2024. *Proc Ling Soc Amer* 9(1). 5729. https://doi.org/10.3765/plsa.v9i1.5729.

Regional differences (or lack thereof) in rendaku in Japanese surnames

Yu Tanaka*

Abstract. No study has thoroughly investigated the regional differences of Japanese compound voicing known as *rendaku*. The present study addresses this issue by conducting a large-scale web-based survey with 492 Japanese speakers as participants and 1,776 compound surnames as stimuli. The results show no clear effects of dialects on rendaku application. This raises a novel theoretical issue for further investigation: Even though pitch accent has been argued to be inversely associated with rendaku (e.g., Sugito 1965; Zamma & Asai 2017), dialects with a variety of accentuation patterns nonetheless exhibit very similar voicing patterns in compound surnames. A tentative account based on foot structure is proposed.

Keywords. rendaku; Japanese; compound surname; pitch accent

1. Introduction. Compound formation in Japanese often involves a voicing alternation known as *rendaku*. As shown in (1) below, when rendaku applies, the initial consonant of the second element becomes voiced, possibly with additional changes in the place and manner.

(1) Rendaku application in compounds

a.	maki	+	s̲uʑi	→	maki-**z**uʑi	‘rolled-sushi’
b.	oɾi	+	k̲ami	→	oɾi-**g**ami	‘folding-paper’
c.	ike	+	h̲ana	→	ike-**b**ana	‘enlivening-flower’

Rendaku is a variable rule and does not occur in every compound. Words composed of similar elements may or may not show voicing, as can be seen in the names of the two Japanese scripts: [hira-**g**̲ana] ‘plain-character’ and [kata-k̲ana] ‘fragmented-character’. Researchers have proposed a number of factors, including lexical idiosyncrasy, in order to account for this irregular morphophonological process.

Even though rendaku is one of the most well-studied phenomena in Japanese, not much is known about its regional differences. Only a few studies have directly addressed the issue; van Bokhorst (2018) and Takemura et al. (2019) independently investigated the rendaku patterns of compound place names across Japan, both showing that, although some minor differences are found, there are no systematic dialectal effects on the way compound voicing occurs.[1] However, several methodological and other challenges, which are listed in (2), make it difficult to take these results at face value.

(2) Problems of using place names for dialectal research
- a. Non-uniform sampling
- b. Effects of non-linguistic factors
- c. Possible differences from local pronunciations

* I thank the audience at the 2024 Annual Meeting of the Linguistic Society of America (New York City; January 4-7, 2024) for their comments. This study was partially supported by JSPS Kakenhi Grants (20K13039; 22K13106).
Author: Yu Tanaka, Doshisha University (yutanak@mail.doshisha.ac.jp; yutanaka@ninjal.ac.jp).
[1] Sugito (1965) also discusses some dialectal differences in rendaku in surnames, but only compares two dialects using surnames with a particular second-element morpheme (田 /ta/ ‘paddy’). For research on rendaku in specific dialects, see, for example, Vance et al. (2014, 2021).

Place name data may be skewed due to non-uniform sampling. Different regions have distinct place names with idiosyncratic rendaku behaviors. For example, voicing is found in *Yodogawa* [jodo-**g**awa] 'stagnant-river' in Osaka but not in *Arakawa* [ara-kawa] 'rough-river' in Tokyo. It is unclear whether this really reflects dialectal differences with respect to rendaku application. The distribution of place names may also be affected by non-linguistic factors. Recently, so-called "municipal amalgamations" resulted in massive mergers of towns and cities. Many place names were thus lost without any obvious linguistic reasons. Lastly, place names registered in the official postal code directory, on which the previous studies are based, can be different from local pronunciations. For instance, *Kobuchisawa* [ko-butɕi-sawa] 'small-abyss-stream', a town in Yamanashi, is registered with the non-rendaku reading; yet some local residents say [ko-butɕi-**z**awa] with voicing, which is allegedly an original pronunciation. This suggests that the postal code directory may be unsuitable for dialectal research.

In order to examine the dialectal differences of rendaku while avoiding the problems above, the present study employs compound surnames as its main data. As in regular compound words and place names, rendaku may apply in compound surnames. Some surnames show rendaku while others do not, and there are also some that variably undergo voicing, as shown in (3). Note that surnames are usually written in *kanji* (Chinese characters) and rendaku application is not reflected in the writing.

(3) Variable rendaku application in compound surnames

a.	oka	+	ta	→	oka-**d**a	'hill-paddy'	岡田
b.	saka	+	ta	→	saka-ta	'slope-paddy'	坂田
c.	naka	+	ta	→	naka-ta ∼ naka-**d**a	'central-paddy'	中田

I conduct an on-line survey with Japanese speakers from various regions as participants and common compound surnames as experimental stimuli. This methodology has several advantages. It guarantees quasi-uniform sampling and takes actual regional pronunciations into account, since it collects data of rendaku judgment about the same or a similar set of surnames from a large number of speakers.[2] Additionally, the sound patterns of surnames have some independent theoretical issues; conducting research on surnames can further our understanding of Japanese phonology, as will be discussed in more detail in Section 3.

2. Experiment. In this section, I report a web-based experiment aiming to investigate the patterns of rendaku in compound surnames across Japanese dialects.[3]

2.1. METHODS.

2.1.1. STIMULI. From existing surname databases (Shirooka & Murayama 2011; Suzaki 2013), I extracted 1,176 common surnames that were composed of two morphemes and had second elements with initial voiceless obstruents (i.e., potential rendaku undergoers). For each participant, 120 items were randomly chosen to be presented as the stimuli. Some examples of the surnames are given in Table 1.[4]

[2] Due to randomization (see the next section for details), the stimulus set was not identical for every speaker. However, with a large number of participants, the overall results would nonetheless show general dialectal differences, if any, with respect to the rendaku applicability of common surnames.

[3] The experiment was originally conducted for a different study that had other purposes. The reader is referred to Tanaka, Y. (to appear) for details.

[4] A full list of the stimuli is provided as the supplementary material of Tanaka, Y. (to appear).

No	UR	Kanji	No	UR	Kanji	No	UR	Kanji
1	sa-to:	佐藤	11	jama-saki	山崎	21	mura-kami	村上
2	suzu-ki	鈴木	12	ike-ta	池田	22	kon-to:	近藤
3	taka-hasi	高橋	13	jama-sita	山下	23	ao-ki	青木
4	i-to:	伊藤	14	isi-kawa	石川	24	huku-ta	福田
5	ko-hajasi	小林	15	naka-sima	中島	25	huzi-hara	藤原
6	ka-to:	加藤	16	mae-ta	前田	26	naka-kawa	中川
7	jama-ta	山田	17	huzi-ta	藤田	27	hara-ta	原田
8	josi-ta	吉田	18	oka-ta	岡田	28	matu-ta	松田
9	sai-to:	斎藤	19	go-to:	後藤	29	en-to:	遠藤
10	jama-kuti	山口	20	o-kawa	小川	30	sai-to:	斉藤

Table 1. Examples of stimulus items

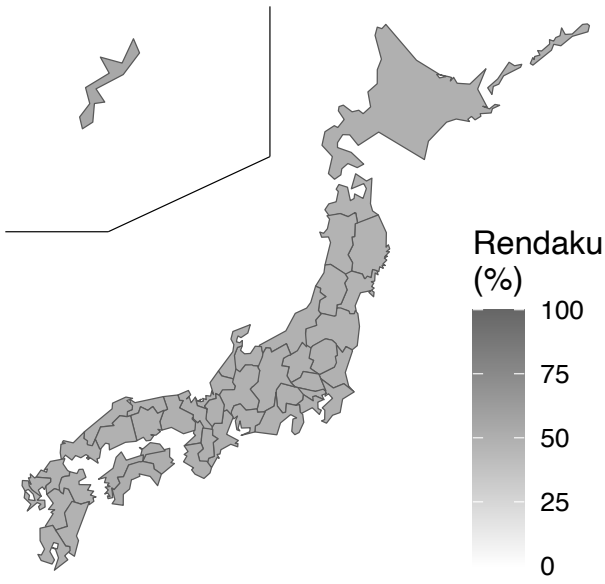
2.1.2. PROCEDURE. The experiment was implemented on the web-based survey platform Experigen (Becker & Levine 2013). At each trial, participants were presented with a compound surname written in kanji (e.g., 山崎 'mountain-cape'), which came with the honorific suffix *-san* in phonographic hiragana and was embedded in a frame sentence ("There is a person called 山崎-san."). They were also given two numbered options: one with the rendaku reading (e.g., 1. *Yamazaki*) and the other with the non-rendaku reading (e.g., 2. *Yamasaki*) of the surname written in hiragana. They were asked to read both of them out loud, and to choose the one that would sound more natural by clicking on a button with the corresponding number. They completed this task 120 times with surnames randomly selected from the stimulus pool. At the end of the experiment, they also voluntarily filled out a questionnaire about their age and home prefecture as well as their knowledge of linguistics.

2.1.3. PARTICIPANTS. Through a crowdsourcing service (*CrowdWorks*), 500 native speakers of Japanese were recruited for a reward of 200 Japanese Yen. As it turned out, eight participants did not provide their home prefecture, and thus their data were discarded. The remaining 492 speakers (mean age: 39.22; SD: 10.59) were further grouped by 10 dialect-based regions, roughly following the dialect divisions proposed by previous studies (see Tojo 1954; Kato 1977). The distributions of the participants' home regions are as follows: *Hokkaido*: 23; *Tohoku*: 35; *Tokyo/Kanto*: 159; *Tokai-Tosan*: 62; *Hokuriku*: 12; *Osaka/Kansai*: 99; *Chugoku*: 31; *Shikoku*: 13; *Kyushu*: 54; *Okinawa*: 4.[5]

2.2. RESULTS.

2.2.1. RENDAKU RATES BY REGION. Figure 1 shows the overall average rendaku application rates by region. They are plotted on a color scale over a map of Japan. The detailed figures are also provided in the table on the right. (*n* represents the number of speakers.) As can be seen, there are no clear regional differences. All the regions show similar colors on the map, and the figures all hover around 50%, as shown in the table. This indicates that the overall

[5] The participant data reported in this study are different from those in Tanaka, Y. (to appear) (n = 463) since the latter had stricter exclusion criteria and did not analyze the data of speakers who had participated in another similar study or had studied linguistics. Those speakers were included here as their data would not affect the results for the purpose of this study.



Region	*n*	Rendaku
Hokkaido	23	49.51%
Tohoku	35	47.88%
Tokyo/Kanto	159	48.62%
Tokai-Tosan	62	48.98%
Hokuriku	12	49.17%
Osaka/Kansai	99	49.38%
Chugoku	31	48.63%
Shikoku	13	50.83%
Kyushu	54	48.31%
Okinawa	4	55.00%

Figure 1. Overall average rendaku rates (%) by region

rendaku application rates do not differ much across participants' dialects.

A mixed-effects logistic regression model was constructed using the *glmer* function of the *lme4* and *lmerTest* packages (Bates et al. 2015; Kuznetsova et al. 2017) on the R software (R Core Team 2021) with Rendaku (applied or not) as its response variable, Region (10 regions) as its predictor, and by-item (first and second elements) and by-participant random intercepts as its random structure. The results of the analysis are summarized in Table 2. The baseline intercept here corresponds to Region: Tokyo/Kanto. As can be seen, the model indicates no statistically significant effects of Region on Rendaku. Although Osaka/Kansai shows the largest effect ($z = 1.854$), it is still not significant at the alpha level of 0.05 ($p = 0.064$). In other words, rendaku application rates do not differ greatly across dialects.

	β	SE	*z*	*p*
(Intercept)	–0.347	0.271	–1.280	0.201
Region: Hokkaido	0.015	0.083	0.183	0.855
Region: Tohoku	–0.013	0.070	–0.192	0.848
Region: Tokai-Tosan	0.019	0.057	0.332	0.740
Region: Hokuriku	0.166	0.109	1.520	0.129
Region: Osaka/Kansai	0.089	0.048	1.854	0.064
Region: Chugoku	–0.111	0.073	–1.519	0.129
Region: Shikoku	0.170	0.109	1.567	0.117
Region: Kyushu	–0.007	0.060	–0.121	0.904
Region: Okinawa	0.225	0.185	1.217	0.224

Table 2. Mixed-effects logistic regression analysis

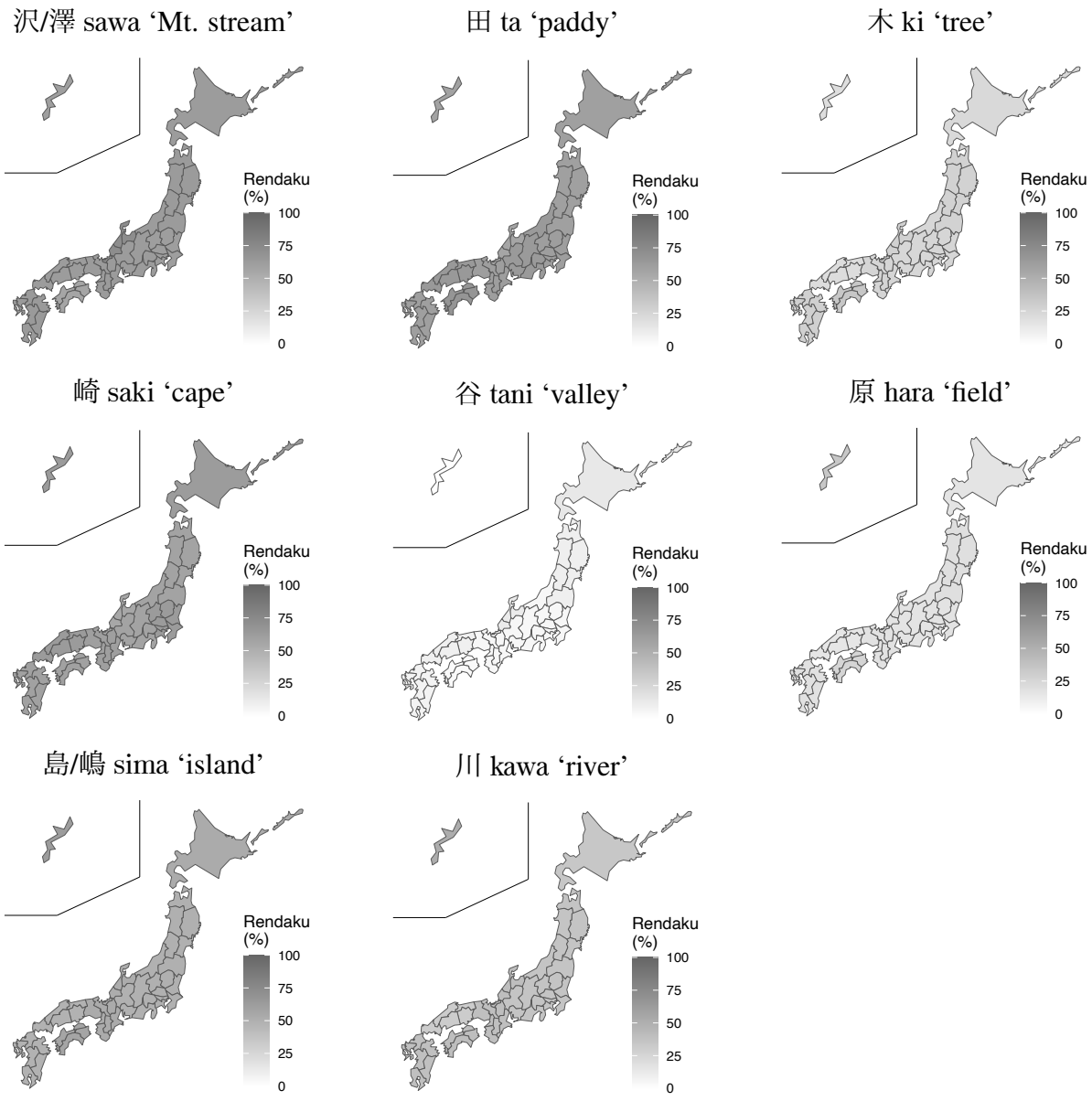


Figure 2. Average rendaku rates (%) by region and E2

2.2.2. RENDAKU RATES BY REGION AND COMMON E2. Following Takemura et al. (2019), the data were further analyzed by second elements ("E2"). Figure 2 (on the previous page) plots the average rendaku application rates by region for eight common E2 morphemes. The average rendaku rates appear to differ by E2; that is, some morphemes are more likely to undergo rendaku (e.g., 崎 /saki/ 'cape') than others (e.g., 谷 /tani/ 'valley'), as has been reported by previous studies on surnames (Sugito 1965; Zamma & Asai 2017). Within each E2, however, differences by region look rather small. When the differences are present, they are mostly due to the skew in datapoints; some regions have relatively few speakers/responses when the data are broken down by second-element type.

2.2.3. RENDAKU RATES BY INDIVIDUAL SURNAME. A closer look at the results suggests that the variability of rendaku in certain surnames may actually be attributable to regional differences. For example, for the surname 山崎 /jama-saki/ 'mountain-cape', the average rendaku rate is 100% in Tokyo/Kanto, while it is 62.5% in Osaka/Kansai (see Iwasaki 2013:42 for an anecdotal remark on this particular surname). To examine the trend in such surnames with a large regional difference, I extracted those that differ in rendaku application rates by more than 25 percentage points between Tokyo/Kanto and Osaka/Kansai. Here, I only focus on the differences between the two regions with the largest population in the country (hence, the largest numbers of participants). Table 3 shows a subset of those surnames.

UR	Kanji	Tokyo/Kanto	Osaka/Kansai
jama-saki	山崎	100.00%	62.50%
mina-kawa	皆川	100.00%	50.00%
take-sawa	竹沢	80.95%	50.00%
mizo-huti	溝渕	73.33%	100.00%
moro-hosi	諸星	61.54%	86.67%
kawa-ta	川田	61.54%	18.18%
aka-hori	赤堀	60.00%	0.00%
kimi-sima	君島	58.33%	100.00%
te-sima	手島	53.85%	71.43%
i-seki	井関	50.00%	16.67%
ko-seki	小関	25.00%	66.67%
to-tuka	戸塚	22.73%	50.00%
saka-ta	阪田	4.76%	33.33%
janagi-hara	柳原	0.00%	40.00%
joko-kawa	横川	0.00%	28.57%

Table 3. Surnames with a large regional difference in rendaku rates

The rendaku patterns of these surnames indicate that there is no consistent directionality in the regional differences. In some surnames, the rendaku rates are higher in Tokyo/Kanto than in Osaka/Kansai, but they are the other way around in others. It should also be noted that these surnames are not large in number. In all, 61 of the 1,176 surnames (5.19%) show differences greater than 25 percentage points. Lastly, surnames that are known for showing variable rendaku behaviors do not necessarily show a large regional difference. For example, the rendaku rates for 中田 /naka-ta/ 'central-paddy' are 26.32% in Tokyo/Kanto and 21.43% in

Osaka/Kansai; for 中嶋/中島 /naka-sima/ 'central-island', the rates are 84.82% in Tokyo/Kanto and 92.10% in Osaka/Kansai.

2.3. SUMMARY. To summarize, no clear regional effects are found on the way rendaku is applied in compound surnames. This is also true when the results are analyzed by common E2 morphemes. Although some individual surnames may show differences that can be attributed to dialects, the trends are not consistent and such surnames are not large in number. Thus, the study mostly replicated the findings of previous studies of rendaku in compound place names (van Bokhorst 2018; Takemura et al. 2019) with a different set of data (surnames) and systematic experimental methods.

3. Discussion. In this section, I discuss the inverse association between rendaku and accent in relation to the results shown in the previous section. I also propose a tentative analysis of the phenomenon.

3.1. RENDAKU-ACCENT ASSOCIATION. The experimental results above, though non-significant, further our understanding of Japanese rendaku. As has been stated, there are no systematic differences of rendaku application caused by regions or dialects as far as compound surnames are concerned. This actually raises a novel theoretical issue with respect to the relationship between voicing and accentuation.

It has been documented that rendaku application and accentedness are inversely associated. In other words, rendaku voicing and pitch accent tend not to co-occur; compounds with rendaku are likely to be unaccented, while those without are often accented. This is exemplified in (4) where surnames that variably undergo rendaku also show different accentuation patterns depending on the occurrence of compound voicing. Here, the lack of an accent mark indicates that the word is unaccented.

(4) No co-occurrence of rendaku and accent

a.	nak**á**-ɕima	vs.	naka-**ʑ**ima	'centeral-island'	中嶋
b.	mij**á**-kawa	vs.	mija-**g**awa	'shrine-river'	宮川

Sugito (1965) first revealed this inverse association by examining the patterns of accentuation as well as rendaku application in compound surnames with /ta/ 'paddy' as E2 (e.g., [s**á**ka-ta] 'slope-paddy' vs. [jama-**d**a] 'mountain-paddy'). Later studies, such as Zamma (2005), Ohta (2013), and Zamma & Asai (2017), have shown that the generalization holds true with other kinds of surnames, even though the degree of association may be weaker depending on the E2 morpheme and there are not a few exceptions. The phenomenon has thus been well-documented; but it remains one of the unresolved issues in formal Japanese phonology.[6]

The question now arises as to why surnames show similar rendaku patterns across dialects. Japanese dialects are known for showing various accentuation patterns (see e.g., Haraguchi 1977; Shibatani 1990; Uwano 1999; Kubozono 2012, 2015 and references therein). If rendaku voicing and pitch accent are strongly associated with each other, dialects with different accent patterns could also show great variability in rendaku application.

[6] It has been suggested that this rendaku-accent inverse association is especially robust in proper names (see Tanaka, S.-i. 2005a,b; Ohta & Tamaoka 2017; Tanaka, Y. 2017b for related discussion). Other than names, it is also attested in complex verbs with some further complications (see, e.g., Okumura 1984; Yamaguchi 2011). It has also been hinted that the phenomenon may not be productive in present-day Japanese (see Ohta & Tamaoka 2017).

3.2. A POSSIBLE ACCOUNT. As a tentative explanation, I propose that it is foot structure, not pitch accent itself, that has a strong relationship with rendaku voicing. Specifically, I argue that application of rendaku requires two consecutive feet and that this foot structure entails unaccentedness in Tokyo Japanese, but not necessarily in other dialects.

Ito & Mester (2016) have proposed a foot-based account of accentuation in Tokyo Japanese, including the emergence of unaccentedness. Their analysis is illustrated in (5) below with loanwords having antepenultimate accent as well as those that are unaccented. Assumed feet are indicated by parentheses.

(5) Antepenultimacy and unaccentedness

a.	(ká̠na) da	'Canada'	(gí̠ri) ɕa	'Greece'
b.	(ame) (ɾika)	'America'	(ita) (ɾia)	'Italy'

Three-mora words often receive antepenultimate accent; they have a bimoraic trochaic foot and a final extrametrical syllable, as shown in (5a), which is commonly assumed for deriving antepenultimacy in many languages. By contrast, four-mora words are often unaccented; they are exhaustively footed with two bimoraic feet, as shown in (5b). Given this foot structure, unaccenteness arises from a tension between two well-known metrical constraints: NONFINALITY(FT'), which bans a head foot from occurring at the right edge of a prosodic word, and RIGHTMOST, which requires a head foot to be the rightmost foot. Note that if a word is exhaustively footed into two consecutive feet, as in "(σσ)(σσ)," placing an accent on either foot would violate one of the above constraints. If the two constraints outrank WORDACCENT, which requires a prosodic word to have a prominence peak, the conflict is resolved by rendering the word unaccented.

While adopting this basic analysis for compound surnames, I posit that the constraint WORDACCENT exerts a stronger force on them, given the observation that proper names in general are more likely to be accented than common nouns (see Tanaka, S.'i. & Kubozono 1999; Tanaka, Y. & Sugawara 2018; Tanaka, Y. 2023). I also assume that each element of a compound surname does not need to have its own foot and that a foot can even span across a morpheme boundary (see Alderete 2015), unlike in the case of a common noun compound (Kubozono 1995, 1997); this is because proper name compounds lack semantic compositionality and behave more like simplex words (see Tanaka Y. 2017a; 2023 for discussion). This ensures that surnames, whether they are three or four moras in length, receive antepenultimate accent by default, as in [(sá̠ka)-ta] 'slope-paddy' and [na(ká̠-ɕi)ma] 'mountain-cape'.[7]

Why, then, do surnames become unaccented when they undergo rendaku? In order to account for this incompatibility of accent and voicing, I further propose a constraint that requires rendaku, which is a feature-sized linking morpheme (Ito & Mester 2003), to be realized only between stems projecting their own feet (see Rosen 2003 for a similar idea). The constraint is grounded in the raison d'être of rendaku: Even though a compound surname may usually behave like a simplex word (see above), it needs to be like a real compound word having two full-fledged stems with their own feet to receive compound voicing. With the high ranking of this constraint, rendaku can be realized only in surnames with two consecutive feet, which in turn results in unaccentedness, as in [(jama)-(da)] 'mountain-paddy' and [(naka)-(ʑima)]

[7] Four-mora surnames with antepenultimate accent thus violate the otherwise higher-ranked constraint INITIALFOOT, which requires the prosodic word to begin with a foot aligned with its left edge.

'central-paddy'. For the sake of clarity, the examples of accented and unaccented surnames are repeated and placed next to each other in (6).

(6) Footing of accented and unaccented surnames

a.	(s**á**ka)-ta	'slope-paddy'	vs.	(jama)-(**d**a)	'mountain-paddy'
b.	na(k**á**-ɕi)ma	'central-island'	vs.	(naka)-(**ʑ**ima)	'centeral-island'

Returning to the lack of dialectal differences in rendaku, I speculate that surnames have similar foot structures across dialects but do not necessarily show identical accent patterns. This is illustrated in (7) below with examples of accentuation contrasting Tokyo/Kanto Japanese and Osaka/Kansai Japanese. For clarity, the tonal melodies of words are also given, with L representing a low tone and H a high tone.

(7) Accentuation and (possible) footing in two dialects

	Tokyo/Kanto			Osaka/Kansai		
a.	(**á**ki)-ta	HLL	vs.	(ak**í**)-ta or a(k**í**-ta)	LHL	'autumn-paddy'
b.	(jama)-(**d**a)	LHH	vs.	(jama)-(**d**a)	LLH	'mountain-paddy'
c.	(ima)-(**d**a)	LHH	vs.	(im**á**)-(**d**a) or (i)(m**á**-**d**a)	LHL	'present-paddy'
				∼ (ima)-(**d**a) ?	LLH	

A three-mora surname without rendaku often has antepenultimate accent in Tokyo/Kanto Japanese; again, this can be analyzed by positing a left-aligned bimoraic foot and an extrametrical syllable following Ito & Mester (2016), as shown in (7a). On the right, the same surname has penultimate accent in Osaka/Kansai Japanese, but it is nonetheless analyzed as having a single foot: penultimacy can be derived from a left-aligned iambic foot or a right-aligned trochaic foot.

As shown in (7b), a surname with rendaku voicing can be unaccented in both dialects, which is analyzed as being exhaustively footed with two bimoraic feet. The melody is not the same (LHH vs. LLH), but this is most probably due to a difference in tonal assignment. What matters here is that the two dialects have the same foot structure: rendaku application requires two consecutive feet, which leads to unaccentedness.

(7c) shows that a surname with rendaku may possibly have different accentuation in the two dialects in question: unaccentedness and penultimate accent. Note that both accent patterns can be derived from the same foot structure. Unaccentedness straightforwardly arises from two feet in a row; with high-ranked NONFINALITY(FT') and RIGHTMOST, neither foot can hold accent (Ito & Mester 2016). Exhaustive footing is also compatible with penultimacy. On the assumption that WORDACCENT is ranked higher as dialectal variation, an iambic foot on the left results in penultimate accent, as in [(im**á**)-(**d**a)]. With trochaic footing, one can still derive antepenultimacy by assuming that a monomoraic foot and a boundary-spanning bimoraic foot, as in [(i)(m**á**-**d**a)], meet the requirement for rendaku application. Additionally, some Osaka/Kansai speakers are reported to pronounce the same surname as unaccented, as in [(ima)-(**d**a)] (Sugito 1965); again, unaccentedness can easily be derived from consecutive feet.

The analysis proposed here is, of course, not comprehensive and its validity should be tested with more data of accentuation in surnames across different dialects. I leave this for future research.

4. Conclusion. To conclude, a large-scale judgment experiment has found no clear regional effects on rendaku application in compound surnames. Although some individual surnames may show different rendaku profiles across dialects, the variability does not seem to be derived in any systematic manner. This raises the question of why rendaku voicing is inversely associated with pitch accent. I have proposed a tentative account based on foot structure. Future studies should investigate not only the rendaku patterns but also the accent patterns of surnames across Japanese dialects to test the validity of the proposed account.

References

Alderete, John. 2015. Updating the analysis of Japanese compound accent. In Short 'schrift for Alan Prince compiled by Eric Baković. https://princeshortschrift.wordpress.com/.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. lme4: Linear mixed-effects models using 'Eigen' and S4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Becker, Michael & Jonathan Levine. 2013. Experigen — An online experiment platform. http://becker.phonologist.org/experigen.

van Bokhorst, Michelle Hermina. 2018. *Rendaku in place names*. Leiden University MA thesis.

Haraguchi, Shôsuke. 1977. *The tone pattern of Japanese: An autosegmental theory of tonology*. Tokyo: Kaitakusha.

Ito, Junko & Armin Mester. 2003. *Japanese morphophonemics*. Cambridge, MA: MIT Press.

Ito, Junko & Armin Mester. 2016. Unaccentedness in Japanese. *Linguistic Inquiry* 47(3). 471–526. https://doi.org/10.1162/LING_a_00219.

Iwasaki, Shoichi. 2013. *Japanese: Revised edition* (London Oriental and African Language Library 17). Amsterdam and Philadelphia: John Benjamins.

Kato, Masanobu. 1977. Hōgen kukaku-ron [Theory of dialect divisions]. In Susumu Ono & Takeshi Shibata (eds.), *Iwanami kōza Nihongo 11: Hōgen [Iwanami courses in Japanese 11: Dialects]*, 41–82. Tokyo: Iwanami Shoten.

Kubozono, Haruo. 1995. Constraint interaction in Japanese phonology: Evidence from compound accent. In Rachel Walker, Ove Lorentz & Haruo Kubozono (eds.), *Phonology at Santa Cruz [PASC]*, 21–38. Santa Cruz: UC Santa Cruz Linguistics Research Center.

Kubozono, Haruo. 1997. Lexical markedness and variation: A nonderivational account of Japanese compound accent. *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL)* 15. 273–287.

Kubozono, Haruo. 2012. Varieties of pitch accent systems in Japanese. *Lingua* 122. 1395–1414. https://doi.org/10.1016/j.lingua.2012.08.001.

Kubozono, Haruo. 2015. Japanese dialects and general linguistics. *Gengo Kenkyū* 148. 1–31.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2017. lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26. https://doi.org/10.18637/jss.v082.i13.

Ohta, Satoshi. 2013. On the relationship between rendaku and accent: Evidence from -kawa/-gawa alternation in Japanese surnames. In Jeroen van de Weijer & Tetsuo Nishihara (eds.), *Current issues in Japanese phonology: Segmental variation in Japanese*, 63–87. Tokyo: Kaitakusha.

Ohta, Satoshi & Katsuo Tamaoka. 2017. Rendaku to akusento: Futsū meishi to muimigo no baai [Rendaku and accent: The cases of common nouns and nonce words]. In Timothy J.

Vance, Emiko Kaneko & Seiji Watanabe (eds.), *Rendaku no kenkyū [Research on rendaku]*, 69–93. Tokyo: Kaitakusha.
Okumura, Mitsuo. 1984. Rendaku. *Nihongogaku* 3(5). 89–98.
R Core Team. 2021. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
Rosen, Eric. 2003. Systematic irregularity in Japanese rendaku: How the grammar mediates patterned lexical exceptions. *Canadian Journal of Linguistics* 48(1–2). 1–37.
Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.
Shirooka, Keiji & Tadashige Murayama. 2011. A database of Japanese surnames and their rankings. http://hdl.handle.net/10297/00025667.
Sugito, Miyoko. 1965. Shibata-san to Imada-san: Tango no chōkakuteki benbetsu ni tsuite no ichi kōsatsu [Shibata-san and Imada-san: An examination of auditory distinction of words]. *Gengo Seikatsu* 165. 64–72.
Suzaki, Haruo. 2013. A private on-line corpus of 111,711 Japanese family names. http://www2s.biglobe.ne.jp/˜suzakihp/index40.html.
Takemura, Akiko, Thomas Pellard, Hyun Kyung Hwang & Timothy J. Vance. 2019. Rendaku in place names across Japanese dialects. *Reports of the Keio Institute of Cultural and Linguistic Studies* 50. 79–89.
Tanaka, Shin-ichi. 2005a. *Akusento to rizumu [Accent and rhythm]*. Tokyo: Kenkyusha.
Tanaka, Shin-ichi. 2005b. Where voicing and accent meet: Their function, interaction, and opacity problems in phonological prominence. In Jeroen van de Weijer, Kensuke Nanjo & Tetsuo Nishihara (eds.), *Voicing in Japanese*, 261–278. Berlin and New York: Mouton de Gruyter.
Tanaka, Shin'ichi & Haruo Kubozono. 1999. *Nihon-go no hatsuon kyōshitsu: Riron to renshū [A course in Japaense pronunciation: Theory and practice]*. Tokyo: Kuroshio Shuppan.
Tanaka, Yu. 2017a. Phonotactically-driven rendaku in surnames: A linguistic study using social media. In Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel & Edward J. Rubin (eds.), *Proceedings of the 34th West Coast Conference on Formal Linguistics*, 519–528. Somerville, MA: Cascadilla Proceedings Project.
Tanaka, Yu. 2017b. *The sound pattern of Japanese surnames*. Los Angeles: University of California dissertation.
Tanaka, Yu. 2023. Phonology of proper names. *Language and Linguistics Compass* 17(5). e12502. https://doi.org/doi.org/10.1111/lnc3.12502.
Tanaka, Yu. to appear. Learning biases in proper nouns. *Phonology*.
Tojo, Misao. 1954. Josetsu [Introduction]. In Misao Tojo (ed.), *Nihon hoōgengaku [Japanese dialectology]*, Tokyo: Yoshikawa Kobunkan.
Uwano, Zendo. 1999. Classification of Japanese accent systems. In Shigeki Kaji (ed.), *Cross-linguistic studies of tonal phenomena, tonogenesis, typology, and related topics*, 151–186. Tokyo: ILCAA.
Vance, Timothy J., Shigeto Kawahara & Mizuki Miyashita. 2021. The diachronic origins of Lyman's Law: Evidence from phonetics, dialectology and philology. *Phonology* 38(3). 479–511. https://doi.org/10.1017/S0952675721000270.

Vance, Timothy J., Mizuki Miyashita & Mark Irwin. 2014. Rendaku in Japanese dialects that retain prenasalization. In *Japanese/Korean linguistics*, vol. 21, 33–42. Stanford: CSLI Publications.

Yamaguchi, Kyoko. 2011. Accentedness and rendaku in Japanese deverbal compounds. *Gengo Kenkyū* 140. 117–133.

Zamma, Hideki. 2005. The correlation between accentuation and rendaku in Japanese surnames: A morphological account. In Jeroen van de Weijer, Kensuke Nanjo & Tetsuo Nishihara (eds.), *Voicing in Japanese*, 157–176. Berlin and New York: Mouton de Gruyter.

Zamma, Hideki & Atsushi Asai. 2017. Sei ni mirareru Sugitō no hōsoku to kakuchō-ban Raiman no hōsoku ni kansuru keitai-teki/onin-teki kōsatsu [Morphological and phonological analyses of Sugito's Law and Strong Lyman's Law in surnames]. In Timothy J. Vance, Emiko Kaneko & Seiji Watanabe (eds.), *Rendaku no kenkyū [Research on rendaku]*, 147–179. Tokyo: Kaitakusha.