submitted
comments welcome

The non-local nature of Lyman's Law revisited

Shigeto Kawahara and Gakuji Kumagai
Keio University and Kansai University

Abstract

Rendaku is a morphophonological process in Japanese in which the first obstruent of a second member of a compound is realized as voiced (e.g. nise+**t**anuki/ → [nise-**d**anuki]). Lyman's Law blocks this voicing process when the second member already contains a voiced obstruent, whether the blocker is in the second syllable (e.g. /zaru+soba/ → [zaru-**s**o**b**a]) or in the third syllable (e.g. /çi+tokage/ → [çi+**t**oka**g**e]). Vance (1979), a seminal experimental study on rendaku, showed that in nonce words, the blockage of rendaku by Lyman's Law is not deterministic; moreover, it found some evidence that the blockage effect tends to be stronger when the blocker consonant is in the second syllable than in the third syllable, i.e. Lyman's Law may be sensitive to a locality effect in nonce words. On the other hand, a naturalness judgment experiment by Kawahara (2012) failed to find this locality effect. To settle these conflicting results from the past studies, with a general issue of the replication crisis in linguistics in mind (Sönning & Werner 2021), we first conducted a large scale forced-choice experiment with 72 stimuli and with 184 native speakers of Japanese. The results show that Lyman's Law is, overall, sensitive to a locality effect. To investigate why Kawahara (2012) failed to find a locality effect, we next replicated Kawahara (2012) with a larger number of speakers (187 participants), which found some evidence that the locality effect is identifiable in a naturalness judgment experiment as well. We conclude that Lyman's Law is indeed sensitive to a locality effect, at least for some speakers (Vance 1979).

Keywords: rendaku, Lyman's Law, dissimilation, locality, replication, experimental phonology
Approximate word count: 5,500

1 Introduction

Dissimilation effects are often sensitive to a distance-and-decay effect: i.e. dissimilative forces are stronger between two closer segments (see Suzuki 1998 for a review; see also Bennett 2015 and Hansson 2001 for other typological studies of dissimilation). For example, in Yimas, rhotic dissimilation applies only when two rhotics are in the adjacent syllables, but not when they are farther apart (Foley 1991, cited by Suzuki 1998). A famous case of similarity-based phonotactic restrictions in Arabic is also more stringent between two adjacent consonants than between two non-adjacent consonants (Frisch et al. 2004). Against this cross-linguistic observation, this paper tests whether Lyman's Law in Japanese—a dissimilation constraint against two voiced obstruents within a morpheme—is stronger between two local consonants than between two non-local consonants, since the past results on this question have been mixed.

Lyman's Law most clearly manifests itself in the blockage of rendaku.[1] Rendaku is a morphophonological process, in which the morpheme-initial obstruent of the second element (henceforth, E2) in a compound undergoes voicing, as in (1) (/h/ surfaces as [b] as a result of voicing, as /h/ in Japanese was historically—or is perhaps underlyingly—/p/: McCawley 1968). Rendaku, however, is blocked when E2 already contains a voiced obstruent, as in (2) and (3). This blockage of Rendaku is known as Lyman's Law after Lyman (1894) (although Lyman is probably not the first scholar who found this generalization: see Vance 2022 for extended discussion).

(1) Examples of rendaku

a. /nise+**t**anuki/ → [nise+**d**anuki] 'fake raccoon'

b. /juki+**k**uni/ → [juki+**g**uni] 'snow country'

c. /hoɕi+**s**oɾa/ → [hoɕi+**z**oɾa] 'starry sky'

d. /oɕi+**h**ana/ → [oɕi+**b**ana] 'dried flower'

(2) Blocking of rendaku by Lyman's Law (by a local consonant)

a. /çito+**t**a**b**a/ → [çito+**t**a**b**a], *[çito+**d**a**b**a] 'one bundle'

b. /omo+**k**a**g**e/ → [omo+**k**a**g**e], *[omo+**g**a**g**e] 'resemblance'

c. /moɾi+**s**o**b**a/ → [moɾi+**s**o**b**a], *[moɾi+**z**o**b**a] 'cold soba'

d. /çito+**h**a**d**a/ → [çito+**h**a**d**a], *[çito+**b**a**d**a] 'people's skin'

(3) Blocking of rendaku by Lyman's Law (by a non-local consonant)

a. /ni+**t**ama**g**o/ → [ni+**t**ama**g**o], *[ni+**d**ama**g**o] 'boiled egg'

[1] A constraint against two voiced obstruents within a morpheme also functions as a phonotactic restriction in native words in Japanese—no native morphemes seem to contain two voiced obstruents; e.g. [ɸuda] 'amulet,' [buta] 'pig' but *[buda] (Ito & Mester 1986). Lyman's Law has been formalized as an OCP constraint on the feature [+voice] (Ito & Mester 1986) or as a locally-conjoined constraint against a voiced obstruent within a morpheme (Ito & Mester 2003). See Kawahara & Zamma (2016) for a review of the theoretical treatments of this restriction.

b. /umi+**k**uɾage/ → [umi+**k**uɾa**g**e], *[umi+**g**uɾa**g**e] 'see jellyfish'

c. /mitɕi+ɕiɾube/ → [mitɕi+ɕiɾube], *[mitɕi+ʑiɾu**b**e] 'guide post'

d. /oo+**h**aɕagi/ → [oo+**h**aɕa**g**i], *[oo+**b**aɕa**g**i] 'big excitement'

In existing words, the blockage of rendaku is almost exception-less and it holds regardless of whether the blocker consonant is in the second syllable, as in (2) or in the third syllable, as in (3). Unambiguous cases of lexical exceptions of Lyman's Law include two local cases ([X-**z**a**b**uɾoo] 'PROPER NAME' and [hun-ʑi**b**aɾu] 'to tightly bind') and one non-local case ([nawa-**b**aɕi**g**o] 'rope ladder').[2] Thus from the lexical patterns, it is not clear whether Lyman's Law is sensitive to a locality restriction or not. In other words, learners of Japanese, who are exposed to the Japanese data, would not know whether Lyman's Law would block rendaku to a stronger degree when the blocker and rendaku-undergoer are in the adjacent syllables, as expected from a cross-linguistic trend of dissimilation (Suzuki 1998).[3]

Vance (1979) is a seminal experimental study on rendaku, which addressed this question using an experimental paradigm. He presented 50 nonce words, each combined with 8 real words, to fourteen native speakers of Japanese and asked whether each compound should undergo rendaku or not. The results showed, first of all, that the blockage of rendaku by Lyman's Law is not deterministic, unlike in real words and hence nonce words can violate Lyman's Law. Moreover, the experiment found that for a number of speakers (eight out of fourteen), the blockage of rendaku is more likely when the blocker and the undergoer are in the adjacent syllables than when they are separated by one intervening syllable.[4] This result would arguably instantiate a case of the emergence of the unmarked (TETU: McCarthy & Prince 1994) in an experimental setting, since, as discussed above, there is very little, if any, lexical evidence for the locality effect on Lyman's Law (see e.g. Coetzee 2009, Shinohara 1997, Wilson 2006 and Zuraw 2007 for similar observations). One could also take this result as a case for the poverty of stimulus argument (Chomsky 1986), because the lexical data from the actual spoken Japanese does not distinguish the local blockage effect and the non-local blockage effect.

However, a later experimental study by Kawahara (2012) failed to replicate this result by Vance (1979). This study was a naturalness judgment experiment, in which the participants were asked, using a 5-point Lickert scale, how natural rendaku-undergoing forms were. That experiment had 36 test items (12 items for three conditions, no Lyman's Law violations, local Lyman's Law vio-

[2] There may be a few other possible cases of exceptions to Lyman's Law, although it is not clear that they are standard pronunciations: see §7.2.4 of Vance (2022) for detailed discussion.

[3] A locality effect on dissimilation is also expected to the extent that dissimilation has a phonetic underpinning, such as avoidance of perceptual confusion (Ohala 1981; Stanton 2019) and/or articulatory difficulty (Alderete & Frisch 2007; Pulleyblank 2002), because such phonetic problems are expected to be worse between local segments than between non-local segments.

[4] One speaker had no rendaku responses in either conditions; four speakers had a very small-size reversal (e.g. 20% vs. 17%); and only one speaker had a fairly clear reversal (44% vs. 14%).

lations and non-local Lyman's Law violations). The data were collected from 54 native speakers of Japanese. In that experiment, forms with the local violation were judged to be slightly less natural than forms with the non-local violation (average naturalness ratings = 2.76 vs. 2.86), but this difference was not significant.

Kawahara (2012) offered the following conjecture regarding where this difference comes from. Another set of experiments reported by Ihara et al. (2009) showed that the locality effect of Lyman's Law decreased from 1984 when they ran their first experiment compared to 2005 when they ran their second experiment. It may have been the case that this trend continued and it has disappeared completely by 2011, when Kawahara run his experiment. In other words, the locality effect of Lyman's Law was fading away, as a part of historical change in Japanese phonology. Vance (2022), which reflects the most updated opinion by Vance, suspects that the fact that Vance (1979) found a locality effect was due to some uncontrolled factors, implying that he now believes that Lyman's Law is not sensitive to a locality effect after all.

To settle these conflicting results from the previous studies, the experiments reported in the current paper revisit this question—is Lyman's Law sensitive to a locality effect after all? We were set out to run a new experiment with a large number of stimuli and a large number of participants, because one reason for why Kawahara (2012) failed to find the locality effect may have been due to a small number of N, i.e., the experiment simply lacked a sufficient statistical power (see e.g. Chambers 2017; Sprouse & Almeida 2017; Vasishth & Gelman 2021; Winter 2019 for discussion on the general lack of statistical power in linguistics and neighboring fields).

One general issue that we had in mind as we revisited this old question, already addressed by these previous studies, was "the replication crisis" (Chambers 2017; Open Science Collaboration 2015; Roettger 2019; Sönning & Werner 2021; Winter 2019), in which many results that are published in previous research cannot be replicated by later studies. One reason behind this general problem is insufficient statistical power, resulting from an insufficient number of N, both in terms of participants and items. For the case at hand, Kawahara (2012) had only three items for each segment type that can undergo rendaku (/t/, /k/, /s/ and /h/, i.e. three items $\times$ four segments for each Lyman's Law violation condition). Another reason behind the replication crisis may be the inappropriate use of frequentist analyses (Chambers 2017). In this respect too, Kawahara (2012) made a mistake of concluding a null effect given a statistically non-significant result using a frequentist analysis, when he says "the locality effect has disappeared by 2011" (p. 1197).

To address these problems, our experiment included 72 stimuli and we collected data from 184 speakers. We also resorted to a Bayesian analysis, as it would allow us to access to what degree we can believe in a null effect (Gallistel 2009), if the results were to show that no differences exist between a local violation of Lyman's Law and a non-local violation of Lyman's Law.

2 Experiment 1

2.1 Method

Following the open science initiative in linguistics as a step toward addressing the replication crisis problem (Cho 2021; Winter 2019), the raw data, the R Markdown file and the Bayesian posterior samples are made available at an Open Science Framework (OSF) repository.[5]

2.1.1 Overall design

The current experiment consisted of three conditions: (1) nonce words whose rendaku would not result in any violations of Lyman's Law (e.g. [taɾuna]→[daɾuna]), (2) nonce words whose rendaku would incur a local violation of Lyman's Law (e.g. [ta**g**uta]→[**d**a**g**uta]), and (3) nonce words whose rendaku would result in a non-local violation of Lyman's Law (e.g. [tatsuga]→[**d**atsu**g**a]). The comparison between the first and the second condition would test the psychological reality of Lyman's Law, which has been confirmed by a number of previous experimental studies (Ihara et al. 2009; Kawahara 2012; Kawahara & Sano 2014a,b; Kawahara & Kumagai 2023; Vance 1979). The comparison between the second condition and the third condition would test the (non-)local nature of Lyman's Law, the main concern of the current experiment.

2.1.2 Stimuli

Table 1 shows the the list of nonce word E2s used in Experiment 1. The experiment tested all four sounds that can undergo rendaku in contemporary Japanese (=/t/, /k/, /s/ and /h/) with 6 nonce items in each cell. These resulted in a total of 72 stimuli (3 conditions × 4 consonant types × 6 items). The stimuli for the first two conditions were adapted from Kawahara & Kumagai (2023).

None of the stimuli becomes a real word after rendaku. The syllable structure of the stimuli was controlled in that none of the stimuli contained a heavy syllable. Since the rendaku probability may be influenced when it results in identical CV mora sequences (Kawahara & Sano 2014a,b), in no forms would rendaku result in CV moras that are identical to those in the second syllables or to those in third syllables. Since we chose to use [nise] 'fake' as E1 (see below), we avoided stimuli that begin with [se] as well.

[5] `https://osf.io/ym79p/?viewonly=ce17de5a39834ae397c44a19e74db082`. We fully acknowledge that making the open science policy is not panacea for the general replication crisis problem, but also note that it is nevertheless a necessary and useful step toward addressing the problem.

Table 1: The list of nonce words used as E2s in Experiment 1. /h/ allophonically becomes [ç] before [i] and [ɸ] before [u].

	No violation	Local violation	Non-local violation
/t/	[tamuma]	[taguta]	[tatsuga]
	[tatsuka]	[tozumi]	[tesago]
	[taɾuna]	[teguɾa]	[tekibi]
	[tonime]	[tazanu]	[takuga]
	[tekeha]	[tegesa]	[tekoʑi]
	[tokeho]	[toboɸu]	[teçigi]
/k/	[kimane]	[kidaku]	[kitebe]
	[kikake]	[kobono]	[kotiba]
	[kotona]	[kabomo]	[kaɕido]
	[kumise]	[kedeɾe]	[kutɕibo]
	[konihe]	[kuʑiha]	[kesodo]
	[kehaɾo]	[kozana]	[katsuba]
/s/	[samaɾo]	[sabaɾe]	[sokabo]
	[sokato]	[sogeha]	[sohogi]
	[sutane]	[sobumo]	[sukabi]
	[samohe]	[sadanu]	[suhode]
	[soɾise]	[sodoka]	[satage]
	[sateme]	[sudaɸu]	[sokebi]
/h/	[honaɾa]	[hobasa]	[hokida]
	[çinumi]	[hazuke]	[hekazu]
	[honiko]	[hogoɾe]	[hetado]
	[hakisa]	[çigiɾo]	[hategi]
	[heɾaho]	[ɸuzumo]	[çisuda]
	[çihonu]	[hedeno]	[ɸuhode]

2.1.3 Participants

The experiment was conducted online using SurveyMonkey (`https://jp.surveymonkey.com`). The participants were collected using a snowball-sampling method, primarily on Twitter, advertised on the first author's account. As a result, 162 speakers, who were native speakers of Japanese and had not heard about rendaku or Lyman's Law, voluntarily completed the online experiment. In addition, the data from 22 additional participants were collected from Keio University, who earned an extra credit for completing the experiment.[6] Combined together, the data from a total of 184 participants were considered in the following statistical analyses.

[6]The data from 17 speakers had to be excluded because their were either a non-native speaker of Japanese or were already familiar with rendaku.

2.1.4 Procedure

During the instructions, the participants were told that when they create a compound in Japanese, some combinations undergo voicing (i.e. rendaku), but this is not always the case. This explanation was given to the participants to remind them that rendaku is not an exception-less process. Three existing examples of rendaku-undergoing forms and non-rendaku-undergoing forms were used to illustrate this nature of rendaku. However, none of these examples involved a potential violation of Lyman's Law, so that these examples used for the illustration would not bias participants about the property of Lyman's Law.

In the main session, the participants were presented with one stimulus item and were asked to combine it with [nise] 'fake' as E1 to make a compound. They were then asked whether the resulting compound would sound more natural with rendaku or without it; a sample question is thus, "given a nonce word [kimane], when it is combined with [nise], which form sounds more natural, [nise-kimane] or [nise-gimane]?"

The stimuli were written in the *hiragana* orthography. This is because rendaku applies primarily to native words, and the *hiragana* orthography is used to write native words in the Japanese orthographic system. Prior to the main session, the participants went through two practice questions with existing compounds to make sure that they understand the task. The stimuli in the main session were presented to the participants as nonce words.[7] Each participant was assigned a uniquely randomized order of stimuli, using the randomization function of Surveymonkey.

2.1.5 Statistical analyses

For statistical analyses, we fit a Bayesian mixed effects logistic regression model, using the `brms` package (Bürkner 2017) and R (R Development Core Team 1993–) (for accessible introduction to Bayesian modeling, see e.g. Franke & Roettger 2019; Kruschke 2014; Kruschke & Liddell 2018; McElreath 2020; Vasishth et al. 2018). Bayesian analyses take both prior distribution (if any) and the obtained data into consideration and produce a range of possible values (a.k.a. posterior distributions) for each parameter that we would like to estimate. One advantage of Bayesian analyses is that we can interpret these posterior distributions as directly reflecting the likely values of these estimates, unlike the 95% confidence intervals that we obtain in a frequentist analysis. Another advantage is that it would allow us to access with how much confidence we can believe in a null effect (Gallistel 2009).

One standard way to interpret the results of Bayesian regression models is to examine the middle 95% of the posterior distribution, known as 95% Credible Interval (henceforth, 95% CrI),

[7] An alternative method is to present the stimuli as obsolete native words (Vance 1979; Zuraw 2000). We chose the former method, because it is simpler and did not cause a particular problem in previous experiments on rendaku that we conducted in the past.

of an estimate parameter. If that interval does not include 0, we can interpret that effect to be meaningful/credible. However, with Bayesian analyses, we do not need to commit ourselves to a "meaningful" vs. "non-meaningful" dichotomy, as in a frequentist "significant" vs. "non-significant" dichotomy. To be more concrete, another way to interpret the results of Bayesian regression models is to calculate how many posterior samples of a particular coefficient are in an expected direction. In what follows we deployed both ways of interpretation.

The details of the model specifications in the current model were as follows. The dependent variable was whether each item was judged to undergo rendaku or not (rendaku-undergoing response =1 and non-rendaku-undergoing response = 0). For independent variables, one main fixed factor was three conditions regarding Lyman's Law (no violation vs. local violation vs. non-local violation). The reference level was set to be the local violation condition, so that we can compare (i) the difference between no-violation and local violation (i.e. the psychological reality of Lyman's Law) and (ii) the local violation and the non-local violation (i.e. the locality of Lyman's Law). Another fixed factor was sound type (i.e. /t/-/k/-/s/-/h/). For this factor, the baseline was arbitrarily set to be /h/, because we had no particular reason to choose one segment over the others. The interaction term between the two factors was also coded, because we wanted to see whether the effects of Lyman's Law, if any, would generalize to all four segments. The model also included a random intercept of items and participants in addition to random slopes of participants for both of the fixed factors and their interaction.

We run four chains with 4,000 iterations and disregarded the first 1,000 iterations as warmups (running only 2,000 iterations resulted in inappropriate effective sample size (ESS) values). For prior specifications, we used a Normal(0, 1) weakly informative prior for the intercept (Lemoine 2019) and a Cauchy prior with scale of 2.5 for all slope coefficients (Gelman et al. 2018). As a result, all the $\hat{R}$-values for the fixed effects were 1.00 and no divergent transitions were detected, i.e. the four chains mixed successfully. Complete details of this analysis are available in the R Markdown file available at the osf repository mentioned above.

2.2 Results

2.2.1 General results

Figure 1 shows the rendaku application rate for each condition in the form of violin plots, in which their widths represent normalized probability distributions. Each facet shows a different segment type. Within each facet, each violin shows the three critical conditions. Transparent circles, jittered slightly to avoid overlap, represent averaged responses from each participant. Solid red circles are the averages in each condition. Abstracting away from the differences among the four segments, the three conditions resulted in the following rendaku application rates from left to right: (1) 60.0% (2) 32.2% (3) 41.2%. See the markdown file for segment-specific average values.

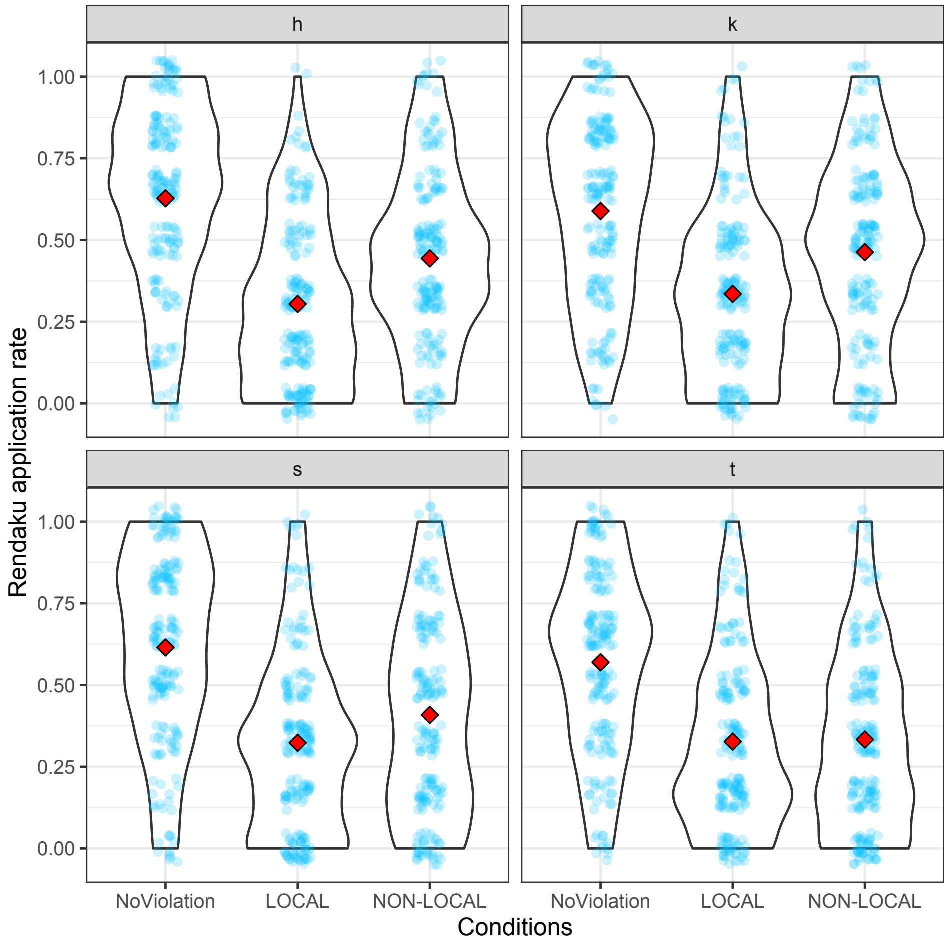

Figure 1: The comparison between the three critical conditions, with each facet showing a different segment type.

We observe that the first condition (no violations of Lyman's Law) showed higher rendaku responses compared to the second condition (the local violation of Lyman's Law), providing support for the psychological reality of Lyman's Law, which was shown by a number of previous studies (Ihara et al. 2009; Kawahara 2012; Kawahara & Sano 2014a,b; Kawahara & Kumagai 2023; Vance 1979).

More interestingly, the second condition (the local violation of Lyman's Law) generally showed lower rendaku responses than the third condition (the non-local violation of Lyman's Law), although this difference is very small in the /t/-facet. Overall, then, the current results support that of Vance (1979), not that of Kawahara (2012)—Lyman's Law does seem to exhibit a locality effect, at least for /h/, /k/ and /s/.

The model summary of the Bayesian mixed effects logistic regression analysis is provided in Table 2. The intercept is negative, as it represents the baseline condition (/h/, local violation), which shows smaller than 50% rendaku responses. As for the sound type (=the coefficients in (b)), for which /h/ serves as the baseline, all of the relevant 95% CrIs for the coefficients include 0, suggesting that differences among the four segment types were not very meaningful. The interaction terms in (d)—interactions between the segment type and the difference between the no-violation and the local violation—were also not very credible, suggesting that the local version of Lyman's Law functions to a comparable degree across the four segments, although for /k/ and /t/, they are leaning toward the negative, i.e., the effects of local Lyman's Law tend to be smaller. The main effect of the difference between the no-violation and the local violation ((c), the top) was very credible, supporting the psychological reality of Lyman's Law.

Table 2: Summary of the Bayesian mixed effects logistic regression model (Experiment 1).

		β	error	95% CrI
(a) intercept	(/h/, local)	-1.00	0.18	[-1.36, -0.66]
(b) sound type	/k/	0.14	0.23	[-0.31, 0.58]
	/s/	0.04	0.23	[-0.40, 0.49]
	/t/	0.08	0.23	[-0.37, 0.54]
(c) condition	no-violation vs. local	1.64	0.24	[1.17, 2.10]
	local vs. non-local	0.69	0.23	[0.24, 1.14]
(d) interactions I	/k/:no-violation vs. local	-0.34	0.32	[-0.96, 0.29]
	/s/:no-violation vs. local	-0.08	0.32	[-0.70, 0.56]
	/t/:no-violation vs. local	-0.39	0.32	[-1.02, 0.24]
(e) interactions II	/k/:local vs. non-local	-0.03	0.31	[-0.66, 0.58]
	/s/:local vs. non-local	-0.24	0.32	[-0.86, 0.39]
	/t/:local vs. non-local	-0.68	0.32	[-1.31, -0.05]

More interestingly, the main effect of the difference between the local violation and non-local violation ((c), the bottom) was also credible, at least at the baseline level /h/. However, the interaction term between the locality effect and /t/ was also credible, suggesting that we should look at the locality effect of Lyman's Law for each segment. Given this set of results, we calculated how many posterior samples of the locality effect was in the expected direction, $p(\beta > 0)$, for each segment type, which represent how likely the non-local Lyman's Law condition induced higher rendaku responses than the local Lyman's Law condition.

The results show that $p(\beta > 0)$ is 0.52 for /t/, 0.994 for /k/, 0.964 for /s/ and 0.998 for /h/. We thus conclude that Lyman's Law is sensitive to a locality effect for all segments but /t/. Statistically speaking, in short, the current results appear to accord better with Vance (1979), than with Kawahara (2012).

To be complete, we also calculated $p(\beta > 0)$ for the difference between the no-violation condition and the local violation condition. The results show that it is 1 for all segments—i.e. the effects of Lyman's Law is undoubtedly present for all segment types.

2.2.2 By speaker analysis

One question that arises regarding the current results, given the variability observed in Figure 1—and also given that Kawahara (2012) failed to find such an effect—is inter-speaker differences. How general does the locality effect hold? With this question in mind, Figure 2 plots, for each participant, the average rendaku application rate for the local violation condition and the non-local violation condition. Those dots above the diagonal axis are those speakers who are sensitive to a locality effect in the expected direction, and there were many of them. However, there are a number of participants who are around the diagonal axis, who are not sensitive to the locality effect. And rather surprisingly, there were also those who are below the diagonal axis, who represent an "anti-locality" effect.

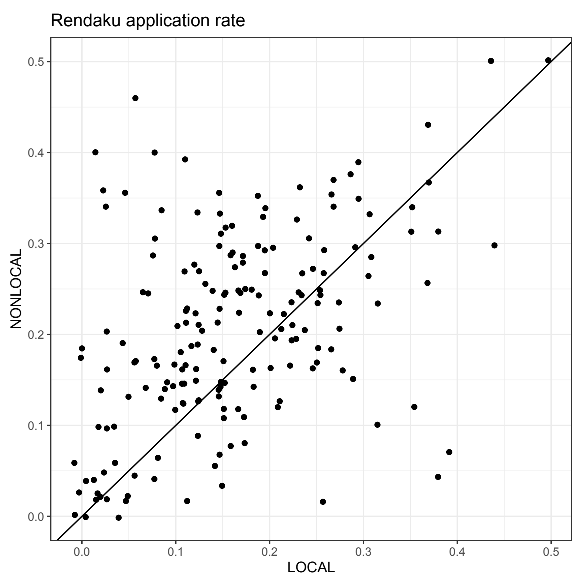

Figure 2: The comparison between the local violation condition and the non-local violation condition by each speaker (Experiment 1).

Given that Vance (1979) found eight out of the fourteen speakers show the locality effect in the expected direction, and that one speaker showed a clear reversal (44% vs. 14%), the current results may be comparable to that of Vance (1979) and thus may not be too surprising.

2.3 Discussion

The first and foremost important finding of the current study is to have shown that Lyman's Law is, at least for many speakers, indeed sensitive to a locality effect, *a la* Vance (1979), for the three segments other than /t/. This is an interesting result especially because, as discussed in the introduction, evidence from the Japanese lexicon does not distinguish the local violation from the non-local violation.

The current finding thus instantiates a case of the emergence of the unmarked (TETU: McCarthy & Prince 1994) in an experimental setting. More broadly speaking, the current result shows that there is an aspect of phonological knowledge of Japanese which cannot be learned from the lexical patterns of rendaku and Lyman's Law. This result supports the role of abstract grammatical knowledge which somehow imposes a locality effect on Lyman's Law, although we admit that it is puzzling that some speakers exhibit such an "anti-grammatical effect."[8]

Some questions arise from the current results, not of all which we can answer in this paper. First, we have no good explanation regarding why /t/ behaves differently from /k/, /s/ and /h/. As far as we know, there is nothing that is special about /t/—or [d]—in Japanese, rendaku-related or otherwise, that would make it exceptional to the locality effect of Lyman's Law. Recall that there is very little evidence for the local nature of Lyman's Law in the Japanese lexicon after all. Second, we are unable to offer a good explanation for why there is a non-trivial degree of interspeaker variability, as in Figure 2; neither are we able to explain why there are speakers who show "anti-locality" effect. Finally, a new question arises regarding why Kawahara (2012) failed to find a difference between the local condition and the non-local condition. We find the last question to be important, partly because it led Vance to consider his old results to an artifact of uncontrolled factors (Vance 2022). Therefore, in the next experiment we attempted to address this last question.

3 Experiment 2

We can consider two possibilities regarding why Kawahara (2012) failed to find a locality effect: (1) a naturalness judgment experiment, for some reason or another, was not a good task to reveal that effect or (2) the experiment by Kawahara (2012) lacked a sufficient statistical power, i.e., the N was too small. Recall that there were only three items for each segment-condition combination. While 54 participants may not be a very small number of speakers for a linguistic experiment, it

[8]The argument in this paragraph rests on the assumption that learners use only rendaku-related evidence to learn the grammatical status of Lyman's Law. It may be possible, however, that the local nature of Lyman's Law can be learned from somewhere else; for instance, there may be more loanwords which incur a non-local violation of Lyman's Law (e.g. [**d**eɾi**d**a] 'Derrida') than a local violation of Lyman's Law (e.g. [hai**d**e**g**aa] 'Heidegger'). With this said, the importance of the current findings remains robust, we believe, whatever the source of the locality effect is.

may nevertheless have been insufficient. To tease apart these two possibilities, we attempted to replicate Kawahara (2012) with a larger number of speakers, that is with N that is comparable to that of Experiment 1.

3.1 Method

Since we used up a pool of participants who can take a rendaku-related experiment (recall that we needed participants who are not familiar with rendaku or Lyman's Law), we resorted to the Buy Response function offered by Surveymonkey, the limitation of which is that we can include only up to 50 questions. Therefore, we limited ourselves to two segments /k/ and /s/, which showed a clear locality effect in Experiment 1.

The methodological details of Experiment 2 were similar to those of Experiment 1, except for a few difference. First, Experiment 2 was a naturalness judgment experiment, in which the participants were asked to rate the naturalness of rendaku-undergoing forms using a 5-point Lickert scale, where 5 was labeled as 'very natural' and 1 was labeled was 'very unnatural.' For statistical analyses, we used a Baysian *ordinal* logical regression with the same random structure as Experiment 1. The baseline for the segmental difference was arbitrarily chosen as /k/. Again the R markdown file available at the osf repository shows complete details of the analysis.

3.2 Results

Figure 3 shows the distribution of naturalness ratings for the three conditions, with the two facets showing the two segment types. We observe that the first condition with no violations of Lyman's Law was generally rated as most natural. The forms with a local violation of Lyman's Law were rated as least natural and those with the non-local violation were rated as intermediate. The grand averages from the left to right were: 3.01, 2.68 and 2.79.

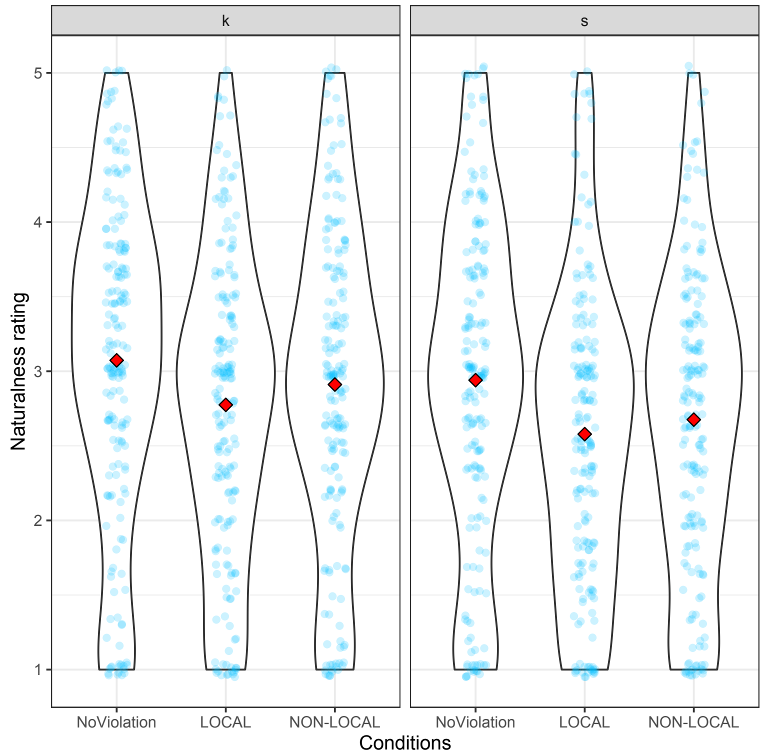

Figure 3: The comparison between the three critical conditions (Experiment 2).

The model summary of the results in Experiment 2 appears in Table 3. The 95% CrI for the segmental difference (coefficient (b)) includes 0, although the distribution is leaning toward the negative, suggesting that [z]-initial forms were rated less natural than [g]-initial items. The 95% CrI for the difference between the no-violation and the local violation (coefficient (c), the top) does not include 0, suggesting the robustness of the effects of (local) Lyman's Law.

Table 3: Summary of the Bayesian mixed effects ordinal logistic regression model (Experiment 2).

		β	error	95% CrI
(a) (baseline = /k/, local)				
intercept[1]		-2.44	0.25	[-2.92, -1.95]
intercept[2]		-0.50	0.25	[-0.98, -0.01]
intercept[3]		1.63	0.25	[1.15, 2.11]
intercept[4]		3.79	0.25	[3.30, 4.29]
(b) segment		-0.47	0.24	[-0.94, -0.00]
(c) condition	no-violation vs. local	0.78	0.25	[0.28, 1.28]
	local vs. non-local	0.34	0.24	[-0.13, 0.82]
(d) interactions	seg:no-violation vs. local	-0.13	0.33	[-0.79, 0.53]
	seg:no-violation vs. local	0.12	0.34	[-0.55, 0.77]

The 95% CrI for the difference between the local and non-local violation conditions (coefficient (c), the bottom) include 0, but it is leaning toward positive values, suggesting that the non-local violation condition tended to induce more natural responses than local responses. For the comparison between non-local and local violation conditions, as with Experiment 1, we thus calculated the probabilities of the β-coefficients being in the expected direction in terms of their posterior distributions.

The difference between the local violation and non-local violation at the baseline level (=/k/) was $p(\beta > 0) = 0.92$. The locality comparison at the level of /s/ was $p(\beta > 0) = 0.81$. Thus, we are at least 80% positive that the local and non-local violation conditions induced different naturalness ratings. These results are not as robust as those found in Experiment 1, but we find the converging results between the two experiments to be encouraging.

The effects of the Lyman's Law—the comparison between no violation and local violation—were clearer: for /k/, ($p(\beta > 0) = 0.99$ and for /s/ as well, ($p(\beta > 0) = 0.99$. These results are compatible with what Kawahara (2012) found.

Figure 4 shows the by-speaker analysis of the results in Experiment 2. Those dots above the diagonal axis represent speakers who show a locality effect, whereas those who are below the diagonal line are those who show an anti-locality effect. As with Experiment 1, we do observe that both types of speakers exist, but more speakers show a locality effect than an anti-locality effect, hence the overall results in Figure 3.

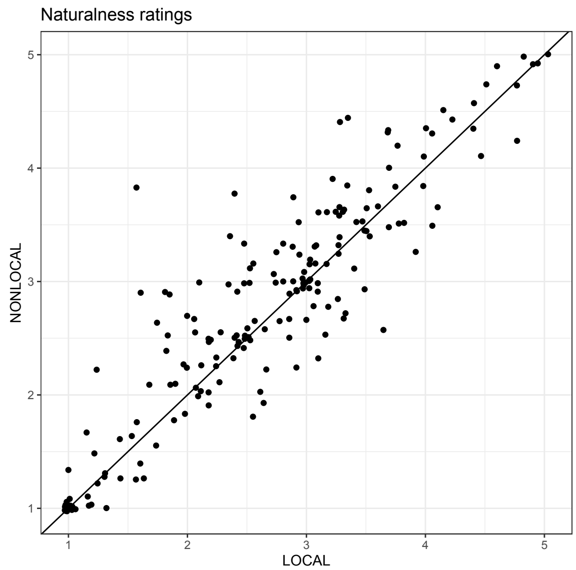

Figure 4: The comparison between the local violation condition and the non-local violation condition by each speaker (Experiment 2).

3.3 Discussion

We thus observe at least modest evidence (80%–90% confidence) that the local violation of Lyman's Law and the non-local violation induce different naturalness ratings—i.e. local violation tend to be judged to be less natural, contrary to the conclusion drawn by Kawahara (2012). We note, however, that Kawahara (2012) did observe a trend in the expected direction and that the sizes of differences were almost identical between Kawahara (2012) and the current experiment (2.76 vs. 2.86 = 0.10 in Kawahara 2012 and 2.68 vs. 2.79 = 0.11 in the current experiment). We also note that if we were using a frequentist analysis and were stuck with a "$p < .05$" threshold, then the current results may have turned out to be "non-significant." The use of Bayesian analyses allowed us to see how confident we can be about the difference between the local condition and the non-local condition.[9]

Having said these, it is also true that the results are less clear-cut in Experiment 2 than in Experiment 1, which suggests that naturalness rating experiments using a Lickert scale may not be an optimal method to identify the locality effect of Lyman's Law. One reason may be that the participants were presented only with one form (i.e. rendaku-undergoing form), whereas in Experiment 1, the participants were allowed to compare rendaku-undergoing forms and non-rendaku-undergoing forms (see Daland et al. 2011; Kawahara 2015; Sprouse & Almeida 2017 for related observations, especially in terms of how these two experimental paradigms can differ).

[9] These considerations led us to reanalyze the data reported in Kawahara (2012) using a Bayesian method. See Appendix.

Another reason may be that some participants may have had difficulty in interpreting what "naturalness" really means, especially when they are given nonce words.

4 Overall discussion

The most important finding of the current experiments, we believe, is empirical: we found that generally speaking, Lyman's Law shows a locality effect in that its dissimilatory force is stronger when the two voiced consonants are in adjacent syllables than when they are not, as Vance (1979) showed. This is not too surprising given that dissimilatory forces tend to function in this manner cross-linguistically (Suzuki 1998). The result, on the other hand, can be taken to be indeed surprising, because the Japanese lexicon does not offer clear evidence for this locality effect of Lyman's Law. Recall that Vance (2022) himself, who found the effect in 1979, later speculated that his finding was due to some uncontrolled factors.

The current results also offer some lessons for experimental phonology in general. First, the fact that Kawahara (2012) failed to find a "statistically significant" difference suggests that using a frequentist analysis as in Kawahara (2012) may not have been an optimal strategy to identify a linguistic effect (see Chambers 2017; Vasishth & Gelman 2021 for related discussion). Second, a naturalness judgment experiment may be a less reliable tool compared to a forced judgment task—it may be easier for naive participants to compare two distinct forms than making naturalness judgments of one form in isolation. These lessons open up an opportunity for future research: to re-examine the aspects of rendaku that have been studied in previous experimental studies (Kawahara 2016), ideally with a large number of speakers using a Bayesian method.

Appendix: Reanalyzing Kawahara's (2012) data

While we acknowledge that it is not desirable to run a statistical test after the results are obtained (Kerr 1998), having seen the results of Experiment II prompted us to see what would happen if we run a Bayesian analysis to the data obtained by Kawahara (2012). Explicitly bearing in mind that this is a post-hoc analysis, whose results should be interpreted with caution, we ran a Bayesian analysis that is similar to the one that was used for our Experiment 2. However, since there were only three items for each segment-condition combination, we dropped the segmental difference as a fixed factor from the model, as a three-level random factor is inappropriate (Snijders & Bosker 2011). There is an R markdown file available on the osf repository which shows the complete details of this reanalysis.

The result of the reanalysis shows that for the difference between the local violation condition and the non-local condition violation, $p(\beta > 0) = 0.938$ even for this old dataset. While this model is incomplete in that we had to drop segment type as a factor, the data obtained by Kawahara (2012) seem to be comparable with what we obtained in Experiment 2.

Funding

This project is supported by JSPS grants #22K00559 to the first author and #19K13164 to the second author.

Conflicts of interest

We declare no conflicts of interest.

Availability of data and material

The data are available at
`https://osf.io/ym79p/?viewonly=ce17de5a39834ae397c44a19e74db082`

Code availability (software application or custom code)

The code is also available at
`https://osf.io/ym79p/?viewonly=ce17de5a39834ae397c44a19e74db082`

Authors' contributions

Both authors contributed to the conception and execution of the experiments. The first author wrote the manuscript and the second author revised it. The statistical analysis was primarily conducted by the first author. The second author checked the details.

Ethics approval

The current experiments were conducted with an approval from the authors' institute.

Consent to participate

The participants read the written consent form before participating in the experiments.

Consent for publication

Both authors approve that the current manuscript be evaluated for publication in the journal.

References

Alderete, John & Stefan Frisch. 2007. Dissimilation in grammar and the lexicon. In Paul de Lacy (ed.), *The Cambridge handbook of phonological theory*, 379–398. Cambridge: Cambrdige University Press.

Bennett, Wm. G. 2015. *The phonology of consonants*. Cambrige: Cambridge University Press.

Bürkner, Paul-Christian. 2017. brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software* 80(1). 1–28.

Chambers, Chris. 2017. *The 7 deadly sins of psychology*. Princeton: Princeton University Press.

Cho, Taehong. 2021. Where we are at: Impact, special collections, open science and registered report at the *journal of phonetics*. *Journal of Phonetics* 89.

Chomsky, Noam. 1986. Rules and representations. *The Behavioral and Brain Sciences* 3. 1–15.

Coetzee, Andries W. 2009. Grammar is both categorical and gradient. In Steve Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*, 9–42. London: Equinox.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis & Ingrid Norrmann. 2011. Explaining sonority projection effects. *Phonology* 28(2). 197–234.

Foley, James. 1991. *The Yimas language of Papua New Guinea*. Stanford: Stanford University Press.

Franke, Michael & Timo B. Roettger. 2019. Bayesian regression modeling (for factorial designs): A tutorial. Ms. https://doi.org/10.31234/osf.io/cdxv3.

Frisch, Stephan, Janet Pierrehumbert & Michael Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22. 179–228.

Gallistel, Randy C. 2009. The importance of proving the null. *Psychological Review* 116(2). 439–453.

Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau & Yu-Sung Su. 2018. A weakly informative default prior distribution for logistic and other regression models. *Annual Applied Statistics* 2(4). 1360–1383.

Hansson, Gunnar Olafur. 2001. *Theoretical and typological issues in consonant harmony*: University of California, Berkeley Doctoral dissertation.

Ihara, Mutsuko, Katsuo Tamaoka & Tadao Murata. 2009. Lyman's Law effect in Japanese sequential voicing: Questionnaire-based nonword experiments. In The Linguistic Society of Korea (ed.), *Current issues in unity and diversity of languages: Collection of the papers selected from the 18th International Congress of Linguists*, 1007–1018. Seoul: Dongam Publishing Co., Republic of Korea.

Ito, Junko & Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17. 49–73.

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

Kawahara, Shigeto. 2012. Lyman's Law is active in loanwords and nonce words: Evidence from naturalness judgment experiments. *Lingua* 122(11). 1193–1206.

Kawahara, Shigeto. 2015. Comparing a wug-test and a naturalness rating test: An exploration using rendaku. *Language Sciences* 48. 42–47.

Kawahara, Shigeto. 2016. Psycholinguistic studies of rendaku. In Timothy Vance & Mark Irwin (eds.), *Sequential voicing in Japanese compounds: Papers from the ninjal rendaku project*, 35–46. Amsterdam: John Benjamins.

Kawahara, Shigeto & Gakuji Kumagai. 2023. Lyman's law counts only up to two. *Laboratory Phonology* .

Kawahara, Shigeto & Shin-ichiro Sano. 2014a. Identity avoidance and Lyman's Law. *Lingua* 150. 71–77.

Kawahara, Shigeto & Shin-ichiro Sano. 2014b. Identity avoidance and rendaku. *Proceedings of Phonology 2013* .

Kawahara, Shigeto & Hideki Zamma. 2016. Generative treatments of rendaku. In Timothy Vance & Mark Irwin (eds.), *Sequential voicing in Japanese compounds: Papers from the NINJAL rendaku Project*, 13–34. Amsterdam: John Benjamins.

Kerr, N.L. 1998. HARKing: Hypothesizing after the results are known. *Personality and Psychology Review* 2(3). 196–217.

Kruschke, John K. 2014. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan.* Waltham: Academic Press.

Kruschke, John K. & Torrin M. Liddell. 2018. The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychological Bulletin and Review* 25. 178–206.

Lemoine, N.P. 2019. Moving beyond noninformative priors: Why and how to choose weakly informative priors in bayesian analyses. *Oikos* 128. 912–928.

Lyman, Benjamin S. 1894. Change from surd to sonant in Japanese compounds. *Oriental Studies of the Oriental Club of Philadelphia* 160–176.

McCarthy, John J. & Alan Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In Merce Gonzalez (ed.), *Proceedings of the North East Linguistic Society 24*, 333–379. Amherst, Mass.: GLSA Publications.

McCawley, James D. 1968. *The phonological component of a grammar of Japanese.* The Hague: Mouton.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan, 2nd edition.* London: Taylor & Francis Ltd.

Ohala, John. 1981. The listener as a source of sound change. In T. Myers, J. Laver & Anderson J. (eds.), *Proceedings of Chicago Linguistic Society 17*, 178–203. Chicago: Chicago Linguistic Society.

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science* 349(aac4716–aac4716).

Pulleyblank, Douglas. 2002. Harmony drivers: no disagreement allowed. In *Proceedings of Berkeley Linguistic Society 28*, 249–267. Berkeley: Berkeley Linguistics Society.

R Development Core Team. 1993–. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing Vienna, Austria.

Roettger, Timo B. 2019. Researcher degree of freedom in phonetic research. *Laboratory Phonology* 10(1). 1, doi `https://doi.org/10.5334/labphon.147`.

Shinohara, Shigeko. 1997. *Analyse phonologique de l'adaptation japonaise de mots etrangers*: Universite de la Sorbonne nouvelle Paris III Thèse pour le doctorat.

Snijders, Tom & Roel Bosker. 2011. *Multilevel analysis: An introduction to basic and advanced multilevel modeling, 2nd ed.* Los Angeles: Sage Publications.

Sönning, Lukas & Valentine Werner. 2021. The replication crisis, scientific revolutions, and linguistics. *Linguistics* 59(5). 1179–1206.

Sprouse, Jon & Diogo Almeida. 2017. Design sensitivity and statistical power in acceptability judgment experiments. *Glossa* 2(1).

Stanton, Juliet. 2019. Constraints on contrast motivate nasal cluster dissimilation. *Phonology* 36(4). 655–694.

Suzuki, Keiichiro. 1998. *A typological investigation of dissimilation*: University of Arizona Doctoral dissertation.

Vance, Timothy. 1979. *Nonsense word experiments in phonology and their application to rendaku in Japanese*: University of Chicago Doctoral dissertation.

Vance, Timothy. 2022. *Irregular phonological marking of Japanese compounds.* Berlin: Mouton de Gruyter.

Vasishth, Shravan & Andrew Gelman. 2021. How to embrace variation and accept uncertainty in linguistic and psycholinguistic data analysis. *Linguistics* 59(5). 1311–1342.

Vasishth, Shravan, Bruno Nicenboim, Mary Beckman, Fangfang Li & Eun Jong Kong. 2018. Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics* 71. 147–161.

Wilson, Colin. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30(5). 945–982.

Winter, Bodo. 2019. *Statistics for linguists.* New York: Taylor & Francis Ltd.

Zuraw, Kie. 2000. *Patterned exceptions in phonology*: University of California, Los Angeles Doctoral dissertation.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language* 83(2). 277–316.