A quantitative study of voiced velar nasalization in Japanese*

Canaan Breiss, Hironori Katsuda, and Shigeto Kawahara

1 Introduction

This paper presents a corpus-based analysis of voiced velar nasalization (VVN) in the standard (Yamanote) dialect of Japanese, in which a voiced velar plosive /g/ becomes nasalized in prosodic-word-medial position. To the best of our knowledge, this is the first quantitative study of the phenomenon, and confirms the impressionistic observation reported in the previous literature; at the same time, our study finds that these generalizations are stochastic. Looking more closely at the determinants of this variation reveals intriguing ways in which phonological grammar and lexicon interact, as well as the role of frequency in shaping phonological variation. Outside of frequency, we also examine factors such as prosodic length, and the effects of the segmental context on VVN.

2 Voiced velar nasalization

In various dialects of Japanese including the standard or prestige Yamanote dialect, [g] and [ŋ] stand in an allophonic relationship, with [g] occurring in prosodic-word-initial position, and [ŋ] occurring in prosodic-word-medial position (e.g. [gama] "toad" vs. [kaŋami] "mirror"). This pattern has been extensively studied in the traditional studies of Japanese (e.g. Kindaichi 1942, Labrune 2012), and also noted in the pre-generative European literature (Trubetskoy 1969). The relevance of this pattern to theoretical phonology was inspired by McCawley (1968) and extended by Ito and Mester (1996, 2003a), which the current study substantially builds upon.

Although VVN is classically described as being a static phonotactic restriction on the distribution of allophones, it has been most extensively discussed in the context of compounds, because the contrast between bound and free realizations of compound members allows VVN to be observed as an alternation. In this paper we study the behavior of compound-medial stem-initial instances of VVN, where the pattern of velar nasalization is reportedly sensitive to various morphophonological factors. When the second member (=N2) begins with [g] and can occur as a free morpheme, as in (1), voiced velar nasalization targeting that [g] is optional:

(1) a. /doku + ga/ → [doku-ŋa] ~[doku-ga]
poison moth, "poison moth"

b. /ga/ → [ga],
"moth"

Data in (1) demonstrate that VVN is not merely a static generalization about surface allophonic patterns, but it in addition manifests itself as an active (morpho)phonological alternation. Ito and Mester (1996) propose that the optionality in (1) arises from the conflict between a language-wide phonotactic constraint that requires VVN, and pressure for the N2 to retain the [g]-initial status of its free form, formalized using output-output constraints (Benua 2000).[1]

A consequence of this argument is that when N2 is bound, as in (2), there is no free [g]-initial form of N2 to enforce faithfulness to, and thus VVN should be obligatory, a prediction that is borne out by actual data. Note that in their analysis Ito and Mester (1996) conflated boundness and propensity to occur as N2 in a compound; it is logically possible that there are nouns which

*Thanks to audiences at the 45th Penn Linguistics Conference, WCCFL 39, AMP 2021, and UCLA for helpful feedback and discussion. The responsibility for all remaining errors rest with the authors. This research was funded in part by NSF Graduate Research Fellowship DGE-1650604 to Canaan Breiss.

[1]Technically Ito and Mester (1996) derive VVN through the interaction of context-free *[g] constraint and position-specific *[ŋ] constraint that prohibits word-initial [ŋ]. Ito and Mester (2003b) present some arguments that VVN should be better characterized in terms of *[VgV] constraint, directly prohibiting intervocalic [g]; this analytical path is also followed by Breiss et al. (2021).

are morphologically bound and must be prosodically dependent on some host, but can occur in both N1 and N2 position; we do not discuss this point further here, however, leaving it for future research.

(2) a. /doku + ga/ → [doku-ŋa], *[doku-ga]
poison fang, "poison fang"

b. /ga-ʒoo/→ [ga-ʒoo]
fang castle, "main castle"

The bulk of this paper focuses on verifying the claims made above and examining how lexical frequency conditions the optionality in VVN, which contributes to a growing body of literature demonstrating that phonological patterns which make reference to non-local surface forms are sensitive to the lexical characteristics of these forms (on which see more below). We also consider a number of other predictors of the alternation in compounds. In the following section we describe how we obtained data for our study.

3 Methods

We used the NHK Pronunciation and Accent Dictionary (NHK 1993) as a representation of the lexicon of the Yamanote dialect of Japanese. From it we extracted all compounds which contained a /g/-initial N2, and recorded their listed pronunciations as having either undergone VVN, as in /hai + gan/ → [hai-ŋan] "lung cancer", or not, as in /noo + geka/ → [noo-geka] "brain surgery". In cases where both pronunciations were listed and variation was noted, we recorded the pronunciation listed as preferred or predominant.

We then used the Balanced Corpus of Contemporary Written Japanese (Maekawa et al. 2014) (BCCWJ) to find part-of-speech tags of N1, N2, and the compound itself. We then excluded all compounds which contained either N1 or N2 with a non-noun tag or were marked as being loan words, and used the BCCWJ to annotate each remaining N1, N2, and compound form for log-frequency and length in mora We hand-annotated words or morphemes absent from the BCCWJ, giving them a log-frequency of 0 and manually calculating mora length. We also manually annotated each compound for the final segment of N1 (vowel or nasal). This left us with 1,323 compounds for analysis.

We also extracted all compounds with /k/-initial N2s but did not annotate them further, because they behaved categorically in undergoing VVN.

4 Results

4.1 Verifying Ito and Mester (1996)'s data

We first assessed the evidence for the observations proposed by Ito and Mester (1996) in (1) and (2) above. Figure 1 depicts the number of compounds undergoing VVN, depending on the status of the N2. All compounds with bound N2s undergo VVN (the left panel), supporting the data in (2). Those compounds with free N2s (the right panel) exhibit substantially more variation ($\sim$ 76% undergoing, $\sim$ 24% not undergoing), supporting the claims which data in (1) exemplify.

4.2 Relative frequency of the compound and its parts

Noting that there is a substantial amount of variation in VVN among compounds with free N2s (right panel of 1), we probed for determinants of this variation, focusing the following analyses on the 1,071 unique compounds in the right panel of 1. Recent work by Steriade and Stanton (2020) and Breiss (2021a,b) has found that other phonological processes that involve reference to surface forms of a paradigm (ex., Lexical Conservatism: Steriade 1997, among others) display experimentally-revealed variability that is governed in part based on those forms' lexical characteristics (frequency,

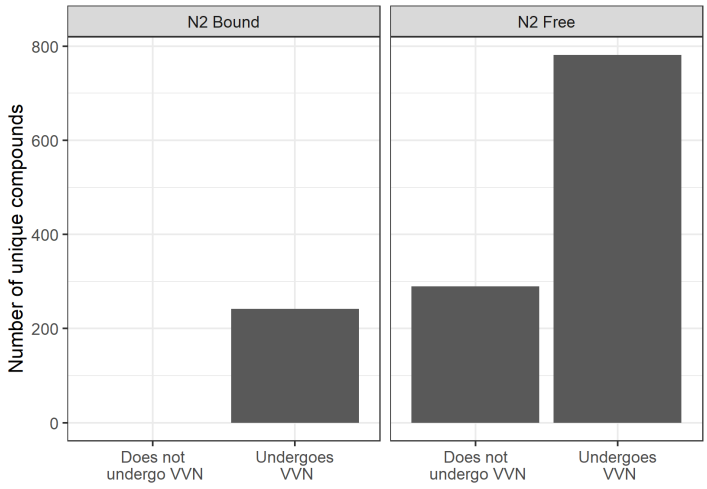

Figure 1: Division of the corpus of compounds according to whether a given compound undergoes VVN or not (horizontal axis), divided by whether or not N2 is able to occur as a free form (panels). The vertical axis plots the absolute number of compounds in each category.

semantic similarity, resting activation, etc). This finding is compatible with a model of the grammar where the lexical representation of the remote surface form is accessed in real time as part of the computation of the outcome of the word-formation process.

We ask the same question here about VVN, as a case of optional paradigm uniformity, examining whether the relative frequency of the compound and its N1 or N2 significantly predicts VVN. Figure 2 displays the behavior of the compounds with free N2s (that is, only those in the right panel of Figure 1) in two ways: on the left is the probability of undergoing VVN (vertical axis) as a function of the difference in relative frequency of N2 and the compound itself (horizontal axis). We examined relative frequency (cf. Hay 2003 and Hay and Baayen 2005 which demonstrate the utility of this measure), obtained by subtracting the log-frequency of either N1 or N2 respectively from the log-frequency of the compound itself, because it best captures the notion of comparative salience between the compound and the free form of its parts that underpins the formalism proposed by Ito and Mester (1996). On the right, the same metrics are plotted for the relative frequency of N1 and the compound. In both cases, we find that increased frequency of the independent noun — N1 or N2 — relative to the compound itself reduces the likelihood of VVN.

4.3 Segmental considerations

We also followed up on the native speaker intuitions of two of the authors that the overall size (in mora) of the compound itself influences the likelihood of that compound undergoing VVN. We find that this intuition corresponds to a robust trend in the corpus, such that the number of mora in the compound is a strong determinant of whether VVN applies, with longer compounds being less likely to undergo VVN compared to shorter ones. As shown in Figure 3, the effect is quite dramatic: compounds shorter than five mora undergo VVN almost without exception, while compounds longer than six mora almost never undergo the process, with compounds in the five- to six- mora range exhibiting robust variation.

Finally, we probed to see whether VVN was affected by the local phonological environment — specifically, the final segment of N1. We had no strong *a priori* hypotheses in this domain, but given the broad sensitivity of phonological phenomena to their local segmental environment, we divided the compounds displaying variability based on whether the preceding segment was a vowel or a nasal consonant (the only licit coda that can precede /g/ in Japanese phonotactics). Indeed, we find a subtle effect (more evident in the model than in the plot in Figure 3) of a decrease in VVN

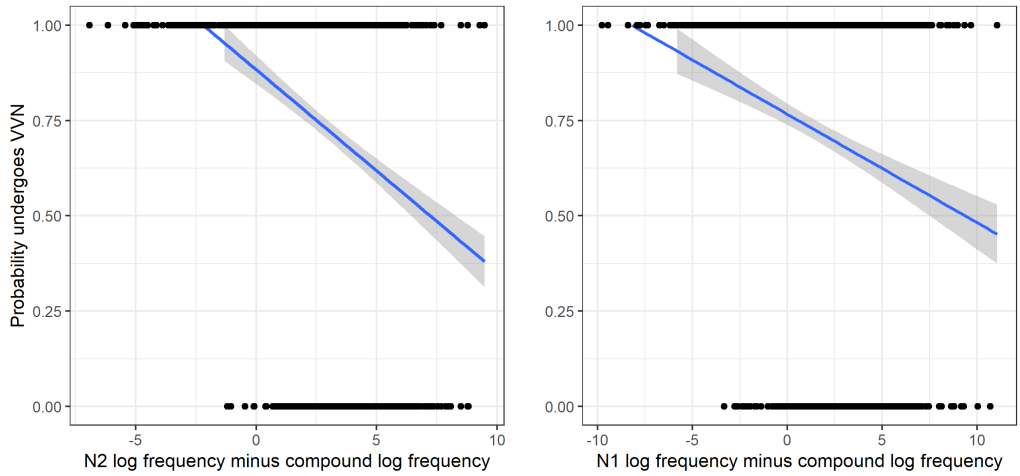

Figure 2: Probability of a compound undergoing VVN (vertical axis, both plots) based on the relative frequency of the compound and its N2 (horizontal axis, left plot) and the relative frequency of the compound and its N1 (horizontal axis, right plot). Each dot represents an individual compound; linear smooths are plotted purely for visualization purposes, using the ggplot2 function `geom_smooth()`; for statistical analyses, see Table 1.

application following vowel-final N1s relative to nasal-final N1s. This suggests that to some extent, VVN is influenced by processes of local assimilation, with nasal realizations of N2's initial segment being more likely adjacent to other nasal sounds.[2]

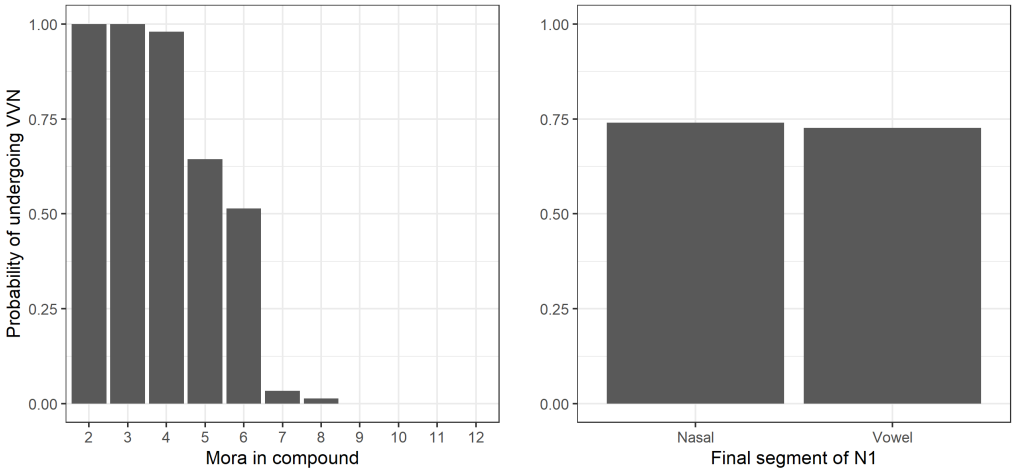

Figure 3: Probability of a given compound undergoing VVN (vertical axis, both plots) divided by its mora count (horizontal axis, left plot) and the final segment of N1 (horizontal axis, right plot).

[2]Though see Kawahara (2013) for evidence that geminate nasals are still somewhat marked in Japanese phonology.

4.4 Statistical analysis

We statistically verified the robustness of these findings using a Bayesian logistic regression model. Bayesian statistical models allow the user to directly estimate likely values for each parameter of interest (the posterior distribution) by integrating prior knowledge of the researcher (if any), and the information contained in the data to be analyzed. In this case, we used minimally-informative default priors, reflecting our prior assumption that there might not be a robust frequency effect, fit the model in R (R Core Team 2021) using the `brms` package (Bürkner et al. 2017). The dependent variable was VVN (not undergoing coded 0, undergoing coded 1), and the model contained fixed effects of relative N1 log frequency, relative N2 log frequency, mora count, (all centered, such that the intercept corresponds to average values of the predictor in the dataset), and preceding segment (nasal or non-nasal). We used the No U-Turn Sampler algorithm to draw 10,000 samples in each of four chains from the posterior distribution over credible values for coefficients associated with these parameters, of which we discarded the first 1,000 as warm-up. We can summarise these samples in terms of the median value, with 95% Credible Intervals (abbreviated "CrI", reported in square brackets), and can interpret these values directly in terms of degree of belief that the true effect lies within the specified interval. We can also assess the evidence in favor of *any* nonzero effect in the direction of the coefficient by looking at what percentage of the posterior distribution lies to one side of zero ($p(|\hat{\beta}|) > 0$); we report both measures in this paper. For more on Bayesian statistics for data-analysis in the context of linguistics, see Nicenboim and Vasishth (2016); for tutorial introductions to Bayesian data-analysis using linguistic data, see Vasishth et al. (2018) and Nalborczyk et al. (2019); for broader uses of Bayesian methods for data analysis, see Kruschke (2014). The results of this model are reported below in Table 1.

Parameter	*Median*	*95% CrI*	$p\,(\|\hat{\beta}\| > 0)$
Intercept: Relative N1 frequency (centered), Relative N2 frequency (centered), Compound mora length = *average values* Preceding segment = *nasal*	2.57	[1.9, 3.2]	
Relative N1 frequency *(scaled 1-unit increase)*	-0.5	[-0.8, -0.2]	0.99
Relative N2 frequency *(scaled 1-unit increase)*	-0.3	[-0.6, -0.0]	0.96
Compound mora length *(scaled 1-unit increase)*	-3.1	[-3.5, -2.7]	≈ 1
Preceding segment = *vowel*	-0.5	[-1.1, 0.1]	0.93

Table 1: Model of conditioning factors in compounds where N2 occurs free (right panel of Figure 1). Coefficients are in log-odds, with negative signs indicating an decrease in probability of a compound being listed as undergoing VVN, relative to the intercept.

4.5 Rendaku and VVN

Finally, a notable feature of VVN is that it stands in a feeding relationship with an independent voicing process, Rendaku (cf. Kawahara 2016: and the extensive literature cited therein). When N2 begins with /k/ which is rendered [g] by Rendaku, voiced velar nasalization is obligatory in the Corpus.

(3) a. /joko + kaki/ → [joko-ŋaki, *joko-gaki, *joko-kaki]
horizontal writing, "horizontal writing"

b. /kaki/ → [kaki], "writing"

As noted in previous studies of Rendaku, there is variation in Rendaku application: even when the N2 is eligible to undergo Rendaku, it sometimes does not ($\sim$ 18% of the time), though in general the tendency is to undergo ($\sim$ 82%).

The obligatoriness of VVN in cases where it is parasitic on Rendaku applying to /k/-initial forms is striking in contrast to the optionality and frequency-conditioning of VVN for non-derived /g/-initial forms. However, by appealing to the terms of Ito and Mester (1996)'s proposal, we can understand where the difference arises from: since the free base of N2 begins with [k] and not [g], faithfulness to a free form could not require that the bound form of N2 in compounds begin with [g].

The facts in (3) are also interesting because they constitute a *saltatory* phonological alternation (Hayes and White 2015): underlying /k/ is realized as [ŋ], skipping the phonetically intermediate [g], while underlying /g/ at least sometimes remains [g] out of a pressure to be faithful to a free N2. Although a recent body of work has demonstrated that saltatory phonological alternations are typologically rare (Hayes and White 2015, Smolek and Kapatsinski 2018) and are difficult for speakers to learn (White 2017, White and Sundara 2014, White 2014, Smolek and Kapatsinski 2018), the corpus data here suggest that — if representative of the synchronic knowledge of speakers of Japanese — the markedness of the alternation might be mitigated by the overwhelming pressure to maintain surface-oriented faithfulness in high-frequency /g/-initial N2s from the lexicon.

5 Discussion and conclusion

Our study is the first to quantitatively verify the observations of Ito and Mester (1996) about the distribution and conditioning of [g] and [ŋ] in compounds in standard Japanese. Beyond these findings, we find that the variation noted by Ito and Mester (1996) in compounds with a free N2 is conditioned by the frequency of both N2 and N1. We argue that the effect of N2 frequency stems from the functional basis of output-output faithfulness constraints being the relative activation or accessibility of its parts: N2s which are more frequent than the compound exert more of an effect than those less frequent; this analysis is elaborated on and formalized in Breiss et al. (2021). Further experimental work by Kawahara et al. (2021) supports this analysis with experimental evidence that the frequency-conditioned variation holds at the level of individual speakers, so the results reported here are unlikely to be an artefact arising from averaging of data from different, categorically-behaving speakers.

The effect of N1 frequency cannot be explained the same way, however — we speculate that as the frequency of the parts decreases relative to the compound as a whole, the compound is treated as less "compositional" and more as a monomorpheme, and thus subject to the markedness phonotactic motivating the allophonic distribution we find in compounds with bound N2s (cf. similar findings by Rebrus and Törkenczy 2017 in the behavior of Hungarian compounds with respect to vowel harmony). This hypothesis requires that we posit that morphological boundaries can be gradient (Hay and Baayen 2005, Kawahara and Tanaka 2020), which can be formalized using Gradient Symbolic Representation (Smolensky and Goldrick 2016); this finding is also broadly compatible with the dual-route storage model (Zuraw 2000, 2010, Zuraw and Peperkamp 2015, Zuraw et al. 2020), or Representational Strength Theory (Moore-Cantwell, *submitted*).

We also examine several non-frequency-based determinants of VVN application. There is a strong effect of prosodic length, such that longer compounds are less likely to undergo VVN than shorter ones, and there is a subtle effect of local segmental environment, with nasal-final N1s being more likely to have N2s that undergo VVN than N1s that are vowel-final. The current study also provides support for the thesis that the saltatory pattern is a marked phenomenon, but it can nevertheless be productive (Hayes and White 2015). Future work might fruitfully probe whether the corpus evidence suggesting the saltatory nature of the alternation is representative of speakers' synchronic knowledge, as judged by, for example, a *wug*-test (in fact, see Kawahara et al. 2021 for an initial attempt toward this goal). Finally, we note that there are conditioning factors on VVN that we cannot assess here: for example, Hibiya (1995) notes that the deployment of VVN can be used to signal a range of sociolinguistic factors, such as generational or class affiliation. Further,

it is unclear how the phonological data reported in the NHK pronouncing dictionary manifest phonetically, though see work by Sano and Ooigawa (2010) on this point.

It is our hope that this work can provide a useful summary of the quantitative trends of VVN in the Yamanote lexicon, and as a jumping-off point for future research on the topic.

References

Benua, Laura. 2000. *Phonological relations between words.* Psychology Press.

Breiss, Canaan. 2021a. Base effects are probabilistic: a case study in lexical conservatism. Talk at the 95th Annual Meeting of the LSA.

Breiss, Canaan, Hironori Katsuda, and Shigeto Kawahara. 2021. Paradigm uniformity is probabilistic: Evidence from velar nasalization in Japanese. In *Proceedings of WCCFL 39*. Cascadilla Press.

Breiss, Canaan Michael. 2021b. Lexical Conservatism in phonology: Theory, experiments, and computational modeling. Doctoral dissertation, University of California, Los Angeles.

Bürkner, Paul-Christian, et al. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software* 80:1–28.

Hay, Jennifer. 2003. *Causes and consequences of word structure.* Psychology Press.

Hay, Jennifer B, and R Harald Baayen. 2005. Shifting paradigms: Gradient structure in morphology. *Trends in cognitive sciences* 9:342–348.

Hayes, Bruce, and James White. 2015. Saltation and the P-Map. *Phonology* 267–302.

Hibiya, Junko. 1995. The velar nasal in Tokyo Japanese: A case of diffusion from above. *Language Variation and Change* 7:139–152.

Ito, Junko, and Armin Mester. 1996. Correspondence and compositionality: The ga-gyō variation in Japanese phonology .

Ito, Junko, and Armin Mester. 2003a. *Japanese morphophonemics: Markedness and word structure*, volume 41. MIT Press.

Ito, Junko, and Armin Mester. 2003b. Lexical and postlexical phonology in Optimality Theory: evidence from Japanese. In *Linguistische Berichte. Sonderheft 11: Resolving Conflicts in Grammars*, volume 11, 183–207.

Kawahara, Shigeto. 2013. Emphatic gemination in Japanese mimetic words: A wug-test with auditory stimuli. *Language sciences* 40:24–35.

Kawahara, Shigeto. 2016. Psycholinguistic studies of rendaku. *Perspectives on rendaku: Sequential voicing in Japanese compounds. Berlin: Mouton* .

Kawahara, Shigeto, Canaan Breiss, and Hironori Katsuda. 2021. Token frequency in the grammar: a wug-test of Japanese voiced velar nasalization. Talk at AMP 2021.

Kawahara, Shigeto, and Yu Tanaka. 2020. Partially activated morpheme boundaries in Japanese surnames. *Proceedings of NELS 52* .

Kindaichi, Haruhiko. 1942. Ga-gyo bionron. *Onin no Kenkyû (Reprinted in Kindaichi 1967)* .

Kruschke, John. 2014. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.* Academic Press.

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

Maekawa, Kikuo, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. 2014. Balanced corpus of contemporary written Japanese. *Language resources and evaluation* 48:345–371.

McCawley, James D. 1968. *The phonological component of a grammar of Japanese.* 2. Mouton.

Nalborczyk, Ladislas, Cédric Batailler, Hélène Lœvenbruck, Anne Vilain, and Paul-Christian Bürkner. 2019. An introduction to Bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard Indonesian. *Journal of Speech, Language, and Hearing Research* 62:1225–1242.

NHK. 1993. Nhk broadcasting culture research institute .

Nicenboim, Bruno, and Shravan Vasishth. 2016. Statistical methods for linguistic research: Foundational Ideas — Part II. *Language and Linguistics Compass* 10:591–613.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rebrus, Péter, and Miklós Törkenczy. 2017. Gradient harmonicity in compounds. In *Proceedings of the Annual Meetings on Phonology*, volume 4.

Sano, Shin-ichiro, and Tomohiko Ooigawa. 2010. Distribution and trichotomic realization of voiced velars in Japanese - An experimental study. In *Eleventh Annual Conference of the International Speech Communication Association.*

Smolek, Amy, and Vsevolod Kapatsinski. 2018. What happens to large changes? Saltation produces well-liked outputs that are hard to generate. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9.

Smolensky, Paul, and Matthew Goldrick. 2016. Gradient symbolic representations in grammar: The case of French liaison. *Rutgers Optimality Archive* 1552.

Steriade, Donca. 1997. Lexical conservatism. *Linguistics in the morning calm* 157–179.

Steriade, Donca, and Juliet Stanton. 2020. Productive pseudo-cyclicity and its significance. Talk at LabPhon 17. URL `https://julietstanton.github.io/files/steriade_stanton_labphon17.pdf`.

Trubetskoy, N. 1969. Principles of phonology [translated from the Russian by C. Baltaxe]. *Berkeley: U. of California Press* .

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

White, James. 2014. Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130:96–115.

White, James. 2017. Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-Map bias. *Language* 93:1–36.

White, James, and Megha Sundara. 2014. Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition* 133:85–90.

Zuraw, Kie. 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language & Linguistic Theory* 28:417–472.

Zuraw, Kie, Isabelle Lin, Meng Yang, and Sharaon Pepperkamp. 2020. Competition between whole-word and decomposed representations of English prefixed words. *Morphology* .

Zuraw, Kie, and Sharon Peperkamp. 2015. Aspiration and the gradient structure of English prefixed words. In *ICPhS*.

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