Rendaku and identity avoidance: Consonantal identity and moraic identity

1 Introduction

1.1 Synopsis

Chapter 2 of this book provides an overview of the previous experiments on rendaku, but no experimental details are provided in that chapter for the sake of readability. To complement Chapter 2, as a case study, this chapter reports a new experiment on rendaku in full detail.

Recent experimental studies of rendaku identified a hitherto unnoticed factor that inhibits rendaku (Kawahara & Sano, 2014a,b): when rendaku results in adjacent identical CV moras, rendaku applicability is reduced (e.g., schematically, */**iga+ga**niro/ from /**iga**/+/**ka**niro/). However, these previous experiments have only tested the Identity Avoidance effects at the CV moraic level.

The current study therefore tests whether Identity Avoidance at the consonantal level (i.e. *[C_i...C_i]) affects the applicability of rendaku. The current study shows that, although its effect is weaker than the moraic Identity Avoidance effect, Japanese speakers do avoid creating identical consonants in adjacent moras (e.g., schematically, */**iga+go**moke/ from /**iga**/+/**ko**moke/). The current study overall offers the following new understanding about Japanese phonology: (i) rendaku is subject to the consonantal Identity Avoidance effect, which is a new descriptive discovery, (ii) a restriction that is operative in many other languages is also operative in Japanese, revealing an intriguing cross-linguistic parallel, (iii) Identity Avoidance at different phonological levels can coexist within a single language, and (iv) the strength of the avoidance effect positively correlates with the degree of similarity.

1.2 Background and the current study

Rendaku is a well-known and well-studied morphophonological phenomenon in Japanese, in which the initial obstruent of the second member of a compound appears as voiced. However, rendaku is in no way an exceptionless "automatic" phonological rule, and many factors affect the applicability

of rendaku (see Vance 2015 for a recent comprehensive overview). For example, one factor that blocks rendaku is Lyman's Law (Lyman 1894 *et seq*; see Chapter XX), according to which if a second element already contains a voiced obstruent, rendaku is almost categorically blocked (/**t**oka**ge**/ 'lizard' → /oo-**t**oka**ge**/ 'big lizard'). Rendaku is also said to be blocked when the element is on a left branch of a compound (e.g. /nise+/**t**anuki+jiru// '[fake [raccoon soup]]') (Otsu, 1980). A research program initiated by the seminal work by Vance (1980) has experimentally investigated whether these rendaku inhibiting factors, including Lyman's Law, are psychologically real (see Chapter 2 for a recent overview of this research program).

Up until recently, the experimental research focused on the factors that are already known to affect rendaku applicability in the existing patterns of Japanese phonology. For example, several works have investigated the psychological nature of Lyman's Law (Ihara et al., 2009; Kawahara, 2012; Vance, 1980), which has confirmed the activity of Lyman's Law in the minds of contemporary Japanese speakers. Some of the experiments have confirmed the psychological reality of rendaku-affecting factors in this way (e.g. Nakamura & Vance 2002; Vance 1980, 2014). However, some experiments did not: for example, neither Kozman (1998) nor Kumagai (2014) succeeded in obtaining the results that confirm the hypothesis by Otsu (1980) that only the elements on the right branch of a compound undergo rendaku (though cf. Ihara & Murata 2006).

One emerging research program that grew out of this research tradition is whether hitherto unknown phonological factors can affect the applicability of rendaku in experimental settings (see Chapter 2). For example, many languages show evidence for Identity Avoidance (Yip, 1998), in which adjacent identical elements are avoided. Recent experimental studies have shown that this identity avoidance effect does reduce the applicability of rendaku. Japanese speakers are less likely to apply rendaku when it results in adjacent identical CV moras: rendaku is less likely when it would violate moraic CV Identity avoidance (e.g. /**iga+ga**niro/ from /**iga**/+/**ka**niro/) than when it would not (e.g. /**iga+da**niro/) (Kawahara & Sano, 2014a,b). Particularly interesting about this finding is that, in terms of the statistical patterns in the Japanese lexicon, there is no evidence for such Identity Avoidance effects related to rendaku, according to a study by Irwin (2014), based on a large corpus of rendaku (Irwin & Miyashita, 2013) (though cf. Sato 1988 and Labrune 2012 who point out some sporadic examples). A general lesson that is emerging from these studies is that we can learn something new about a phonological pattern by way of experimentation, which is otherwise difficult to identify just by looking at existing patterns in the lexicon.

Although this finding by Kawahara & Sano (2014a,b) is interesting, one limitation of these studies was that they tested only the CV moraic identity effect. This choice was not without a reason, because CV moraic sequences constitute an important phonological unit in Japanese (Kubozono, 1989; Labrune, 2012). A question nevertheless remains whether the Identity Avoidance effect at the consonantal level is also operative in the phonology of Japanese. This is an im-

portant question to address, because consonantal Identity Avoidance effects are observed in many different languages, most famously in various Semitic languages (Frisch et al., 2004; Greenberg, 1950), but also in languages like English (Raffelsiefen, 1999), French (Zuraw, 2015), Mandarin (Yip, 1998), and others (Alderete & Frisch, 2007; Yip, 1998; Zuraw & Lu, 2009). This paper takes up this task of addressing whether Identity Avoidance exists at the consonantal level in the phonology of Japanese.

2 Method

2.1 Task

The current experiment used a two-way forced-choice wug-test (Berko 1958 *et seq.*). Within each trial, the participants were given two elements (E1 and E2), and were provided with a compound form with rendaku and one without rendaku. They were then asked to choose the better resulting compounding form. For example, they were asked: "given /iga/ 'thorny' and /kaniro/ (nonce), what would be the better outcome, /igakaniro/ or /igaganiro/? Please choose the one that sounds more natural to you." The stimuli were presented in the form of Japanese orthography, although the participants were encouraged to subvocalize the stimuli before answering each question.

Our previous experiments showed that using nonce words for both E1 and E2 can impose too much psycholinguistic burden on native speakers, at least during wug-tests about rendaku. Therefore, real words were used for E1 and nonce words were used for E2 in the current experiment. The participants were told to treat E2 as old animal names that used to inhabit in Japan. This procedure was used because rendaku applies only to native words and not to loanwords (e.g. Vance 2015; see also Chapter XX), and this technique allows the participants to conceive the nonce word stimuli as (old) native words (see Kawahara 2012; Vance 1980; Zuraw 2000 for this technique).

2.2 Stimuli

The stimuli consisted of two sets: Set 1 tested the effect of Identity Avoidance at the moraic level; Set 2 tested the effect of Identity Avoidance at the consonant level. Set 1 and Set 2 used the same set of E1 and a similar set of E2 with all the consonants that can potentially undergo rendaku (/t, k, s, h/). Three different nonce words were created for each type of the four consonants for E2, while controlling for the quality of the last two moras ($4 \times 3 = 12$ types of E2). All the stimuli had only CV light syllables, where one CV syllable coincides with one mora in the phonology of Japanese (Kubozono, 1989; Labrune, 2012).

In both sets, all the factorial combinations of E1 and E2 were included, in order to test the effect of *combination* of moras and consonants at the morphological juncture, while controlling for

potential lexical-specific effects of E1 or E2. Within the factorial combinations, some combinations resulted in the violation of Identity Avoidance, whiles other combinations did not.

The experimental items for Set 1 are provided in Table 1. In Set 1, in one condition, the two moras across the morpheme boundary were identical except for voicing of the onset consonant (e.g., /i**ga**/+/**ka**niro/); in this condition, rendaku would result in two adjacent identical CV moras (i.e., /i**ga+ga**niro/). In the other condition, the first obstruent in E1 differed in place and/or manner with the voiced obstruent in E2 (e.g., /i**ga**/+/**ta**niro/); rendaku would not result in two identical moras or consonants (i.e., /i**ga+da**niro/). E1 always contained a voiced obstruent, thereby controlling for the potential effect of the presence of a voiced obstruent in E1 (Ito & Mester, 2003; Kawahara & Sano, 2014c; Unger, 1975).

Table 1: The list of the stimuli used in Set 1. All combinations of E1 and E2 (4 * 12 =48) were tested. Some combinations resulted in the violation of the moraic Identity Avoidance. The others did not. Gloss: /iga/='thorn'; /aza/='bruise'; /kuda/='pipe'; /kaba/='hipopotumus'.

E1		E2		
/iga/	*	/kaniro/	/kamoke/	/karimo/
/aza/		/saniro/	/samoke/	/sarimo/
/kuda/		/taniro/	/tamoke/	/tarimo/
/kaba/		/haniro/	/hamoke/	/harimo/

The stimuli for Set 2 are presented in Table 2. The basic structure is the same as Set 1, but in this set, some combinations would result in identical consonants in adjacent moras after rendaku, not adjacent identical CV moras (i.e., /i**ga+ge**niro/). Other combinations did not violate either consonantal Identity Avoidance or moraic Identity Avoidance (i.e., /i**ga+de**niro/).

Table 2: The list of the stimuli used in Set 2. All combinations of E1 and E2 (4 * 12 =48) were tested. Some combinations resulted in the violation of the consonantal Identity Avoidance; the other did not.

E1		E2		
/iga/	*	/keniro/	/komoke/	/korimo/
/aza/		/seniro/	/somoke/	/sorimo/
/kuda/		/teniro/	/tomoke/	/torimo/
/kaba/		/heniro/	/homoke/	/horimo/

2.3 Procedure and participants

The participants were undergraduate students at Okayama Prefectural University.[1] The experiment was run online using Surveymonkey (for the reliability of online experimentation in psychology and linguistic research, see Reips 2002, Sprouse 2011, and Yu & Lee 2014). The participants were first told what rendaku is, and then were asked to go through three practice questions using /nise/ 'fake' as E1 and real words as E2 in order to familiarize themselves with rendaku and the task of the current experiment. Although the stimuli were presented online using the Japanese orthography, the participants were reminded for each question that they should choose the more natural *sounding* choice.[2]

The stimuli for Set 1 and Set 2 were mixed together in one block, including all the 96 stimulus items. The order of the stimuli was randomized per participant by Surveymonkey. There were no time limits for answering questions. Forty-three native speakers of Japanese completed this study. They received extra credit for participation.

2.4 Statistics

Since the response was binary (yes-rendaku or no-rendaku), logistic linear-mixed model analyses were run to analyze the results (Baayen, 2008). Subjects and items (both E1 and E2) were encoded as random factors. Both slopes and intercepts of random effects were included in the models to have the maximal random structure (Barr et al., 2013).

3 Results

Figure 1 shows the proportions of rendaku application for each condition, with error bars representing 95% confidence intervals, calculated over all the participants. The first bar is a case which violates moraic Identity Avoidance. The second bar is the control condition in Set 1, which shares the same E2 with the first condition, but does not involve Identity Avoidance effect. The third bar is from Set 2, and violates consonantal Identity Avoidance. The fourth bar is the control condition

---

[1]The majority of the participants were therefore from the areas around Okayama. This limitation was practical rather than theoretically-motivated, and we do not mean to over-generalize our finding to speakers of Tokyo Japanese (or for speakers of other dialects of Japanese). Dialectal differences in terms of rendaku are in fact an understudied area of research, although there are some recent work (Vance et al., to appear). At any rate, we believe that it suffices, for the current purpose, to show that both moraic Identity Avoidance and consonantal Identity Avoidance hold in *some* dialect of Japanese. A follow-up experiment using Tokyo Japanese speakers would be interesting and informative.

[2]It would be interesting to replicate the experiment with auditory stimuli. Most if not all the experiments on rendaku use a paper-based format, however, and it is a task for rendaku-related experiments in general to test rendaku using auditory stimuli (see Chapter 2). See Kawahara (2013) for a set of experiments addressing this issue, using geminate devoicing found in Japanese loanwords

in Set 2, which again shares the same E2 with the third condition, but without Identity Avoidance effect.
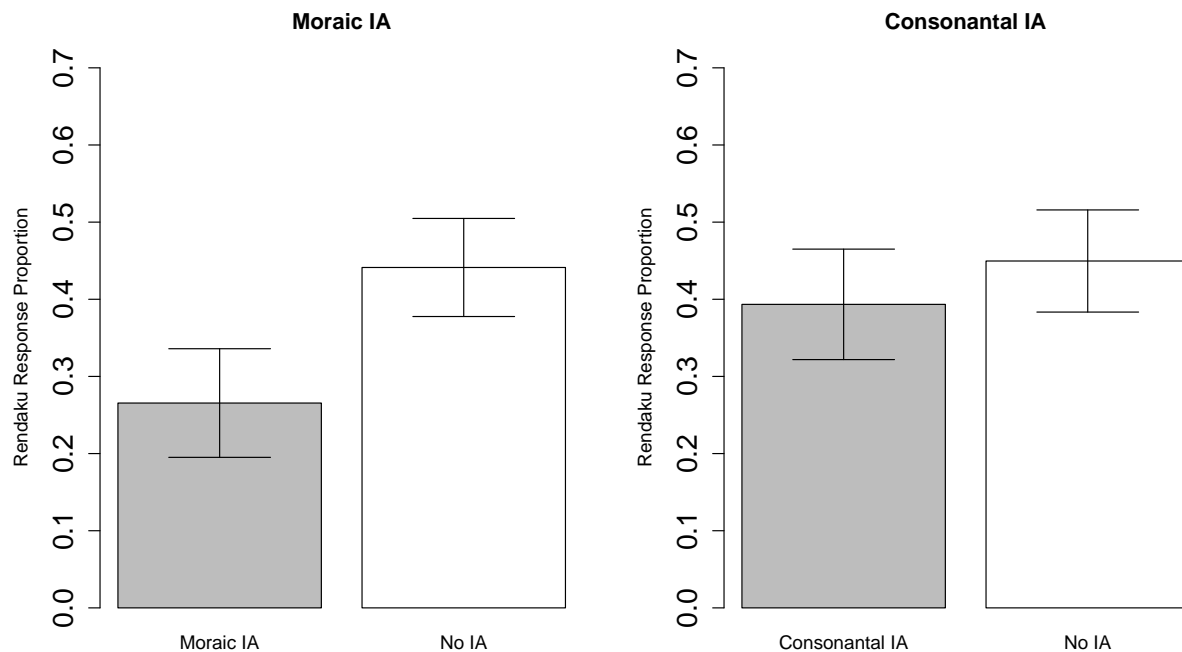

Figure 1: The proportions of rendaku application for each condition. The error bars represent 95% confidence intervals, calculated over all the participants. Rendaku was less likely when it violates either moraic Identity Avoidance or consonantal Identity Avoidance. The effect of Identity Avoidance is stronger at the moraic level than at the consonantal level.

The statistical results show, first of all, that the moraic Identity Avoidance shows a strong effect: the comparison between the first and the second bars show a significant difference (0.27 vs. 0.44; $z = 5.318, p < .001$). Second, the comparison of the results in Set 2 shows that those that violate consonantal Identity Avoidance shows lower rendaku proportion than the control condition (0.39 vs. 0.45: $z = 2.227, p < .05$).

We also observe that the effects of Identity Avoidance are stronger at the moraic level (the first bar) than at the consonantal level (the third bar). The difference between the moraic Identity Avoidance violation and the consonantal Identity Avoidance violation is statistically significant ($z = 4.55, p < .001$).

4 Discussion

4.1 The effect of moraic Identity Avoidance

The current study has identified a strong effect of rendaku blockage due to moraic Identity Avoidance: the degree of blockage by moraic Identity Avoidance is about 17%. This effect was already shown by our previous experiments (Kawahara & Sano, 2014a,b), but it is a good replication, given that the existence of the moraic Identity Avoidance effect was not identified by Irwin (2014) in the existing patterns of rendaku in a large corpus of rendaku (Irwin & Miyashita, 2013).

The fact that we replicate this effect in slightly different experimental settings with different sets of stimuli and different sets of speakers shows that this effect of moraic Identity Avoidance may hold generally among contemporary Japanese speakers. Taken together with Irwin's (2014) study that there is no evidence for such Identity Avoidance effects in the contemporary Japanese lexicon, the results may instantiate a case of a grammatical effect that emerges in experimental settings, beyond the lexical patterns (see e.g. Berent et al. 2007; Moreton 2002 for other works showing this emergence of the grammatical effects).[3]

4.2 The effect of consonantal Identity Avoidance

Next, moving on to the consonantal Identity Avoidance effect, rendaku has been much studied in detail in both the traditional studies of Japanese and the theoretical literature (see the Rendaku Bibliography Chapter). However, despite this research tradition, the effect of consonantal Identity Avoidance has not been known, to the best of our knowledge.

Our result therefore offers a new descriptive discovery to the study of rendaku. We can further conclude that a linguistic experiment is a useful methodology that complements the traditional approach to phonology based on dictionaries and introspection, in that it can reveal aspects of phonological knowledge that are difficult to access otherwise.

It is also interesting that the consonantal Identity Avoidance effect that we identified in this experiment is also commonly found in many other languages (Frisch et al., 2004; Greenberg, 1950; Zuraw & Lu, 2009). In Arabic and many other languages, Identity Avoidance effect functions as phonotactic restrictions, and in other languages, Identity Avoidance causes dissimilation. Our results show that similar constraints can block rendaku in Japanese. Therefore this study reveals an intriguing cross-linguistic parallel between Japanese and other genetically-unrelated languages. It is possible that similarity avoidance has its roots in speech processing (Alderete & Frisch, 2007; Frisch et al., 2004), and may thus be shared by speakers of different languages.

---

[3]This conclusion is based on the assumption that the database used by Irwin (2014)—Irwin & Miyashita (2013)—is comparable to the dataset that the participants of the current experiment are exposed to through the course of language acquisition. This assumption may not strictly hold, however.

This finding also highlights a related, and perhaps equally important, point—the importance of cross-linguistic examination of phonological patterns. Traditional Japanese linguists would not have looked for the effect of consonantal Identity Avoidance effects, because of their strong tendency to think in terms of moras rather than in terms of segments (for which see Labrune 2012). Therefore, a cross-linguistic study, in which we attempt to address whether effects observed in other languages also exist in Japanese, was crucial in helping us identify this effect at the sub-moraic, consonantal level.

4.3 Coexistence and granularity of Identity Avoidance effects

The current experiment shows that Identity Avoidance at different phonological levels can coexist within a single language, and the strength of the avoidance effect positively correlates with the degree of similarity; moraic Identity Avoidance is stronger than consonantal Identity Avoidance, because the former involves a larger phonological unit or involves more similar segments (moraic Identity Avoidance involves two segments, whereas consonantal Identity Avoidance involves only one).

This observation—the correlation between degrees of similarity and the extent of avoidance—is in line with the finding by some recent work on the effect of similarity avoidance (Frisch et al., 2004), but goes beyond that observation in a sense that the current experiment shows that the degrees of similarity in string sequences matter, whereas the previous studies were about degrees of similarity within a segment.

5 Summary

The current study has revealed two Identity Avoidance effects within a single language, taking Japanese rendaku as a case study. In addition to its new descriptive discovery, the current study has identified an intriguing cross-linguistic parallel between Japanese and other languages.

Acknowledgements

References

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

Baayen, Harald R. (2008) *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Barr, Dale J., Roger Levy, Christoph Scheepers, & Harry J. Tily (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68**: 255–278.

Berent, Iris, Donca Steriade, Tracy Lennertz, & Vered Vaknin (2007) What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* **104**: 591–630.

Berko, Jean (1958) The child's learning of English morphology. *Word* **14**: 150–177.

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

Fukuda, Suzy & Shinji Fukuda (1994) To voice or not to voice: The operation of rendaku in the Japanese developmentally language-impaired. *McGill Working Papers in Linguistics* **10**: 178–193.

Greenberg, Joseph (1950) The patterning of root morphemes in Semitic. *Word* **6**: 162–181.

Ihara, Mutsuko & Tadao Murata (2006) Nihongo-no rendaku-ni kansuru ikutsuka-no jikken [Some experiments on sequential voicing]. *On-in Kenkyuu [Phonological Studies]* **9**: 17–24.

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

Irwin, Mark (2014) Rendaku across duplicate moras. *NINJAL Research Papers* **7**: 93–109.

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

Ito, Junko & Armin Mester (2003) *Japanese Morphophonemics*. Cambridge: MIT Press.

Ito, Junko & Armin Mester (2008) Lexical classes in phonology. In *The Oxford Handbook of Japanese Linguistics*, Shigeru Miyagawa & Mamoru Saito, eds., Oxford: Oxford University Press, 84–106.

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 (2013) Testing Japanese loanword devoicing: Addressing task effects. *Linguistics* **51**(6): 1271 – 1299.

Kawahara, Shigeto (2015) Can we use rendaku for phonological argumentation? *Linguistic Vanguard* : 1–12.

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 & Shin-ichiro Sano (2014c) Testing Rosen's Rule and Strong Lyman's Law. *NINJAL Research Papers* **7**: 111–120.

Kobayashi, Yuki, Yoko Sugioka, & Takane Ito (2014) Rendaku (Japanese sequential voicing) as rule application: An ERP study. *NeuroReport* **25**(16): 1296–1301.

Kozman, Tam (1998) The psychological status of syntactic constraints on *rendaku*. In *Japanese/Korean Linguistics 8*, David Silva, ed., Stanford: CSLI, 107–120.

Kubozono, Haruo (1989) The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech* **32**: 249–278.

Kumagai, Gakuji (2014) The psychological status of the Right-Branch Condition on rendaku: An experiment with specific contexts. *Studies in Language Sciences* **13**: 124–145.

Labrune, Laurence (2012) *The phonology of Japanese*. Oxford: Oxford University Press.

Lyman, Benjamin S. (1894) Change from surd to sonant in Japanese compounds. *Oriental Studies of the Oriental Club of Philadelphia* : 1–17.

Moreton, Elliott (2002) Structural constraints in the perception of English stop-sonorant clusters. *Cognition* **84**: 55–71.

Nakamura, Kumiko & Timothy Vance (2002) *Rendaku* in noun+ verb-stem compounds: A production task. Talk presented at LP 2002, Urayasu.

Otsu, Yukio (1980) Some aspects of rendaku in Japanese and related problems. In *MIT Working Papers in Linguistics*, vol. 2, Ann Farmer & Yukio Otsu, eds., Cambridge, Mass.: Department of Linguistics and Philosophy, MIT, 207–228.

Raffelsiefen, Renate (1999) Phonological constraints on English word formation. *Yearbook of Morphology* **8**: 225–287.

Reips, Ulf-Dietrich (2002) Standards for internet-based experimenting. *Experimental Psychology* **49**(4): 243–256.

Sato, H. (1988) Hukugougo-ni okeru akusento kisoku-to rendaku kisoku [Rules on compound accents and rendaku]. *Nihongo no Onsei to On'in* **2**: 233–265.

Sprouse, Jon (2011) A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior and Research Methods* **43**(1): 155–167.

Unger, Marshall (1975) *Studies in Early Japanese Morphophonemics*. Doctoral dissertation, Yale University.

Vance, Timothy (1980) The psychological status of a constraint on Japanese consonant alternation. *Linguistics* **18**: 245–267.

Vance, Timothy (2007) Have we learned anything about *rendaku* that Lyman didn't already know? In *Current issues in the history and structure of Japanese*, Bjarke Frellesvig, Masayoshi Shibatani, & John Carles Smith, eds., Tokyo: Kurosio, 153–170.

Vance, Timothy (2014) If rendaku isn't a rule, what in the world is it? In *Usage-Based Approaches to Japanese Grammar: Towards the Understanding of Human Language,*, Kaori Kabata & Tsuyoshi Ono, eds., Amsterdam: John Benjamins, 137–152.

Vance, Timothy (2015) Rendaku. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Berlin: Mouton de Gruyter, 397–441.

Vance, Timothy, Mizuki Miyashita, & Mark Irwin (to appear) Rendaku in Japanese Dialects that Retain Prenasalization. In *Japanese/Korean Linguistics*, S Nam, J. Jun, & Ko. H., eds., CSLI.

Yip, Moira (1998) Identity avoidance in phonology and morphology. In *Morphology and its Relation to Phonology and Syntax*, Steven G. Lapointe, Diane K. Brentari, & Patrick M. Farrell, eds., Stanford: CSLI Publications, 216–246.

Yu, Alan & Hyunjung Lee (2014) The stability of perceptual compensation for coarticulation within and across individuals: A cross-validation study. *Journal of the Acoustical Society of*

America **136**(1): 382–388.

Zuraw, Kie (2000) *Patterned Exceptions in Phonology*. Doctoral dissertation, University of California, Los Angeles.

Zuraw, Kie (2015) Allomorphs of French *de* in coordination: A reproducible study. *Linguistic Vanguard* .

Zuraw, Kie & Yu-An Lu (2009) Diverse repairs for multiple labial consonants. *Natural Language and Linguistic Theory* **27**(1): 197–224.