends with *ta.* She finds that surnames that end with *ta* often have an accent on the first mora in the Tokyo region, while surnames with *da* are often accentless. This is often referred to as 'Sugito's Law'.

Zamma (2005) continued Sugito's research by investigating Japanese surnames other than *ta*. In total he investigated 347 names and found that non-rendaku names were accented in 62,1% of his data, which is in accordance with Sugito's Law. However, he also found that within names that did have rendaku, 58,2% were accented, which contradicts Sugito's Law. Therefore he states that the behavior of *ta* is not representative of Japanese surnames in general. Zamma (2005) argues that individual morphemes are specified for both accent/non-accent and rendaku sensitivity. For instance, the morphemes 原 *hara* 'field', 下 *shita* 'under', 谷 *tani* 'valley', 瀬 *se* 'shallows' and 坂 *saka* 'slope' are specified for both accent and non-rendaku, and thus never occur with rendaku but without accent in standard Japanese.

The morphemes 口 *kuchi* 'mouth' and 林 *hayashi* 'forest' are specified for both accent and rendaku, since almost all of these names undergo rendaku and have an accent. The only exception Zamma (2005) mentions is *Kobayashi*, which is accentless. Another pattern he found was related to the morpheme 倉 *kura* 'storehouse', which is accentless and specified for non-rendaku. The only exceptions he found were *Ogura* and *Kuma'kura*.

He also argues that there are names that are only specified for one of the properties. For instance, in 沢 *sawa* 'swamp', the morpheme is specified as accentless, while rendaku seems to be determined by the last onset of the base. More about this will be discussed in the next section. Other names that show a similar behavior are names with 島 *shima* 'island', 塚 *tsuka* 'mound', 埼 *saki* 'cape' and 畑 *hata* 'farm'. Lastly, he argues that there are names that are specified for neither. In names like these, both rendaku and accent seem to be assigned randomly. Examples of morphemes he gives are 橋 *hashi* 'bridge and 川 *kawa* 'river'. Regarding *kawa*, he states that the behavior of the morpheme is different when it refers to a river, in which case it always undergoes rendaku.

4.2 Individual segments

In the same publication as Zamma (2005), Kubozono (2005) also published a research based on the data from Sugito (1967). Sugito (1967) observed that if the mora preceding the second morpheme or a name contains a /d/, /b/, /g/, /dz/ or a /j/, the latter morpheme is pronounced as *ta*. On the other hand, if the preceding mora contains a /t/, /s/, /n/ or /m/, it is pronounced as *da*. This can be seen as a stronger version of Lyman's Law (Sugito 1967, Kubozono 2005). This strong version means that not an obstruent within the same morpheme that undergoes rendaku excert influence (as in (1) in the previous chapter), but that the final consonant of the preceding mora influences the occurrence of rendaku.

Kubozono (2005) expands this observation by arguing that the occurrence of rendaku can be attributed to natural classes, without considering accents. The rules he stated are repeated in (4), with the examples he proposed in (5).

(4) a. /da/ is preferred after voiceless obstruents and nasals.

b. /da/ is prohibited after voiced obstruents.

c. /ta/ is preferred after liquids.

(5) a. fuku-da, kashi-da, kusu-da, ashi-da, kase-da, kaku-da, shima-da, naka-da (or naka-ta), kata-da

b. fugu-ta, kaji-ta, kuzu-ta, aji-ta, kaze-ta, kagu-ta, shiba-ta, kubo-ta, naga-ta, sugi-ta, kado-ta

(From Kubozono 2005, 4. Slight moderations are mine)

He argues that /ta/ is chosen when the preceding mora has the feature [+voice], while /da/ may be chosen if the mora has the feature [-voice] or is unspecified (Kubozono 2005). Furthermore, he argues that this is a case of the OCP (Obligatory Contour Principle, which prohibits two similar features from following each other), since rendaku does not occur if the preceding mora already

contains a voiced obstruent. Therefore, he argues that Lyman's Law is working here, albeit across morpheme boundaries and thus according to the strong variant. However, similar to Sugito (1965) and Zamma (2005), he found many exceptions as well, especially regarding names with /k/ in the first morpheme (6).

(6) a. /ta/: iku-ta, aki-ta, oki-ta, kaki-ta, maki-ta, saka-ta

b. /da/: fuku-da, oka-da, taka-da, toku-da, oku-da, kaku-da, ike-da, take-da, fuka-da

(From Kubozono 2005, 5. Slight alterations are mine)

/k/ seems to behave different from other voiceless obstruents, since it allows rendaku more often than for instance /s/ and /t/.

Tanaka (2017) expanded the research of Kubozono (2005) by creating his own database using *mixi*, a Japanese social network website. On this website, users can provide both their name in Japanese kanji and in romaji or kana, which makes it possible to determine the pronunciation of a surname. He established the rendaku rate of each name and found for instance that *Nakata* has a rendaku rate of approximately 20%. Tanaka (2017) argues that rendaku is triggered due to *Identity Avoidance*, a restriction which prevents similar consonants to follow in adjacent syllables. Since /s/ and /t/ are both voiceless alveolar obstruents, rendaku is more likely to be triggered when these follow each other in a name, such as in *Asa-da*, in order to create a larger contrast. /k/ on the other hand is a voiceless velar obstruent and thus differs in the place of articulation. Therefore, the contrast is larger and rendaku less likely to be triggered.

4.3 Diachronic variation

There has been little research regarding changes within surnames, but there has been some research regarding the diachronic change of rendaku within place names. In this section I will discuss the research of Shirooka (2014) about place names starting with 小 *o/ko* "small" and 大

oo “large”. Based on place name dictionaries, he investigated and compared the pronunciation of surnames from the Meiji, Taisho and Heisei periods to examine whether and how rendaku in place names has been changing over the ages. For this, he used the book “地名索引” *Chimei Sakuin* ‘Place name index’ for names in the Meiji period, “市町村文字読方名彙” *Shichōson moji yomikata meii* ‘Municipal character reading vocabulary’ for names in the Taisho period, and a ZIP code book for the Heisei period.

His results indicated that there have been several changes in the behavior of morphemes regarding rendaku. For instance, he noticed that the rendaku tendency of morphemes such as 澤 *sawa*, 島 *shima* and morphemes with three mora has become stronger. Where many of the names starting with 小 *ko/o* approximately had an equal amount of rendaku and non-rendaku varieties, in the Heisei period very few names remained without rendaku. As for names starting with 大 *oo*, he concluded that the rendaku rate was not very high in the Meiji period and has only become lower in the Heisei period.

Shirooka (2014) also noticed that rendaku is not the only thing that has changed in these place names. He discovered that the pronunciation of the morpheme 小 has changed from both *o* and *ko* in the Meiji period to a preference for *o* in the Heisei period. For instance, he argues that 小国 could be pronounced as both *Kokuni* and *Okuni* in the Meiji period, was mainly pronounced as *Okuni* in the Taisho period and as *Oguni* in the Heisei period. Shirooka (2014) explains the change in rendaku by arguing that with the change of *ko* to *o,* names starting with 大 *oo* and 小 *o* were only different regarding the length of the vowel. Therefore rendaku started to occur more often in word starting with *o*, to create a clearer distinction between for instance 大倉 *Ookura* and 小倉 *Ogura*.

4.4 Synchronic variation

Up until now, most research focused on the nature of rendaku in names itself, investigating mainly accents and individual segments. There also has been some research regarding the historical change of rendaku. However, when you investigate variation within single names from a

synchronic view, there is very little research. Many scholars have mentioned in their research that for some names multiple pronunciations might be possible, although few have tried to explain this.

Zamma (2005) observed in his research that names can have alternate pronunciations. Five speakers of standard Japanese voiced the names in his research; out of the 347 names there were 25 names where the name was pronounced both with and without rendaku. However, he does not state the names or tries to explain the variation. Tanaka (2017) observed that the rendaku rate can vary for names. For instance, he noticed that the rendaku rate for *Nakata* was around 20%. There were more names in his database with a low rendaku rate and thus internal variation, such as *Fukuta* and *Takikawa*.

The *Nikkei Shinbun* reported in an informal research based on entries in the telephone dictionary that *Nakajima* and *Yamazaki* occur more often in the eastern part of Japan, while *Yamasaki* and *Nakashima* occurs more often in Kyushu (*Nikkei Shinbun* 4/20/2012, through Iwasaki 2013).

Van Bokhorst (2017) showed that there was regional variation within surnames as well. 231 participants filled out a questionnaire in which they had to write down the hiragana pronunciation of 80 surnames, which were given in kanji. This research mainly focused on the Hokkaido, Kanto, Kansai and Kyushu areas. The rendaku rate for *Nakata* was similar to the one Tanaka (2017) observed, namely around 20%. This research however indicated that these occurrences of *Nakada* where only in the Kanto area (Van Bokhorst 2017). In the Hokkaido, Kansai and Kyushu areas, the only form of the name that occurred was *Nakata*. Similar to the *Nikkei Shinbun*, the study also found that *Nakajima* and *Yamazaki* are more frequent in the eastern part of Japan, but start to slightly decrease within the Kansai region, and significantly in the Kyushu region. The name *Tajima* showed a similar pattern as well. Interestingly, for names such as *Tanigawa* and *Miyagawa*, the rendaku rate was lower in the eastern part of Japan than in the western part. Therefore, it was not possible to conclude that rendaku occurs more often in the eastern part of Japan and less often in Kyushu. However, this research only focused on a limited set of surnames and only four regions, therefore further research is necessary.

Shirooka (2014) also mentioned the synchronic change in place names starting with 小 *ko*/*o*. He did not explicitly research the influence of the region on place names, but he did notice for some names that rendaku seemed to be more frequent in certain parts of Japan. He argues that in some regions, the rendaku rate increased slower from the Meiji period onwards than in other regions. Shirooka combined the results of 小川 *o-kawa/ko-kawa*, 小島 *o-shima/ko-shima*, 小林 *o-hayashi/ko-hayashi*, 小倉 *o-kura/ko-kura* and 小原 *o-hara/ko-hara* and listed in how many village names they did not undergo rendaku (table 1) or did undergo rendaku (table 2). Since Shirooka only present the old country names, I added the modern names as well for easy comparison.

Table 1: Regions with many village names without rendaku

11 villages	Ecchu (Toyama)
6 villages	Echigo (Niigata)
5 villages	Kaga (Ishikawa), Ise (Mie), Yamato (Nara)
4 villages	Tosa (Kochi)
3 villages	Bizen (Okayama), Iwaki (Fukushima, Miyagi), Shinano (Nagano), Oumi (Shiga), Kii (Wakayama, Mie), Iwashiro (Fukushima), Shimousa (Chiba, Ibaraki)
2 villages	Uzen (Yamagata), Ugo (Akita, Yamagata), Sado (Niigata), Shimotsuke (Tochigi), Mikawa (Aichi), Mino (Gifu), Tanba (Kyoto, Hyogo), Settsu (Osaka, Hyogo), Harima (Hyogo), Iyo (Ehime), Bungo (Oita)
1 village	Rikuchu (Iwate, Akita), Rikuzen (Miyagi, Iwate), Kozuke (Gunma), Hitachi (Ibaraki), Echizen (Fukui), Wakasa (Fukui), Yamashiro (Kyoto), Izumi (Osaka), Kawachi (Osaka), Bingo (Hiroshima), Awa (Tokushima), Inaba (Tottori), Iwami (Shimane), Chikuzen (Fukuoka), Chikugo (Fukuoka), Hizen (Saga, Nagasaki), Hyuga (Miyazaki, Kagoshima), Satsuma (Kagoshima)

(From Shirooka 2014, 40. Translation and modifications are mine)

Table 2: Regions with many village names with rendaku

9 villages	Musashi (Saitama, Tokyo, Kanagawa)
5 villages	Omi (Shiga)
4 villages	Kozuke (Gunma), Hitachi (Ibaraki), Totoumi (Shizuoka), Yamato (Nara)
3 villages	Shimotsuke (Tochigi), Mino (Gifu), Ise (Mie), Aki (Hiroshima)

2 villages	Iwaki (Fukushima, Miyagi), Kazusa (Chiba), Noto (Ishikawa), Mikawa (Aichi), Kii (Wakayama, Mie), Tanba (Kyoto, Hyogo), Mimasaka (Okayama), Iyo (Ehime), Hoki (Tottori), Osumi (Kagoshima)
1 village	Mutsu (Aomori, Iwate, Miyagi, Fukushima, Akita), Ugo (Akita, Yamagata), Iwashiro (Fukushima), Shimousa (Chiba, Ibaraki), Awa (Chiba), Sagami (Kanagawa), Suruga (Shizuoka), Ecchu (Toyama), Echizen (Fukui), Wakasa (Fukui), Bicchu (Okayama), Settsu (Osaka), Kawachi (Osaka), Harima (Hyogo), Tajima (Hyogo), Awa (Tokushima), Iwami (Shimane), Inaba (Tottori), Bungo (Oita), Hizen (Saga, Nagasaki), Higo (Kumamoto)

(From Shirooka 2014, 41. Translation and modifications are mine)

The data from Shirooka (2014) shows that rendaku mainly occurred in the Kanto region. Interestingly, this data also illustrates that names in the Kanto region almost only occurred with rendaku, especially in the Tokyo, Kanagawa, Chiba and Saitama prefectures. Shirooka (2014) suggests that it might have been possible that rendaku spread from the capital, which was the city of Edo at the time, to the outward regions. He also argues that the area which is now the Hokuriku area formed a strong front where rendaku did not occur very often in names starting with *o*/*ko*. However, Shirooka's research has only focused on names that started with *o* and *oo*, and thus cannot be extended to names in general. Furthermore, it must be noted that he did not include the total amount of villages with these names, regardless of rendaku. In order to judge whether rendaku really occurs more often, percentages or the total amount of villages would be necessary.

This chapter explored several studies regarding rendaku in Japanese surnames and place names from a wide variety of angles. In the following chapter, the research questions and hypotheses will be proposed.

5. Research questions and hypotheses

5.1 Rendaku according to a core periphery model

Since the effect of the Tokyo dialect is very strong, it is often difficult to see whether there is no dialectal variation across regions according to rendaku in common nouns, or whether it is the

influence of the Tokyo dialect. Surnames do show variation, although there are no sources where both the place of a person with the surname and the pronunciation are given. Furthermore, people can move or change the pronunciation of their name, making it difficult to make predictions about the occurrence of rendaku in relation to a region. Place names do not have this problem, since the place is fixed, just as the pronunciation in most cases. Therefore, it is very interesting to investigate place names to discover more about the rendaku sensitivity in Japan.

Based on the data of Shirooka (2014) as presented in table 1 and 2, it is possible to argue that rendaku in Japanese place names has spread from the cultural center of Edo towards the outer regions. This would be in accordance with the core periphery theory of Yanagita (1930), as discussed in paragraph 2. Furthermore, previous literature (Iwasaki 2013, Van Bokhorst 2017) suggest that it does not only apply to the names Shirooka (2014) investigated, but also to *Yamazaki* and *Nakajima*, amongst others. Therefore, the first research question this thesis tries to answer is given in (7) and the hypothesis in (8).

(7) **Research question 1:**

How is rendaku in place names distributed across Japan?

(8) **Hypothesis 1:**

The rendaku rate in Japanese place names is higher around a cultural and/or political center and decreases in the peripheral areas.

More specifically, this research expects to find differences in rendaku rate for place names in different regions, similarly to Shirooka (2014) and Van Bokhorst (2017). Japan has had many capitals, although many linguistic phenomena seem to be clustered around the old capital of Kyoto (Shibatani 1990) and rendaku in surnames around Tokyo (Shirooka 2014, Van Bokhorst 2017), this research will assume that rendaku in place names will be clustered around these central area of Japan as well, although it is difficult to predict exactly where the core will be. Furthermore, it is

expected that the rendaku rate will be lower in the areas west of the Kinki region, in the Hokuriku region (based on both Shirooka 2014 and Shibatani 1990) and east from the Kanto region. The null hypothesis would be that there is no difference between the center and the peripheral areas, meaning that the rendaku rate is similar in each place or shows variation, but does not slowly decreases the more you move away from the center.

However, it must be noted that since this research only focusses on place names in contemporary Japan, it is not possible to argue whether rendaku has really originated in a cultural centre and spread to the periphery. In order to draw a conclusion regarding the historical change, a diachronic research similar to Shirooka (2014) is necessary, although impossible within the scope of this research.

5.2 Morphemes and rendaku sensitivity

Based on the data of Rosen (2003) and Zamma (2005), it is possible to argue that there is a connection between names ending with the same morpheme and the rendaku rate of the name. Van Bokhorst (2017) also discovered a tendency for names ending with *shima*, such as *Nakajima* and *Tajima*, to behave similarly regarding rendaku. *Yamazaki* behaved similar as well, which would be in accordance with Zamma (2005), who also argues that both *saki* and *shima* were specified as accentless and did not have a specification for rendaku. The current research is not very concerned with accent, although based on both Rosen (2003) and Zamma (2005) it does seem plausible that these morphemes share similar specifications. According to Van Bokhorst (2017), the names *gawa* and *ta* underwent rendaku through their own distinct pattern, which is also argued in Zamma (2005). Therefore, the second research question (9) and hypothesis (10) of this thesis are as following:

(9) **Research question 2:**

Is there a connection in rendaku occurrence between names ending with the same morpheme?

(10) **Hypothesis 2:**

The rendaku rate per region depends on the morpheme.

This study expects to find that place names ending with the same morpheme will behave similarly regarding rendaku in various regions as well. For instance, it is expected that not only *Nakajima* will occur with less rendaku in the Kyushu area, but other names ending with *shima* will have a similar pattern. Furthermore, it is predicted that names ending with different morphemes might behave differently as well. The null hypothesis would be true if all names show the same pattern or if names ending with the same morpheme show different patterns.

In the following chapter, the selection criteria for the place names, chosen dialect division and the research method will be discussed.

6. Method

In order to answer the research questions, it is necessarily to investigate place names and where they are pronounced in more detail. Due to time constraints, there was decided to only include place names of which it is known that there is variation in the occurrence of rendaku rate. In order to make this selection, the list *Zenkoku no myōji* "Surnames in the entire country" (Suzaki 1999) was used, which is a list of the 1000 most frequent Japanese surnames. Based on this list, all of the names ending on the morphemes *hara*, *hata*, *kaki*, *kawa*, *saki*, *sawa*, *sato*, *shima*, *ta* and *tsuka,* were noted. The frequency of place names might differ from surnames, but since many surnames are also used as place names, it was decided to use this data. Names in which the second morpheme will not undergo rendaku due to Lyman's Law were omitted from the list. Furthermore, names that have multiple possible character combinations were counted as one when the meaning was similar. Examples are 太田, 多田 and 大田 for *Oo-ta* or 中島 and 中嶋 for *Naka-shima*.

Since the assumption in this research is that there is a difference in the rendaku rate between certain place names, it was decided to only include names which occurred more than ten

times in the data and of which at least five times differed regarding the occurrence of rendaku. The website *Japan Postal Code List* is based on a database from the Japan Post (30 April 2018) and allows for a quick search of surnames. Thus by inserting the place names in kanji plus the kana (for instance, "中島 なかじま"), it was possible to make a rough estimate for each name how often it occurs in Japan and what the rendaku rate should be.

Some morphemes, such as *hata* and *kaki*, did not have names with much variation in the pronunciation, therefore these names were ultimately omitted from this research. The morpheme *sato* only had one name that occurred frequently and showed internal variation. However, it will still be interesting to see if there are regional influences visible in this name, hence it is included in the experiment. Names such as *Ta-shima* and *Ko-shima* did not match the selection criteria, but were included to be contrasted to earlier research. Ultimately 39 place names were selected. These names are presented in table 3.

Table 3: Place names that show internal variation regarding rendaku

Kawa	Nishi-kawa, Miya-kawa, Kuro-kawa, Taki-kawa, Mae-kawa, Yoko-kawa
Saki	Yama-saki, Iwa-saki, Matsu-saki, Shima-saki
Sawa	Taki-sawa, Oo-sawa, Tomi-sawa
Sato	Naka-sato
Shima	Naka-shima, Ko-shima, Fuku-shima, Oo-shima, Tera-shima, Ta-shima
Ta	Yoshi-ta, Mae-ta, Oku-ta, Fuku-ta, Oo-ta, Taka-ta, Yasu-ta, Naka-ta, Sawa-ta, Miya-ta, Tomi-ta, Saka-ta, Toyo-ta, Kawa-ta, Tsuru-ta, Shio-ta, Ari-ta, Ya-ta, Fuka-ta

On 30 April 2018 the Japan Post published the newest edition of the address database. This database contains the place names written in kanji and katakana, and also the city and prefecture in which this name occurs. By using this list, the names in table 3 were investigated and the rendaku rate for each prefecture was calculated. Names were not counted if they occurred on Okinawa or one of the other Ryukyu islands, since these regions had very few data and are located far from the other parts of Japan as well. When a certain neighborhood had multiple

instances of the same name, for instance 西中島 *Nishi-Nakajima* and 東中島 *Higashi-Nakajima*, the name was only counted once. The results are presented in the following paragraph.

The names were divided into ten different regions, based on the discussion as presented in chapter 2. There was chosen for the classification of Koto (1989) (figure 1), with a few small exceptions. Since the Umpaku dialect only covers a very small area, it was decided to include names from this dialect region in the Chugoku region. Furthermore, based on Shibatani (1990) his observation regarding the distinction between eastern and western language traits in the Tokai-Tozan dialect area, it was decided to divide this into two different regions: East Chubu (consisting of Niigata prefecture, Nagano prefecture, Shizuoka prefecture and Yamanashi prefecture) and West Chubu (the Gifu and Aichi prefectures). The division can be seen in figure 2.

Figure 2: classification of regions

7. Results

7.1 Influence of the region

In this chapter the results of the research will be presented. A summary of the data can be found in table 4, which contains all of the surnames ending with the same morpheme grouped together and the rendaku rates for each region. The number in front of the slash is the amount of place names that underwent rendaku, the number behind the slash is the total amount of place names and the number between brackets is the rendaku rate in percentages. Only the investigated names were included, therefore the results will not be applicable to Japanese place names in general. However, since the research does consist of the names which have the largest variation in rendaku rate, it might show some interesting results.

Table 4: Rendaku rate of morphemes per region

Region/E2	***Kawa***	***Saki***	***Sato***	***Sawa***	***Shima***	***Ta***	**Total**
Hokkaido	0/6 (0,0)	1/2 (50,0)	0/7 (0,0)	0/6 (0,0)	18/23 (78,3)	6 /17 (35,3)	25/61 (41,0)
Tohoku	0/24 (0,0)	25/33 (75,8)	10/18 (55,6)	13/52 (25,0)	39/54 (72,2)	155/249 (62,2)	242/ 430 (56,3)
Kanto	3/16 (18,8)	21/29 (72,4)	29/30 (96,7)	4/18 (22,2)	53/78 (67,9)	101/165 (61,2)	211/336 (62,8)
East Chubu	8/18 (44,4)	17/24 (70,8)	4/5 (80,0)	3/22 (13,6)	72/84 (85,7)	82/111 (73,9)	186/264 (70,5)
Hokuriku	5/9 (55,6)	4/9 (44,4)	0/0 -	1/4 (25,0)	26/39 (66,7)	30/71 (42,3)	66/132 (50,0)
West Chubu	9/14 (64,3)	11/19 (57,9)	0/1 (0,0)	2/5 (40,0)	30/48 (62,5)	65/106 (61,3)	117/193 (60,6)
Kinki	16/26 (61,5)	13/31 (41,9)	5/6 (83,3)	3/9 (33,3)	37/67 (55,2)	123/183 (67,2)	197/322 (61,2)
Chugoku	4/8 (50,0)	4/11 (36,4)	1/1 (100)	0/4 (0,0)	8/29 (27,6)	76/104 (73,1)	93/157 (59,2)
Shikoku	8/13 (61,5)	1/7 (14,3)	1/1 (100)	0/0 -	6/20 (30,0)	32/51 (62,7)	48/92 (52,2)
Kyushu	6/8 (75,0)	10/29 (34,5)	8/9 (88,9)	0/0 -	6/50 (12,0)	60/107 (56,1)	90/203 (44,3)
Total	59/142 (41,5)	107/194 (55,2)	58/78 (74,4)	26/120 (21,7)	295/492 (60,0)	730/1164 (62,7)	1275/2190 (58,2)

When looking at table 4, a few observations can be made. The first point is that the rendaku rate for each morpheme seems to differ within each region. For instance, the rendaku rate of *shima* in East Chubu is 85,7%, while in West Chubu it is 62,5%. Furthermore, in each region the rendaku

rate also seems to differ for different morphemes. For instance, in East Chubu the rendaku rate of *shima* is 85,7%, while for *kawa* it is only 44,4%.

When looking at the *total* row, it can be seen that both the rendaku rate and the actual number of place names differ per morpheme. Some morphemes tend to appear more often than others and also seem to resist rendaku less often, such as *ta* which occurs 1164 times with a rendaku rate of 62,7%. On the other hand, s*awa* only occurs 120 times with a rendaku rate of 21,7%. However, these results are heavenly influenced by the names included in the research. Interestingly, the results in the *total* column do not only show that the total rendaku rate differs for each region, it also shows that it decreases in the peripheral areas as predicted by the core periphery theory. In order to investigate whether these results are significant, several logistic regressions were conducted in SPSS. The results can be found in table 5, in which $p \leq .050$ is considered to be statistically significant. These cells are marked in grey.

Table 5: Results of the logistic regressions across regions

	Hok-kaido	Toho-ku	Kanto	East Chubu	Hoku-riku	West Chubu	Kinki	Chu-goku	Shi-koku	Kyu-shu
Hokkaido		0.026	0.002	0.000	0.372	0.004	0.004	0.016	0.176	0.644
Tohoku	0.026		0.069	0.000	0.084	0.146	0.178	0.522	0.472	0.005
Kanto	0.002	0.069		0.049	0.003	0.958	0.669	0.449	0.066	0.000
East Chubu	0.000	0.000	0.049		0.000	0.079	0.019	0.019	0.002	0.000
Hokuriku	0.372	0.084	0.003	0.000		0.008	0.008	0.050	0.518	0.525
West Chubu	0.004	0.146	0.958	0.067	0.008		0.756	0.528	0.098	0.000
Kinki	0.004	0.178	0.669	0.019	0.008	0.756		0.683	0.112	0.000
Chugoku	0.016	0.522	0.449	0.019	0.050	0.528	0.683		0.278	0.005
Shikoku	0.176	0.472	0.066	0.002	0.518	0.098	0.112	0.278		0.212
Kyushu	0.644	0.005	0.000	0.000	0.525	0.000	0.000	0.005	0.212	

The results show a very clear pattern which is in accordance with the core periphery theory as discussed in chapter 2. Based on these results, the core seems to be in the eastern part of the Chubu area, since the rendaku rate in East Chubu is significantly higher compared to every other region in Japan, with West Chubu being the only exception.

The rendaku rate in both the Kanto and Kinki area is around 60% (p = .669), meaning that there is no significant difference in the rendaku rate. The rendaku rates in Tohoku, Chugoku and Shikoku are lower, but these results are not significant as well compared to both Kanto and Kinki.

On the other hand, the data shows that there are very clear peripheral areas as well. Both Hokkaido and Kyushu show a significant difference with the more central Kanto, East Chubu, West Chubu and Kinki regions, where rendaku occurs less in Hokkaido and Kyushu. Areas which are between the outer peripheral areas and the core areas have a significantly higher rendaku rate as well compared to the periphery. Tohoku shows a significant difference with Hokkaido (p = .026) and Kyushu (p = .005) and Chugoku does as well (p = .016 for Hokkaido and p = .005 for Kyushu).

The difference with the Shikoku region is not statistically significant, but Shikoku does not show significant differences with other regions as well, with the exception of East Chubu (p = .002). The rendaku rate in the Hokuriku region is significantly lower compared to Kanto (p = .003), East Chubu (p < .001), West Chubu (p = .008), Kinki (p = .008) and Chugoku (p = .050). Therefore, it is possible to argue that based on this data, the rendaku rate is lowest in the peripheral Hokkaido, Hokuriku and Kyushu regions and highest in the East Chubu region.

7.2 Influence of morphemes

The second aim of the research dealt with the rendaku rate in names in relation to the second morpheme. Similar to the regions in paragraph 7.1, logistic regressions were performed to determine whether the rendaku rate was significantly different across morphemes as well. The results are given in table 6.

Table 6: Results of the logistic regressions across morphemes

	Kawa	***Saki***	***Sato***	***Sawa***	***Shima***	***Ta***
Kawa		0.004	0.000	0.002	0.000	0.000
Saki	0.004		0.004	0.000	0.250	0.046
Sato	0.000	0.004		0.000	0.016	0.041
Sawa	0.002	0.000	0.000		0.000	0.000
Shima	0.000	0.250	0.016	0.000		0.291
Ta	0.000	0.046	0.041	0.000	0.291	

As can be seen in the table, most of the p-values are significant, meaning that there is a difference in the rendaku rate of almost every morpheme. The only names which did not have a significant difference were names ending with *shima* and names ending with *saki* (p = .250) on one hand, and names ending with *shima* and *ta* (p = .291) on the other hand. However, the rendaku rate is largely dependent on the names included in the research. There were only three names ending with *sawa*, of which *Oo-sawa* occurred the most often. This name is usually pronounced without rendaku, thus leading to a lower rendaku rate for names ending with *sawa* in general as well.

When looking at eventual patterns for specific morphemes based on table 4, it becomes clear that not every morpheme behaves the same. For *sawa* the rendaku rate is the highest in the West Chubu region, which contrasts the results of the other morphemes. The rendaku rate of *ta* is surprisingly highest in the Chugoku region, which can be considered a more peripheral area. The morphemes *saki*, *shima*, *sato* and *kawa* do not seem to have a central area with a higher rendaku rate, but show a gradual decline in rendaku from one outer end of Japan to another. In case of *saki* and *shima*, the rendaku rate becomes lower in the western parts of Japan, while for *sato* and *kawa* it becomes lower in the eastern parts of Japan.

In the following sections, the data will be discussed based on the different morphemes to see whether a (similar) pattern as in paragraph 7.1 is visible as well. If there is a pattern amongst morphemes, individual place names are expected to behave in accordance with the average pattern of that morpheme. However, since there was not enough data to maintain the ten regions used in the previous section, it was decided to divide the names into four regions (figure 3).

Figure 3: Division into four regions

These regions are the eastern region consisting of the Hokkaido and Tohoku areas, which are considered to be peripheral areas. The second region is the east central region including Kanto and the eastern part of the Chubu region. According to the data from 7.1, this can be considered a central area. The third region is the western central region, which includes the western part of the Chubu region, the Hokuriku region and the Kinki region. Both western Chubu and Kinki can also be considered as core areas according to the data from section 7.1. Hokuriku is a peripheral area, but has too little data to form its own group. Therefore, it has been included in the western central area based on the division Shibatani (1990) proposed of east and west Chubu. Lastly, the western area consists of the peripheral areas of Shikoku, Chugoku and Kyushu.

7.3 Names ending with kawa

Table 7: Rendaku in names ending with kawa

Name	East	Central East	Central West	West	Total
Miya-kawa	0/4 (0,0)	10/11 (90,0)	13/14 (92,9)	3/3 (100)	26/31 (81,3)

Mae-kawa	0/3 (0,0)	0/3 (0,0)	4/5 (80,0)	3/3 (100)	7/14 (50,0)
Nishi-kawa	0/7 (0,0)	0/2 (0,0)	7/12 (58,3)	4/5 (80,0)	11/26 (42,3)
Yoko-kawa	0/3 (0,0)	1/11 (90,9)	3/3 (100)	3/3 (100)	7/20 (35,0)
Kuro-kawa	0/13 (0,0)	0/7 (0,0)	3/15 (20,0)	5/15 (33,3)	8/50 (16,0)
Total	0/30 (0,0)	11/23 (32,4)	30/49 (61,2)	18/29 (62,1)	59/142 (41,5)

When looking at the data in table 7, there seems to be a very distinctive pattern. For every name, the rendaku rate is highest in the western areas and lowest in the eastern areas of Japan. After running several logistic regressions, it was found that the rendaku rate is significantly higher in the East Central area compared to both the West Central area ($p = .011$) and the West area ($p = .020$). Surprisingly, the East area did not have any significant results even though the rendaku rate was 0,0%. As observed in section 7.2 as well, it is possible to argue that rendaku occurs less frequently in the eastern part of Japan, although there does not seem to be a core around which it is clustered. Instead, there seems to be a gradual transition from non-rendaku areas in the east to rendaku-areas in the west.

When investigating individual place names names, the name *Miya-kawa* occurs significantly more with rendaku than the other place names ending with *kawa* ($p = .003$ for *Nishi-kawa*, $p < .001$ for *Kuro-kawa*, $p = .036$ for *Mae-kawa* and $p < .001$ for *Yoko-kawa*). Furthermore, *Kuro-kawa* occurs significantly less often with rendaku than the other names ($p < .001$ for *Nishi-kawa*, *Miya-kawa* and *Mae-kawa*, $p = .045$ for *Yoko-kawa*). However, since every name shows a similar pattern regardless of total rendaku rate, I would argue that names ending with *kawa* behave similarly, thus there being a very distinctive pattern for names ending with this morpheme.

7.4 Names ending with saki

A total of four names that ended with *saki* were included in the research. The results are presented in Table 8.

Table 8: Rendaku in names ending with saki

Name	East	Central East	Central West	West	Total
Matsu-saki	6/7 (85,7)	7/10 (70,0)	2/5 (40,0)	13/17 (76,5)	28/39 (71,8)
Yama-saki	20/21 (95,2)	24/27 (88,9)	21/34 (61,8)	1/16 (6,3)	66/98 (67,3)
Shima-saki	0/0 -	2/2 (100)	3/8 (37,5)	1/3 (33,3)	6/13 (46,2)
Iwa-saki	0/7 (0,0)	5/14 (35,7)	2/12 (16,7)	0/11 (0,0)	7/44 (15,9)
Total	26/35 (74,3)	38/53 (71,7)	28/59 (47,5)	15/47 (31,9)	107/194 (55,2)

Similar to the data in section 7.2, logistic regressions for the names ending with *saki* illustrate that the rendaku rate is significantly higher in the Eastern region compared to the Central West area ($p = .013$) and West area ($p < .001$), and also significantly higher in the Central East area compared to the Central West area ($p = .010$). It must be noted however, that half of the data is from the name *Yama-saki*, which might largely influence the total occurrence of rendaku in different areas.

As for the difference in rendaku rate, *Iwa-saki* has a significantly lower rendaku rate compared to *Matsu-saki* ($p < .001$), *Yama-saki* ($p < .001$) and *Shima-saki* ($p = .029$). The results for the other names turned out not be statistically significant, thus suggesting that there is little variation in rendaku rate or not enough data, which is most likely the case of *Shima-saki*.

When looking at the individual place names, *Yama-saki* seems to have a very distinctive pattern, where the rendaku rate is higher in the eastern part of Japan and gradually decreases in the western area. However, the other three names do not show this pattern. The data in table 8 illustrates that for *Matsu-saki* rendaku is blocked more often in the Central West region compared to other areas, for *Iwa-saki* there seems to be a core in the central east areas around which rendaku is clustered, while for *Shima-saki* there is not enough data to draw a conclusion. Therefore, it is not possible to conclude that names ending with *saki* behave similar in rendaku.

7.5 Names ending with sato

There was only one name included in the research that ended with *sato*, therefore it is not possible to draw a conclusion regarding names ending with *sato* in general. However, as can be seen in table 9, there does seem to be a pattern regarding rendaku in *Naka-sato*.

Table 9: Rendaku in names ending with sato

Name	East	Central East	Central West	West	Total
Naka-sato	10/25 (40,0)	33/35 (94,3)	5/7 (71,4)	10/11 (90,9)	58/78 (74,4)

A logistic regression showed that the rendaku rate is significantly lower in the eastern area as opposed to Central East ($p < .001$) and West ($p = .016$). The Central West area has less rendaku according to table 9, but these results were not significant. It can be concluded however, that the rendaku rate of *Naka-sato* is lower in east Japan.

7.6 Names ending with sawa

Table 8: Rendaku in names ending with sawa

Name	East	Central East	Central West	West	Total
Tomi-sawa	6/12 (50,0)	3/3 (100)	2/2 (100)	0/1 (0,0)	11/18 (61,1)
Taki-sawa	7/12 (58,3)	4/8 (50,0)	0/0 -	0/0 -	11/20 (55,0)
Oo-sawa	0/34 (0,0)	0/29 (0,0)	4/16 (25,0)	0/3 (0,0)	4/82 (4,9)
Total	10/25 (40,0)	33/35 (94,3)	5/7 (71,4)	10/11 (90,9)	58/78 (74,4)

Table 8 contains the data for the surnames ending on *sawa.* However, there was too little data to draw adequate conclusions as well. A logistic regression was performed for the different place names, which showed that *Oo-sawa* behaves different from both *Taki-sawa* and *Tomi-sawa* ($p < .001$). Therefore, it is possible to say that there is a difference within *sawa* based solely on the rendaku rate.

When comparing the rendaku rate in different regions, logistic regressions did not show any significant results however. Based on table 8 rendaku in both *Tomi*-sawa and *Oo-sawa* seems to be clustered around the central areas. Since names ending with *sawa* did not occur in the Shikoku and Kyushu regions, it is not possible to draw conclusions here as well.

7.7 Names ending with shima

There were more names ending with *shima* that show internal variation compared to *saki*. The results of the six names included in the research are presented in table 9.

Table 9: Rendaku in names ending with shima

Name	East	Central East	Central West	West	Total
Ko-shima	3/3 (100)	13/13 (100)	16/17 (94,1)	6/8 (75,0)	38/41 (92,7)
Ta-shima	2/2 (100)	25/25 (100)	6/7 (85,7)	2/9 (22,2)	35/43 (81,4)
Naka-shima	51/54 (94,4)	56/56 (100)	57/73 (78,1)	6/36 (16,7)	170/219 (77,6)
Tera-shima	0/1 (0,0)	6/8 (75,0)	5/7 (71,4)	1/4 (25,0)	12/20 (60,0)
Fuku-shima	1/15 (6,7)	14/24 (58,3)	5/19 (26,3)	1/13 (7,7)	21/71 (29,6)
Oo-shima	0/2 (0,0)	11/36 (30,6)	4/31 (12,9)	4/29 (13,8)	19/98 (19,4)
Total	57/77 (74,0)	125/162 (77,2)	93/154 (60,4)	20/99 (20,2)	295/492 (60,0)

When looking at all the names in total, there is a clear distinction in rendaku rate between the eastern region on one hand and the western region on the other. The rendaku rate in eastern Japan was significantly higher than in the central west and western areas ($p = .042$ and $p < .001$) respectively). Furthermore, the rendaku rate in the central eastern region was also significantly higher than in the western area ($p < .001$). Only the difference between the central east and central west areas was not significant. However, based on the data in table 9, there seems to be a general tendency for names to undergo rendaku more often in the eastern part of Japan and least often in the western part of Japan.

The rendaku rate varied across names as well. Based on several logistic regressions, it is possible to argue that *Ko-shima* occurs significantly more often with rendaku compared to *Naka-shima* ($p = .037$), *Fuku-shima* ($p < .001$), *Oo-shima* ($p < .001$) and *Tera-shim*a ($p = .005$). Oo-shima has the lowest rate and differs significantly with every other name except for *Fuku-shima* ($p < .001$ for every other name). Most of the other names were significantly different as well.

Not all names seem to behave according to the pattern as described above. Similar to *Yama-saki* in section 7.4, *Naka-shima* also occurs far more often than other names ending with *shima*, thus most likely distorting the results. The names *Ko-shima* and *Ta-shima* seem to behave similarly to *Naka-shima* based on the data in table 9. However, the names *Oo-shima*, *Fuku-shima* and *Tera-shima* seem to behave more similar to the core periphery model as seen in section 7.1, with a core in the Central East region and a decrease in the rendaku rate in both the western and the eastern regions of Japan. However, since there were not enough names, it was not possible to determine whether this decline is also statistically significant.

7.8 Names ending with ta

In total there were 19 names in the data that ended with *ta*. The rendaku rate of the individual names per region are given in table 10.

Table 10: Rendaku in names ending with ta

Name	East	Central East	Central West	West	Total
Yoshi-ta	13/16 (81,3)	40/40 (100)	32/36 (88,9)	47/47 (100)	132/139 (95,0)
Oka-ta	10/10 (100)	15/18 (83,3)	13/15 (86,7)	11/11 (100)	49/54 (90,7)
Fuku-ta	18/18 (100)	23/24 (95,8)	11/20 (55,0)	29/30 (96,7)	81/92 (88,0)
Mae-ta	38/50 (76,0)	8/11 (72,7)	39/39 (100)	17/17 (87,2)	102/117 (87,2)
Sawa-ta	19/25 (76,0)	7/7 (100)	8/8 (100)	4/7 (57,1)	38/47 (80,9)
Yasu-ta	4/9 (44,4)	2/2 (100)	13/17 (76,5)	9/9 (100)	28/37 (75,7)
Fuka-ta	2/2 (100)	0/0 -	8/11 (72,7)	2/5 (40,0)	12/18 (66,7)
Kawa-ta	2/3	6/9	7/10	1/3	16/25

	(66,7)	(66,7)	(70,0)	(33,3)	(64,0)
Toyo-ta	9/21 (42,9)	16/17 (94,1)	11/18 (61,1)	5/9 (55,6)	41/65 (63,1)
Naka-ta	12/18 (66,7)	10/17 (58,8)	14/25 (56,0)	6/8 (75,0)	42/26 (61,8)
Taka-ta	25/27 (92,6)	24/32 (75,0)	20/39 (51,3)	4/24 (16,7)	73/122 (59,8)
Ari-ta	0/1 (0,0)	0/1 (0,0)	5/5 (100)	3/8 (37,5)	8/15 (53,3)
Tomi-ta	1/14 (7,1)	10/18 (55,6)	15/23 (65,2)	5/7 (71,4)	31/63 (50,0)
Shio-ta	3/3 (100)	6/6 (100)	0/10 (0,0)	1/5 (20,0)	10/24 (41,7)
Ya-ta	3/3 (100)	3/6 (50,0)	7/21 (33,3)	3/9 (33,3)	16/39 (41,0)
Miya-ta	2/9 (22,2)	6/12 (50,0)	6/14 (42,9)	5/13 (38,5)	19/48 (39,6)
Tsuru-ta	0/5 (0,0)	1/4 (25,0)	0/3 (0,0)	5/7 (71,4)	6/19 (31,6)
Saka-ta	0/1 (0,0)	3/11 (27,3)	1/7 (14,3)	4/8 (50,0)	8/27 (29,6)
Oo-ta	0/31 (0,0)	3/41 (7,3)	8/39 (20,5)	7/35 (20,0)	18/146 (12,3)
Total	161/266 (60,5)	183/276 (66,3)	218/360 (60,6)	168/262 (64,1)	730/1164 (62,7)

When looking at the data, there does not really seem to be a pattern. The rendaku rate of all the names taken together is between 60% and 67% for each region. Logistic regressions showed that the difference in rendaku rate in these regions was not statistically significant.

There were many significant differences between individual names as well. It would be very tedious to describe every significant difference, but as can be seen in the table, the rendaku rate varies greatly among names. While *Yoshi-ta* and *Oka-ta* have a rendaku rate of more than 90%, the name *Oo-ta* only occurs with rendaku in around 12% of the names.

When further examining the individual names, there does not seem to be a pattern at first glance. However, rendaku in names does occur according to various patterns. For instance, for *Oka-ta* and *Fuku-ta* the rendaku rate is lower in the central east and central west areas compared to the peripheral areas. This is interesting, since this research predicted that the rendaku rate would be higher in the central areas, since rendaku can be seen as the 'newer' variety. In *Oka-ta* and *Fuku-ta*, it seems that the opposite is happening: non-rendaku is higher in the central areas,

while the rendaku rate increases in the peripheral areas. The name *Naka-ta* shows a similar pattern as well.

Toyo-ta does seem to occur in Japan according to the core periphery theory as predicted, with a core in the central east area. Not only *Toyo-ta* has this pattern, *Miya-ta* and *Sawa-ta* behave similarly as well. The name *Taka-ta* also shows a very distinct pattern in the sense that the rendaku rate is higher in the eastern and central eastern regions, and especially lower in the western region. However, statistics were not conducted, so no conclusions can be drawn based on this.

This chapter presented the data of the place names based on the database of the Japan Post. First there was investigated what the influence of the region was, and after that the morphemes and individual names were considered. In the following chapter, the data will be further evaluated and explained where possible.

8. Discussion

8.1 Regional influences

The first aim of this thesis was to discover how rendaku is spread across Japan. The hypothesis was that rendaku is clustered around a political/cultural center, most likely in the Kanto or Kinki areas. Based on the data in chapter 7 it seems that the rendaku rate is dependent on the region where a place name is located. However, there is not a clear pattern as initially predicted. When all of the names are combined, the rendaku rate seems to be clustered around the Chubu area, which can be considered as one of the metropolitan areas of Japan and thus a center area. Furthermore, the data also showed that the rendaku rate is significantly lower in the peripheral areas of Hokkaido, Hokuriku and Kyushu. Therefore, for all the data combined, rendaku occurs according to the core periphery theory.

Regarding the Hokuriku region, Shirooka (2014) also pointed out that already in the Meiji area, this region seemed to block rendaku more often than other regions. For a follow up research, it would be interesting to see whether rendaku in other names and/or common nouns occur also with less rendaku in the dialects spoken in the Hokuriku region.

It must be noted that the obtained pattern is largely dependent on the names included in the research, as became clear by the data regarding morphemes and individual names. This research included many names ending with *ta*, which might have influenced the results as well. Investigating other names will most likely lead to a different image. However, for many names the rendaku rate did seem to decrease in one or more peripheral areas.

The data illustrated that when combining all of the names belonging to one morpheme, there are several patterns as well. However, these patterns differed per morpheme and name. For instance, names ending with *kawa* had a higher rendaku rate in the western part of Japan, while names ending with *shima* had a higher rendaku rate in the eastern part of Japan. One of the implications the findings of these two reverse patterns has, is that it is not possible to claim that one dialect is less sensitive to rendaku. It really seems to depend on the names, which will be the central theme of the following section.

8.2 Influence of morphemes

The second aim of this research was to investigate whether the second morpheme of a name influenced the occurrence of rendaku. It was expected that rendaku in place names ending with the same morpheme would be spread in a similar manner across Japan. However, the results showed that there does not seem to be a very clear relationship between the morpheme on which a place name ends and the occurrence of rendaku, with the exception of *kawa*. Not only does the rendaku rate of individual names regardless of location varies, the pattern in which rendaku within these names occur differs as well.

In order to have a better grip on the data of individual names, the place names will be categorized based on their pattern. Based on the results in chapter 7, it is possible to distinguish four different categories (11).

(11) a. **Core in a central area**

Iwa-saki, Tomi-sawa, Oo-sawa, Fuku-shima, Oo-shima, Tera-shima, Sawa-ta, Toyo-ta, Miya-ta

b. **Reversed core in a central area**

Matu-saki, Oka-ta, Fuku-ta, Naka-ta

c. **Gradual decline in the west**

Yama-saki, Naka-shima, Ko-shima, Ta-shima, Fuka-ta, Taka-ta, Kawa-ta, Ya-ta

d. **Gradual decline in the east**

Nishi-kawa, Kuro-kawa, Mae-kawa, Yoko-kawa, Miya-kawa, Naka-sato, Yasu-ta, Tomi-ta, Yosi-ta, Tsuru-ta, Mae-ta, Oo-ta

e. **Not enough data**

Shima-saki, Taki-sawa, Ari-ta, Shio-ta, Saka-ta

Within these categories, it was investigated whether these names shared similarities in the first morpheme in order to explain the results.

Regarding the names that occur according to a core periphery theory (11a), there does not seem to be a very apparent center area which is the same for every name. For *Iwa-saki* the core seems to be in the Central East region, while for *Oo-sawa* it seems to be in the Central West area. There were only three names ending with *sawa* included in the research of which one has too little data, although the other two names do seem to occur according to a core periphery theory as well. However, more names ending with *sawa* should be investigated in order to draw any conclusions regarding this observation. There are also three names ending with *shima* that occur according to this pattern, although no similarities could be found. Lastly, the three names ending with *ta* that could be placed in this category all contained a glide in the first morpheme. Only the name *Kawa-ta* was classified in a different group.

As for the category with a reversed core (11b), it is meant that the rendaku rate is actually lower in the core area compared to the other areas of Japan. Interestingly, there are three names

that end on *ta* which look very similar and can be classified according to this pattern. The rendaku rate was lowest for *Oka-ta* in the central areas, while for *Fuku-ta* and *Naka-ta* it was lowest in the West Central area. These names are similar in the sense that the last mora of the first morpheme contains a /k/, although other similar names such as *Fuka-ta* and *Taka-ta* showed a different pattern.

The third observed pattern shows a gradual decrease in rendaku the more a name is located in the western part of Japan (11c). Both *Fuka-ta* and *Taka-ta* mentioned above show a very clear decrease in rendaku rate. The names *Yama-saki* and *Naka-shima* also occur according to this pattern, just as was found in the *Nikkei Shinbun* (Iwasaki 2013) and Van Bokhorst (2017). According to Van Bokhorst (2017), the name *Ta-shima* had a similar pattern as well. In the study it was found that not only *Ta-shima*, but also *Ko-shima* shows a gradual decrease in rendaku rate in the western areas of Japan. Both names only consist of one mora in the first element, and the only other name in the data that also starts with a mora of one morpheme (*Ya-ta*) also belongs to this pattern, although this might be a coincidence.

Lastly, there was also a pattern where the rendaku rate decreased more in the eastern areas of Japan (11d). Every name ending with *kawa* behaved according to this pattern. The only name ending with *sato* also had a far lower rendaku rate in the east, although there is not enough data regarding this name. There were a few names ending with *ta* as well, although there are few similarities between these names. *Mae-ta* and *Oo-ta* are similar in the sense that they both have a long vowel in the first morpheme, while *Yosi-ta* and *Yasu-ta* are the only names ending with *ta* with an /s/ in the first morpheme. The only other name with an /s/ in the second mora of the first morpheme that appeared in the data was *Nishi-kawa*, which also falls under this group.

There were a few names of which there was little data and of which it was thus very difficult to classify them based on the percentages only (11e). These will not be further discussed. To sum up the above findings, names ending with *kawa* show a clear pattern regarding rendaku rate. Names ending with *shima* can be divided into two distinctive patterns, although this falsifies the hypothesis that names ending with the same morpheme all have a similar pattern as well. For *sawa* and *sato* more data is necessary, and names ending with *ta* or *saki* have a lot of variation

between individual names. Furthermore, it was examined whether the second mora in the first morpheme of the name influences the rendaku rate, although there is not enough data to draw any conclusions regarding this matter.

8.3 Names with the same first morpheme

The data consisted of a few place names that have the same first morpheme. Similar to individual segments in the previous section, there was also investigated whether there might be a relationship between the first morpheme and the rendaku rate. The names that share the same initial morpheme are listed in (12).

(12)

Oo	Oo-sawa (core), Oo-shima (core), Oo-ta (more west)
Naka	Naka-shima (more east), Naka-ta (reversed core)
Fuku	Fuku-shima (core), Fuku-ta (reversed core)
Miya	Miya-kawa (more west), Miya-ta (core)
Tomi	Tomi-sawa (core), Tomi-ta (more west)
Mae	Mae-kawa (more west), Mae-ta (more west)

As for names beginning with *oo*, *Oo-sawa* and *Oo-shima* were classified according to the core periphery model, while *Oo-ta* showed a decrease in the eastern areas of Japan but had a very high rendaku rate in the west. Similarly, the patterns for names beginning with *naka*, *fuku*, *miya* and *tomi* differed as well. Only the names starting with *mae* behaved similar, with both *Mae-kawa* and *Mae-ta* having a lower rendaku in the eastern parts of Japan. However, there were only two names included in the research that started with *mae*, therefore this might very well have been a coincidence. An interesting observation is that *Fuku-shima* and *Fuku-ta* actually show a complete opposite pattern. The spread of rendaku in both names seems to be according to the core periphery theory, although in *Fuku-shima* the rendaku rate is highest in the core area, while in *Fuku-ta* it is lowest in the core area.

Based on the initial morphemes, there was no clear pattern regarding the initial morpheme and the occurrence of rendaku in the second morpheme as well. Therefore, the initial morpheme does not seem to exert influence on the variation within the rendaku rate of Japanese place names.

8.4 Comparison with previous research

Chapter 4 discussed earlier research regarding rendaku in Japanese surnames and place names. In this section, the results of the earlier research and the current research will be compared.

Shirooka (2014) investigated the change in rendaku rate in names starting with *oo* and *o* from the Meiji period until the Heisei period. He found that names starting with *oo* usually do not undergo rendaku, while names starting with *o*/*ko* usually do not occur without rendaku. The results of the current investigation showed a similar trend: *Oo-sawa*, *Oo-shima* and *Oo-ta* all had very low rendaku rates (4,9%, 19,4% and 12,3% respectively). On the contrary, *Ko-shima* had a very high rendaku rate (92,7%). Other names starting with *o*/*ko* were not investigated, since they barely occurred without rendaku.

Interestingly, some names behave different then in Tanaka (2017) and Van Bokhorst (2017). The rendaku rate of *Naka-ta* in particular was surprising. When you look at surnames, both Tanaka (2017) and Van Bokhorst (2017) showed that the rendaku rate is less than 20%. However, this study points out that the total rendaku rate of *Naka-ta* is actually 66,7% when you investigate place names, which is a very large difference. Furthermore, Van Bokhorst (2017) showed that *Nakada* was pronounced mainly in the Kanto region, while the results of the current research showed that the central areas have a lower rendaku rate than the peripheral areas.

Lastly, the *Nikkei Shinbun* (Iwasaki 2013) and Van Bokhorst (2017) showed that the rendaku rate of *Naka-shima* and *Yama-saki* decreased in the western areas. This was found in the current research as well. Similarly to Van Bokhorst (2017), *Ta-shima* showed a similar pattern to *Naka-shima*. The opposite pattern for *kawa*, with the rendaku rate being higher in the west, was also found in the current research and even expanded.

8.5 Problems and further research

Because this research focused on place names, there is fewer data than when surnames would be considered. Therefore, there were many names for which it was difficult to draw adequate conclusions. Preferably more names should be included. It would also be interesting to investigate how people pronounce the names such as Van Bokhorst (2017) did, to see whether the productivity of rendaku differs across different dialect regions. One name that might suggest this is the occurrence (or rather non-occurrence) of *Tani-kawa* as a place name in the Kanto area. Van Bokhorst (2017) found that this name was pronounced far more often with rendaku in the western part of Japan compared to the eastern part. This is similar to the results if *kawa* within the current research. Interestingly, the place name *Tani-kawa* only occurs with rendaku in Japan, although the participants from the Kanto area almost unanimously did not apply rendaku (even though they did in the Kansai and Kyushu areas). Therefore, there could be argued that the rendaku productivity of *kawa* is lower in the eastern parts of Japan, although further research with place names, common nouns or non-words would be necessary.

Furthermore, it is not clear whether the kana writing of a name and the actual pronunciation is the same. For instance, someone pointed out to me that the JR 尾久駅 in Tokyo is pronounced as *Oku-eki*, while all the other stations and places in the neighborhood are pronounced as *Ogu*. Therefore, the people living in the neighborhood always pronounce *Oku-eki* as *Ogu-eki*, even though it is officially written with a *k*. There might be more place names in Japan which are written differently from how they are pronounced, although for the scope of this research, it was impossible to verify the pronunciation.

Concerning the method of this research, there were several difficulties with the classification of place names as well. For instance, in names such as *Nakatashima*, it was impossible to know whether it was *Naka Tashima* or *Nakata shima*. In case of *Nakadashima* or *Nakadajima* it was clear that *Nakada* was one element since *ta* as a second morpheme underwent rendaku, thus easy to include in the data. However, this might have led to a higher rendaku rate, since only the names without rendaku were ambiguous and thus left out of the research.

Furthermore, the grouping of prefectures might also have influenced the results. In order to have results which are as accurately as possible, there was chosen to use ten regions for the general analysis. By using ten regions, it is less likely that some interesting phenomenon is overseen, while the regions still have sufficient data for analysis. However, this was not possible for the analysis of morphemes or individual names since there was not enough data to maintain ten groups. The consequence is that distinctions within certain regions disappear, leading to an image of the spread which is less accurate.

Lastly, there has been a lot of research regarding rendaku in names related to accent as well, as was mentioned in chapter 4. For this research, I have not investigate the accent patterns. For a follow up research it would be very interesting to investigate whether the accent patterns of place names with and without rendaku differs and if it coincides with the accent pattern of the region. Zamma (2005) described accent tendencies for names ending with several morphemes. It would be interesting to see whether the accent patterns he found can be applied across the entire country, or whether the accent also differs across regions.

9. Conclusion

The literature regarding rendaku in Japan is extensive, although few scholars investigated the occurrence of rendaku in Japanese dialects. The few studies that were conducted showed that within common nouns, there was little to no variation across dialects. However, Japanese surnames and place names do show variation in rendaku across the country. Therefore, this research investigated whether the rendaku rate varies in different areas of Japan.

The first research question dealt with the geographical distribution of rendaku. It was predicted that rendaku would occur according to a core periphery model, where the rendaku rate would be highest in the core area. The findings suggest that in case there was a core, this was usually somewhere in the Kanto, Chubu or Kinki areas, which coincides with the three large metropolitan areas of Japan. Furthermore, the rendaku rate had a tendency to be lower in the peripheral areas such as Hokkaido, Hokuriku and Kyushu. Thus, the results indicated that the

region has an influence on the rendaku rate, although the pattern is not that clear as predicted. For all the data combined the rendaku rate seemed to be highest in the Chubu area and decreased in the predicted peripheral areas, but there were also names where the rendaku rate was actually higher in the peripheral areas and slowly decreased in the other parts of Japan.

The second part of this research had to do with the relationship between morphemes and the occurrence of rendaku. Previous literature suggested that the final morpheme in a compound can exert influence on the occurrence of rendaku as well. Some morphemes might have a larger tendency to block rendaku or might be specified for rendaku or accent. It was predicted that the rendaku rate of names ending with the same morpheme would behave similar with regard to rendaku in various regions in Japan. The results showed that names ending with *kawa* all behaved similarly with the rendaku rate being the highest in Kyushu and gradually decreases in the direction of Hokkaido. However, other morphemes did not have a clear pattern or had multiple distinctive patterns.

Contrary to the research based on common nouns, this research shows that there are differences in the rendaku rate in different regions. There are regions, such as the peripheral areas, where rendaku tends to occur less, though this is heavily dependent on individual names. Since accent is often said to be strongly connected to rendaku, further research should focus on the accent patterns of these names across Japan to see whether there is a relation between the accent and the occurrence of rendaku as well.

Bibliography

D-maps.com. Map Japan. https://d-maps.com/carte.php?num_car=365&lang=en. Accessed June 2018.

Irwin, Marc and Timothy Vance. 2015. Rendaku across Japanese Dialects. *Phonological Studies* 18, 19-26.

Iwasaki, Shoichi. 2013. *Japanese. Revised edition.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Japan Post. http://www.post.japanpost.jp/zipcode/download.html. Accessed May 2018.

Japan Postal Code List. http://https://yubin.senmon.net/en/. Accessed May 2018.

Kadooka, Ken-Ichi. 2007. A Note on the Division of the Japanese Dialects. *The Ryukoku Journal of Humanities and Sciences* 29 (1), 157-173.

Katō, Masanobu. 1989. Nihongo [Japanese]. In *Gengogaku Daijiten* [Encyclopaedia of Linguistics]. Tokyo: Sanseidō.

Kubozono, Haruo. 2005. Rendaku: It's domain and linguistic conditions. In *Voicing in Japanese*, edited by Jeroen van de Weijer, Kensuke Nanjo, and Tetsuo Nishihara, 5-24. Berlin: Mouton de Gruyter.

Rosen, Eric. 2003. Systematic Irregularity in Japanese Rendaku: How to grammar mediates patterned lexical exceptions. *The Canadian Journal of Linguistics* 48 (1/2), 1-37.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

Shirooka, Keiji. 2014. O/ko (小) to ō (大) ga chimei no rendaku ni ataeru eikyō ni tsuite: meijiki sonmei to sono go no onpenka kara rendaku/hirendaku no keikō to kisokusei wo yomitoru [About the influence of o/ko and oo on rendaku of place names: names from places in the Meiji period and the later phonetic change leading to the tendency and rules of rendaku/non-rendaku]. *Jinbun ronshū* 65, 27-62.

Sugito, Miyoko. 1967. *Goon to akusento [Word sounds and accents]*. *Imada san to Shibata san* [Iwada and Shibata]. Osaka: Izumi Shōin.

Suzaki, H. 1999. *Zenkoku no myōji [Surnames in the entire country]*. Available at: http://www2s.biglobe.ne.jp/~suzakihp/index40.html (last accessed February 2017).

Tanaka, Yu. 2017. Phonotactically-driven rendaku in surnames: A linguistic study using social media. In *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL)* 34.

Tamaoka, Katsuo & Fumiko Ikeda. 2010. Whiskey or Bhiskey?: Influence of First Element and Dialect region on Sequential Voicing of *shoochuu*. *Gengo Kenkyu* 137, 65-79.

Vance, Timothy. 1987. *An introduction to Japanese phonology*. New York: Stat University of New York Press.

Van Bokhorst, Michelle. 2017. Rendaku in Japanese surnames. Unpublished Bachelor thesis. Leiden University.

Yanagita, Kunio. 1930. *Kagyūkō* [On snails]. Tokyo: Tōkō Shoin.

Zamma, Hideki. 2005. Correlation between accentuation and rendaku in Japanese surnames: With particular attention to morphemes. In *Voicing in Japanese*, edited by Jeroen van de Weijer, Kensuke Nanjo, and Tetsuo Nishiara, 157-176. Berlin: Mouton de Gruyter.