

# Geminate devoicing in Japanese loanwords: Theoretical and experimental investigations\*

## Abstract

This paper provides an overview of theoretical and experimental investigations into voiced geminates in Japanese. Active discussion was initiated by Nishimura's (2003) discovery that in Japanese loanword phonology, voiced geminates can be devoiced, when they co-occur with another voiced obstruent (e.g. /**doggu**/ → /**dokku**/ 'dog'). This context-sensitive devoicing of geminates has been analyzed within several different theoretical frameworks. The phonetic and psycholinguistic natures of voiced geminates have also been explored, in tandem with corpus-based analyses and computational modeling. This devoicing pattern of voiced geminates in Japanese therefore has had substantial impacts on the recent phonological literature. The empirical focus of this paper is on one simple devoicing phenomenon in Japanese, but implications for general linguistic theories are discussed throughout.

## 1 The basic generalizations

### 1.1 Prohibition against voiced geminates in native phonology

Japanese employs a singleton-geminate contrast to convey lexical differences (e.g. /*kata*/ 'frame' vs. /*katta*/ 'bought'), but not all kinds of geminates (=long consonants) are allowed.<sup>1</sup> In the native phonology of Japanese, neither approximant geminates (/rr, ww, jj/) nor voiced obstruent geminates (/bb, dd, gg, zz/) are allowed (Ito & Mester, 1999). It is the phonological behavior of voiced obstruent geminates that this paper focuses on (henceforth, simply "voiced geminates").

Not only do voiced geminates fail to make lexical contrasts, some evidence from phonological alternations shows that voiced geminates are actively avoided in the native phonology. For example, the adverb-forming suffix /-ri/ causes gemination of root-final consonants, as in (1). However,

---

\*Critical yet constructive comments from two anonymous reviewers were extremely helpful in preparing this article. Thanks to Yoonjung Kang for her editorial help as well as Hellen Stickney for careful reading of this draft. This paper is supported by JSPS Kakenhi grants #26770147 and #26284059. Remaining errors are mine.

<sup>1</sup>This paper uses abstract phonemic transcriptions for the sake of simplicity.

when the root-final consonant is a voiced obstruent, gemination is blocked, and a nasal is inserted instead, as in (2) (Ito & Mester, 1999).<sup>2</sup>

- (1) Gemination associated with /-ri/
  - a. /uka+ri/ → /ukkari/ ‘absent-mindedly’
  - b. /biku+ri/ → /bikkuri/ ‘surprised’
- (2) Gemination does not target voiced obstruents
  - a. /syobo+ri/ → /syombori/ ‘disappointed’
  - b. /uza+ri/ → /unzari/ ‘sick of something’

For other types of evidence of the avoidance of voiced geminates from phonological alternations, see Ito & Mester (1996) and Nasu (1999).

## 1.2 Voiced geminates in loanwords

Despite the lack of voiced geminates in the native phonology, voiced geminates do appear in some loanwords. Word-final consonants preceded by a lax vowel are often borrowed as geminates, as shown in (3) (Kubozono 2015 and many references cited therein). This adaptation process created voiced geminates in the loanword sector of the lexicon (Ito & Mester, 1999), as in (3c-d).

- (3) Gemination of word-final consonants in loanword adaptation
  - a. let → /retto/
  - b. pick → /pikku/
  - c. red → /reddo/
  - d. big → /biggu/

## 1.3 Devoicing of voiced geminates

Although a voicing contrast became contrastive in geminates in loanwords, as in (3), researchers observed that some voiced geminates can optionally be pronounced as devoiced (e.g. Ito & Mester 1999; Vance 1987). One big puzzle, however, was that not all voiced geminates seem to be devoicable. Ito & Mester (1999) proposed to treat devoicable geminates as contained in “assimilated foreign items” and non-devoiceable geminates as contained in “unassimilated foreign items”. This quasi-etymological distinction, however, was ad hoc and circular, because this distinction was not independently motivated.

---

<sup>2</sup>Voiced geminates occur in emphatic forms, even in the native phonology (e.g. /hiddoi/ ‘very awful’ from /hidoi/). This emphatic gemination is generally non-structure preserving in that it can create any kind of geminates (Kawahara, 2002).

In 2003, Nishimura identified a *phonological* condition which makes devoicing of geminates possible. Concretely, the devoicing of geminates occurs only when there is another voiced obstruent within the same morpheme, as in (4). In other words, devoicing of geminates is caused by a restriction against two voiced obstruents within the same morpheme. This condition can be understood as a version of a well-known general phonological constraint, Obligatory Contour Principle (OCP) (McCarthy, 1986). This restriction against two voiced obstruents—OCP(voice)—had also been known as Lyman’s Law in the native phonology of Japanese (Ito & Mester, 1986).

- (4) Optional devoicing of OCP-violating geminates: /d...dd/ → /d...tt/
- a. /**b**addo/ → /**b**atto/ ‘bad’
  - b. /**b**aggu/ → /**b**akku/ ‘bag’
  - c. /**d**oggu/ → /**d**okku/ ‘dog’

Nishimura (2003) contrasted OCP-violating geminates in (4) with non-OCP-violating geminates as in (5), and OCP-violating singletons as in (6). For these, devoicing seems impossible.

- (5) Non-OCP-violating geminates: /...dd/ → \*/...tt/
- a. /sunob**b**u/ → \*/sunoppu/ ‘snob’
  - b. /reddo/ → \*/retto/ ‘red’
  - c. /eggu/ → \*/ekku/ ‘egg’
- (6) OCP-violating singletons: /d...d/ → \*/d...t/
- a. /**g**ibu/ → \*/**g**ipu/ ‘give’
  - b. /**b**agu/ → \*/**b**aku/ ‘bug’
  - c. /**d**agu/ → \*/**d**aku/ ‘Doug’

The patterns illustrated in (4)-(6) initiated extensive theoretical debate, which is reviewed in section 2.

Before moving on, one remark is in order: we need to distinguish between loanword adaptation (=3)) and loanword phonology (=4)). The former refers to the phase in which Japanese speakers borrow these words from the source languages; the latter references what happens to these words after the adaptation. This distinction, which is sometimes neglected in the theoretical literature, is important, because Kaneko & Iverson (2009) showed that voiced geminates are not necessarily borrowed as voiceless in the presence of another voiced obstruent. Thus the devoicing of geminates in (4) occurs in loanword phonology rather than in loanword adaptation.

## 2 Phonological analyses of geminate devoicing

This section provides a critical overview of different theoretical analyses of the data set in (4)-(6) in a chronological order. As stated above, devoicing of OCP-violating geminates is optional, but the following analysis abstracts away from this optionality. See section 6 for more on the optionality.

### 2.1 A local conjunction analysis: Nishimura (2003)

The devoicing pattern instantiates a case of “a gang-effect” in that neither being a geminate nor violating the OCP(voice) alone suffices to cause devoicing; only the simultaneous violation of the two conditions (violating OCP and being a voiced geminate) results in devoicing. This patterning is challenging for Optimality Theory (OT: Prince & Smolensky 2004), since OT, in its standard form, does not predict this sort of effect, because of strict domination: a violation of a constraint that is ranked higher takes priority over any amount of violations of lower ranked constraints.

To illustrate the problem that patterns in (4)-(6) present to OT, let us consider three basic constraints, shown in (7):

- (7) Three constraints posited by Nishimura and subsequent work
- FAITH(VOICE): Devoicing is not allowed.
  - OCP: A morpheme cannot contain two voiced obstruents (=Lyman’s Law).
  - \*VOIOBSGEM: A voiced obstruent geminate is prohibited.

The first faithfulness constraint prohibits devoicing, which is necessary because a voicing contrast is contrastive in Japanese phonology in general. The second constraint, OCP, is theoretical instantiation of Lyman’s Law, which prohibits any morpheme containing two voiced obstruents. The final constraint, \*VOIOBSGEM, captures the prohibition against voiced geminates in the native phonology. Given these constraints, in the loanword phonology, the faithfulness constraint must dominate the two markedness constraints, as the tableaux in (8) and (9) show.

- (8) FAITH(VOICE)  $\gg$  OCP: no devoicing of singletons

/dagu/ ‘Doug’	FAITH(VOICE)	OCP	*VOIOBSGEM
→ /dagu/		*	
/daku/	*!		

- (9) FAITH(VOICE)  $\gg$  \*VOIOBSGEM: no devoicing of voiced geminates

/eggu/ ‘egg’	FAITH(VOICE)	OCP	*VOIOBSGEM
→ /eggu/			*
/ekku/	*!		

In other words, since devoicing does not occur either with OCP-violating singletons (= (6)) or non-OCP-violating geminates (= (5)), FAITH(VOICE) must be ranked at the top. However, this top-ranking of the faithfulness constraint blocks the devoicing of geminates, even when the geminates violate OCP as well, because of strict domination, as illustrated in (10).

(10) The top ranking of FAITH(VOICE) prevents devoicing of geminates

/doggu/ ‘dog’	FAITH(VOICE)	OCP	*VOIOBSGEM
→ /doggu/		*	*
/dokku/	*!		

To solve this problem, Nishimura (2003) proposed to deploy the mechanism of local conjunction (Smolensky 1993 *et seq.*). A locally-conjoined constraint consists of two sub-constraints, and is violated if and only if both of these sub-constraints are violated. By conjoining OCP and \*VOIOBSGEM within the domain of stem ( $=\{\text{OCP}\&\text{*VOIGEM}\}_{\text{stem}}$ ), and ranking the conjoined constraint above FAITH(VOICE), Nishimura (2003) obtained the right outcome, as in (11).<sup>3</sup>

(11) Gang-effect: The function of  $\{\text{OCP}\&\text{*VOIGEM}\}_{\text{stem}}$

/doggu/ ‘dog’	$\{\text{OCP}\&\text{*VOIGEM}\}_{\text{stem}}$	FAITH(VOICE)	OCP	*VOIOBSGEM
/doggu/	*!		*	*
→ /dokku/		*		

## 2.2 A split-faithfulness analysis: Kawahara (2006)

Kawahara (2006) argued that the local conjunction analysis of Nishimura (2003) is too powerful in the sense that two seemingly irrelevant constraints are conjoined in a domain as large as a stem (McCarthy, 2003; Padgett, 2002). If we allow local conjunction of two different constraints within a domain of a stem, then we predict the existence of a language that prohibits the co-occurrence of two totally irrelevant structures within a stem (say, a labial consonant and a voiced geminate), which is undesirable.

Instead, Kawahara (2006) proposed that FAITH(VOICE) should be split in such a way that singletons and geminates are subject to different FAITH(VOICE) constraints. Once we posit two faithfulness constraints, we can do away with the local conjunction constraint. In this analysis,  $\text{FAITH(VOICE)}_{\text{sing}}$  is ranked above OCP, which in turn dominates  $\text{FAITH(VOICE)}_{\text{gem}}$ . This ranking allows OCP to devoice geminates, but not singletons, as shown in (12)-(13):

(12)  $\text{FAITH}_{\text{sing}} \gg \text{OCP}$ : No devoicing of singletons

<sup>3</sup>As an anonymous reviewer pointed out, the domain of Lyman’s Law is better characterized as morphemes than stems, and it makes more sense to postulate morpheme as the domain of this local conjunction. Here, however, I follow Nishimura’s formulation of the constraint.

/b <del>ag</del> u/ ‘bug’	FAITH(VOICE) <sub>sing</sub>	OCP	FAITH(VOICE) <sub>gem</sub>
→ /b <del>a</del> gu/		*	
/b <del>a</del> ku /	*!		

(13) OCP  $\gg$  FAITH<sub>gem</sub>: Devoicing of geminates

/d <del>og</del> gu/ ‘dog’	FAITH(VOICE) <sub>sing</sub>	OCP	FAITH(VOICE) <sub>gem</sub>
/d <del>og</del> gu/		*!	
→ /d <del>o</del> kku/			*

Though ranked below OCP, FAITH<sub>gem</sub> dominates \*VOIOBSGEM to prevent context-free devoicing of geminates, as in (14):

(14) FAITH<sub>gem</sub>  $\gg$  \*VOIOBSGEM: No context-free devoicing of geminates

/e <del>gg</del> u/ ‘egg’	FAITH(VOICE) <sub>sing</sub>	OCP	FAITH(VOICE) <sub>gem</sub>	*VOIOBSGEM
→ /e <del>gg</del> u/				*
/e <del>k</del> ku/			*!	

This split-faithfulness approach thus can model the devoicing patterns without resorting to the complex locally-conjoined constraint. Furthermore, as discussed further in section 3, the splitting of faithfulness constraints can be—and perhaps should be—considered to be grounded in the perceptibility differences of voicing contrasts in singletons and geminates.

### 2.3 An approach based on the theory of contrast: Rice (2006)

Rice (2006), as a reply to Kawahara (2006), offered a different interpretation of why singletons and geminates behave differently with respect to OCP. As discussed in section 3, Kawahara (2006) derived the phonological difference between singletons and geminates from their phonetic differences; Rice (2006) suggested that something else is responsible.

Within the framework of the theory of contrast and markedness (Dresher, 2010), Rice (2006) attempted to derive a difference between singletons and geminates from the contrastiveness in the native phonology. A voicing difference is contrastive only in singletons in the native phonology (Ito & Mester, 1999); as a result, a [+voice] feature is projected only for singletons, not for geminates. Since voiced geminates do not have a [+voice] feature, they are more likely to be devoiced than singletons. This analysis thus shares the same spirit with Kawahara (2006) in that they both capitalized on the phonological “devoicability” difference between singletons and geminates—Kawahara (2006) tried to find its root in phonetics; Rice (2006) instead resorted to the contrastiveness in the native phonology.

One problem of this approach, however, is the fact that voicing *is* contrastive for geminates in the loanword phonology. Therefore, it is necessary for voiced geminates to have a [+voice] feature.

Moreover, this approach fails to explain why voiced geminates devoice only in response to OCP, not anywhere else—this approach predicts devoicing everywhere.

## 2.4 A Harmonic Grammar analysis: Pater (2009)

Pater (2009) presented a reanalysis of the devoicing phenomenon within the framework of Harmonic Grammar (HG), in which constraints are weighted instead of ranked. HG is similar to OT, but instead of ranked constraints, it uses a set of weighted constraints (Legendre et al. 1990 *et seq.*). Based on the weights assigned for each constraint, a harmonic score of each candidate is calculated as follows:  $H(cand_j) = \sum w_i * c_i(cand_j)$ , where  $w_i$  represents weight assigned to  $constraint_i$  and  $c_i(cand_j)$  violation profiles of a particular  $candidate_j$  with respect to  $constraint_i$ . In short, harmonic scores are the weighted sums of all constraint violations. The candidate with the highest harmonic score wins.

Pater (2009) used only the three basic constraints in (7). In this analysis, FAITH(VOICE) should have a higher weight than OCP and \*VOIOBSGEM; for example, the weight of FAITH(VOICE) can be set to 3, whereas those of OCP and FAITH can be set to 2 and 2. These weighting relations prevent devoicing of OCP-violating singletons and context-free devoicing of geminates, as in (15)-(16).

(15)  $w(\text{FAITH}) > w(\text{OCP})$ : no devoicing of OCP-violating singletons

/dagu/ ‘Doug’	FAITH (3)	OCP (2)	*VOIOBSGEM (2)	H-Score
→ dagu		-1		-2
daku	-1			-3

(16)  $w(\text{FAITH}) > w(*\text{VOIOBSGEM})$ : no context-free devoicing of geminates

/eggu/ ‘egg’	FAITH (3)	OCP (2)	*VOIOBSGEM (2)	H-Score
→ eggu			-1	-2
ekku	-1			-3

However, as long as the sum of the weight of OCP and that of \*VOIOBSGEM is higher than that of FAITH(VOICE) (e.g.  $2 + 2 = 4 > 3$ ), devoicing occurs to satisfy both OCP and \*VOIOBSGEM, as in (17). A gang-effect occurs because one violation of FAITH(VOICE) simultaneously satisfies the two lower-weighted markedness constraints.

(17) The gang-effect in HG

/doggu/ ‘dog’	FAITH (3)	OCP (2)	*VOIOBSGEM (2)	H-Score
doggu		-1	-1	$-2 + -2 = -4$
→ dokku	-1			-3

This analysis is appealing in that it analyzes the patterns of loanword devoicing using only the three basic constraints in (7), without additional theoretical machineries such as local conjunction or splitting of faithfulness constraints.<sup>4</sup>

### 3 Phonetics of voiced geminates

#### 3.1 Acoustics

The split-faithfulness analysis presented by Kawahara (2006) (reviewed in section 2.2) triggered interests in the phonetics of voiced geminates. The question raised in that work was why there are different faithfulness constraints for singletons and geminates. Descriptively speaking, Japanese speakers devoice only geminates, not singletons. To explain this observation, Kawahara (2006) used the P-Map theory (Steriade, 2008), which posits that a phonological change that causes a larger perceptual change is considered to be worse. In this view, for the case of Japanese, speakers neutralize a voicing contrast in geminates because it is not perceptually salient, whereas devoicing singleton is perceptually too conspicuous, as schematically illustrated in (18).

(18) The predicted perceptual map

$$\begin{array}{l} /dd/ \text{---} > /tt/ \\ /d/ \text{-----} > /t/ \end{array}$$

The specific prediction is thus that devoicing is perceptually less noticeable in geminates than in singletons. To test this prediction, Kawahara (2006) first conducted an acoustic study, which found that Japanese voiced geminates are semi-devoiced, as shown in Figure 1. For a singleton [d], voicing vibration continues throughout the closure—the continuation of voicing is observed both on the waveform as well as the voice bar, the low frequency energy observed at the bottom of the spectrogram. On the other hand, voicing in geminates is ceased at an early phase of closure for a geminate [dd] (shown with an arrow on the spectrogram).

This semi-devoicing of geminates has a well-known aerodynamic root (Hayes & Steriade, 2004; Ohala, 1983). In order to maintain the vibration of the glottis, the intraoral airpressure ( $P_o$ ) must be lower than the subglottal airpressure ( $P_s$ ). However,  $P_o$  automatically rises as the air goes into the oral cavity—to put it plainly, speakers cannot keep sending air into the oral cavity

---

<sup>4</sup>Coetzee & Pater (2011) briefly mentioned an analysis based on a MaxEnt grammar (Hayes & Wilson, 2008), which can also model the devoicing patterns with the three basic constraints. A MaxEnt grammar has two general properties: (i) it uses weighted, rather than ranked, constraints, so that it predicts gang-effects, just like the HG analysis illustrated here; (ii) it inherently predicts variation, and therefore explains the optional nature of OCP-driven devoicing. For details of a MaxEnt grammar, see Coetzee & Pater (2011) and Hayes & Wilson (2008). See also Tesar (2007), who raised various issues that arise from a model of grammar using weighted constraints rather than ranked constraints.



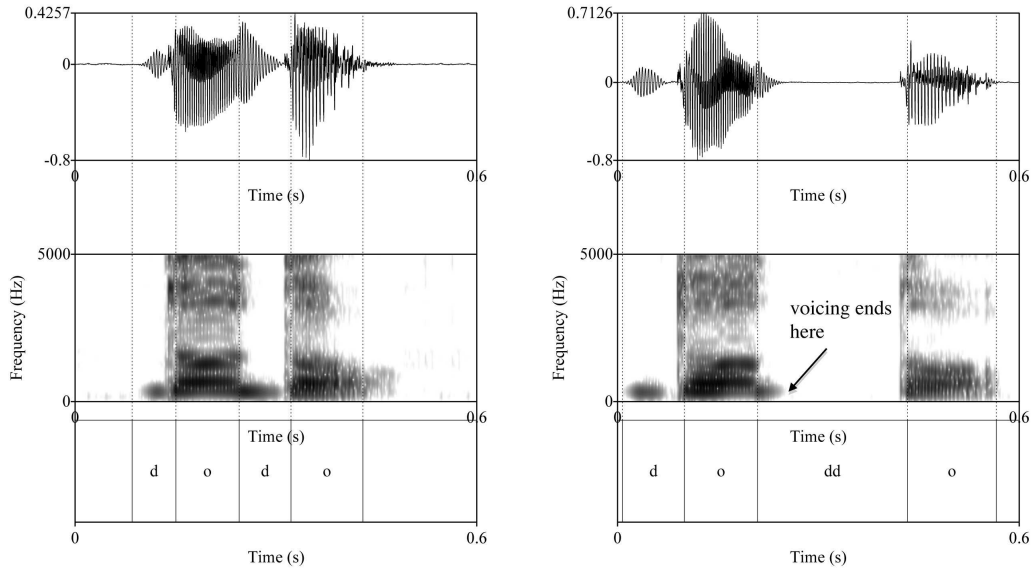


Figure 1: The left figure shows a singleton [d]; the right figure shows a geminate [dd]. Tokens based on Kawahara (2013). This figure is reproduced with the permission from de Gruyter.

when the mouth is closed. As a result, it becomes increasingly hard to maintain voicing during stop closure. Japanese speakers therefore give up on keeping voicing during geminate closure.<sup>5</sup>

The overall results of Kawahara (2006), based on the production of the three speakers, show almost 100% of closure voicing during singleton stops. On the other hand, voiced geminates show only 40% of closure voicing. Since voicing during closure is an important perceptual cue for voicing (Lisker, 1978; Ohala, 1981; Raphael, 1981), Kawahara (2006) hypothesized that voiced geminates are perceptually less clearly voiced than voiced singletons.

### 3.2 Perception

Kawahara (2006) also conducted a perception experiment to directly address the prediction by the P-map hypothesis (Steriade, 2008). The stimuli were covered by multi-layered cocktail party noise to avoid ceiling effects. Native speakers of Japanese judged the voicing quality of intervocalic consonants. To analyze the results,  $d'$ -values, which represent a perceptual distance for each type of voicing contrast, were calculated (Macmillan & Creelman, 2005), one for the singleton pair and one for the geminate pair. The result shows that the average  $d'$  is 3.79 for the singleton pair and .71 for the geminate pair. The perception experiment thus shows that a voicing contrast is less perceptible in geminates than in singletons, as predicted by the P-map theory, illustrated in (18).

To summarize, the phonological observation is that a voicing contrast is more likely to neutral-

<sup>5</sup>Kawahara (2006) stimulated phonetic research on voiced geminates in non-standard dialects of Japanese, which shows that some of Japanese dialects show full voicing during geminates (Matsuura, 2012; Takada, 2013).

ize in geminates than in singletons, and perceptually, the contrast is less perceptible in geminates. Taken together, there is a correlation between phonetic perceptibility and phonological devoicability: the smaller the perceptual change that a phonological change causes, the more likely it occurs. This correlation is exactly what the P-map theory predicts. Kawahara (2006) argues therefore that the Japanese OCP-driven devoicing pattern is phonetically natural in the sense that the threshold of devoicability is determined by perceptibility. Overall, this analysis was taken to be evidence that phonology is non-trivially affected by phonetic factors, such as semi-devoicing due to the aerodynamic difficulty of voiced obstruents and the perceptibility of voicing contrasts.

One complication is that, while the devoicability difference between singletons and geminates may be phonetically natural, the cause of devoicing, OCP(voice), may not be phonetically natural (Kawahara, 2008). OCP(voice), or more descriptively speaking, dissimilation in voicing, is cross-linguistically rare, and always historically arose from dissimilation of other contrasts, such as aspiration or prenasalization (Ohala, 1981, 1993). These observations are accounted for under Ohala's theory of dissimilation, in which dissimilation arises from misperception of a phonological contrast whose phonetic cues are spread out over several segments. Dissimilation in voicing is unexpected—or unnatural—from this perspective, because cues for voicing contrasts are localized, and not spread-out (see Ohala 1981, 1993 and Kawahara 2008 for discussion).

Dissimilation in voicing thus has only arisen historically from dissimilation in other features, prenasalization in the case of Japanese (Vance, 2005). To the extent that dissimilation in voicing does not make phonetic sense, it means that the trigger of the devoicing of geminates in Japanese is phonetically unnatural. Based on these arguments, Kawahara (2008) advanced a view that phonetic naturalness and unnaturalness can coexist within a single phonological system, a view that is further explored by Hayes et al. (2009) (see also references cited therein).

## **4 Psycholinguistics: Judgment experiments**

All the theoretical work after Nishimura (2003) took it for granted that the data in (4)-(6) were correct. However, the examples were based on the intuitions of Nishimura (2003) and Kawahara (2006), the authors of the papers themselves. Kawahara (2011b) raised a concern about this methodology—using data based on the intuitions of the authors themselves.

The issue of the quality of intuition-based data has recently been much discussed especially in syntax (Schütze, 2011), but also in phonology (see Kawahara 2011b for an overview). To briefly summarize the potential concerns, first, some “phonological patterns” have been shown to be non-productive with experiments using nonce words (Ohala, 1974; Sanders, 2003; Vance, 1987). Second, it is questionable whether the data based on an author's intuition can be generalized

to the general population of Japanese speakers.<sup>6</sup> Third, the inner sensations of Nishimura (2003) and Kawahara (2006) cannot be observed from outside, hence cannot be replicated objectively. Fourth, linguists may unconsciously oversimplify a pattern when they report data based on their own intuition (Watanabe, 2009). Finally, it is not clear whether Nishimura (2003) and Kawahara (2006) were completely unbiased when they provided the data (Gibson & Fedorenko, 2010). These concerns are important to address, because generative theories are often developed using intuition-based data, and if such data are not reliable, then the theory would lose its empirical foundation.

To address these concerns, a series of judgement studies have been run using naive native speakers (Kawahara, 2011a,b, 2013). To take Kawahara (2011a) as an example for illustration, the experiment asked the participants to rate the naturalness of devoicing in four contexts: (1) OCP-violating geminates (/d...dd/), (2) non-OCP-violating geminates (/...dd/), (3) OCP-violating singletons (/d...d/), and (4) non-OCP-violating singletons (/...d/). The participants were given one form (e.g. /**doggu**/) and the other variant form with devoicing (i.e. /**dokku**/) and were asked how natural that second form is as a pronunciation of the first form. The experiment was thus a naturalness judgment experiment on a phonological process. The study used a 5-point naturalness scale from “very natural” to “very unnatural”.

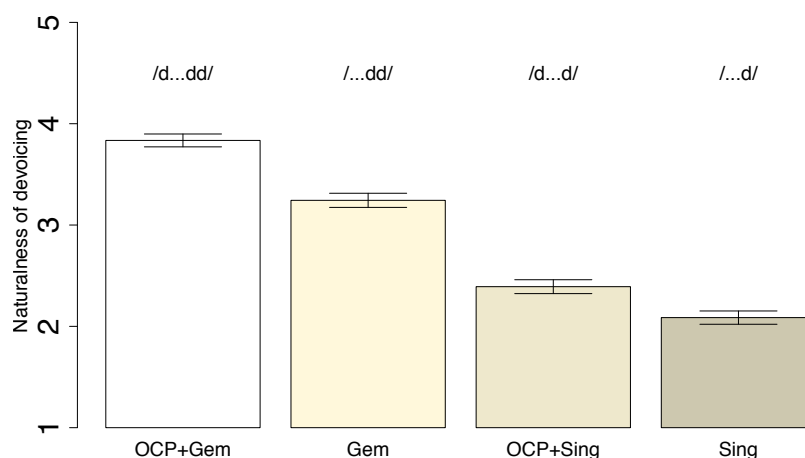


Figure 2: The average naturalness ratings of the devoicing of the four conditions (Kawahara, 2011a). Reprinted with permission from Springer. The error bars are 95% confidence intervals.

Figure 2 shows that results of the rating study by Kawahara (2011a). Japanese speakers found

<sup>6</sup>For generative linguistics, it may not perhaps matter whether a pattern generalizes to the whole speech community or not, as long as there is a single individual that instantiates a particular pattern, because every individual grammar must have arisen from Universal Grammar. While this logic itself seems sound, I find it worrisome when the whole theory is built on data from a single individual (cf. Kelkar’s Hindi: Kelkar 1968, cited and discussed by Hayes 1995, pp.276-278 and Prince & Smolensky 2004, pp.47-50, especially footnote 22).

the devoicing of OCP-violating geminates most natural (the leftmost bar), which shows that the intuitions provided by Nishimura (2003) and Kawahara (2006) were not ungrounded. However, the story was not as simple as that. First, the Japanese speakers found the devoicing of non-OCP-violating geminates more natural than that of OCP-violating singletons (2nd vs. 3rd bar). Second, OCP made devoicing of singletons more natural too (3rd vs. 4th bar), although devoicing singleton consonants were judged to be unnatural overall. Importantly, there was no clear line that divides the continuum into two categories, “grammatical devoicing” and “ungrammatical devoicing”, contra what Nishimura (2003) and Kawahara (2006) claimed. This non-dichotomous distinction among the four conditions is observed even when the participants used a binary yes/no response format in a follow-up experiment (Kawahara, 2013).

It thus turned out that native speakers’ judgment patterns are more gradient than the “grammatical” vs. “ungrammatical” dichotomy, contra the common assumption in generative grammar; the classic example is that *brick* and *blick* are both grammatical, whereas *bnick* is not, but no further distinctions are posited. However, this beyond-binary distinction in judgment pattern is in fact well-attested cross-linguistically in phonotactic judgment patterns (see Pierrehumbert 2001 for an overview). The results of the judgment studies revised in this section show that gradient patterns hold for judgment patterns of a phonological process as well (i.e. when native speakers judge the naturalness of devoicing in several environments). These experiments thus lend support to the view that linguistic knowledge cannot be modeled as a matter of the grammatical vs. ungrammatical dichotomy.

In addition, Kawahara (2011a) found that various linguistic factors other than OCP and geminacy impact the naturalness of devoicing. For example, the presence of multiple triggers (e.g. /*deibiddo*/ ‘David’) made devoicing more natural. Second, speakers’s ratings were lower when devoicing resulted in the merger of two lexical items; e.g. devoicing /*baggu*/ ‘bag’ would be homophonous with /*bakku*/ ‘back’. Third, speakers rated the devoicing of more frequent items as more natural, as shown in Figure 3. This aspect of devoicing is more fully addressed in section 6.

All of these results show that the characterization of the devoicing described by Nishimura (2003) and Kawahara (2006) involved oversimplification (see also Watanabe 2009), which in turn highlights the importance of experimentation in phonological research.

## 5 Corpus studies

All the judgment experiments show that Japanese speakers judge OCP-violating geminates to be the most natural. The results of the judgment experiments thus lend some credibility to the intuition-based data presented in Nishimura (2003) and Kawahara (2006). Nevertheless, a question still remained because speakers’ intuition and their actual speech behavior do not always

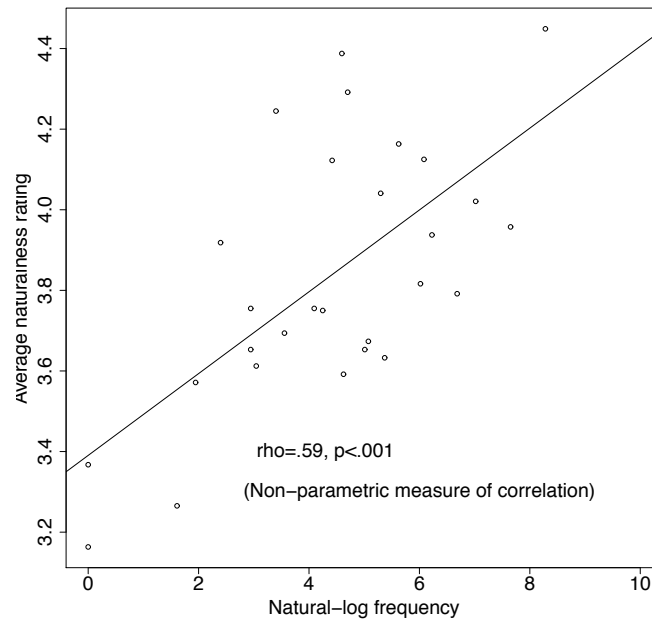


Figure 3: The correlation between average naturalness ratings (y-axis) and lexical frequency (x-axis) (Kawahara, 2011a). Reprinted with permission from Springer.

match (Labov, 1996)—what do native speakers of Japanese actually do? Kawahara & Sano (2013) and Sano & Kawahara (2013) took up this issue by exploring the behavior of voiced geminates in actual utterances.

These studies used the Corpus of Spontaneous Japanese (the CSJ) (NINJAL, 2008). This database is a large database of spoken Japanese, which comes with a rich annotation system. The annotation provides both underlying forms and surface forms, which allows us to assess whether voiced geminates are devoiced or not. The corpus studies confirmed that OCP-violating geminates appear more often as devoiced (about 40%) than do non-OCP violating geminates (about 5%). Kawahara & Sano (2013) also (more or less) confirmed the frequency effect found by Kawahara (2011a), shown in Figure 3.

Kawahara & Sano (2013) also found an effect of place of articulation on the devoicability of geminates as well: the further back the place, the more likely the geminates are to devoice. This patterning is in accordance with the well-known aerodynamic difficulty hierarchy of voiced stops (Hayes & Steriade, 2004; Ohala, 1983). The further back the place, the smaller and less flexible the oral cavity behind the constriction is, the harder it is to maintain voicing.

In addition to these grammatical factors, Sano & Kawahara (2013) found that non-grammatical factors impact the likelihood of devoicing. For example, female speakers were found to devoice geminates more often than male speakers. Other non-grammatical factors that were found to im-

pact the devoicability of geminates include age (younger speakers devoice more), speech style (devoicing is more likely in informal speech), education level (people with higher education devoice less), and others. See also Watanabe (2009) for a related observation.

To summarize, the corpus analyses revealed that several factors affect the devoicability of geminates, both grammar-internal and grammar-external. As with the grammaticality judgment experiments summarized in section 4, the corpus-based studies show that devoicing of voiced geminates is not as monolithic as it was once thought to be. This practice also raises an alarming message about intuition-based data; the original description reported by Nishimura (2003) and Kawahara (2006) was very much oversimplified. Idealization is necessary in linguistic theorization, but this research shows that idealization is merely a starting point (Riemer, 2009).

## **6 Modeling: Lexical frequency effects on phonology**

Finally, Coetzee & Kawahara (2013) proposed a model which makes one step toward incorporating such complications of actual phonological patterns into linguistic theorization. Recall that there is a correlation between devoicing and lexical frequency, both in the judgment patterns as well as in the patterning in the corpus. This correlation is an example of an old observation; the probability of an optional process actually applying correlates with the lexical frequency of that particular item (Bybee, 2006). However, generative models were not good at dealing with this observation; phonological theories have often set aside this observation, sometimes in the name of idealization, or sometimes by relegating it to a matter of performance.

Coetzee & Kawahara (2013) proposed a noisy Harmonic Grammar model in which the weights of faithfulness constraints are scaled for each lexical item based on its lexical frequency. Details aside, this system assigns higher weights to non-frequent items and lower weights to frequent items (where the precise values were determined based on  $\beta$ -distributions). As a result, more frequent items are more likely to undergo phonological processes. Their modeling shows a significantly better fit with actual data once frequency effects are incorporated into grammar.

This proposal shows that generative grammatical models can incorporate the effect of lexical frequencies on phonological patterns. It demonstrates that maybe generative phonology is now at a point where we can broaden our empirical coverage, without relying too much on idealization or relegating the frequency effects as a matter of performance.

## **7 A remaining challenge: /p/ causes devoicing**


One remaining challenge is the behavior of singleton /p/. All of the previous studies assumed that it is only voiced obstruents that trigger devoicing of geminates. However, recent studies point out

that /p/ can trigger devoicing as well. Some examples are shown in (19).

(19) /p/-driven devoicing

- a. /kyuup**iddo**/ → /kyuup**itto**/ ‘cupid’
- b. /**p**iram**iddo**/ → /**p**iram**itto**/ ‘pyramid’
- c. /a**i**p**addo**/ → /a**i**p**atto**/ ‘i-pad’

Kawahara & Sano (2014) showed that singleton /p/ indeed causes devoicing of geminates, based on a corpus study and judgment experiments. The challenge is that none of the theoretical analyses reviewed in section 2 predict this /p/-driven devoicing, because all of the analyses, in some way or another, assume that the trigger of devoicing is OCP(voice), but the co-occurrence of /p/ and voiced geminates should not violate OCP(voice). To solve this problem, Fukazawa et al. (2015) point out that neither singleton /p/’s nor voiced geminates are allowed in the native phonology (Ito & Mester, 1999) and that as a result these segments are the two most infrequent sounds in the whole Japanese lexicon. They argue that Japanese speakers disfavor the co-occurrences of two unfamiliar sounds within a morpheme.

Kawahara (2015) entertained an alternative analysis of this /p/-driven devoicing based on Japanese orthography: voiced obstruents and /p/ are shown with diacritic marks, the former with *dakuten* and the latter with *han-dakuten* [INSERT THE COMMENT HERE]  therefore /p/, voiced obstruents, and voiced geminates are all written with an orthographic diacritic. OCP(voice) may then actually be OCP(diacritic), which accounts for both devoicing driven by /p/ and devoicing driven by a voiced obstruent. This analysis is radical in that it (partly) shifts the burden of explanation from sounds to letters. Although Kawahara (2015) ultimately argued against this orthography-based explanation based on several evidence regarding the behavior of *rendaku*, Kawahara (2015) also suggested that phonologists may need to pay more attention to extra-grammatical factors like orthography.

## 8 Conclusion

This paper has reviewed how the phonology of voiced geminates in Japanese loanwords has been analyzed from different perspectives. This review has shown that we can take one phonological phenomenon and tackle it from various perspectives: theoretical, experimental, corpus-based, and computational. These approaches can reveal how phonology interacts with other factors (phonetics, lexical, and sociolinguistic), having ramifications in phonological theorization as well as in related fields. This research program has raised various important issues for phonological theorization in general; e.g., how to model constraint interaction, how the phonetics-phonology interface works, how the intuition-based data should be complemented with quantitative studies, and how

orthography may or may not affect phonological knowledge, etc. This research therefore highlights the importance of studying one pattern in depth from different perspectives.

## References

- Bybee, Joan (2006) *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- Coetzee, Andries W. & Shigeto Kawahara (2013) Frequency biases in phonological variation. *Natural Language and Linguistic Theory* **30**(1): 47–89.
- Coetzee, Andries W. & Joe Pater (2011) The place of variation in phonological theory. In *The Handbook of Phonological Theory, 2nd Edition*, John A. Goldsmith, Jason Riggle, & Alan Yu, eds., Oxford: Blackwell-Wiley, 401–431.
- Dresher, B. Elan (2010) *The Contrastive Hierarchy in phonology*. Cambridge: Cambridge University Press.
- Fukazawa, Haruka, Shigeto Kawahara, Mafuyu Kitahara, & Shin-ichiro Sano (2015) Two is too much: [p]-driven geminate devoicing in Japanese. *On-in Kenkyu [Phonological Studies]* **18**.
- Gibson, Edward & Evelina Fedorenko (2010) Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* **14**(6): 233–234.
- Hayes, Bruce (1995) *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.
- Hayes, Bruce & Donca Steriade (2004) Introduction: The phonetic bases of phonological markedness. In *Phonetically Based Phonology*, Bruce Hayes, Robert Kirchner, & Donca Steriade, eds., Cambridge: Cambridge University Press, 1–33.
- Hayes, Bruce & Colin Wilson (2008) A Maximum Entropy model of phonotactics and phonotactic learning. *Linguistic inquiry* **39**: 379–440.
- Hayes, Bruce, Kie Zuraw, Péter Siptár, & Zsuzsa Londe (2009) Natural and unnatural constraints in Hungarian vowel harmony. *Language* **85**(4): 822–863.
- 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 (1996) Stem and word in Sino-Japanese. In *Phonological Structure and Language Processing: Cross-Linguistic Studies*, T. Otake & A. Cutler, eds., Berlin: Mouton de Gruyter, 13–44.
- Ito, Junko & Armin Mester (1999) The phonological lexicon. In *The Handbook of Japanese Linguistics*, Natsuko Tsujimura, ed., Oxford: Blackwell, 62–100.
- Kaneko, Emiko & Gregory Iverson (2009) Phonetic and other factors in Japanese on-line adaptation of English final consonants. In *Studies in Language Sciences 8: Papers from the eighth annual conference of the Japanese Society for Language Science*, Shunji Inagaki & Makiko Hirakawa, eds., Tokyo: Kuroshio Publications, 179–195.
- Kawahara, Shigeto (2002) Similarity among Variants: Output-Variant Correspondence, BA thesis, International Christian University, Tokyo Japan.
- Kawahara, Shigeto (2006) A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* **82**(3): 536–574.
- Kawahara, Shigeto (2008) Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* **17**(4): 317–330.



- Kawahara, Shigeto (2011a) Aspects of Japanese loanword devoicing. *Journal of East Asian Linguistics* **20**(2): 169–194.
- Kawahara, Shigeto (2011b) Japanese loanword devoicing revisited: A rating study. *Natural Language and Linguistic Theory* **29**(3): 705–723.
- 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*.
- Kawahara, Shigeto & Shin-ichiro Sano (2014) /p/-driven geminate devoicing in Japanese: Corpus and experimental evidence. Ms. Keio University.
- Kawahara, Shigeto & Shinichiro Sano (2013) A corpus-based study of geminate devoicing in Japanese: Linguistic factors. *Language Sciences* **40**: 300–307.
- Kelkar, A.R. (1968) *Studies in Hindi-Urdu: Introduction and word phonology*. Poona: Deccan College.
- Kubozono, Haruo (2015) Loanword phonology. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton.
- Labov, William (1996) When intuitions fail. In *Proceedings of Chicago Linguistic Society 32: Papers from the parasession on theory and data in linguistics*, Lisa McNair, Kora Singer, Lise Dolbrin, & Michelle Aucon, eds., Chicago: Chicago Linguistics Society, 77–106.
- Legendre, Géraldine, Yoshiro Miyata, & Paul Smolensky (1990) Harmonic Grammar – A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, Mahwah, NJ: Lawrence Erlbaum Associates, 388–395.
- Lisker, Leigh (1978) On buzzing the English /b/. *Haskins Laboratories Status Report on Speech Research* **SR-55/56**: 251–259.
- Macmillan, Neil & Douglas Creelman (2005) *Detection Theory: A User's Guide. 2nd Edition*. Mahwah: Lawrence Erlbaum Associates Publishers.
- Matsuura, Toshio (2012) Yuusei sogai jyuushion-no onsei jitsugen-niokeru chiikisa-ni kansuru yobitekibunseki [A preliminary analysis on regional variation in phonetic realization of voiced obstruent geminates]. Talk presented at the Phonetic Society of Japan, Sept 29th.
- McCarthy, John J. (1986) OCP Effects: Geminata and antigeminata. *Linguistic Inquiry* **17**: 207–263.
- McCarthy, John J. (2003) Comparative markedness. *Theoretical Linguistics* **29**: 1–51.
- Nasu, Akio (1999) Chouhukukei onomatope no kyouchou keitai to yuhyousei [Emphatic forms of reduplicative mimetics and markedness]. *Nihongo/Nihon Bunka Kenkyuu [Japan/Japanese Culture]* **9**: 13–25.
- NINJAL (2008) Nihongo hanashikotoba koopasu [The Corpus of Spontaneous Japanese], Ver. 2. Database.
- Nishimura, Kohei (2003) *Lyman's Law in loanwords*. MA thesis, Nagoya University.
- Ohala, John J. (1974) Experimental historical phonology. In *Historical Linguistics II: Theory and Description in Phonology. Proceedings of the First International Linguistic Conference on Historical Linguistics*, J. M. Naderson & Charles Jones, eds., New York: Elsevier, 353–389.
- Ohala, John J. (1981) The listener as a source of sound change. In *Proceedings of Chicago Linguistic Society 17*, T. Myers, J. Laver, & Anderson J., eds., Chicago: Chicago Linguistic Society, 178–203.

- Ohala, John J. (1983) The origin of sound patterns in vocal tract constraints. In *The Production of Speech*, Peter MacNeilage, ed., New York: Springer-Verlag, 189–216.
- Ohala, John J. (1993) The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*, C. Jones, ed., London: Longman Academic, 237–278.
- Padgett, Jaye (2002) Constraint conjunction versus grounded constraint subhierarchies in Optimality Theory. Ms. University of California, Santa Cruz.
- Pater, Joe (2009) Weighted constraints in generative linguistics. *Cognitive Science* **33**: 999–1035.
- Pierrehumbert, Janet B. (2001) Stochastic phonology. *GLOT* **5**: 1–13.
- Prince, Alan & Paul Smolensky (2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden and Oxford: Blackwell.
- Raphael, Lawrence (1981) Duration and contexts as cues to word-final cognate opposition in English. *Phonetica* **38**: 126–47.
- Rice, Keren (2006) On the patterning of voiced stops in loanwords in Japanese. *Toronto Working Papers in Linguistics* **26**: 11–22.
- Riemer, Nicholas (2009) Grammaticality as evidence and as prediction in a Galilean linguistics. *Language Sciences* **31**: 612–633.
- Sanders, Nathan (2003) *Opacity and Sound Change in the Polish Lexicon*. Doctoral dissertation, University of California, Santa Cruz.
- Sano, Shinichiro & Shigeto Kawahara (2013) A corpus-based study of geminate devoicing in Japanese: The role of the OCP and external factors. *Gengo Kenkyu [Journal of the Linguistic Society of Japan]* **144**: 103–118.
- Schütze, Carlson (2011) Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Reviews: Cognitive Science* **2**(2): 206–211.
- Smolensky, Paul (1993) *Optimality, markedness, and underspecification*. Paper presented at the Rutgers University Optimality Workshop, New Brunswick, NJ.
- Steriade, Donca (2008) The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The nature of the word*, Kristin Hanson & Sharon Inkelas, eds., Cambridge: MIT Press, 151–179.
- Takada, Mieko (2013) Regional differences in sound patterns during the closure of Japanese voiced geminates. Paper presented at 3rd ICPP at NINJAL, Japan.
- Tesar, Bruce (2007) A comparison of lexicographic and linear numeric optimization using violation difference ratios. Ms. Rutgers University.
- Vance, Timothy (1987) *An Introduction to Japanese Phonology*. New York: SUNY Press.
- Vance, Timothy (2005) Sequential voicing and Lyman’s Law in Old Japanese. In *Polymorphous Linguistics: Jim McCawley’s Legacy.*, Salikoko S. Mufwene, Elaine J. Francis, & Rebecca S. Wheeler, eds., Cambridge: MIT Press, 27–43.
- Watanabe, Seiji (2009) *Cultural and educational contributions to recent phonological changes in Japanese*. Doctoral dissertation, University of Arizona.