Papers on geminate devoicing in Japanese

September, 2015

Shigeto Kawahara

The Keio Institute of Cultural and Linguistic Studies

Preface

This volume “Papers on geminate devoicing in Japanese” is a collection of papers on geminate devoicing found in the loanword phonology of Japanese. Some of the papers are co-authored with my colleagues; detailed publication information is found at the beginning of each paper.

As I mention in some of the papers contained here, the first person who found this pattern is Kohei Nishimura. I believe that he presented a talk on this devoicing pattern in 2001 at Phonological Association of Kansai (PAIK). In 2003, I finished a manuscript of what later became Kawahara (2006), published in Language, and that’s the same year when Nishimura wrote an MA thesis on this phenomenon, his first official publication on this pattern. Although Kawahara (2006) became the standard citation of this phenomenon, Nishimura (2003) should be given as much credit.

Why put these papers in one place? The reason is that I have done a wide range of analyses on geminate devoicing from different perspectives (both theoretical and experimental), and it is convenient to have all the papers in one place (I hope). The first paper contained in this volume “Geminate Devoicing in Japanese Loanwords: Theoretical and Experimental Investigations” should provide a decent guide about what each paper contained in this volume is about.

I also wrote an overview paper in Japanese:

Kawahara, Shigeto (2012) 外来語有声促音の無声化-理論的貢献-. 音韻研究 15: 93-104.

which is not included in this volume. If you’re interested, please find the paper on my website, although the content may be a bit outdated.

Shigeto Kawahara
September, 2015

Kawahara, Shigeto (2015) Geminate devoicing in Japanese loanwords: Theoretical and experimental investigations. Language and Linguistic Compass. 9(4): 168–182.

Language and Linguistics Compass 9/4 (2015): 168–182, 10.1111/lnc3.12130

Geminate Devoicing in Japanese Loanwords: Theoretical and Experimental Investigations[1]

Shigeto Kawahara*
Keio University, The Keio Institute of Cultural and Linguistic Studies Tokyo, Japan

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. /**d**o**gg**u/ → /do**kk**u/ '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 the paper.

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., /ka**t**a/ 'frame' vs. /ka**tt**a/ 'bought'), but not all kinds of geminates (=long consonants) are allowed.[2] In the native phonology of Japanese, neither approximant geminates (/rr, ww, jj/) nor voiced obstruent geminates (/bb, dd, gg, zz/) are allowed (Ito and 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, when the root-final consonant is a voiced obstruent, gemination is blocked, and a nasal is inserted instead, as in (2) (Ito and Mester 1999).[3]

(1) Gemination associated with /-ri/

a. /u**k**a+ri/ → /u**kk**ari/ 'absent-mindedly'
b. /bi**k**u+ri/ → /bi**kk**uri/ 'surprised'

(2) Gemination does not target voiced obstruents

a. /syo**b**o+ri/ → /syo**mb**ori/ 'disappointed'
b. /u**z**a+ri/ → /u**nz**ari/ 'sick of something'

For other types of evidence of the avoidance of voiced geminates from phonological alternations, see Ito and 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 and Mester 1999), as in (3c–d).

(3) Gemination of word-final consonants in loanword adaptation

a. le**t** → /re**tt**o/
b. pi**ck** → /pi**kk**u/
c. re**d** → /re**dd**o/
d. bi**g** → /bi**gg**u/

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 and Mester 1999; Vance 1987). One big puzzle, however, was that not all voiced geminates seem to be devoicable. Ito and Mester (1999) proposed to treat devoicable geminates as contained in 'assimilated foreign items' and non-devoicable geminates as contained in 'unassimilated foreign items'. This quasi-etymological distinction, however, was ad hoc and circular, because this distinction was not independently motivated.

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 and Mester 1986).

(4) Optional devoicing of OCP-violating geminates: /d…dd/ → /d…tt/

a. /**b**a**dd**o/ → /**b**a**tt**o/ 'bad'
b. /**b**a**gg**u/ → /**b**a**kk**u/ 'bag'
c. /**d**o**gg**u/ → /**d**o**kk**u/ '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. /suno**bb**u/ → */suno**pp**u/ 'snob'
b. /re**dd**o/ → */re**tt**o/ 'red'
c. /e**gg**u/ → */e**kk**u/ 'egg'

(6) OCP-violating singletons: /d…d/ → */d…t/

a. /**gib**u/ → */**gip**u/ 'give'
b. /**b**a**g**u/ → */**b**a**k**u/ 'bug'
c. /**d**a**g**u/ → */**d**a**k**u/ '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 and 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:

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 and 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

a. FAITH(VOICE): Devoicing is not allowed.
b. OCP: A morpheme cannot contain two voiced obstruents (=Lyman'sLaw).
c. *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) shows.

(8) FAITH(VOICE)≫ OCP: no devoicing of singletons

/**d**agu/ 'Doug'	**FAITH(VOICE)**	**OCP**	***VOIOBSGEM**
→ /**d**a**g**u/		*	
/**d**a**k**u/	*!		

(9) FAITH(VOICE)≫ *VOIOBSGEM: no devoicing of voiced geminates

/**egg**u/ 'egg'	**FAITH(VOICE)**	**OCP**	*VOIOBSGEM
→ /e**gg**u/			*
/e**kk**u/	*!		

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.

/**d**o**gg**u/ 'dog'	**FAITH(VOICE)**	**OCP**	* VOIOBSGEM
→ /**d**o**gg**u/		*	*
/**d**o**kk**u/	*!		

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 (={OCP&*VOIOBSGEM}$_{stem}$), and ranking the conjoined constraint above FAITH(VOICE), Nishimura (2003) obtained the right outcome, as in (11).[4]

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

/**d**o**gg**u/ 'dog'	**{OCP&*VOIOBSGEM}$_{stem}$**	**FAITH(VOICE)**	**OCP**	***VOIOBSGEM**
/**d**o**gg**u/	*!		*	*
→ /**d**o**kk**u/		*		

2.2. A SPLIT-FAITHFULNESS ANALYSIS:

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, FAITH(VOICE)$_{sing}$ is ranked above OCP, which in turn dominates FAITH(VOICE)$_{gem}$. This ranking allows OCP to devoice geminates, but not singletons, as shown in (12)–(13):

(12) FAITH(VOICE)$_{sing}$ ≫ OCP: No devoicing of singletons

/**b**a**g**u/ 'bug'	FAITH(VOICE)$_{sing}$	OCP	FAITH(VOICE)$_{gem}$
→ /**b**a**g**u/		*	
/**b**a**k**u/	*!		

(13) OCP ≫ FAITH(VOICE)$_{gem}$: Devoicing of geminates

/**d**o**gg**u/ 'dog'	FAITH(VOICE)$_{sing}$	OCP	FAITH(VOICE)$_{gem}$
/**d**o**gg**u/		*!	
→ /**d**o**kk**u/			*

Though ranked below OCP, FAITH(VOICE)$_{gem}$ dominates *VOIOBSGEM to prevent context-free devoicing of geminates, as in (14):

(14) FAITH(VOICE)$_{gem}$ ≫ *VOIOBSGEM: No context-free devoicing of geminates

/e**gg**u/ 'egg'	FAITH(VOICE)$_{sing}$	OCP	FAITH(VOICE)$_{gem}$	*VOIOBSGEM
→ /e**gg**u/				*
/e**kk**u/			*!	

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), 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 (see 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 and 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 capitalize 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) 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: $H(cand_j)=\Sigma w_i{*}c_i(cand_j)$, 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 *VOIOBSGEM 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*(FAITH(VOICE)) > *w*(OCP): no devoicing of OCP-violating singletons

/**d**a**g**u/ 'Doug'	FAITH(VOICE) (3)	OCP (2)	*VOIOBSGEM (2)	H-Score
→ **d**a**g**u		−1		−2
da**k**u	−1			−3

(16) *w*(FAITH(VOICE)) > *w*(*VOIOBSGEM): no context-free devoicing of geminates

/e**gg**u/ 'egg'	FAITH(VOICE) (3)	OCP (2)	*VOIOBSGEM (2)	H-Score
→ e**gg**u			−1	−2
e**kk**u	−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 shown 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

/**d**o**gg**u/ 'dog'	FAITH(VOICE) (3)	OCP (2)	*VOIOBSGEM (2)	H-Score
do**gg**u		−1	−1	−2+−2 = −4
→ **d**o**kk**u	−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.[5]

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

/dd/—>/tt/
/d/————————>/t/

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], glottal 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 and 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 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.[6]

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.

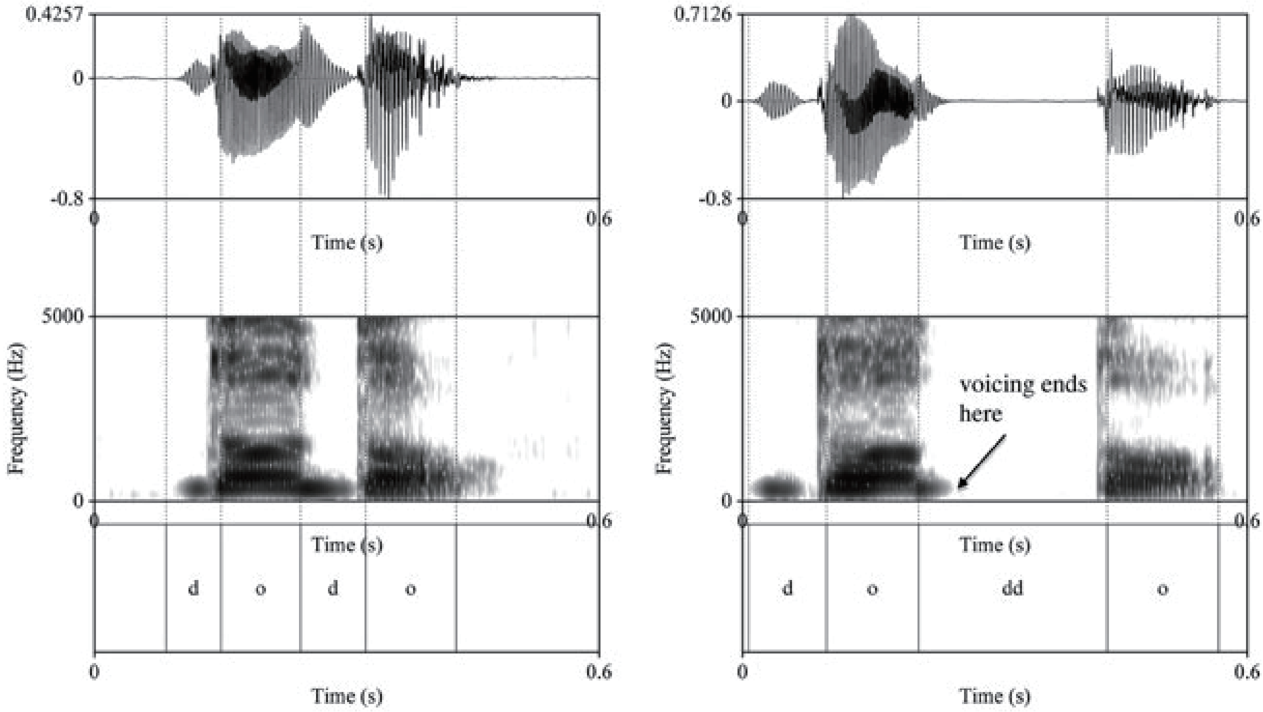


Fig. 1. The left figure shows a singleton [d]; the right figure shows a geminate [dd]. Tokens based on Kawahara (2013). This figure is reproduced with the permission from de Gruyter.

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 and Creelman 2005), one for the singleton pair and one for the geminate pair. The result shows that the average *d'*-value 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 neutralize 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.[7] 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 and 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 judgment studies have been run using naive native speakers (Kawahara 2011a, 2011b, 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. /**d**o**gg**u/) and the other variant form with devoicing (i.e. /**d**o**kk**u/) 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 shows the results of the rating study by Kawahara (2011a). Japanese speakers found 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

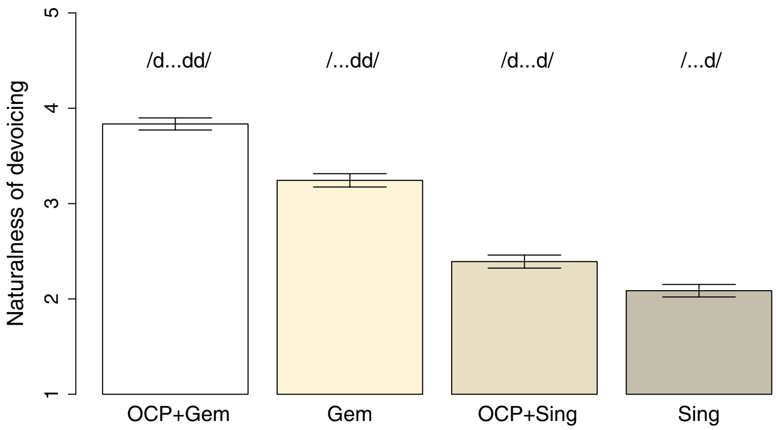


Fig. 2. The average naturalness ratings of the devoicing of the four conditions (Kawahara 2011a). Reprinted with permission from Springer. The error bars are 95% confidence intervals.

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 reviewed 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. /**d**ei**b**i**dd**o/ 'David') made devoicing more natural. Second, speakers' ratings were lower when devoicing resulted in the merger of two lexical items; e.g., devoicing /ba**gg**u/ 'bag' would be homophonous with /ba**kk**u/ 'back', and such patterns were rated as less natural. 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 match (Labov 1996) – what do native speakers of Japanese actually do? Kawahara and Sano (2013) and Sano and 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.

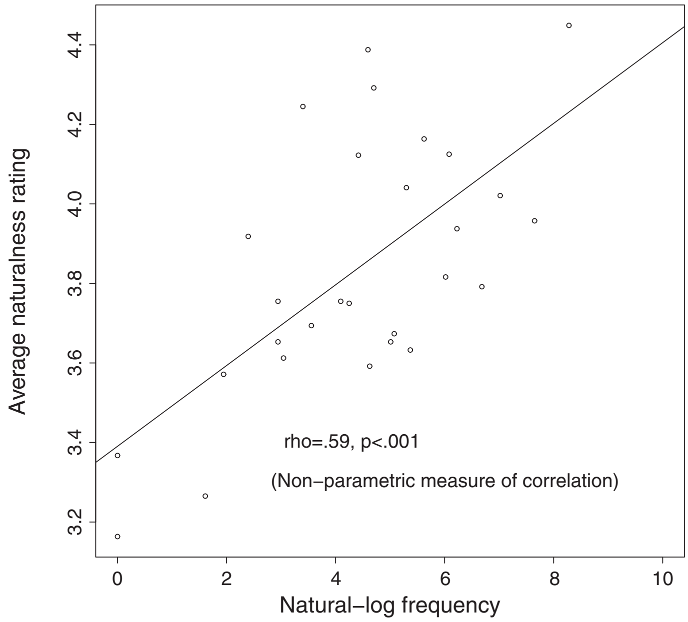


Fig. 3. The correlation between average naturalness ratings (y-axis) and lexical frequency (x-axis) (Kawahara 2011a). Reprinted with permission from Springer.

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 and Sano (2013) also (more or less) confirmed the frequency effect found by Kawahara (2011a), shown in Figure 3.

Kawahara and 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 and 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 and 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 impact 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 and 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 (see 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 and 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 β-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 of Geminates*

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 geminate devoicing

a. /kyuu**piddo**/ → /kyuu**pitt**o/ 'cupid'
b. /**p**irami**dd**o/ → /**p**irami**tt**o/ 'pyramid'
c. /ai**p**a**dd**o/ → /ai**p**a**tt**o/ 'i-pad'

Kawahara and 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 and 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* (e.g., ば=[ba], ぱ=[pa], は=[ha]). 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 argues against this orthography-based explanation based on several pieces of evidence regarding the behavior of rendaku, Kawahara (2015) also suggests 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.

Short Biography

Shigeto Kawahara is currently a Senior Assistant Professor at the Keio Institute of Cultural and Linguistic Studies. Before he moved to Keio University, he taught at University of Georgia, Rutgers University, and as a visiting professor, at International Christian University. He has worked both on phonetics and phonology, as well as their interface, and has published articles in many academic journals, including Language, Natural Language and Linguistic Theory, Journal of Phonetics, and others.

Notes

* Correspondence address: Prof. Shigeto Kawahara, Keio University, The Keio Institute of Cultural and Linguistic Studies, 2-15-45 Minatoku, Mita, Tokyo, 108-8345, Japan. E-mail: kawahara@icl.keio.ac.jp

[1] 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 Helen Stickney for careful reading of this draft. This paper is supported by JSPS Kakenhi grants #26770147 and #26284059. Remaining errors are mine.

[2] This paper uses abstract phonemic transcriptions for the sake of simplicity.

[3] Voiced geminates occur in emphatic forms, even in the native phonology (e.g., /hi**dd**oi/ 'very awful' from /hi**d**oi/). This emphatic gemination is generally non-structure preserving in that it can create any kind of geminates (Kawahara, 2002).

[4] 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.

[5] Coetzee and Pater (2011) briefly mention an analysis based on a MaxEnt grammar (Hayes and 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 and Pater (2011) and Hayes and Wilson (2008). See also Tesar (2007), who raised various issues that arise from a model of grammar using weighted constraints rather than ranked constraints.

[6] 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).

[7] 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 and Smolensky 2004, pp. 47–50, especially footnote 22).

Works Cited

Bybee, Joan. 2006. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Coetzee, Andries W., and Shigeto Kawahara. 2013. Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 30(1). 47–89.

Coetzee, Andries W., and Joe Pater. 2011. The place of variation in phonological theory. *The handbook of phonological theory*, 2nd edn., ed. by J. A. Goldsmith, J. Riggle, and A. Yu, 401–31. Oxford: Blackwell-Wiley.
Dresher, B. Elan. 2010. *The contrastive hierarchy in phonology*. Cambridge: Cambrdige University Press.
Fukazawa, Haruka, Shigeto Kawahara, Mafuyu Kitahara, and Shin-ichiro Sano. 2015. Two is too much: [p]-driven geminate devoicing in Japanese. *On-in Kenkyu [Phonological Studies]* 18. 3–10.
Gibson, Edward, and Evelina Fedorenko. 2010. Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* 14(6). 233–4.
Hayes, Bruce. 1995. *Metrical stress theory: principles and case studies*. Chicago: The University of Chicago Press.
Hayes, Bruce, and Donca Steriade. 2004. Introduction: the phonetic bases of phonological markedness. *Phonetically based phonology*, ed. by B. Hayes, R. Kirchner, and D. Steriade, 1–33. Cambridge: Cambridge University Press.
Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440.
Hayes, Bruce, Kie Zuraw, Péter Siptár, and Zsuzsa Londe. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85(4). 822–63.
Ito, Junko, and Armin Mester. 1986. The phonology of voicing in Japanese: theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17. 49–73.
——. 1996. Stem and word in Sino-Japanese. *Phonological structure and language processing: cross-linguistic studies*, ed. by T. Otake, and A. Cutler, 13–44. Berlin: Mouton de Gruyter.
——. 1999. The phonological lexicon. The handbook of Japanese linguistics, ed. by N. Tsujimura, 62–100. Oxford: Blackwell.
Kaneko, Emiko, and Gregory Iverson. 2009. Phonetic and other factors in Japanese on-line adaptation of English final consonants. *Studies in language sciences 8: papers from the eighth annual conference of the Japanese society for language science*, ed. by S. Inagaki, and M. Hirakawa, 179–95. Tokyo: Kuroshio Publications.
Kawahara, Shigeto. 2002. Similarity among variants: output-variant correspondence. BA thesis, International Christian University, Tokyo Japan.
——. 2006. A faithfulness ranking projected from a perceptibility scale: the case of [+voice] in Japanese. *Language* 82(3). 536–74.
——. 2008. Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* 17(4). 317–30.
——. 2011a. Aspects of Japanese loanword devoicing. *Journal of East Asian Linguistics* 20(2). 169–94.
——. 2011b. Japanese loanword devoicing revisited: a rating study. *Natural Language and Linguistic Theory* 29(3). 705–23.
——. 2013. Testing Japanese loanword devoicing: addressing task effects. *Linguistics* 51(6). 1271–99.
——. 2015. Can we use rendaku for phonological argumentation? *Linguistic Vanguard*.
Kawahara, Shigeto, and Shin-ichiro Sano. 2014. /p/-driven geminate devoicing in Japanese: corpus and experimental evidence. Ms. Keio University.
Kawahara, Shigeto, and Shinichiro Sano. 2013. A corpus-based study of geminate devoicing in Japanese: linguistic factors. *Language Sciences* 40: 300–7.
Kelkar, A.R. 1968. *Studies in Hindi-Urdu: introduction and word phonology*. Poona: Deccan College.
Kubozono, Haruo. 2015. Loanword phonology. *The handbook of Japanese language and linguistics: phonetics and phonology*, ed. by Haruo Kubozono. Berlin: Mouton de Gruyter. 313–361.
Labov, William. 1996. When intuitions fail. *Proceedings of Chicago Linguistic Society 32: Papers from the Parasession on Theory and Data in Linguistics*, ed. by L. McNair, K. Singer, L. Dolbrin, and M. Aucon, 77–106. Chicago: Chicago Linguistics Society.
Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: theoretical foundations. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 388–95. Mahwah, NJ: Lawrence Erlbaum Associates.
Lisker, Leigh. 1978. On buzzing the English /b/. *Haskins Laboratories Status Report on Speech Research* SR-55/56. 251–9.
Macmillan, Neil, and Douglas Creelman. 2005. *Detection theory: a User's guide*, 2nd edn., 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: gemination and antigemination. *Linguistic Inquiry* 17. 207–63.
——. 2003. Comparative markedness. *Theoretical Linguistics* 29. 1–51.
Nasu, Akio. 1999. Chouhukukei onomatope no kyouchou keitai to yuuhyousei [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. *Historical linguistics II: theory and description in phonology. Proceedings of the first international linguistic conference on historical linguistics*, ed. by J. M. Naderson and C. Jones, 353–89. New York: Elsevier.
——. 1981. The listener as a source of sound change. *Proceedings of Chicago linguistic society 17*, ed. by T. Myers, J. Laver, and J. Anderson, 178–203. Chicago: Chicago Linguistic Society.
——. 1983. The origin of sound patterns in vocal tract constraints. The production of speech, ed. by P. MacNeilage, 189–216. New York: Springer-Verlag.
——. 1993. The phonetics of sound change. Historical linguistics: problems and perspectives, ed. by C. Jones, 237–78. London: Longman Academic.
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, and 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. 612–33.
Sanders, Nathan. 2003. Opacity and sound change in the Polish lexicon. Doctoral dissertation, University of California, Santa Cruz.
Sano, Shinichiro, and 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. 103–18.
Schütze, Carlson. 2011. Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Reviews: Cognitive Science* 2(2). 206–11.
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. *The nature of the word*, ed. by K. Hanson, and S. Inkelas, 151–79. Cambridge: MIT Press.
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.
——. 2005. Sequential voicing and Lyman's Law in old Japanese. Polymorphous linguistics: Jim McCawley's legacy, ed. by S. S. Mufwene, E. J. Francis, and R. S. Wheeler, 27–43. Cambridge: MIT Press.
Watanabe, Seiji. 2009. Cultural and educational contributions to recent phonological changes in Japanese. Doctoral dissertation, University of Arizona.

Kawahara, Shigeto (2006) A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. Language 82(3): 536-574.

A FAITHFULNESS RANKING PROJECTED FROM A PERCEPTIBILITY SCALE: THE CASE OF [+VOICE] IN JAPANESE

SHIGETO KAWAHARA

University of Massachusetts, Amherst

Within the framework of optimality theory (Prince & Smolensky 2004), Steriade (2001a,b) proposes the P-map hypothesis, whose fundamental tenet is that the rankings of faithfulness constraints are grounded in perceptual-similarity rankings. This article provides empirical support for this hypothesis. In Japanese loanword phonology, a voiced geminate, but not a singleton, devoices to dissimilate from another voiced obstruent within a single stem. Based on this observation, I argue that the [+voice] feature is protected by two different faithfulness constraints, IDENT(+voi)$_{Sing}$ and IDENT(+voi)$_{Gem}$, and they are ranked as IDENT(+voi)$_{Sing}$ » IDENT(+voi)$_{Gem}$ in Japanese. I further argue that this ranking is grounded in the relative perceptibility of [+voice] in singletons and geminates, and this claim is experimentally supported. The general theoretical implication is that phonetic perceptibility can directly influence patterns in a phonological grammar.*

1. INTRODUCTION. The degree to which phonetics can affect phonology has been an oft-discussed topic in phonological theory. Much work in phonology has reported recurrent phonological patterns that are motivated by phonetics. The idea that phonology is at least partly driven by phonetics has many antecedents in the literature, including Jakobson 1941, Chomsky & Halle 1968:Ch. 9, Stampe 1973 (natural phonology), Hooper 1976 (natural generative phonology), and Archangeli & Pulleyblank 1994 (grounded phonology), among many others. With the advent of OPTIMALITY THEORY (OT; Prince & Smolensky 2004), this question has received renewed attention, as OT provides a novel way to express phonetic naturalness directly in the grammar (see e.g. contributions in Hayes et al. 2004, Myers 1997).

One of the ways in which phonology can be affected by phonetics has to do with perceptibility. In particular, a number of recent proposals have argued that phonological distinctions are prone to neutralization in a position where their cues are not saliently perceived (Boersma 1998, Côté 2004, Guion 1998, Hura et al. 1992, Jun 2004, Kohler 1990, Padgett 2002, Steriade 1995, 1997, Zhang 2000, among many others). Place distinctions, for instance, are often neutralized in codas, correlating with the fact that perceptual cues for those distinctions are not salient there (Benkí 2003, Fujimura et al. 1978, Jun 2004).

Building on these observations, Steriade (2001a,b) proposes the P-MAP HYPOTHESIS within the framework of OT. The P-map is 'the repository of speakers' knowledge, rooted in observation and inference, that certain contrasts are more discriminable than others' (Steriade 2001a:236). From this knowledge of similarity, a faithfulness constraint ranking is projected: among alternations that involve different degrees of perceptibility changes, the more perceptible the change an alternation involves, the higher-ranked the faithfulness constraint it violates. For example, a voicing contrast is more saliently perceived prevocalically than preconsonantally—that is, the contrast between

* Many thanks to Leah Bateman, Ben Gelbart, Becca Johnson, Ian Maddieson, Caren Rotello, Taka Shinya, the participants of HUMDRUM 2004, and the participants of MIT Phonology Circle for discussing several aspects of this article. I am especially grateful to José Benkí, Kathryn Flack, Brian Joseph, John Kingston, John McCarthy, Joe Pater, Donca Steriade, and an anonymous referee for their extensive comments on earlier versions. All errors are mine.

[pa] and [ba] is more salient and perceptible than the contrast between [apta] and [abta] (Steriade 1997). Based on this knowledge of similarity, speakers project the faithfulness ranking $\text{FAITH(voi)}_{/\_V}$ » $\text{FAITH(voi)}_{/\_C}$. As a consequence, a voicing contrast is more prone to neutralization in preconsonantal position than in prevocalic position, because $\text{FAITH(voi)}_{/\_C}$ is ranked low. Put in theory-neutral terms, the gist of the P-map hypothesis is that an alternation that involves a less perceptible change is more likely to occur than an alternation that involves a more perceptible change.

The primary aim of this article is to provide empirical support for the P-map hypothesis. A devoicing phenomenon in the loanword phonology of Japanese shows that voiced geminates are more prone to devoicing than voiced singletons are. As exemplified by the data in 1, when voiced geminates occur in a word with another voiced obstruent, they undergo optional devoicing (Haraguchi 2006, Nishimura 2003). By contrast, voiced singletons do not devoice even when they cooccur with another voiced obstruent, as seen in 2.

(1) Optional devoicing of voiced geminates

gebberusu	~	gepperusu	'Göbbels'
beddo	~	betto	'bed'
baggu	~	bakku	'bag'

(2) Singletons do not devoice

bagii	*pagii	*bakii	'buggy'
dagu	*tagu	*daku	'Doug'
gibu	*kibu	*gipu	'give'

I argue that a satisfactory account of this asymmetry between singletons and geminates requires that faithfulness constraints for voicing (i.e. FAITH(voi)) be differentiated into two constraints, one for singletons ($\text{FAITH(voi)}_{\text{Sing}}$) and one for geminates ($\text{FAITH(voi)}_{\text{Gem}}$), and that $\text{FAITH(voi)}_{\text{Sing}}$ be ranked higher than $\text{FAITH(voi)}_{\text{Gem}}$. The P-map hypothesis predicts that Japanese speakers have this ranking because a voicing contrast is more perceptible in singletons than in geminates. Just as preconsonantal voicing is more prone to neutralization because of its lower salience, voicing in geminates can be neutralized because its cues are not saliently perceived. I report on two experiments which show that this prediction of the P-map hypothesis is indeed borne out. This article thus overall provides empirical endorsement for the P-map's central claim. A larger implication is that neutralization patterns in phonology are closely tied to phonetic perceptibility.

This discussion unfolds as follows. I first present a description of the phonological patterns of the [+voice] feature in the loanword phonology of Japanese and follow that with a formal phonological analysis within the framework of OT. I show that $\text{FAITH(voice)}_{\text{Gem}}$ is ranked lower than $\text{FAITH(voice)}_{\text{Sing}}$. The P-map hypothesis predicts that this ranking holds because a [+voice] feature is less perceptible in geminates than in singletons. I next turn to experimental evidence that supports this claim, and finally I integrate the results of the phonological analysis and phonetic experiments and offer empirical predictions that emerge from this study.

2. [+voice] IN JAPANESE LOANWORD PHONOLOGY.

2.1. THE PHONOLOGICAL DATA. Although voiced geminates are prohibited in the native vocabulary of Japanese (Itô & Mester 1995, 1999, Kuroda 1965), they are allowed in recent loanwords.[1] In recent Japanese loanwords from foreign languages (mainly

[1] I assume, following Itô and Mester (1995, 1999), that there is stratification of foreign and native vocabulary in the Japanese lexicon. Native speakers of Japanese readily differentiate between native and foreign

English), coda consonants that follow a lax vowel in the source language are often borrowed as geminates (Katayama 1998). Also, since loanwords enter Japanese through written materials more frequently than through spoken language (Lovins 1973, Miura 1993, Smith 2006), consonants spelled with two letters are often borrowed as geminates (e.g. *slugger* is borrowed as [suraggaa]). This gemination process has created voiced geminates in the loanword phonology of Japanese, and a [±voice] distinction is thus contrastive in geminates. Some near-minimal pairs of voiced and voiceless geminates are given in 3.[2] Throughout this article, those words that contain voiced geminates but no other voiced obstruents are schematically referred to as TVDDV words.

(3) TVDDV words and their near-minimal pairs

webbu	'web'	wippu	'whipped (cream)'
sunobbu	'snob'	sutoppu	'stop'
habburu	'Hubble'	kappuru	'couple'
kiddo	'kid'	kitto	'kit'
reddo	'red'	autoretto	'outlet'
heddo	'head'	metto	'helmet'
suraggaa	'slugger'	surakkaa	'slacker'
eggu	'egg'	tʃekku	'check'
furaggu	'flag'	furakku	'Flack (proper name)'

In the TVDDV words shown in 3, devoicing of a voiced geminate is impossible, suggesting that a voicing contrast is phonemic in geminates. However, Haraguchi (2006) and Nishimura (2003) have pointed out that when voiced obstruent geminates appear with another voiced obstruent they can undergo optional devoicing. Some illustrative data are given in 4. These words, which contain voiced geminates and additional voiced obstruents, are referred to as DVDDV words in the subsequent discussion.

(4) DVDDV words: voiced geminates may devoice when they appear with another voiced obstruent

gebberusu	~	gepperusu	'Göbbels'
guddo	~	gutto	'good'
beddo	~	betto	'bed'
doreddo	~	doretto	'dreadlocks'
deddobooru	~	dettobooru	'dead ball (baseball term)'
baddo	~	batto	'bad'
deibiddo	~	deibitto	'David'
budda	~	butta	'Buddha'
doggu	~	dokku	'dog'
baggu	~	bakku	'bag'
doraggu	~	dorakku	'drug'
biggu	~	bikku	'big'

words, as reflected by the fact that they use different orthographic systems for these two lexical classes. See also Moreton & Amano 1999 and Gelbart & Kawahara 2005 for evidence from perceptual experiments that supports the psychological reality of lexical stratification in Japanese.

[2] The Japanese data reported in this article were collected by the author with the help of native speaker informants. [b] tends to resist gemination, as in [nobu], *[nobbu], 'knob', so data that include [bb] are rare (Katayama 1998).

Nishimura (2003) supports the productivity of this devoicing phenomenon through a corpus study using a database compiled by the National Institute for Japanese Language, Communication Research Laboratory and Tokyo Institute of Technology. This database contains approximately eighty-six hours of spoken Japanese, including both formal and spontaneous speech. In the spontaneous speech transcribed in the corpus, out of fifty-four DVDDV words, which canonically have voiced geminates, thirty-four appear with voiceless geminates (57.4%). By contrast, only one out of twenty-seven TVDDV words (3.7%) appears with a voiceless geminate. Furthermore, in Nishimura's (2003) web-based search using Google (http://www.google.co.jp), he found that DVDDV words are also frequently transcribed with voiceless geminates (86,670 out of 448,192 tokens: 19.3%), whereas TVDDV words very rarely are (2,187 out of 408,225 tokens: 0.5%).[3]

In addition to these arguments put forth by Nishimura, an informal survey of four Japanese speakers confirmed that devoicing of the geminates in DVDDV words is acceptable, while devoicing of TVDDV words is not. One speaker commented that she in fact more commonly pronounces the DVDDV words with voiceless geminates than with voiced geminates.

By contrast, devoicing is impossible when there are two singleton voiced obstruents in a word. This is illustrated by the words in 5 (henceforth DVDV words), in which neither of the two voiced singleton consonants can be devoiced. Japanese speakers clearly reject the pronunciation of the DVDV words with a devoiced singleton consonant.

(5) DVDV words: words with two singletons do not undergo devoicing

bagii	'buggy'	bogii	'bogey'
bobu	'Bob'	bagu	'bug'
dagu	'Doug'	daibu	'dive'
daiyamondo	'diamond'	doguma	'dogma'
giga	'giga (10^9)'	gaburieru	'Gabriel'
gibu	'give'	gaidansu	'guidance'

The phonology of [+voice] obstruents in Japanese loanwords is summarized in Table 1.[4]

		POSSIBILITY OF DEVOICING	EXAMPLES
TVDDV WORDS	one voiced geminate	impossible	[eggu] *[ekku] [webbu] *[weppu]
DVDV WORDS	two voiced singletons	impossible	[dagu] *[daku], *[tagu] [giga] *[kiga], *[gika]
DVDDV WORDS	one voiced singleton and one geminate	possible	[doggu] ~ [dokku] [beddo] ~ [betto]

TABLE 1. Summary of the phonology of [+voice] obstruents in Japanese loanwords.

2.2. EXPERIMENT 1: IS DEVOICING CATEGORICAL?[5] Before developing an analysis of the patterns summarized in Table 1, I address the question of whether the de-

[3] Transcription in spelling might not be a completely reliable indicator of the actual pronunciations. See §2.2 for acoustic evidence that supports the devoicing pattern described here.

[4] There are no words that contain two voiced geminates because of an independent condition that bans two geminates within a single word (Itô & Mester 2003:49–51, Spaelti 1997, Tsuchida 1995).

[5] Thanks to José Benkí for raising the question addressed in this subsection. I am also grateful to Michael Kenstowicz for his suggestions about how this question could be tested.

voicing that occurs in DVDDV words is a categorical phonological phenomenon or a gradient phonetic process. This issue is important to address here, because voiced geminates undergo context-free phonetic devoicing almost obligatorily (see below). This raises the question of whether the devoicing in DVDDV words is simply a reflex of such context-free phonetic devoicing, rather than a categorical phonological phenomenon.

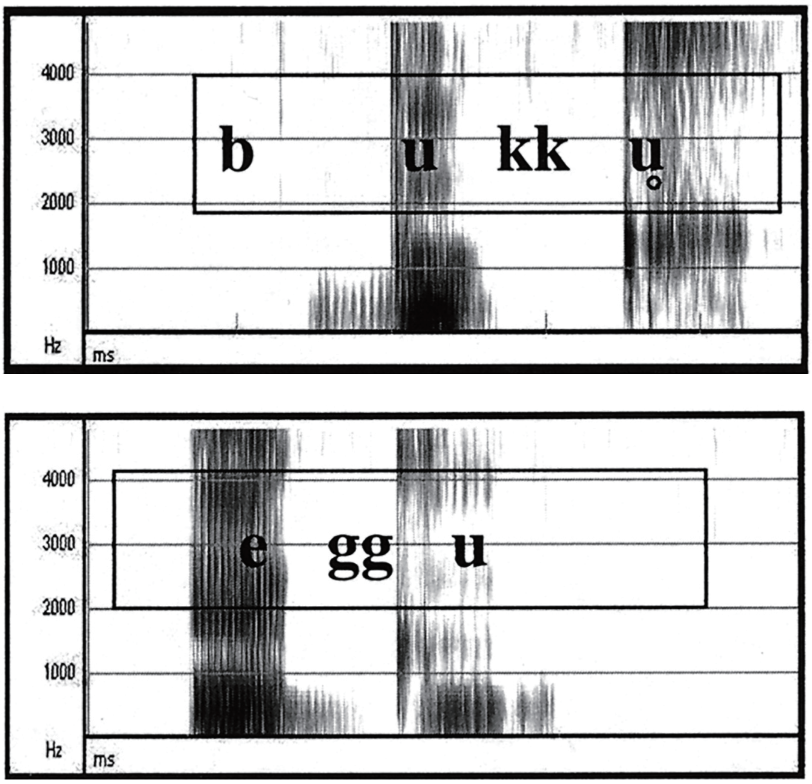
There are several pieces of evidence showing that the devoicing in DVDDV words is a categorical neutralization, different from context-free phonetic devoicing; high vowels following the geminates provide evidence for this conclusion. In Japanese, a high vowel is devoiced word-finally after a voiceless consonant, but devoicing does not take place after a voiced consonant (see Tsuchida 1997 for an overview and references to earlier work). Devoiced geminates in DVDDV words can induce devoicing of following vowels, just like underlyingly voiceless consonants, while voiced geminates do not cause such devoicing. The remainder of this subsection reports an acoustic experiment that demonstrates this claim.

A male native speaker of Tokyo Japanese was recorded. He was in his early thirties and was paid for his time. The speaker was naive to the purpose of this study. His speech was recorded through a microphone (MicroMic II C420 by AKG) by a CD-recorder (TASCAM CD RW-700) at a 44.1 KHz sampling rate, in a sound-attenuated booth at University of Massachusetts, Amherst. The recorded tokens were then downsampled to 22.050 KHz and 16-bit quantization level when they were saved on a PC. To determine the voicing of word-final high vowels after voiced and voiceless geminates, the set of stimuli listed in 6 was used. In order to elicit devoicing, all of the stimuli were existing loanwords.

(6)	a. [. . . kku#]		b. [. . . ggu#]		c. [D . . . ggu#]	
	bukku	‘book’	eggu	‘egg’	biggu	‘big’
	bakku	‘back’	furaggu	‘flag’	doraggu	‘drug’
	pakku	‘pack’	foggu	‘fog’	doggu	‘dog’

In addition to these target words, another fifteen real words were included as fillers. Each stimulus was embedded in a different frame sentence. The words that followed the target words began with a voiceless consonant in order to facilitate devoicing—for example, [bukku katte] ‘please buy a book’. The speaker was first instructed to read all of the stimuli five times in a natural style of speech. Then he was asked to read the stimuli five more times, but this time in a fast and casual register, as if he were talking to his friends. This was done in order to elicit devoicing of voiced geminates, which is more likely to take place in casual speech than in formal speech. The recording session took about thirty minutes. In this experiment, the speaker pronounced devoiced variants of the DVDDV words in 6c once or twice for each item: one instance of [dorakku], and two instances of [bikku] and [dokku]. A few tokens were mispronounced by the speaker and hence excluded from further consideration.

The devoicing pattern of word-final vowels supports the claim that devoicing in DVDDV words in 6c is categorical. First, devoicing of high vowels takes place after underlyingly voiceless geminates (= 6a) but not after voiced geminates (= 6b), as illustrated in Figure 1, the spectrograms of [bukku] ‘book’ and [eggu] ‘egg’. In Fig. 1a, the word-final [u] in [bukku] has aperiodic energy and shows no acoustic reflexes of glottal pulses, indicating that the [u] is devoiced. By contrast, in Fig. 1b, the word-final [u] after [gg] does show acoustic reflexes of glottal pulsing (i.e. the vowel is not devoiced).



(a)

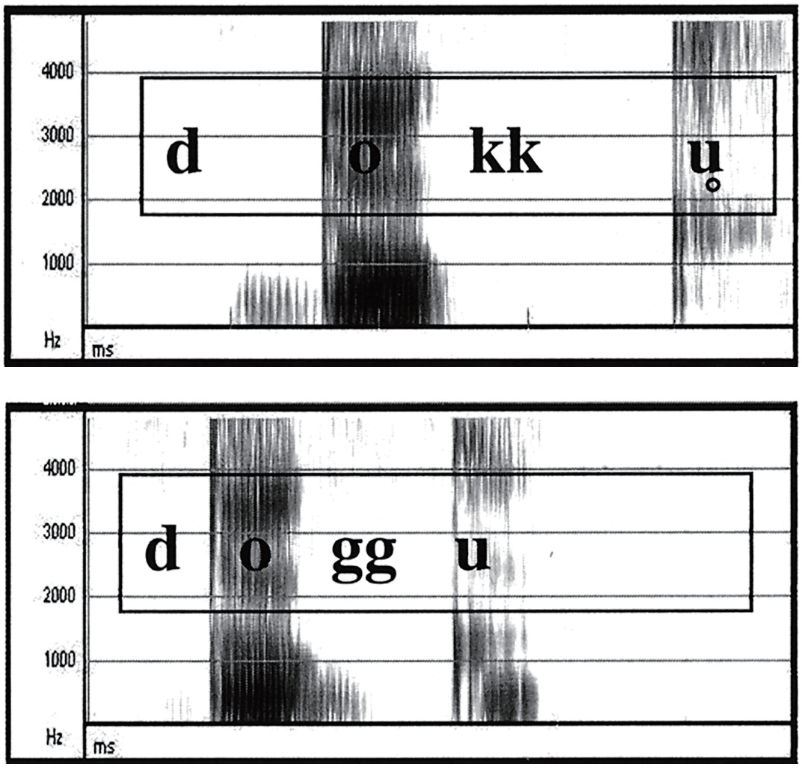
(b)

FIGURE 1. Spectrograms illustrating devoicing of word-final high vowels. Time scale is the same (750 ms). A high vowel devoices after [kk] (a), but not after [gg] (b).

In DVDDV words like those in 6c, the [u] following the geminate is devoiced if and only if the geminate is fully devoiced. Figure 2 shows spectrograms of the pronunciation of /doggu/ with and without devoicing of the voiced geminate. Comparing the two spectrograms in Fig. 2, we observe that the word-final [u] is devoiced only in Fig. 2a, where the geminate is also devoiced. To summarize, both [kk] derived from /gg/ and [kk] derived from /kk/ cause devoicing of a following high vowel. This overall patterning of high-vowel devoicing thus provides evidence that the devoicing of geminates in DVDDV words neutralizes the voicing contrast of underlying geminates.

It can be seen in Fig. 1b and Fig. 2b that much of the oral-closure interval in [gg] has no voicing—that is, the final portion of [gg] is devoiced (see §4.2 for further discussion). Despite this partial devoicing, however, the following vowel is still voiced. This observation suggests that devoicing of high vowels is not merely a continuation of the phonetic voicelessness of the preceding constriction, but rather that devoicing is induced by a PHONOLOGICALLY voiceless preceding consonant (see Tsuchida 1997 for further arguments that devoicing of high vowels is phonologically conditioned). The fact that the geminates in 6c can, when devoiced, induce devoicing of the following vowels thus supports the claim that these geminates are phonologically devoiced, a process that is distinct from the context-free phonetic gradient devoicing observed in [gg] (see Cohn 1993, Keating 1996, Pierrehumbert 1990, Tsuchida 1997, and Zsiga 1995, 1997 among others for the distinction between gradient phonetic processes and categorical phonological processes).

Additional acoustic evidence for neutralization of underlying /gg/ to [kk] is found in the closure duration and closure voicing duration. A comparison of the spectrograms



(a)

(b)

FIGURE 2. Spectrograms illustrating the presence/absence of devoicing of [u] after the two variant pronunciations of a voiced geminate. A high vowel devoices after a devoiced geminate (a), but not after a voiced geminate (b).

in Fig. 1 reveals that closure duration is longer in [kk] than in [gg]. Furthermore, closure voicing duration (acoustically realized as a voice bar) is shorter in [kk] than in [gg]. The same patterns emerge from the spectograms in Fig. 2: devoiced geminates have longer closure duration but shorter closure voicing duration than voiced geminates. This parallel between [kk] derived from /kk/ and [kk] derived from /gg/ again suggests that devoicing of a voiced geminate is complete neutralization. To verify this observation quantitatively, the closure durations and closure voicing durations were measured for [gg] derived from /gg/, [kk] derived from /gg/, and [kk] derived from /kk/. In measuring these values, the boundaries between the consonants and the flanking vowels were set at the points where F3 disappears and reemerges. The results are summarized in Figure 3 and Figure 4.

As seen in Fig. 3, the closure duration of phonologically devoiced /gg/ (the third bar) is indistinguishable from that of [kk] derived from /kk/ (the fourth bar); an independent-sample t-test reveals no significant difference ($t(22) = 0.35$, $p = 0.73$). Furthermore, phonologically devoiced /gg/ (the third bar) is longer than voiced /gg/ (the first and second bars), and the difference is statistically significant ($t(28) = 7.77$, $p < 0.001$). In short, [kk] derived from /gg/ patterns with [kk] derived from /kk/, not with [gg] derived from /gg/.

The same pattern emerges in closure voicing duration, summarized in Fig. 4. There is no difference in closure voicing duration between devoiced /gg/ (the third bar) and [kk] derived from /kk/ (the fourth bar) ($t(22) = 1.09$, $p = 0.29$). Furthermore, closure voicing is longer for voiced /gg/ (the first and second bars) than for devoiced /gg/ (the third bar) ($t(28) = 4.73$, $p < 0.001$). All of these acoustic observations again suggest

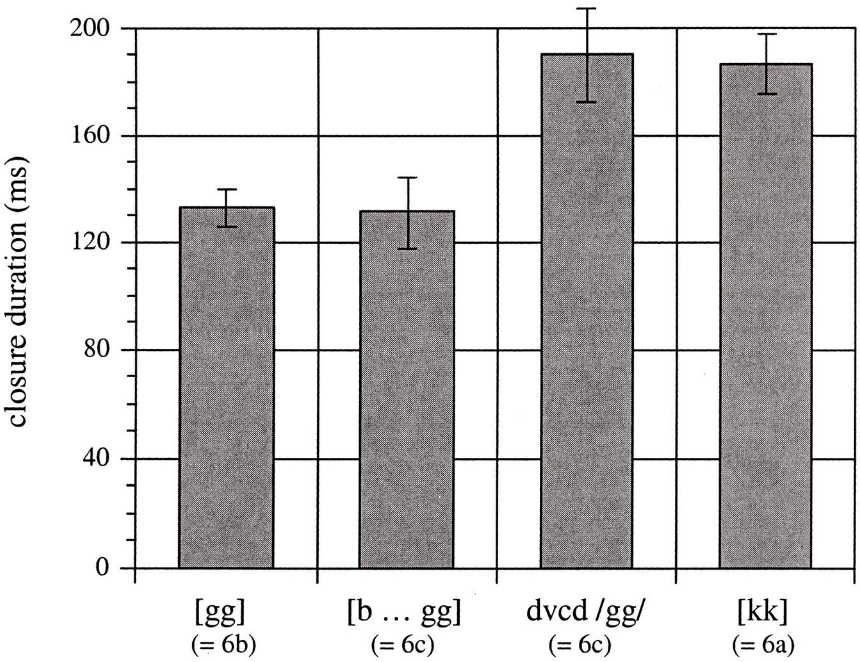


FIGURE 3. Mean closure duration of the words in 6. Error bars represent 95% confidence intervals. First bar represents [gg] in TVDDV words, the second [gg] in DVDDV words, the third [kk] in DVDDV words, and the fourth [kk] derived from /kk/.

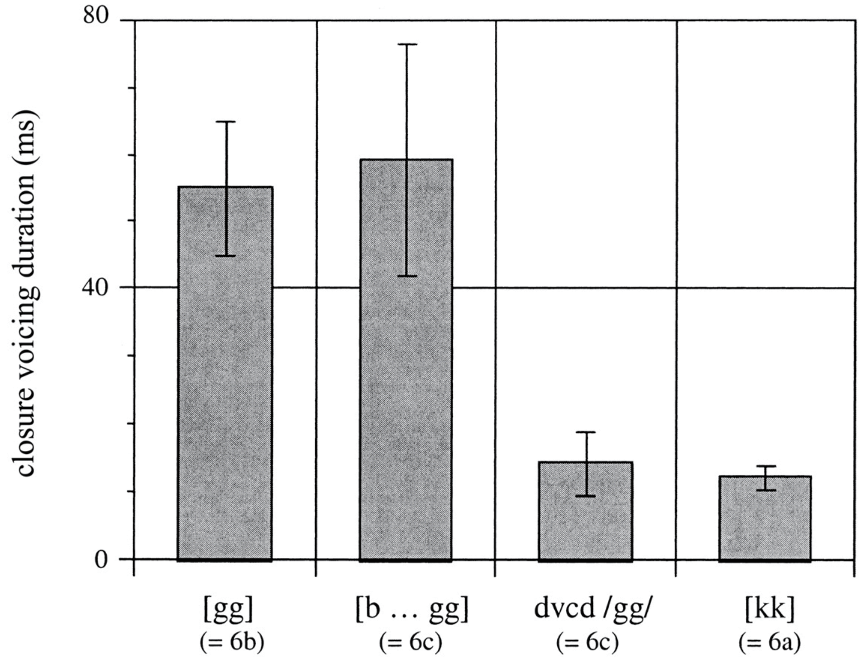


FIGURE 4. Mean closure voicing duration of the words in 6.

that the devoicing of the geminates in 6c, as shown in Fig. 2b, is phonological neutralization, because devoiced /gg/ behaves just like [kk] derived from /kk/.

To summarize, complete devoicing of voiced geminates is acoustically demonstrated by (i) the shortening of closure voicing duration, (ii) the longer closure duration, and (iii) the devoicing of following vowels. These results are in line with the evidence put forward by Nishimura (2003) as well as with the intuitions of native speakers. Taken together, these pieces of evidence strongly suggest that complete devoicing of voiced geminates in DVDDV words is a categorical phonological alternation.

3. Proposal and analysis.

3.1. Background. Having established that the devoicing of voiced geminates observed in DVDDV words is phonological in nature, I present in this section a phonological analysis of the behavior of voiced consonants in Japanese loanwords. I first discuss some assumptions crucial to the proposed analysis.

First, the analysis is framed within optimality theory (Prince & Smolensky 2004), which captures phonological patterns through the interaction of conflicting and violable constraints. Phonological prohibitions against particular structures, expressed in terms of markedness constraints, are therefore violable and may not hold absolutely within a language. This fundamental characteristic of OT allows us to model the fact that voiced geminates, though prohibited in the native phonology of Japanese, can appear in the loanword phonology (Itô & Mester 1995, 1999) and the fact that the devoicing of geminates in DVDDV words is optional.

In addition, OT is suitable here because of the centrality of faithfulness in the theory. While markedness constraints govern output structures' wellformedness, faithfulness constraints prohibit disparities between input forms and output forms. As I argue below, the notion of faithfulness plays an important role in distinguishing the behavior of singletons and geminates.

The analysis also crucially relies on the theory of positional faithfulness (Beckman 1998, Casali 1997), which claims that phonological features in different positions can be protected by different constraints. Drawing on this, I argue that in Japanese, the faithfulness constraints that protect a [+voice] feature must be relativized to singletons and geminates, and that this relativization has its basis in the perceptibility of a [+voice] feature in different contexts by virtue of the P-map (Steriade 2001a,b).

Finally, the focus of this article is not the process of gemination as loanwords enter the language and are adapted to native patterns, but rather the behavior of a [+voice] feature in that sector of the lexicon that can be identified (e.g. through speaker perceptions, by the way they are written in Japanese orthography, and other such criteria) as synchronic loanwords, words clearly of foreign (and specifically English) origin (see n. 1). Therefore, in the phonological analysis that follows, I use input forms that have already been borrowed such that they contain voiced geminates. For further details of the gemination process in loanword adaptation into Japanese, see, for example, Katayama 1998, Lovins 1973, Takagi & Mann 1994, and Tsuchida 1995.

3.2. Proposal and analysis. To analyze the patterns described in §2, I argue that there must be two faithfulness constraints regarding the feature [+voice]: one that applies when the [+voice] feature is hosted by a singleton consonant and another that applies when [+voice] is hosted by a geminate consonant. The intuitive idea behind this split is that neutralizing [+voice] to [−voice] in geminates is regarded as a 'perceptually tolerated articulatory simplification' (Guion 1998, Hura et al. 1992, Kohler 1990). Since [+voice] in geminates is not well perceived, it is protected only by a low-ranked faithfulness constraint. By contrast, [+voice] is well perceived in singletons, and it is therefore protected by a higher-ranked constraint (see Boersma 1998, Côté 2004, Fleischhacker 2001, Jun 2004, Padgett 2002, 2006, Steriade 1997, 2001a,b, Zhang 2000, and Zuraw 2005 for related proposals).

To formally express the proposal, I employ the IDENT family of constraints which regulate featural changes (McCarthy & Prince 1995). The IDENT constraints are formalized in 7.

(7) Two IDENT(+voi) constraints:
Let S_1 be an input string and S_2 be an output string, and let S_1 and S_2 stand in correspondence.
a. IDENT(+voi)$_{\text{Sing(leton)}}$: Let y $\in$ S_2 such that y is a singleton consonant. For all x $\in$ S_1 where x is a correspondent of y, if x is [+voi] then y is [+voi].
b. IDENT(+voi)$_{\text{Gem(inate)}}$: Let y $\in$ S_2 such that y is a geminate consonant. For all x $\in$ S_1 where x is a correspondent of y, if x is [+voi] then y is [+voi].

IDENT(+voi)$_{\text{Sing}}$ prohibits a change from [+voice] to [−voice] (i.e. it prohibits devoicing) when the [+voice] feature is hosted by a singleton consonant in the output, whereas IDENT(+voi)$_{\text{Gem}}$ prohibits a change from [+voice] to [−voice] when the feature is hosted by a geminate. Crucially, the ranking IDENT(+voi)$_{\text{Sing}}$ » IDENT(+voi)$_{\text{Gem}}$ holds in Japanese loanwords, and it is grounded in the perceptibility of [+voice] in Japanese singletons and geminates. See §6.2 for discussion of the (non)universality of the ranking IDENT(+voi)$_{\text{Sing}}$ » IDENT(+voi)$_{\text{Gem}}$.

Some remarks on the formulation in 7 are in order. First, the IDENT(+voi) constraints prohibit devoicing but not voicing, unlike a more general IDENT(voi), which prohibits both (see Itô & Mester 2003:Ch. 7, Pater 1999, and Walker 2001 for related discussion). In theory, either formulation works; however, IDENT(+voi) was chosen to reflect the fact that this article deals primarily with the change of [+voice] into [−voice].[6]

Second, the constraints are formulated in such a way that they are sensitive to a geminacy distinction in the output, rather than in the input. This captures the intuition that the differentiation of these faithfulness constraints is grounded in a difference in the perceptibility of [+voice] in singletons and geminates: perceptibility is a property of surface representations, and the percept of [+voice] in different contexts can depend on how it is phonetically implemented (see §§6.1 and 6.2 for further discussion).

In addition to the differentiation of IDENT(+voi) into two constraints, I argue that the source of devoicing in DVDDV words is a constraint against two voiced obstruents within a single stem. This constraint is well motivated in the native phonology of Japanese (see e.g. Itô & Mester 1986 and much subsequent work). There are no native stems that contain more than one voiced obstruent (e.g. *[buda]; cf. [huda] 'amulet' and [buta] 'pig'). In addition, this restriction manifests itself in the fact that rendaku is blocked. Rendaku is a phenomenon wherein the initial obstruent in the second stem of a compound becomes voiced, as in /nise-tako/ → [nise-dako] 'fake octopus'. But if the second stem already contains a voiced obstruent, voicing is blocked, as in /nise-taba/ → [nise-taba], *[nise-daba], 'fake bill'.

[6] The choice of IDENT(+voi) over IDENT(voi) is also motivated by the results of the perceptual experiment reported in §5, in which the [+voice] percept in geminates suffers significantly from misidentification, whereas the [−voice] percept does not.

Itô and Mester (1986) formalize the restriction against two voiced obstruents as an effect of the OBLIGATORY CONTOUR PRINCIPLE (OCP; Goldsmith 1976, Leben 1973, McCarthy 1986, Odden 1986, and much subsequent work) for [+voice] obstruents, which requires that there be no more than one voiced obstruent within a stem. I refer to this constraint as OCP(+voi).

(8) OCP(+voi): Two voiced obstruents cannot cooccur within a single stem.

With OCP(+voi) and the faithfulness constraints IDENT(+voi)$_{Sing}$ and IDENT(+voi)$_{Gem}$, the patterns presented in §2.1 can be accounted for easily. First, since singleton consonants do not devoice under the duress of OCP(+voi) in DVDV words, IDENT(+voi)$_{Sing}$ dominates OCP(+voi), as shown in 9.

(9) IDENT(+voi)$_{Sing}$ ›› OCP(+voi)

/bagii/	IDENT(+voi)$_{Sing}$	OCP(+voi)
a. ☞ [bagii]		*
b. [bakii]	*!	
c. [pagii]	*!	
d. [pakii]	*!*	

In contrast, in DVDDV words in which devoicing takes place, the winning candidate satisfies OCP(+voi) while violating IDENT(+voi)$_{Gem}$. Thus, OCP(+voi) must be ranked higher than IDENT(+voi)$_{Gem}$ when devoicing occurs, as shown in 10. Since devoicing is optional, however, OCP(+voi) and IDENT(+voi)$_{Gem}$ can be left unranked; see, for example, Anttila 2002 for discussion of unranked constraints in OT.

(10) OCP(+voi) ›› IDENT(+voi)$_{Gem}$

/baggu/	IDENT(+voi)$_{Sing}$	OCP(+voi)	IDENT(+voi)$_{Gem}$
a. [baggu]		*!	
b. ☞ [bakku]			*
c. [paggu]	*!		
d. [pakku]	*!		*

Finally, the fact that both voiced singletons and voiced geminates are independently allowed in Japanese loanwords follows if both IDENT(+voi)$_{Sing}$ and IDENT(+voi)$_{Gem}$ are ranked above the markedness constraint that prohibits voiced obstruents, *VOIOBS. These ranking arguments are illustrated in 11.

(11) IDENT(+voi)$_{Sing}$, IDENT(+voi)$_{Gem}$ ›› *VOIOBS

/eggu/	IDENT(+voi)$_{Sing}$	IDENT(+voi)$_{Gem}$	*VOIOBS
a. ☞ [eggu]			*
b. [ekku]		*!	
/bagu/			
a. ☞ [bagu]			**
b. [pagu]	*!		*
c. [baku]	*!		*
d. [paku]	*!*		

In sum, the ranking IDENT(+voi)$_{Sing}$ » OCP(+voi) » IDENT(+voi)$_{Gem}$ » *VOIOBS accounts for all of the patterns of [+voice] in the loanword phonology of Japanese.

3.3. AN ALTERNATIVE ANALYSIS. I have argued that the difference in how voiced singletons and voiced geminates react to OCP(+voi) arises because singletons and geminates are governed by different faithfulness constraints. In this section, I critically assess Nishimura's (2003) analysis, which elaborates on the inventory of markedness constraints instead (see Haraguchi 2006 for a similar line of analysis).

Nishimura's analysis uses *VOIOBSGEM, which directly prohibits voicing in geminates. Since voiced geminates are not independently devoiced, IDENT(+voi) must be ranked above *VOIOBSGEM. And since two voiced singletons are allowed within a single stem, IDENT(+voi) must also outrank OCP(+voi). With these rankings, however, geminates cannot devoice to satisfy OCP(+voi), because IDENT(+voi) is undominated. This problem is shown in 12 where the wrong winner is indicated by '(☞)'.

(12) Desired candidate [bakku] fails to win

/baggu/	IDENT(+voi)	OCP(+voi)	*VOIOBSGEM
a. (☞) [baggu]		*	*
b. [bakku]	*!		
c. [paggu]	*!		*
d. [pakku]	*!*		

Nishimura attempts to overcome this dilemma by locally conjoining OCP(+voi) and *VOIOBSGEM (Smolensky 1995). A locally conjoined constraint is violated if and only if both conjuncts are violated in a particular domain. Thus OCP(+voi) and *VOI

ObsGem are conjoined in the domain of stems, yielding {OCP(+voi) & *VoiObsGem}$_{\text{Stem}}$, and this constraint is ranked above Ident(+voi), as shown in 13.

(13) {OCP(+voi) & *VoiObsGem}$_{\text{Stem}}$ » Ident(+voi)

/baggu/		{OCP(+voi) & *VoiObsGem}$_{\text{Stem}}$	Ident(+voi)	OCP(+voi)	*VoiObsGem
a.	[baggu]	*!		*	*
b. ☞	[bakku]		*		
c.	[paggu]		*		*!
d.	[pakku]		**!		

The conjoined constraint causes devoicing only when a voiced geminate violates OCP(+voi). Although this approach is able to account for the asymmetrical behavior of singletons and geminates, it has some problems.

First, in the local-conjunction framework, the constraint {OCP(+voi) & *VoiObs Gem}$_{\text{Stem}}$ is produced by recursive conjunction, because OCP(+voi) is the self-conjunction of *VoiObs (Alderete 1997, Haraguchi 2006, Itô & Mester 1997, 1998, 2001, 2003), and *VoiObsGem is the local conjunction of *VoiObs and *Gem. Thus, {OCP (+voi) & *VoiObsGem}$_{\text{Stem}}$ has the internal structure shown in Figure 5.

{{*VoiObs & *VoiObs}$_{\text{Stem}}$ & {*VoiObs & *Gem}$_{\text{Seg}}$}$_{\text{Stem}}$

{*VoiObs & *VoiObs}$_{\text{Stem}}$ {*VoiObs & *Gem}$_{\text{Seg}}$

*VoiObs *VoiObs *VoiObs *Gem

Figure 5. Recursive local conjunction.

This sort of recursive local conjunction is too powerful. For example, no grammar is known to prohibit three or more occurrences of a particular structure—languages do not seem to 'count' beyond two (Chomsky 1965, McCarthy & Prince 1986). By recursively self-conjoining *X, however, it is possible to derive a language that counts the instances of X within a domain D via a constraint like {{*X & *X}$_{\text{D}}$ & *X}$_{\text{D}}$ (Itô & Mester 1998:n. 17). This prediction is undesirable.

The second problem is that this approach allows nonidentical constraints to be conjoined within the domain of a stem. Allowing this sort of conjunction predicts the existence of unattested constraints (see McCarthy 1999, 2003 for further discussion of this problem). Some examples are given in 14.

(14) Predicted conjoined constraints that are unattested

{*Lab & NoCoda}$_{\text{Stem}}$: A labial and a coda consonant cannot cooccur within a stem.

{Max & NoCoda}$_{\text{Stem}}$: No codas are allowed when deletion has occurred in a stem/no deletion is allowed if there is a coda in a stem.

{Ident(+voi) & Dep}$_{\text{Stem}}$: Devoicing and epenthesis cannot both occur within a stem.

In short, the local-conjunction approach seems too powerful, and the theory proposed in §3.2 obviates its need.

3.4. CONSEQUENCES. I now turn to the theoretical consequences of the proposal presented here, focusing on its typological predictions. Optimality theory is inherently typological, since the set of constraints is assumed to be universal and thus all variation between languages comes from the language-particular rankings of these constraints. Therefore, it is necessary to consider crosslinguistic consequences that arise from the proposed differentiation of IDENT(+voi) into two distinct constraints.

First, given the two IDENT(+voi) constraints, IDENT(+voi)$_{Gem}$ must never outrank IDENT(+voi)$_{Sing}$. This is necessary because no known languages have voiced geminates unless they also have voiced singletons (Hayes & Steriade 2004). If the ranking IDENT(+voi)$_{Gem}$ » *VOIOBS » IDENT(+voi)$_{Sing}$ were allowed, then this unattested pattern would result. To avoid this potential overgeneration, IDENT(+voi)$_{Sing}$ must be ranked either higher than or as high as IDENT(+voi)$_{Gem}$. Since, as I argue below, the ranking between these constraints is grounded in the perceptibility of [+voice] in singletons and geminates, this ranking restriction should be as well. See §6.2 for further discussion of this point.

Second, with the elaboration of faithfulness constraints proposed here, it is desirable in terms of theoretical parsimony to eliminate *VOIOBSGEM, which prohibits voiced geminates but not voiced singletons (Hayes & Steriade 2004, Itô & Mester 1995, 1999, Nishimura 2003, among others). This constraint can be replaced with a general *VOIOBS: languages with only singleton voiced consonants would have the ranking IDENT(+voi)$_{Sing}$ » *VOIOBS » IDENT(+voi)$_{Gem}$. With this ranking, any underlying voiced geminates would be devoiced in the output.

One might wonder whether this simplification is indeed possible in light of the fact that voiced geminates are repaired not just by devoicing, but by other processes as well. For example, in the native phonology of Japanese, the [-ri] suffix contains a floating mora μ, and it causes gemination of the second consonant in a mimetic root as shown in 15a. However, when the second consonant is a voiced obstruent, as in 15b, coda nasalization takes place instead (Kawahara 2006a, Kuroda 1965).

(15) Mimetic gemination in Japanese
- a. /tapu+μ+ri/ → [tappuri] *[tampuri] 'a lot of'
  /kapa+μ+ri/ → [kappari] *[kampari] 'opening'
- b. /zabu+μ+ri/ → [zamburi] *[zabburi] 'splashing'
  /ʃobo+μ+ri/ → [ʃombori] *[ʃobbori] 'depressed'

While *VOIOBSGEM can account for the pattern in 15 by directly penalizing a voiced geminate, so can *VOIOBS, if we assume that a geminate violates it twice (following Baković's (2000) general theory of assessing markedness violations at a segmental level). More concretely, let IDENT(nas)$_{Coda}$ be the faithfulness constraint that militates against coda nasalization. If *VOIOBS dominates IDENT(nas)$_{Coda}$, we obtain the desired result, as illustrated in 16.

(16) *VOIOBS ›› IDENT(nas)$_{Coda}$

/ʃobo+μ+ri/	*VOIOBS	IDENT(nas)$_{Coda}$
a. [ʃobbori]	**!	
b. ☞ [ʃombori]	*	*

Candidate (a), which has a voiced geminate, loses because the geminate has two voiced segments (and hence *VoiObs is violated twice).

This line of analysis essentially regards gemination of a voiced obstruent as the addition of a coda voiced obstruent, which automatically incurs an additional violation of *VoiObs. Therefore, this analysis can be falsified if there is a language that allows the addition of a coda voiced obstruent, but not the creation of a voiced geminate, both of which incur violations of *VoiObs. The proposed elimination of *VoiObsGem therefore predicts that such a pattern does not exist. It is beyond the scope of this article to argue for the absence of such a process with certainty, or to reanalyze all of the cases that have been analyzed using *VoiObsGem. However, with the elaboration of faithfulness constraints proposed here, the simplification of the markedness constraint inventory seems possible.[7]

3.5. Discussion. Based on the evidence from the behavior of [+voice] in singletons and geminates with respect to OCP(+voi), I have argued that the loanword phonology of Japanese has the ranking Ident(+voi)$_{Sing}$ » Ident(+voi)$_{Gem}$. It is worth emphasizing here that there is nothing in the native phonology of Japanese that motivates this proposed ranking. As seen in 15b, voiced geminates are resolved by coda nasalization, not by devoicing, in the native phonology, so Japanese speakers exposed only to native vocabulary should not know this ranking. The question that immediately arises is how Japanese speakers established this ranking when they incorporated loanwords.

The P-map hypothesis provides an answer to this question. The P-map hypothesis asserts that speakers have knowledge of faithfulness rankings that go beyond what can be inferred from their native phonology—faithfulness rankings can be derived from perceptual-similarity rankings, rather than merely from language exposure (Steriade 2001a,b; see also Fleishhacker 2001, Kawahara 2006b, and Zuraw 2005 for related proposals). For the case at hand, if [+voice] is less perceptible in geminates than in singletons, the P-map hypothesis predicts that speakers infer the ranking Ident(+voi)$_{Sing}$ » Ident(+voi)$_{Gem}$. To the extent that the faithfulness ranking is grounded in a perceptibility scale, the prediction goes the other way too: given Ident(+voi)$_{Sing}$ » Ident(+voi)$_{Gem}$, there should be a difference in the perceptibility of [+voice] between singletons and geminates. Thus, one prediction of the P-map hypothesis that can be empirically tested is that [+voice] is less perceptible in geminates than in singletons in Japanese.

The next two sections report experiments that test this prediction. These experiments show that the perceptibility of [+voice] does indeed differ in singletons and geminates, supporting the P-map hypothesis. First, a production experiment was conducted to investigate the set of acoustic cues for a [±voice] distinction in Japanese. In light of the predictions of the P-map hypothesis, it is expected that some of the acoustic cues are weakened in geminates, and this prediction is supported by the results. The second experiment was a perceptual study, an identification task in a noisy environment. The results show that, given attenuation of a [±voice] distinction in geminates, [+voice] is indeed less perceptible in geminates than in singletons.

It is worth mentioning at this point that the support for the P-map hypothesis provided here is necessarily limited—I show that [+voice] is less perceptible in geminates than

[7] Thanks to the anonymous referee whose suggestion led me to this conclusion. One remaining concern in eliminating the constraint *VoiObsGem is that there is a well-motivated aerodynamic reason for why voiced geminates are articulatorily challenging (e.g. Hayes & Steriade 2004, Jaeger 1978, Ohala 1983, Westbury 1979).

in singletons, and that phonologically, [+voice] is more easily neutralized in geminates than in singletons. I make no attempt, however, at a cross-categorical comparison (i.e. comparing voicing and other features in terms of their perceptibility). Such a cross-categorical comparison is possible given the original P-map hypothesis advanced by Steriade (2001a,b); for example, the difference in perceptibility between a voicing change and a nasality change is projected onto a fixed faithfulness ranking. But this position is challenged by the fact that there is a language like Japanese which nasalizes coda consonants to repair underlying voiced geminates (see 15), as well as a language like Endegeň which devoices underlying voiced geminates (Leslau 1976:146); the ranking between FAITH(nas) and FAITH(voi) thus does not seem to be universally fixed. This observation is a counterexample to the original P-map hypothesis, and the proposal advanced here is therefore limited to a within-category comparison, namely that of [+voice], where the prediction of the P-map hypothesis seems most secure.

4. EXPERIMENT 2: ACOUSTICS OF [±voice] IN JAPANESE.[8] To examine whether a voicing contrast is indeed harder to perceive in geminates than in singletons, I investigated in this experiment (i) what kinds of acoustic correlates are associated with a voicing contrast in Japanese, (ii) how these acoustic correlates manifest differently in singletons and in geminates, and (iii) whether some of these acoustic correlates are weaker in geminates than in singletons, as predicted by the P-map hypothesis.

Some remarks on terminology are in order. First, in what follows, LENGTH or GEMINACY is used to refer to a phonological contrast that distinguishes singletons from geminates, while DURATION refers to a phonetic measure indicating how long a particular phonetic event lasts. Second, we are interested in the perceptibility of a phonological voicing contrast, which is associated not only with glottal vibration, but also with other acoustic properties (see below). Therefore, VOICING is used to refer to actual glottal vibration (and its acoustic manifestation), whereas [±VOICE] or A VOICING CONTRAST is used to refer to a phonological distinction between voiced and voiceless consonants.

4.1. METHODS. In this experiment, words containing four kinds of stops (voiceless singletons, voiced singletons, voiceless geminates, and voiced geminates) were recorded. Three native speakers of Japanese were recruited from the University of Massachusetts, Amherst. They were all female and in their mid-twenties. The dialects the subjects spoke were Hiroshima Japanese, Shizuoka Japanese, and Tokyo Japanese.[9] The frame sentence used in the experiment was Standard (Tokyo) Japanese, and the subjects were asked to read the sentences in Standard Japanese as well. An informed consent form was obtained from each speaker in accordance with the University of Massachusetts human research subjects guidelines. The speakers were all paid for their time. The speech was recorded through a microphone (MicroMic II C420 by AKG) by a CD-recorder (TASCAM CD RW-700) at a 44.1 KHz sampling rate, in a sound-attenuated booth. The recorded tokens were then downsampled to 22.050 KHz and 16 bit quantization level when they were saved on a PC. Including short breaks between repetitions, the recording session for each speaker lasted about forty-five minutes.

[8] I am grateful to José Benkí for his extensive comments on §§4 and 5 as a referee, and to John Kingston for his assistance in many aspects of these experiments. For further details of these experiments, see Kawahara 2005.

[9] It might be possible that gender and dialectal factors could have affected the results, but there has as yet been no report of gender or dialectal differences in the pronunciation of geminate consonants in Japanese.

The stimuli consisted of thirty-six words, which were mostly nonce words.[10] Thirty-six additional nonce words were included as fillers. The target words were all disyllabic: the first consonant was [k], the second consonant was the target ([p], [t], [k], [pp], [tt], [kk], [b], [d], [g], [bb], [dd], and [gg]), and three different vowels ([a], [e], [o]) were used in both syllables. Some examples are [kepe], [kabba], [kete], [koddo], [kaga], [kekke], and so forth. The speakers were asked to pronounce these tokens with a HL tonal contour, which is the default accent pattern for loanword and nonce-word pronunciation.

Each word was written on an index card; katakana orthography, conventionally used for loanwords, was employed because voiced geminates are found only in loanwords. Six repetitions of each set were recorded, with a short break between each repetition. The order of the stimuli was randomized after each repetition. In order to elicit natural utterances and avoid domain-edge strengthening effects on target words (Fougeron & Keating 1997), the stimuli were embedded in the frame sentence given in 17.

(17) Jaa __ de onegai.
then __ with please
'Please (do something) with __, then.'

In order to avoid extensive hyperarticulation of the materials, the speakers were encouraged to produce sentences in a natural speech style. Specifically, they were instructed to imagine a situation in which they were preparing for a party and they wanted their friend to fetch the things named by the target words.

All measurements were done using Praat (Boersma & Weenink 2005). In line with the past literature on acoustic and perceptual correlates of a [±voice] contrast (Kingston & Diehl 1994, Lisker 1986, Raphael 1981, Stevens & Blumstein 1981), the following values were measured: (i) duration of closure voicing, (ii) duration of the preceding vowel (V1), (iii) closure duration, (iv) F0 of the surrounding vowels, and (v) F1 of the surrounding vowels. Voice onset time (VOT) is known to cue [±voice] in other languages (Lisker & Abramson 1964 et seq.), but it was not measured because [±voice] in Japanese is not signaled by aspiration. The way that each acoustic property was measured is illustrated in Figure 6, with a representative spectrogram of [kobbo]. The onset of V1 was set where F3 becomes visible after the preceding [k], and the onset of consonantal closure was set where F3 of V1 disappears. The duration of closure voicing was measured based on the presence of low-frequency periodic energy near the bottom of the spectrograms. The offset of the consonantal-closure interval was set at the release of the consonant, which was signaled by the burst noise. The closure durations reported below do not include the duration of the burst noise. F0 and F1 of both the preceding vowel (V1) and the following vowel (V2) were also measured. The measurement points were the last periodic wave before closure for V1 and the first periodic wave after the burst for V2. F1 was measured using Praat's LPC analysis, with the number of LPC coefficients left at the default value of 10. F0 was measured using autocorrelation. Sometimes voiced singleton consonants were spirantized, in which case they were excluded from acoustic analyses.

To statistically analyze the acoustic measures obtained, an ANOVA was performed with voicing contrast (2-level) and consonantal length (2-level) as independent variables. These variables were chosen because we are interested in how a [±voice] differ-

[10] It was impossible to completely exclude real words: [kaba] and [kakka] are real words. However, since the stimuli were written in katakana orthography, [kakka], which is normally written in hiragana, was not recognized as a real word.

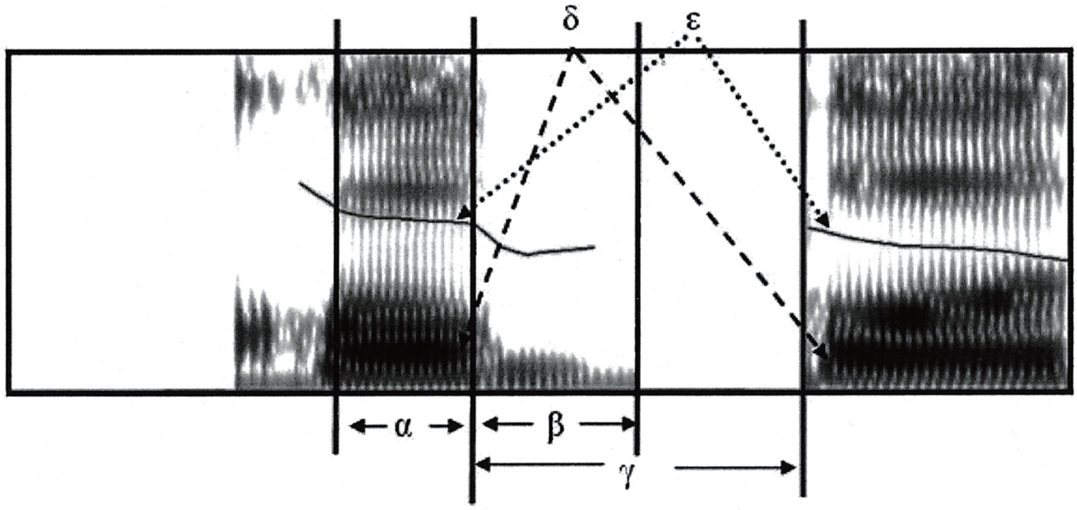


FIGURE 6. Illustrative spectrogram of [kobbo] showing how each acoustic value was measured. α = V1 duration, β = closure voicing duration, γ = closure duration, δ = F1 at vowel edges, ϵ = F0 at vowel edges.

ence manifests itself in these acoustic values, and how the acoustic values vary in singleton and geminate environments. For the sake of exposition, I abstract away from interspeaker differences; Kawahara 2005 reports observed individual differences in detail.

4.2. RESULTS. The overall results show that a phonemic voicing difference is maintained in both singletons and geminates, but that some cues are weakened in geminates. I first report the acoustic differences between voiced and voiceless consonants, and then show that there are at least three ways in which the [$\pm$voice] contrast in geminates is attenuated.

One of the most important correlates of [$\pm$voice] is CLOSURE VOICING DURATION, the extent to which voicing continues into the closure (Lisker 1978, 1986, Raphael 1981, Stevens & Blumenstein 1981). Closure voicing is acoustically realized as a voice bar—that is, low-frequency periodic energy during closure. The results of the measurements of closure voicing duration are summarized in Figure 7. Here and throughout in summary figures, the first pair of bars represents singleton values and the second pair geminate values. Within each pair, the first bar represents voiced consonants and the second bar voiceless consonants. Error bars represent 95% confidence intervals (CIs), calculated as the critical value of $t_{.05}$ associated with an appropriate degree of freedom ($n - 1$) multiplied by standard error of the mean (s.e.). Even though four CIs were calculated simultaneously, no familywise Type 1 error adjustment was applied. The error bars are provided to give an idea of the accuracy of the mean estimations, not for the sake of post-hoc multiple comparisons.

As seen in Fig. 7, voiced consonants, whether singletons or geminates, have on average about 40 ms of closure voicing. Voiceless consonants, by contrast, have about 10 ms of closure voicing. This difference is statistically significant ($F(1,611) = 720.41$, $p < 0.001$). But there does not seem to be any effect of geminacy on closure voicing ($F(1,611) < 1$), nor is the interaction between the two variables significant ($F(1,611) < 1$). The results thus show that closure voicing duration is longer in voiced than in voiceless consonants, and the size of the differences is about the same between singleton and geminate pairs.

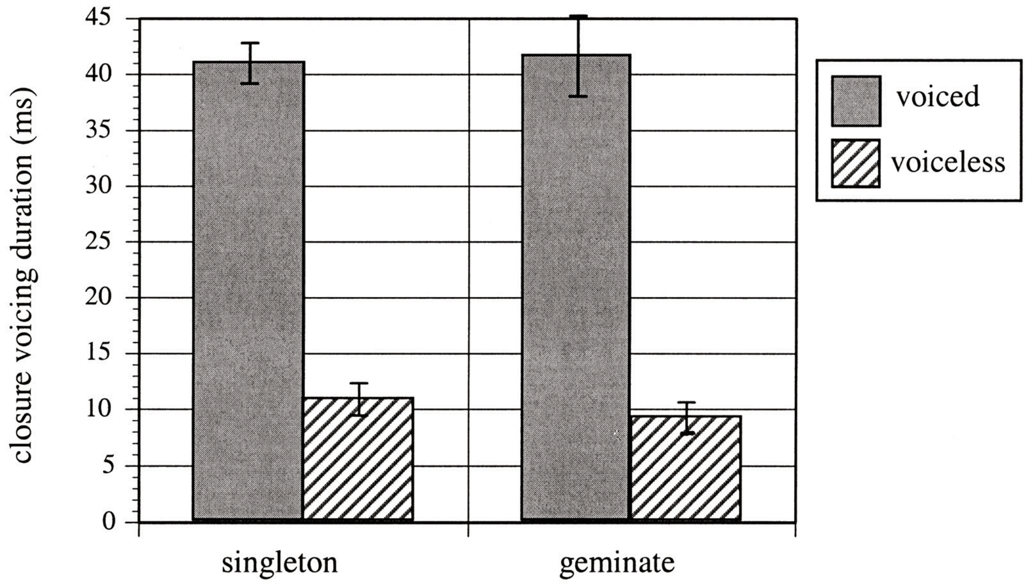


FIGURE 7. Mean closure voicing duration (in milliseconds). Error bars represent 95% confidence intervals, calculated as $t(\mathrm{df})_{.05} \times$ s.e.

However, as implied by the size of the error bars in Fig. 7, closure voicing duration is more variable in voiced geminates than in voiced singletons (the standard deviation is 10.41 for singletons and 23.16 for geminates). This difference in variability is statistically significant, according to a Brown-Forsythe test, which compares absolute deviation of scores around the median in the two groups ($t(290) = 5.49$, $p < 0.001$).[11] The fact that closure voicing duration is more variable in geminates than in singletons implies that it might provide a less reliable cue for a [±voice] distinction in geminates.

The second phonetic difference that correlates with a voicing contrast is the duration of the immediately preceding vowel (V1 DURATION). The results are summarized in Figure 8. As expected from reports for other languages (Chen 1970, Kingston & Diehl 1994, Raphael 1981), vowels are longer before voiced than before voiceless consonants (the difference is on average 12.69 ms; $F(1{,}603) = 166.34$, $p < 0.001$). Next, as previously reported by Han (1994), vowels are also longer before geminates than before singletons (the difference is 20.85 ms; $F(1{,}603) = 453.18$, $p < 0.001$). The interaction of these two variables is significant ($F(1{,}603) = 19.49$, $p < 0.001$). This significant interaction effect arises because the V1 duration difference is larger before geminates than before singletons (by about 8.5 ms). This larger difference before geminates might provide an advantage for signaling a [±voice] distinction in geminates, contra the expectation of the P-map hypothesis; however, it is shown below that there are a number of other cues that are weakened in geminates (see §4.3 for further discussion).

The third acoustic correlate of a [±voice] difference is CLOSURE DURATION, how long the consonantal closure lasts. The results appear in Figure 9. Again, as expected from reports for other languages (Kingston & Diehl 1994, Lisker 1957, Ohala 1983, Westbury 1979), voiceless consonants are longer than voiced (16.54 ms; $F(1{,}603) = 182.94$, $p < 0.001$). Geminates are on average longer than singletons by 69.77 ms

[11] This heterogeneity of variances between these two groups might have inflated Type I errors in the ANOVA (Myers & Well 2003:221). However, since the F-ratio of the effect of [±voice] is very large, this should not be too problematic.

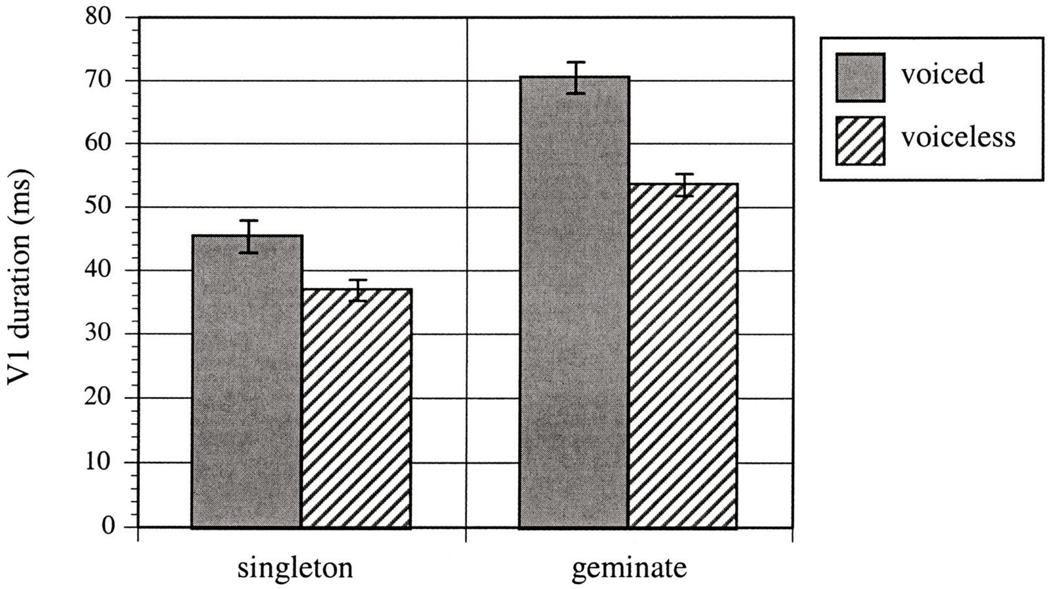


FIGURE 8. Mean V1 duration (ms).

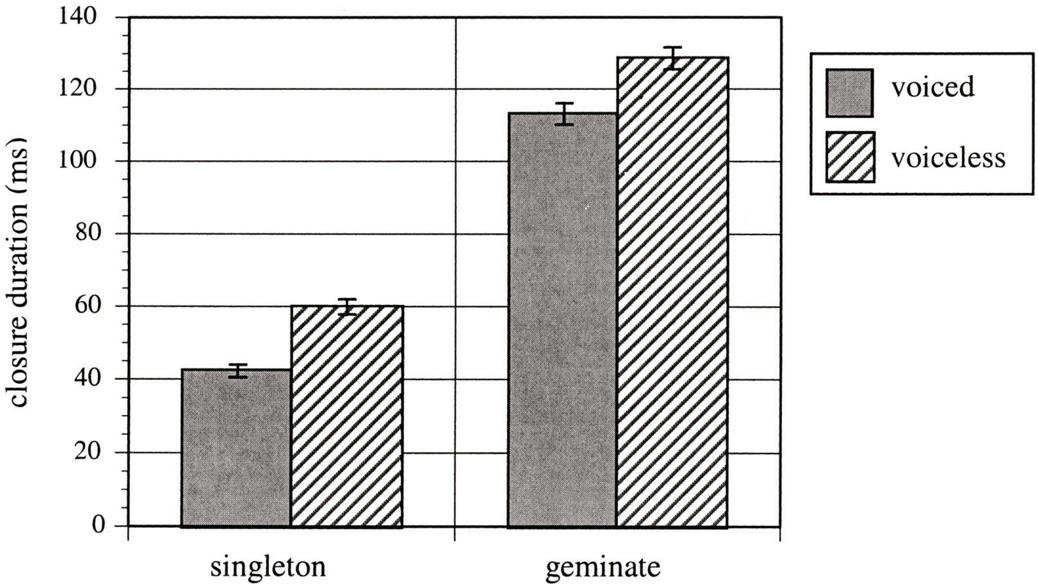


FIGURE 9. Mean closure duration (ms).

($F(1{,}603) = 3220.48, p < 0.001$). The interaction of the two variables is not significant ($F(1{,}603) < 1$). The lack of a significant interaction effect indicates that the closure-duration difference due to a [±voice] difference is about the same size between singleton and geminate pairs.

We have seen that voiced consonants exhibit shorter closure duration and longer V1 duration. These opposite effects of [+voice] create quite different C/V-duration ratios (CV-RATIOS) for [+voice] and [−voice] consonants, which demonstrably constitute an important perceptual cue for the [±voice] distinction (Kingston & Diehl 1994, Port & Dalby 1982). The C/V duration ratios for the Japanese case at hand are summa-

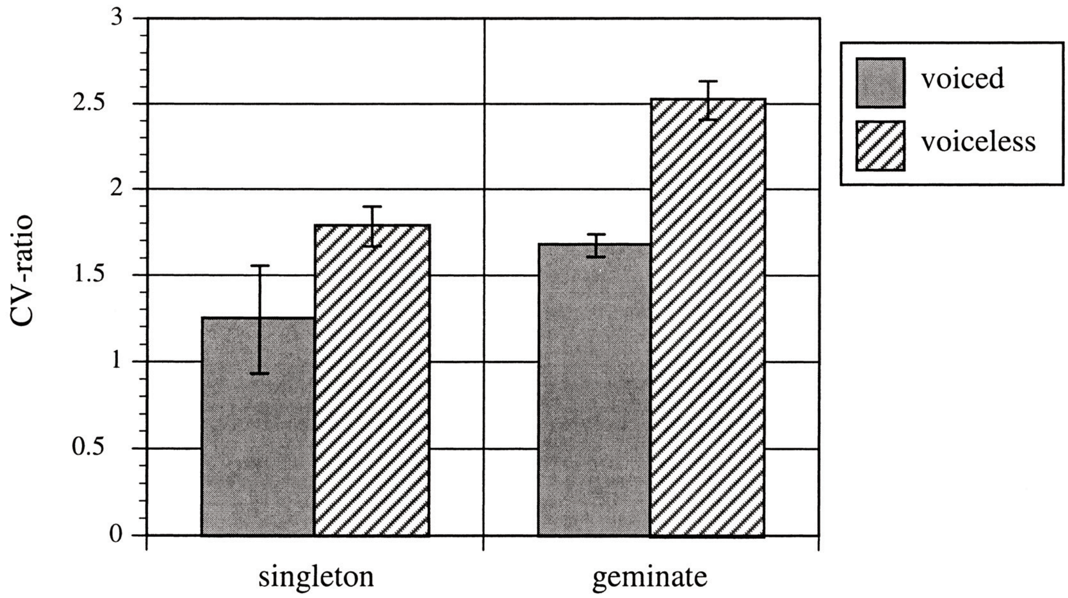


FIGURE 10. Mean CV-ratio, calculated as duration ratio of consonant/vowel.

rized in Figure 10. An ANOVA shows that geminates have a higher CV-ratio than singletons (on average, singletons = 1.52; geminates = 2.10; $F(1{,}587) = 59.89$, $p < 0.001$). Voiceless consonants, as expected, also have a higher CV-ratio than voiced consonants (voiceless = 2.15; voiced = 1.46; $F(1{,}587) = 84.13$, $p < 0.001$). The interaction is marginally significant ($F(1{,}587) = 3.78$, $p = 0.052$), and it reflects the tendency for the difference in CV-ratios to be larger for geminate pairs than for singleton pairs (by about .3). One might suspect that this tendency enhances the perceptual distinction of [±voice] in geminates. However, this suspicion must remain tentative. If we follow Kohler's (1979) suggestion that a voicing distinction should be made by the duration ratio of vowel/(vowel + consonant), then the ratio difference due to [±voice] is larger for singletons than for geminates (singletons: vcd = .50, vls = .38, difference = .12; geminates: vcd = .38, vls = .29, difference = .09), and this difference is statistically significant ($t(586) = 2.33$, $p < 0.05$). Therefore, whether a voicing cue is indeed enhanced in geminates in terms of duration ratio depends on which ratio is the most relevant perceptual cue for Japanese speakers. No evidence is currently available to settle this matter. Yet it is shown in §5 that overall, the [+voice] percept in Japanese geminates is indeed attenuated, so a larger CV-ratio in geminates does not falsify the prediction of the P-map hypothesis.

Finally, F0 and F1 frequencies are known to be higher before and after voiceless consonants than voiced consonants (Kingston & Diehl 1994 and references cited therein). First, in V2, as expected, F0 is higher after voiceless consonants, as illustrated in Figure 11. F0 is on average 20.32 Hz higher after voiceless consonants than after voiced ($F(1{,}604) = 175.95$, $p < 0.001$). By contrast, F0 is on average 32.36 Hz lower after geminates than after singletons ($F(1{,}604) = 365.28$, $p < 0.001$). This is presumably because the tonal contour of the recorded tokens is HL; given longer closure, the F0 fall is more drastic after geminates, as there is more time to implement the HL fall. The interaction is not significant ($F(1{,}604) = 1.80$, $p = 0.18$), which indicates that the F0 frequency difference due to a [±voice] contrast is more or less constant between postsingleton and postgeminate positions.

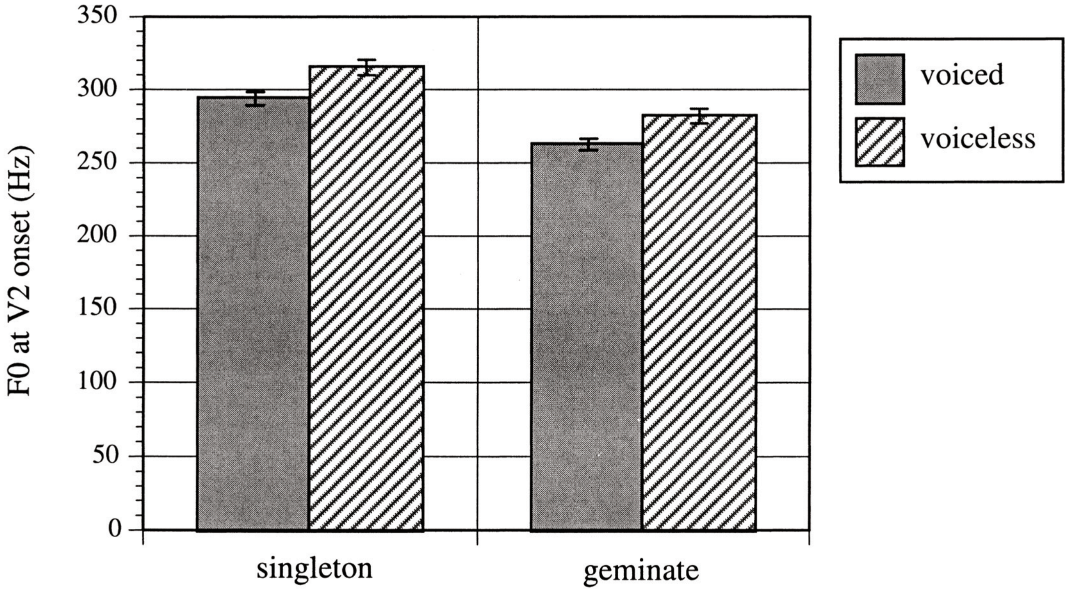


FIGURE 11. Mean F0 at V2 onset (Hz).

F1 is also higher after voiceless consonants than after voiced, as illustrated in Figure 12.

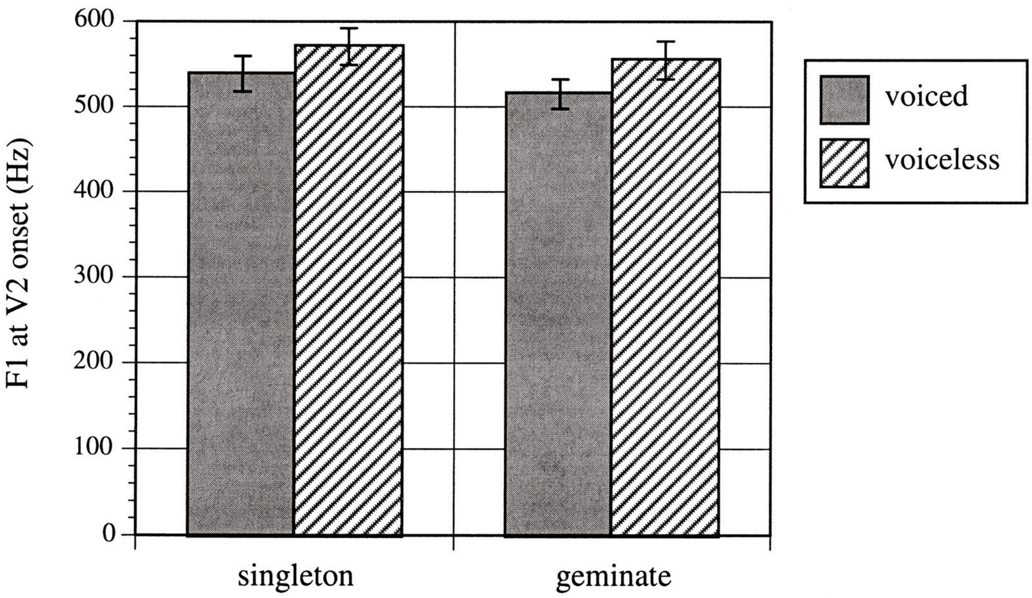


FIGURE 12. Mean F1 at V2 onset (Hz).

The difference due to a [±voice] difference is about 35.68 Hz, which is statistically significant ($F(1,600) = 14.56$, $p < 0.001$). The effect of geminacy is only marginally significant ($F(1,600) = 3.18$, $p = 0.075$); on average, F1 is 19.56 Hz lower after geminates. The interaction is not significant ($F(1,600) < 1$), indicating that F1 differences due to a [±voice] contrast are about the same size in postsingleton and postgeminate environments.

Finally, F0 and F1 are expected to be higher before voiceless consonants than before voiced consonants. This prediction is borne out for F0, as illustrated in Figure 13.

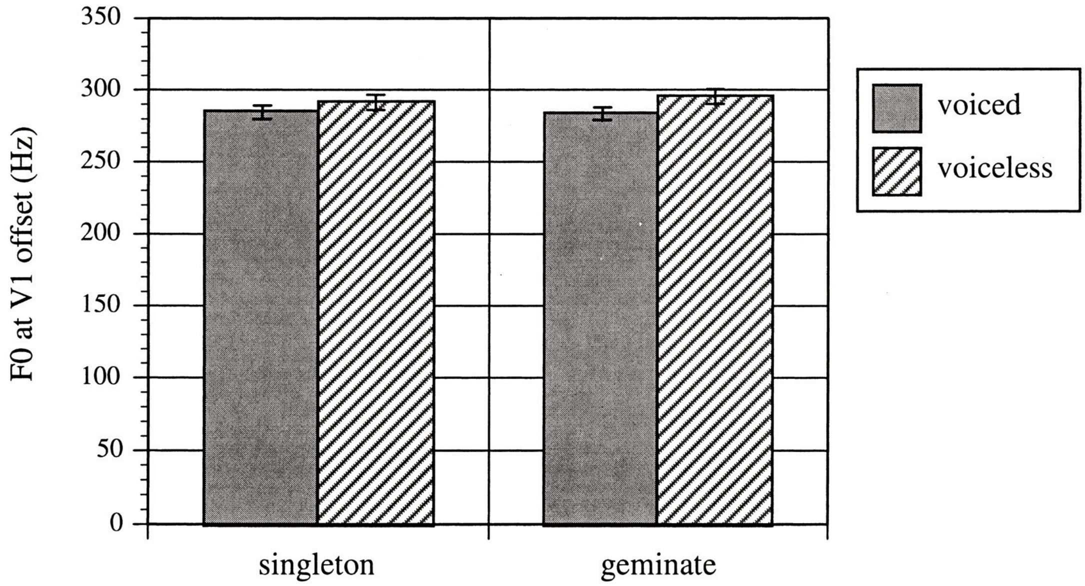


FIGURE 13. Mean F0 at V1 offset (Hz).

F0 is 9.30 Hz higher before voiceless consonants than before voiced consonants ($F(1,601) = 71.29$, $p < 0.001$). F0 seems slightly higher before geminates than before singletons, but this effect of geminacy is only barely significant ($F(1,601) = 4.19$, $p < 0.05$). The interaction between these two factors is not significant ($F(1,601) < 1$), indicating that the size of F0 differences due to a [±voice] contrast does not significantly differ between presingleton and pregeminate positions.

F1, unlike F0, does not show any differences at V1 offset; neither a [±voice] contrast nor geminacy affects F1 values (the *F*-ratios are both below 1). The interaction is not significant either ($F(1,600) < 1$). These are illustrated in Figure 14.

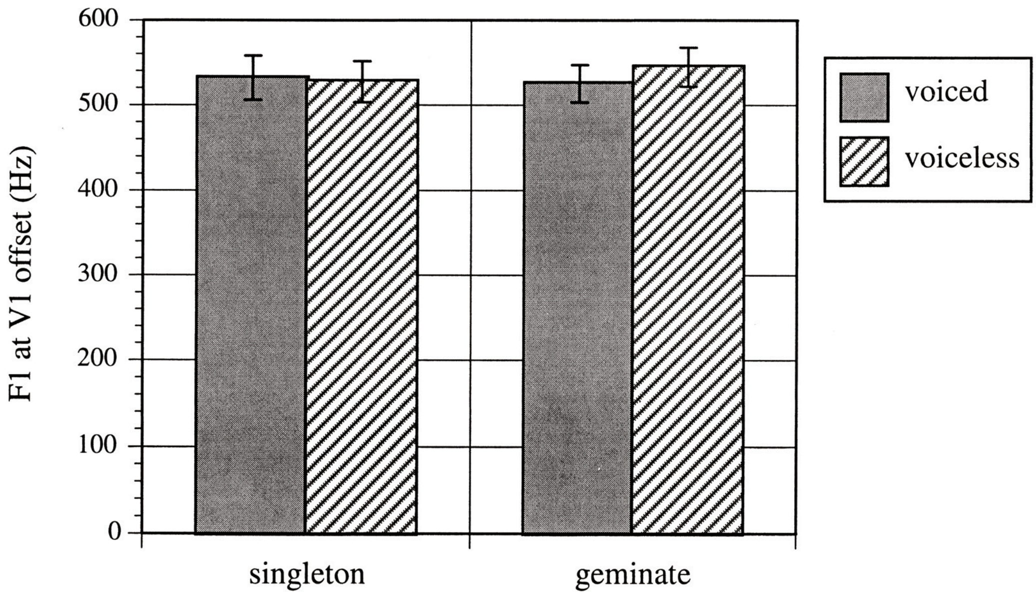


FIGURE 14. Mean F1 at V1 offset (Hz).

One might suspect from Fig. 14 that a difference between voiced and voiceless consonants might emerge before geminates; however, a simple effect analysis of [±voice] using only the geminate data does not reveal any significant effect ($F(1,322) = 1.47$, $p = 0.23$).

We have seen that a voicing contrast in Japanese is associated with many of the cues that are known to signal a [±voice] distinction crosslinguistically. Recall, however, that since [+voice] in Japanese is more easily neutralized to [−voice] in geminates than in singletons, the P-map hypothesis predicts that acoustic cues in geminates might be weakened in some dimensions. There are at least three reasons to suspect that this prediction might be true.

First, glottal vibration stops in the middle of closure for voiced geminates, but not for singletons. In other words, geminates are partially devoiced, whereas singleton consonants are fully voiced. The partial devoicing of Japanese voiced geminates is illustrated by the two spectrograms in Figure 15. As seen in Fig. 15, while voicing is fully maintained in the singleton [b] (top), partial devoicing is observed after the arrow in the geminate [bb] (bottom). This asymmetry between singletons and geminates is consistent across the three speakers. All of the speakers maintain full voicing for almost all singleton tokens, but they rarely produce fully voiced geminates: out of fifty-four tokens of voiced geminates, one speaker produced two tokens of fully voiced [bb], another produced one fully voiced [gg], and the third produced no fully voiced geminates.

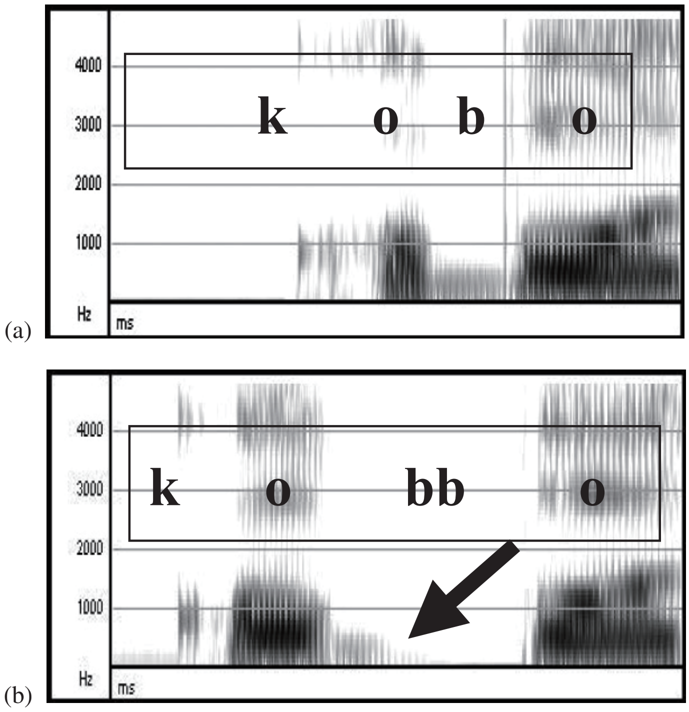


FIGURE 15. Spectrograms of a singleton [b] (a) and geminate [bb] (b). Time scales are the same (350 ms).

To quantify the degree of partial devoicing, the proportion of closure voicing duration to closure duration was calculated. The results appear in Figure 16. Closure voicing is maintained for only about 40% of the entire closure interval in geminates, whereas voicing is fully maintained in singletons. This extensive partial devoicing of voiced geminates is due to the aerodynamic difficulty of maintaining voicing during obstruent

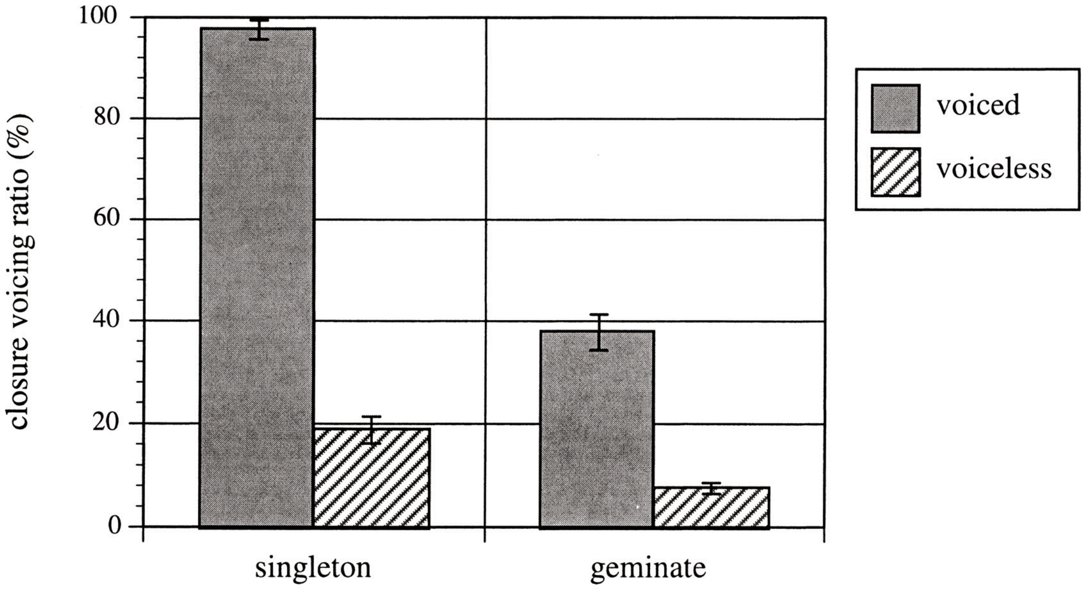


FIGURE 16. Closure voicing ratio, calculated as closure voicing duration divided by corresponding closure duration.

closure: intraoral air pressure goes up quickly, and as a consequence it becomes difficult to maintain a transglottal air pressure drop sufficient to produce voicing (Hayes & Steriade 2004, Jaeger 1978, Ohala 1983, Westbury 1979).

The fact that the last 60% of a voiced geminate is phonetically voiceless may undermine the percept of [+voice] in geminates, since closure voicing is one of its most important cues (Lisker 1978, Parker et al. 1986, Raphael 1981). In particular, Lisker (1978) has shown that consonants with 120 ms closure duration and 40 ms closure voicing, which closely resemble Japanese partially devoiced geminates, are perceived by English speakers as voiceless about 70% of the time, even when other cues such as V1 duration are in favor of a [+voice] percept. Furthermore, since Japanese voiced geminates are acoustically voiceless at the time of release, this should also attenuate the overall [+voice] percept as well, because it is known that onset cues have primacy over offset cues (Raphael 1981, Slis 1986, Steriade 1997).

The second possible source of attenuation of a [±voice] distinction concerns a closure-duration difference. Recall that both singleton and geminate voiceless consonants are longer than their corresponding voiced versions, and the size of the difference is about the same for singleton and geminate pairs (Fig. 9). However, since geminates are inherently longer than singletons, geminate pairs are more similar to each other than singleton pairs—analogically speaking, 20 and 21 are more similar to each other than 1 and 2 are, even though for both of the pairs, the difference is 1. To quantify the degree of similarity between singleton and geminate pairs in terms of closure duration, the proportion of voiced consonants' closure duration to voiceless consonants' closure duration was calculated.

The result is that the average ratio is much higher for geminates: .89 for geminates and .71 for singletons. The standard errors for these estimates are .02 and .03, respectively, which were calculated as $\sqrt{\frac{p(1-p)}{n}}$ by approximation to a Gaussian distribution, where p stands for a voiced/voiceless ratio and n the number of data points ($n =$

289 for singletons and $n = 323$ for geminates). The difference between the two ratios is statistically significant ($z = 5.60, p < 0.001$). This difference in the voiced/voiceless ratios suggests that, proportionally, geminate minimal pairs are more similar to each other than singleton minimal pairs in terms of closure duration. Since it is known that closure duration affects perception of [±voice] in at least some environments (Lisker 1957, 1978, 1981, Parker et al. 1986), the larger voiced/voiceless ratios might make a [±voice] distinction harder to hear in geminates. For a similar line of reasoning see Sanders 2003, which argues that nasal vowels have a disadvantage in signaling a phonemic length contrast compared to oral vowels, because nasal vowels are inherently longer than oral vowels.

The third reason to suspect that a [+voice] difference is weaker for geminates than for singletons is the presence/absence of spirantization. Intervocalic spirantization often occurs in voiced singletons, but voiced geminates resist it. This contrast is illustrated in Figure 17. As seen in Fig. 17, a singleton /g/ is lenited almost to an approximant (top), as the visible formant energy during the constriction indicates, whereas no lenition occurs for geminates (bottom). Since voiceless consonants never spirantize, a [±voice] contrast in singleton pairs can also be signaled by the presence/absence of frication noise (i.e. the contrast is often phonetically realized as a difference between [ɣ] and [k]). Geminate pairs, however, do not have this advantage because whether voiced or voiceless they do not spirantize, and therefore [+voice] geminates are not distinguished from [−voice] geminates in terms of frication noise.

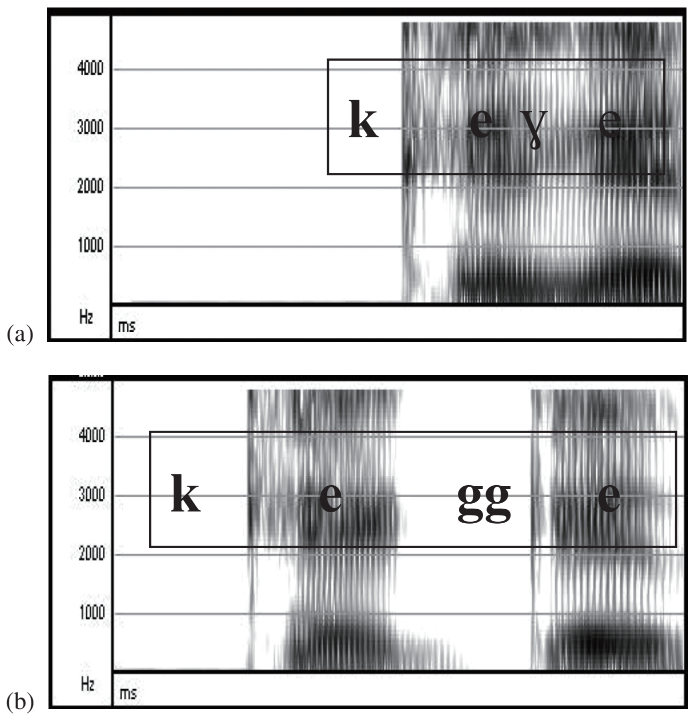


FIGURE 17. Spectrograms of a singleton /g/ realized as [ɣ] (a) and a geminate /gg/ realized as [gg] (b). Time scales are the same (350 ms).

4.3. DISCUSSION. For the three reasons discussed above, the [±voice] distinction seems to be less perceptible in geminates than in singletons. There is a small complication here, however. The speakers show signs of attempting to make up for the inherently attenuated [±voice] contrast in geminates. For example, one of the three speakers shows a larger V1 duration difference before geminates than before singletons (as reflected in Fig. 8). Another speaker has a larger F0 difference in V2 after geminates. Yet, these attempts to make up for weakened cues are speaker-specific and not observed consistently across all the speakers. The size of these enhancements is small as well (e.g. the F0 enhancement is 8 Hz). It is therefore unlikely that attenuated cues in geminates are sufficiently compensated for, and the perceptual experiment presented next in fact shows that a [±voice] distinction is less well perceived in geminates than in singletons. See Kawahara 2005 for a detailed report on the relevant data and discussion.

To summarize, I have shown that the Japanese [±voice] contrast is signaled by a number of the acoustic parameters that are known to cue a [±voice] distinction crosslinguistically. In addition, as predicted by the P-map hypothesis, there are reasons to suspect that a [±voice] distinction is less perceptible in geminates than in singletons. First, voiced geminates are partially devoiced; the consonants are phonetically voiceless during the last 60% of the closure as well as at the time of release. Second, due to their inherently long closure duration, the closure-duration difference is proportionally much smaller in geminates than in singletons. Finally, the lack of spirantization in geminates weakens an acoustic distinction between voiced and voiceless consonants because the presence/absence of frication noise does not cue a [±voice] difference in geminates. Overall, therefore, it seems appropriate to conclude that a [±voice] difference is attenuated in geminates. The perceptual experiment reported in the next section more directly supports this conclusion.

5. EXPERIMENT 3: PERCEPTUAL EXPERIMENT. In order to test more directly the hypothesis that the [+voice] feature is harder to perceive in geminates than in singletons, a perceptual experiment was conducted. In order to replicate most accurately the situation in which Japanese speakers hear [+voice] in geminates and singletons, the natural tokens recorded in the acoustic experiment were used. Had natural tokens and nothing else been used, however, Japanese speakers might have performed at ceiling. To overcome this problem, the stimuli were covered by cocktail party noise to confuse the listeners.

Given the observation from the acoustic experiment that a [±voice] distinction in geminates is acoustically attenuated, the prediction is that [+voice] is less perceptible in geminates than in singletons. The results of this experiment show that this prediction is borne out. This provides support for the hypothesis that the ranking IDENT(+voi)$_{Sing}$ » IDENT(+voi)$_{Gem}$ in Japanese is related to the perceptibility of [+voice] in singletons and geminates.

5.1. METHODS. From the pool of tokens obtained in the acoustic experiment, one representative example of each type of stimulus was chosen. There were thirty-six types of stimuli (three vowels × three places of articulation × two consonantal lengths × two [±voice] types), each of which was pronounced by three speakers. The total number of stimuli was therefore 108. Tokens that contained phonetic irregularities (such as audible clicks or devoiced V1) or spirantization were not used. For the case of singleton [g], which almost always underwent spirantization, tokens with the least spirantization were chosen. Among the tokens of voiced geminates at each place of articulation, those whose closure voicing duration was closest to the average for that place of articulation were used. See the appendix for the acoustic values of the tokens used in this experiment.

Cocktail party noise was used to cover the tokens. This particular kind of noise was used because in order to mask voicing, speech-like noise with energy in low spectra ranges was necessary: Miller and Nicely (1955) found that voicing is not masked well by white noise. To obtain this noise, a party was recorded using a SONY TCD-D8 portable DAT recorder. The recorded sound was divided into three-second noise stretches. Six such stretches were randomly chosen and superimposed on top of one another. This process was repeated twelve times, and twelve such files were created. The amplitudes of all stimuli were equalized by Praat to 0.50 Pascal for the stimuli and to 0.45 Pascal for the noise. As a result, the average amplitude of the stimuli and that of the noise became 71.90 dB and 72.35 dB, respectively. Thus the signal-to-noise ratio (S/N ratio) was -0.45. Finally, one noise file was randomly chosen and was superimposed on each stimulus. All stimuli were approximately 1.5 seconds long, including the frame sentence.

The subjects were seventeen native speakers of Japanese recruited from the University of Massachusetts, Amherst community. They were all in their twenties or early thirties. The speakers who participated in the acoustic experiment were excluded. All of the subjects had normal hearing and were free of any speech disorders. Some were recruited from an undergraduate introductory linguistics course and hence had a basic knowledge of linguistics, but none of them had extensive phonetic training. The range of dialects that the subjects spoke was diverse, including Chiba Japanese, Ibaragi Japanese, Osaka Japanese, Shizuoka Japanese, and Tokyo Japanese. However, no report has been made of a difference in the behavior of voiced geminates among these dialects, so this dialectal variation was not expected to have an impact on the results. Two listeners were native bilingual speakers of Japanese and English, but their results were similar to the results of the other subjects and hence are included in the results reported below. All of the subjects were either paid or given extra credit for linguistics courses. An informed consent form was obtained from each subject.

The experiment was conducted in a sound-attenuated booth. Superlab Pro software (by Cedrus) was used for audio and visual presentation of each stimulus. This software automatically randomized the order of presentation. The subjects listened to one stimulus at a time over headphones (DT 250 by Beyerdynamic). As soon as a listener heard a stimulus, two choices appeared on a computer screen. These were two possible orthographic representations of the stimulus, minimally different in [$\pm$voice] of the second consonant—for example, for the auditory stimulus [kappa], the two visual choices were katakana representations of *kappa* and *kabba*. The task was to make a judgment about whether the auditory stimuli contained voiced or voiceless segments. Katakana orthography was used so that the subjects would be encouraged to perceive the stimuli as nonnative words (recall that voiced geminates are allowed only in loanwords). In order to ensure that the subjects responded to all of the stimuli, no time limits were enforced. The subjects were not given feedback about the correctness of their responses.

Before the main testing sessions, the subjects had a practice session in which they performed the same task for one token of each stimulus pronounced by one speaker. In the practice session, however, the stimuli were not covered by noise, and the subjects were given feedback about the correctness of their answers. They were also instructed to adjust the volume to a comfortable listening level during the practice session.

One testing session consisted of three blocks, each of which contained all stimuli pronounced by one speaker. One block thus contained thirty-six tokens, and one session 108 tokens. Each session lasted only a few minutes. The entire experiment consisted of eight sessions. The subjects thus heard each stimulus twenty-four times (three speakers $\times$ eight sessions). The subjects were encouraged to take short breaks once or twice during the experiment. Including the instructions at the beginning and the postexperiment debriefing explanation, the entire experiment lasted about one hour.

5.2. Results. In order to analyze the perceptibility of a [± voice] contrast for singleton and geminate consonants, a sensitivity measure (*d′*) was computed for each subject: *d′* is a measurement of sensitivity in signal-detection theory (MacMillan & Creelman 2005) that directly represents the perceptual distance between two stimuli—that is, the perceptibility of the contrast between the two stimuli. The advantage of using *d′* instead of the perhaps more familiar 'percent correct (p(c))' analysis is that it distinguishes between overall sensitivity (= perceptual distance; perceptibility) and biases (= the subjects' predisposition toward one response or the other). See MacMillan & Creelman 2005 for more detailed discussion of sensitivity and bias; the biases observed in this experiment are discussed shortly below.

D′ is based on *z*-transformed scores of hit and false-alarm rates, where 'hit' is the probability of the listeners' correctly identifying voiced consonants as voiced, and 'false alarm' is the probability of the listeners' falsely identifying voiceless consonants as voiced. *D′* is defined by $z(\text{hit}) - z(\text{false alarm})$,[12] and therefore *d′* is positive when the hit rate exceeds the false-alarm rate. A *d′* of zero indicates that hit and false-alarm rates are the same, which means that the distinction between the two stimuli is not perceptible at all.

The average *d′* for singletons across all seventeen listeners is 3.79, which is significantly different from zero ($t(16) = 34.15$, $p < 0.001$), and the average *d′* for geminates is 0.71, also significantly different from zero ($t(16) = 11.47$, $p < 0.001$). These results show that for Japanese listeners, [+ voice] and [− voice] segments are perceptually distinct in both singletons and geminates. However, the perceptibility of [± voice] is much higher for singletons than for geminates; a paired *t*-test comparing *d′* for singletons and geminates reveals a significant difference ($t(16) = 27.27$, $p < 0.001$). This finding is exactly as predicted by the P-map hypothesis: the P-map hypothesis is thus supported experimentally.

Furthermore, interesting differences are observed among the three places of articulation regarding the perceptibility of [+ voice]. Voiced geminates' *d′* values are on average 0.82 for labials, 0.64 for coronals, and 0.15 for dorsals. These values indicate that the [± voice] distinction in geminates is most perceptible for labials, less so for coronals, and least so for dorsals.

These perceptibility differences among the three places of articulation are at least partially reflected in the likelihood of phonological devoicing of voiced geminates due to OCP(+ voi). There is only one DVDDV word that contains /bb/ ([gebberusu] 'Göbbels'), so it is thus difficult to make any conclusive generalizations about /bb/. However, analyzing Nishimura's (2003) web-based data to compare the likelihood of devoicing /dd/ and /gg/ reveals that /gg/ is more frequently devoiced (24.6%; 51,131 out of 216,440 tokens) than /dd/ is (15.3%; 35,539 out of 231,752 tokens). This negatively correlates with the *d′* values obtained above; the lesser the perceptibility of the consonant at a given place, the more likely the geminate at that place is to devoice. The correla-

[12] As *z*-scores of 0 and 1 are negative and positive infinity, respectively, I added or subtracted the equivalent of half of one response (i.e. $\sqrt{\frac{1}{2 * n}}$) from each perfect score (Macmillan & Creelman 2005:8). For example, if a listener identified [− voice] geminates as [− voice] 100% of the time, the proportion was adjusted to $1 - \sqrt{\frac{1}{2 * 216}} = 0.998$, where 216 is the number of [− voice] geminate tokens the listeners heard. There was one perfect listener for voiceless singletons, and one perfect listener for voiceless geminates.

tion between *d′* and devoicing likelihood further supports the view that the likelihood of devoicing is closely tied to the perceptibility of [+voice]. The lesser perceptibility of [+voice] in /gg/ compared to that of /dd/ gives rise to the ranking IDENT(+voi)$_{dd}$ » IDENT(+voi)$_{gg}$, which leads to the higher probability of phonological devoicing of /gg/.

Let us now turn our attention to the bias observed in the results of the experiment. One interesting aspect of the data obtained is that although [+voice] geminates are often misidentified as [−voice] (71.3%), listeners rarely misidentified [−voice] geminates as [+voice] (12.3%). No such asymmetries are observed for singletons (the misidentification of [+voice] as [−voice] = 3.6%; the misidentification of [−voice] as [+voice] = 4.1%). This asymmetry indicates that Japanese speakers are biased against hearing [+voice] in geminates, but not in singletons: in other words, the listeners prefer [−voice] responses for geminate stimuli.

Such a perceptual bias can be quantified using the bias function *c* (McMillan & Creelman 2005). The bias function *c* is the sum of the *z*-scores of the hit and false-alarm rates multiplied by −0.5. Recall that ‘hit’ is the probability of identifying voiced consonants as voiced, and ‘false alarm’ is the probability of identifying voiceless consonants as voiced. Since *z*-scores are negative when their probabilities are less than 0.5, positive *c* values (= negative sums of the *z*-scores) can be obtained when listeners prefer a [−voice] response for both [+voice] and [−voice] stimuli.

The mean *c* for singletons is 0.08, which does not significantly deviate from zero ($t(16) = 1.01$, $p = 0.33$). By contrast, the mean *c* for geminates is 1.08, which is significantly different from zero ($t(16) = 5.57$, $p < 0.001$). These results show that there is a perceptual bias against giving a [+voice] response when the stimuli are geminates, but not when the stimuli are singletons.

5.3. DISCUSSION. The results of the perceptual experiment have shown that [+voice] is less perceptible in geminates, and that there is a perceptual bias against hearing [+voice] in geminates. The second point implies that in addition to weakening of acoustic cues, there are some perceptual factors biasing against voiced geminates, such as lexical frequency and/or phonological constraints. Since voiced geminates are allowed only in loanwords, they are much less frequent than voiced singletons in the Japanese lexicon. This is confirmed by a survey using Amano and Kondo’s (2000) database, which is based on issues of *Asahi Shinbun* ‘Asahi Newspaper’ from 1985 to 1998. The type and token frequencies of voiceless singletons, voiced singletons, voiceless geminates, and voiced geminates in Amano & Kondo 2000 are shown in Table 2.

	VOICED		VOICELESS		VCD/VLS RATIO	
SINGLETON	84,732,417	(122,616)	255,086,803	(276,164)	33.2%	(44.4%)
GEMINATE	24,587	(505)	4,274,451	(11,792)	0.6%	(4.3%)

TABLE 2. Frequency of voiceless singletons, voiced singletons, voiceless geminates, and voiced geminates in Amano & Kondo 2000. The numbers represent token frequencies; type frequencies are given in parentheses.

In terms of token frequency, voiced singletons are 33.2% as frequent as voiceless singletons, but voiced geminates are only 0.6% as frequent as voiceless geminates. In terms of type frequency, voiced singletons are 44.4% as frequent as voiceless singletons, whereas voiced geminates are only 4.3% as frequent as voiceless geminates. As a consequence, [+voice] in geminates, which is much less frequent than [+voice] in singletons, might be at a disadvantage in being perceived: it is well established that there is perceptual bias toward hearing an acoustically ambiguous signal as the more frequent possibility rather than the less frequent possibility (see Hay et al. 2003 for a recent overview). Further, grammatical constraints antagonistic to voiced geminates might also be at work: there are perceptual biases against hearing phonologically illegal sounds or sound sequences (Moreton 2002). From the results of the experiment alone, it is not clear which factor(s) is responsible for the bias against the [+voice] percept given geminate stimuli.

6. General discussion and conclusion.

6.1. Theoretical implications. In this article, I first showed that a voicing contrast is more easily neutralized in geminates than in singletons in Japanese loanwords. Based on this observation, I argued that [+voice] is protected by two different faithfulness constraints, Ident$(+\text{voi})_{\text{Sing}}$ and Ident$(+\text{voi})_{\text{Gem}}$, and that these are ranked as Ident$(+\text{voi})_{\text{Sing}}$ » Ident$(+\text{voi})_{\text{Gem}}$. I further argued that this ranking is grounded in the relative perceptibility of [+voice] in singletons and geminates, and this claim has been supported experimentally. A general implication of this conclusion is that phonetic perceptibility can directly influence patterns in a phonological grammar.

Furthermore, the lesser perceptibility of [+voice] in Japanese geminates is likely due, at least in part, to a Japanese-specific way of phonetically implementing voiced geminates. In particular, the low perceptibility of [+voice] in Japanese geminates is partly due to their context-free partial devoicing, but this partial devoicing is not observed in every language. As reported in Kawahara 2006a, for instance, Egyptian Arabic maintains full voicing in voiced geminates, as illustrated by the spectrogram in Figure 18.[13]

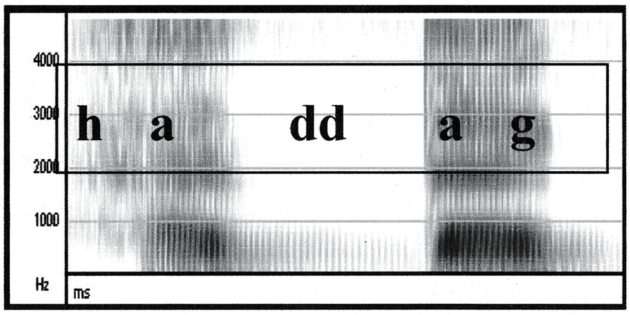


Figure 18. Spectrogram of a nonce Arabic word [haddag] pronounced by a female native speaker of Egyptian Arabic.

To the extent that the lesser perceptibility of [+voice] in Japanese geminates is due to partial devoicing, this suggests that a language-specific phonetic detail (e.g. partial devoicing) can affect a phonological pattern (e.g. categorical devoicing of geminates). This conclusion contributes to a growing body of work that claims that phonology can call on phonetic details, which are rarely if ever contrastive in phonology (Boersma 1998, Browman & Goldstein 1989, Flemming 1995, Gafos 2002, Kirchner 1997, Padgett 2006, Steriade 1997, 2000, Zhang 2000, 2004).

6.2. On the (non)universality of the proposed ranking. Another issue, related to the discussion above, is whether the ranking Ident$(+\text{voi})_{\text{Sing}}$ » Ident$(+\text{voi})_{\text{Gem}}$ is universal. The P-map hypothesis suggests that it is not. As discussed above, the lesser

[13] Cohn et al. (1999) also report that voicing is maintained throughout the closure in voiced geminates in three Austronesian languages: Buginese, Madurese, and Toba Batak.

perceptibility of [+voice] in Japanese voiced geminates might be due to the language-specific way that Japanese speakers phonetically implement them. Lexical-frequency bias against voiced geminates and the fact that voiced geminates are not allowed in the native vocabulary are also specific to Japanese.

These characteristics of Japanese should be contrasted with those of a language like Arabic. As seen above in Fig. 18, closure voicing is fully maintained in this language. Furthermore, there are no lexical or morphological biases against voiced geminates. In Wehr's (1971) Arabic dictionary, there are 1,028 roots whose second consonant is a voiceless obstruent, and of these, 433 occur in the verb pattern in which the second consonant is geminated (= 42.1%). There are also 811 roots whose second consonant is a voiced obstruent, and 325 of these occur in the verb pattern in which the second consonant is geminated (= 40.0%). The difference between these two ratios is not statistically significant (by approximating to a normal distribution, $z = 0.89$, $p = 0.27$). Therefore, there seems to be no evidence for frequency or grammatical biases against voiced geminates in Arabic.

Given a language like Arabic, then, a voicing contrast might be equally well perceived in singletons and in geminates (which must of course be empirically tested in future research). In such a case, the P-map hypothesis predicts that IDENT(+voi)$_{\text{Sing}}$ and IDENT (+voi)$_{\text{Gem}}$ are ranked in the same position: a phonological split between [+voice] in singletons and geminates, like the one found in Japanese loanwords, would not be observed. This prediction remains to be tested.

A further prediction of the theory advanced here is that [+voice] cannot be more perceptible in geminates than in singletons. If that were the case, the P-map hypothesis could generate the ranking IDENT(+voi)$_{\text{Gem}}$ » IDENT(+voi)$_{\text{Sing}}$, and *VOIOBS could be sandwiched between these two faithfulness constraints. The result is a language that permits voiced geminates but not voiced singletons, and no such language exists (Hayes & Steriade 2004). Therefore, IDENT(+voi)$_{\text{Sing}}$ must be universally ranked either as high as or higher than IDENT(+voi)$_{\text{Gem}}$. This conclusion implies that the perceptibility of [+voice] can never be more salient in geminates than in singletons.

To summarize, the P-map hypothesis makes two testable predictions: (i) in languages where [+voice] is equally perceptible in singletons and in geminates (of which Arabic may be an example), IDENT(+voi)$_{\text{Gem}}$ is ranked as high as FAITH(+voi)$_{\text{Sing}}$, and (ii) [+voice] is never more perceptible in geminates than in singletons. Whether these predictions are borne out should be tested crosslinguistically by way of experimentation, but this task is left for future research.

6.3. OTHER ISSUES FOR FUTURE RESEARCH. In addition to the issues discussed above, two other issues are raised here for future research. One is wider testing of the predictions of the P-map hypothesis. I have shown that a faithfulness ranking can indeed reflect a perceptibility scale, and that this can be verified experimentally. Therefore, other faithfulness scales that are claimed to be grounded in perceptibility scales can and should be tested experimentally. This includes various faithfulness scales proposed in the original P-map works of Steriade (2001a,b) and elsewhere (Adler 2006, Howe & Pulleyblank 2004, Kawahara 2006b, Padgett 2002, Zuraw 2005).

Another issue is to investigate how other faithfulness dimensions interact with a geminacy distinction. It is possible that faithfulness for featural dimensions other than [±voice] is governed by a different set of constraints for singletons and geminates. To the extent that the general theme advanced here is correct, it predicts that for a contrast that is more reliably perceived in geminates, unlike the case of

[±voice] in Japanese, the faithfulness constraint for geminates could be ranked higher than the one for singletons. Whether this prediction is borne out remains to be tested.

6.4. Overall conclusion. I have argued here that in the loanword phonology of Japanese, voiced geminates are more prone to categorical devoicing than voiced singletons are. Further, I have claimed that this observation requires differentiation of Ident (+voi) into two kinds, Ident(+voi)$_{Sing}$ and Ident(+voi)$_{Gem}$, and that they are ranked as Ident(+voi)$_{Sing}$ » Ident(+voi)$_{Gem}$. The P-map hypothesis predicts that this ranking originates from the different perceptibility of [+voice] in singletons and geminates. Experimentation demonstrated that [+voice] is indeed less perceptible in geminates than in singletons, and that the lesser perceptibility of [+voice] in geminates is likely to be the cause of the low ranking of Ident(+voi)$_{Gem}$. I have thus provided empirical support for the P-map hypothesis, according to which a faithfulness ranking can be projected from a perceptibility scale.

Appendix: Acoustic values of tokens used in experiment 2.

Speaker 1	voicing duration (ms)	closure duration (ms)	V1 duration (ms)	F0 at V1 (Hz)	F1 at V1 (Hz)	F0 at V2 (Hz)	F1 at V2 (Hz)
kapa	15	54	29	286	825	311	768
kepe	15	72	35	309	533	308	546
kopo	19	63	25	290	594	306	574
kaba	57	57	39	297	714	284	723
kebe	45	45	38	301	495	298	530
kobo	45	45	31	292	420	296	502
kappa	16	132	50	295	809	259	796
keppe	12	127	52	289	542	273	539
koppo	17	133	52	304	590	278	519
kabba	31	93	50	271	811	266	742
kebbe	36	113	72	308	571	267	512
kobbo	44	123	55	285	535	255	513
kata	17	62	30	300	660	320	621
kete	17	63	42	311	516	317	515
koto	24	63	38	295	443	309	489
kada	61	61	43	273	511	288	575
kede	46	46	51	283	354	284	454
kodo	43	43	38	283	421	288	462
katta	2	134	62	299	595	261	555
kette	29	159	62	290	431	312	480
kotto	24	146	62	288	471	267	475
kadda	41	105	74	295	601	284	588
kedde	37	121	76	285	464	266	448
koddo	45	149	79	169	428	274	457
kaka	8	39	33	290	613	321	781
keke	8	42	43	298	338	346	418
koko	2	55	41	295	458	331	538
kaga	36	36	42	286	483	279	630
kege	43	43	73	285	296	287	321
kogo	51	51	58	308	421	288	481
kakka	9	129	63	308	699	263	790
kekke	5	115	69	334	374	284	399
kokko	15	120	51	327	466	273	531
kagga	34	89	61	304	609	302	592
kegge	52	130	83	287	333	285	392
koggo	40	92	52	304	471	265	527

Speaker 2	Voicing duration (ms)	Closure duration (ms)	V1 duration (ms)	F0 at V1 (Hz)	F1 at V1 (Hz)	F0 at V2 (Hz)	F1 at V2 (Hz)
kapa	6	74	33	270	439	283	569
kepe	17	81	52	248	449	282	498
kopo	21	66	27	250	480	266	482
kaba	43	43	31	257	561	207	547
kebe	46	46	21	244	386	259	429
kobo	46	46	34	255	504	258	502
kappa	17	115	53	259	532	249	542
keppe	16	136	59	251	494	243	505
koppo	14	131	56	264	484	254	478
kabba	30	139	83	354	505	219	574
kebbe	38	141	84	263	410	247	471
kobbo	42	120	51	249	488	237	454
kata	0	62	30	259	595	276	582
kete	11	88	50	255	459	254	442
koto	8	55	41	258	512	284	511
kada	26	26	40	258	600	281	538
kede	38	38	63	252	450	252	450
kodo	32	32	50	278	502	273	460
katta	21	107	58	259	600	284	564
kette	14	115	68	268	536	252	472
kotto	0	144	42	257	494	277	469
kadda	40	103	74	266	613	260	528
kedde	35	93	98	262	515	249	457
koddo	24	102	90	270	523	238	465
kaka	10	70	42	255	543	278	565
keke	9	52	49	251	485	297	472
koko	0	51	58	249	431	276	512
kaga	41	41	60	253	503	268	526
kege	39	39	85	254	339	257	370
kogo	52	52	55	244	360	249	319
kakka	7	126	54	260	541	252	548
kekke	0	130	59	241	462	240	456
kokko	9	140	52	260	467	246	457
kagga	41	139	98	242	418	212	504
kegge	46	122	73	243	424	225	427
koggo	44	122	88	261	354	239	403
Speaker 3							
kapa	2	95	33	295	828	275	820
kepe	13	77	36	344	641	335	610
kopo	23	76	26	333	642	339	488
kaba	53	53	47	313	933	320	906
kebe	50	50	64	337	623	339	593
kobo	46	46	39	321	466	338	440
kappa	19	147	47	287	855	268	776
keppe	11	149	56	328	573	286	536
koppo	10	136	61	330	388	310	553
kabba	23	121	63	300	879	259	769
kebbe	37	87	104	234	661	294	409
kobbo	36	115	85	328	670	252	474
kata	0	81	25	288	848	270	810
kete	13	61	52	318	501	335	523
koto	11	62	42	336	557	346	597
kada	33	33	42	305	857	299	730
kede	38	38	62	304	443	314	400
kodo	31	31	71	275	491	266	495

SPEAKER 3	VOICING DURATION (ms)	CLOSURE DURATION (ms)	V1 DURATION (ms)	F0 AT V1 (Hz)	F1 AT V1 (Hz)	F0 AT V2 (Hz)	F1 AT V2 (Hz)
katta	7	142	34	314	893	275	795
kette	8	130	66	328	461	307	482
kotto	0	138	66	317	470	295	566
kadda	35	121	39	307	917	265	820
kedde	34	116	79	318	377	264	477
koddo	35	116	74	349	647	281	525
kaka	0	56	34	343	666	359	951
keke	0	56	47	335	442	343	406
koko	11	66	46	340	566	356	618
kaga	40	40	59	307	594	324	761
kege	29	29	88	299	411	286	410
kogo	44	44	58	280	403	272	501
kakka	0	120	38	320	777	288	907
kekke	11	124	61	353	402	336	419
kokko	12	123	59	322	443	306	533
kagga	37	116	76	323	681	262	665
kegge	27	77	89	335	771	317	428
koggo	43	132	84	328	409	277	511

REFERENCES

ADLER, ALLISON N. 2006. Faithfulness and perception in loanword adaptation: A case study from Hawaiian. *Lingua* 116.1024–45.

ALDERETE, JOHN. 1997. Dissimilation as local conjunction. *North Eastern Linguistic Society* 27.17–32.

AMANO, SHIGEAKI, and TADAHISA KONDO. 2000. *NTT database series: Lexical properties of Japanese: The 2nd release*. Tokyo: Sanseido.

ANTTILA, ARTO. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20.1–42.

ARCHANGELI, DIANA, and DOUGLAS PULLEYBLANK. 1994. *Grounded phonology*. Cambridge, MA: MIT Press.

BAKOVIĆ, ERIC. 2000. *Harmony, dominance and control*. New Brunswick, NJ: Rutgers University dissertation.

BECKMAN, JILL. 1998. *Positional faithfulness*. Amherst: University of Massachusetts, Amherst dissertation.

BENKÍ, JOSÉ. 2003. Analysis of English nonsense syllable recognition in noise. *Phonetica* 60.129–57.

BOERSMA, PAUL. 1998. *Functional phonology: Formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

BOERSMA, PAUL, and DAVID WEENINK. 2005. Praat (version 4.3.31). Amsterdam: University of Amsterdam. Online: http://www.fon.hum.uva.nl/praat.

BROWMAN, CATHERINE, and LOUIS GOLDSTEIN. 1989. Articulatory gestures as phonological units. *Phonology* 6.201–52.

CASALI, RODERIC. 1997. Vowel elision in hiatus contexts: Which vowel goes? *Language* 73.493–533.

CHEN, MATTHEW. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22.129–59.

CHOMSKY, NOAM. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

CHOMSKY, NOAM, and MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper and Row.

COHN, ABIGAIL. 1993. Nasalization in English: Phonology or phonetics. *Phonology* 10.43–81.

COHN, ABIGAIL; WILLIAM HAM; and ROBERT PODESVA. 1999. The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. *International Conference of Phonetic Sciences* 99.587–90.

Côté, Marie-Hélène. 2004. Syntagmatic distinctness in consonant deletion. *Phonology* 21.1–41.

Fleischhacker, Heidi. 2001. Cluster-dependent epenthesis asymmetries. *UCLA Working Papers in Linguistics* 7.71–116.

Flemming, Edward. 1995. Phonetic detail in phonology: Evidence from assimilation and coarticulation. *Proceedings of the 1995 Southwestern Workshop on Optimality Theory*, ed. by Keiichiro Suzuki and Dirk Elzinga, 39–50. Tucson: University of Arizona.

Fougeron, Cecile, and Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 106.3728–40.

Fujimura, Osamu; Marian Macchi; and L. A. Streeter. 1978. Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language and Speech* 21.337–46.

Gafos, Adamantios. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20.269–337.

Gelbart, Ben, and Shigeto Kawahara. 2005. Psychological reality of sublexica in Japanese. Amherst: University of Massachusetts, Amherst, ms.

Goldsmith, John. 1976. *Autosegmental phonology*. Cambridge, MA: MIT dissertation.

Guion, Susan. 1998. The role of perception in the sound change of velar palatalization. *Phonetica* 55.18–52.

Han, Mieko. 1994. Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* 96.73–82.

Haraguchi, Shosuke. 2006. A theory of voicing. Urayasu: Meikai University, ms.

Hay, Jennifer; Janet Pierrehumbert; and Mary Beckman. 2003. Speech perception, well-formedness, and the statistics of the lexicon. *Papers in laboratory phonology 6*, ed. by John Local, Richard Ogden, and Rosalind Temple, 58–74. Cambridge: Cambridge University Press.

Hayes, Bruce; Robert Kirchner; and Donca Steriade (eds.) 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.

Hayes, Bruce, and Donca Steriade. 2004. Introduction: The phonetic bases of phonological markedness. In Hayes et al., 1–33.

Hooper, Joan Bybee. 1976. *An introduction to natural generative phonology*. New York: Academic Press.

Howe, Darin, and Douglas Pulleyblank. 2004. Harmonic scales as faithfulness. *Canadian Journal of Linguistics* 49.1–49.

Hura, Susan; Björn Lindblom; and Randy Diehl. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35.59–72.

Itô, Junko, and Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17.49–73.

Itô, Junko, and Armin Mester. 1995. Japanese phonology. *The handbook of phonological theory*, ed. by John Goldsmith, 817–38. Oxford: Blackwell.

Itô, Junko, and Armin Mester. 1997. German truncation and sympathy theory. *University of Maryland Working Papers in Linguistics* 5.117–39.

Itô, Junko, and Armin Mester. 1998. Markedness and word structure: OCP effects in Japanese. Santa Cruz: University of California, Santa Cruz, ms.

Itô, Junko, and Armin Mester. 1999. The phonological lexicon. *The handbook of Japanese linguistics*, ed. by Natsuko Tsujimura, 62–100. Oxford: Blackwell.

Itô, Junko, and Armin Mester. 2001. Structure preservation and stratal opacity in German. *Segmental phonology in optimality theory*, ed. by Linda Lombardi, 261–95. Cambridge: Cambridge University Press.

Itô, Junko, and Armin Mester. 2003. *Japanese morphophonemics*. Cambridge, MA: MIT Press.

Jaeger, Jeri. 1978. Speech aerodynamics and phonological universals. *Chicago Linguistic Society* 14.311–29.

Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell.

Jun, Jongho. 2004. Place assimilation. In Hayes et al., 58–86.

Katayama, Motoko. 1998. *Loanword phonology in Japanese and optimality theory*. Santa Cruz: University of California, Santa Cruz dissertation.

KAWAHARA, SHIGETO. 2005. Voicing and geminacy in Japanese: An acoustic and perceptual study. *University of Massachusetts Occasional Papers in Linguistics* 31.87–120.

KAWAHARA, SHIGETO. 2006a. Sonorancy and geminacy. *University of Massachusetts Occasional Papers in Linguistics* 32, to appear.

KAWAHARA, SHIGETO. 2006b. Half rhymes in Japanese rap lyrics and knowledge of similarity. *Journal of East Asian Linguistics*, to appear.

KEATING, PATRICIA. 1996. The phonology-phonetics interface. *UCLA Working Papers in Phonetics* 92.45–60.

KINGSTON, JOHN, and RANDY DIEHL. 1994. Phonetic knowledge. *Language* 70.419–55.

KIRCHNER, ROBERT. 1997. Constrastiveness and faithfulness. *Phonology* 14.83–111.

KOHLER, KLAUS. 1979. Dimensions in the perception of fortis and lenis plosives. *Phonetica* 36.332–43.

KOHLER, KLAUS. 1990. Segmental reduction in connected speech: Phonological facts and phonetic explanation. *Speech production and speech modeling*, ed. by William Hardcastle and Alain Marchals, 69–92. Dordrecht: Kluwer.

KURODA, S-Y. 1965. *Generative grammatical studies in the Japanese language*. Cambridge, MA: MIT dissertation.

LEBEN, WILL. 1973. *Suprasegmental phonology*. Cambridge, MA: MIT dissertation.

LESLAU, WOLF. 1976. The triradicals in the Gurage dialect of Endegeň. *Israel Oriental Studies* 14.96–109.

LISKER, LEIGH. 1957. Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* 33.42–49.

LISKER, LEIGH. 1978. On buzzing the English /b/. *Haskins Laboratories Status Report on Speech Research* SR-55/56.251–59.

LISKER, LEIGH. 1981. On generalizing the rapid-rabid distinction based on silent gap duration. *Haskins Laboratories Status Report on Speech Research* SR-65.251–59.

LISKER, LEIGH. 1986. 'Voicing' in English: A catalog of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29.3–11.

LISKER, LEIGH, and ARTHUR ABRAMSON. 1964. A cross-linguistic study of voicing in initial stops: Acoustical measurements. *Word* 20.384–422.

LOVINS, JULIE. 1973. *Loanwords and the phonological structure of Japanese*. Chicago: University of Chicago dissertation.

MACMILLAN, NEIL, and DOUGLAS CREELMAN. 2005. *Detection theory: A user's guide*. 2nd edn. Mahwah, NJ: Lawrence Erlbaum.

MCCARTHY, JOHN. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17.207–63.

MCCARTHY, JOHN. 1999. Sympathy and phonological opacity. *Phonology* 16.331–99.

MCCARTHY, JOHN. 2003. Comparative markedness. *Theoretical Linguistics* 29.1–52.

MCCARTHY, JOHN, and ALAN PRINCE. 1986. Prosodic morphology. Amherst: University of Massachusetts, and New Brunswick, NJ: Rutgers University, MS.

MCCARTHY, JOHN, and ALAN PRINCE. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers in Linguistics* 18.249–384.

MILLER, GEORGE, and PATRICIA NICELY. 1955. An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* 27.338–52.

MIURA, AKIRA. 1993. *English in Japanese*. New York: Weatherhill.

MORETON, ELLIOT. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84.55–71.

MORETON, ELLIOT, and SHIGEKI AMANO. 1999. Phonotactics in the perception of Japanese vowel length: Evidence for long-distance dependencies. *EUROSPEECH '99*.2679–82.

MYERS, SCOTT. 1997. Expressing phonetic naturalness in phonology. *Derivations and constraints in phonology*, ed. by Iggy Roca, 125–52. Oxford: Oxford University Press.

MYERS, JEROME, and ARNOLD WELL. 2003. *Research design and statistical analysis*. Mahwah, NJ: Lawrence Erlbaum.

NISHIMURA, KOHEI. 2003. Lyman's law in loanwords. Tokyo: University of Tokyo, MS.

ODDEN, DAVID. 1986. On the role of the obligatory contour principle. *Language* 62.353–83.

OHALA, JOHN. 1983. The origin of sound patterns in vocal tract constraints. *The production of speech*, ed. by Peter MacNeilage, 189–216. New York: Springer.

PADGETT, JAYE. 2002. Constraint conjunction versus grounded constraint hierarchy in optimality theory. Santa Cruz: University of California, Santa Cruz, MS.

PADGETT, JAYE. 2006. Russian voicing assimilation, final devoicing, and the problem of [v]. *Natural Language and Linguistic Theory*, to appear.

PARKER, ELLEN; RANDY DIEHL; and KEITH KLUENDER. 1986. Trading relations in speech and non-speech. *Perception and Psychophysics* 39.129–42.

PATER, JOE. 1999. Austronesian nasal substitution and other NC effects. *The prosody-morphology interface*, ed. by Rene Kager, Harry van der Hulst, and Wim Zonneveld, 310–43. Cambridge: Cambridge University Press.

PIERREHUMBERT, JANET. 1990. Phonological and phonetic representation. *Journal of Phonetics* 18.375–94.

PORT, ROBERT, and JOHN DALBY. 1982. Consonant/vowel ratio as a cue for voicing in English. *Perception and Psycholinguistics* 32.141–52.

PRINCE, ALAN, and PAUL SMOLENSKY. 2004. *Optimality theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

RAPHAEL, LAWRENCE. 1981. Duration and contexts as cues to word-final cognate opposition in English. *Phonetica* 38.126–47.

SANDERS, NATHAN. 2003. *Opacity and sound change in the Polish lexicon*. Santa Cruz: University of California, Santa Cruz dissertation.

SLIS, IMAS. 1986. Assimilation of voice in Dutch as a function of stress, word boundaries and sex of speaker and listener. *Journal of Phonetics* 14.311–26.

SMITH, JENNIFER. 2006. Loan phonology is not all perception: Evidence from Japanese loan doublets. *Japanese/Korean Linguistics 14*, ed. by Timothy J. Vance and Kimberly A. Jones, 63–74. Stanford, CA: CSLI Publications.

SMOLENSKY, PAUL. 1995. On the structure of the constraint component Con of UG. Paper presented at University of California, Los Angeles.

SPAELTI, PHILIP. 1997. *Dimensions of variation in multi-pattern reduplication*. Santa Cruz: University of California, Santa Cruz dissertation.

STAMPE, DAVID. 1973. *A dissertation on natural phonology*. Chicago: University of Chicago dissertation.

STERIADE, DONCA. 1995. Positional neutralization. Los Angeles: University of California, Los Angeles, MS.

STERIADE, DONCA. 1997. Phonetics in phonology: The case of laryngeal neutralization. Los Angeles: University of California, Los Angeles, MS.

STERIADE, DONCA. 2000. Paradigm uniformity and the phonetics/phonology boundary. *Papers in laboratory phonology 5*, ed. by Michael Broe and Janet Pierrehumbert, 313–34. Cambridge: Cambridge University Press.

STERIADE, DONCA. 2001a. Directional asymmetries in place assimilation. *The role of speech perception in phonology*, ed. by Elizabeth Hume and Keith Johnson, 219–50. San Diego: Academic Press.

STERIADE, DONCA. 2001b. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Los Angeles: University of California, Los Angeles, MS.

STEVENS, KENNETH, and SHEILA BLUMSTEIN. 1981. The search for invariant acoustic correlates of phonetic features. *Perspectives on the study of speech*, ed. by Peter Eimas and Joanne Miller, 1–38. Hillsdale, NJ: Erlbaum Associates.

TAKAGI, NAOYUKI, and VIRGINIA MANN. 1994. A perceptual basis for the systematic phonological correspondences between Japanese loan words and their English source words. *Journal of Phonetics* 22.343–56.

TSUCHIDA, AYAKO. 1995. English loans in Japanese: Constraints in loanword phonology. *Working Papers of the Cornell Phonetics Laboratory* 10.145–64.

TSUCHIDA, AYAKO. 1997. *Phonetics and phonology of Japanese vowel devoicing*. Ithaca, NY: Cornell University dissertation.

WALKER, RACHEL. 2001. Round licensing, harmony, and bisyllabic triggers in Altaic. *Natural Language and Linguistic Theory* 19.827–78.

WEHR, HANS. 1971. *A dictionary of modern written Arabic*. 3rd edn. New York: Spoken Language Services.

WESTBURY, JOHN. 1979. *Aspects of the temporal control of voicing in consonant clusters in English*. Austin: University of Texas, Austin dissertation.

ZHANG, JIE. 2000. Non-contrastive features and categorical patterning in Chinese diminutive suffixation: Max[F] or Ident[F]? *Phonology* 17.427–78.

ZHANG, JIE. 2004. The role of contrast-specific and language-specific phonetics in contour tone distribution. In Hayes et al., 157–90.
ZSIGA, ELIZABETH. 1995. An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. *Papers in laboratory phonology 4*, ed. by Bruce Connell and Amalia Arvaniti, 282–302. Cambridge: Cambridge University Press.
ZSIGA, ELIZABETH. 1997. Features, gestures, and Igbo vowels: An approach to the phonology-phonetics interface. *Language* 73.227–74.
ZURAW, KIE. 2005. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog reduplication. Los Angeles: University of California, Los Angeles, MS.

Department of Linguistics
226 South College
University of Massachusetts, Amherst
Amherst, MA 01003
[kawahara@linguist.umass.edu]

[Received 28 May 2004;
revision invited 11 February 2005;
revision received 11 May 2005;
accepted 8 November 2005]

Kawahara, Shigeto (2008) Phonetic naturalness and unnaturalness in Japanese loanword phonology. Journal of East Asian Linguistics 17(4): 317-330.

J East Asian Linguist (2008) 17:317–330
DOI 10.1007/s10831-008-9030-z

Phonetic naturalness and unnaturalness in Japanese loanword phonology

Shigeto Kawahara

Received: 14 April 2007 / Accepted: 6 July 2008 / Published online: 12 September 2008

Abstract This paper argues that phonetic naturalness and unnaturalness can interact within a single grammatical system. In Japanese loanword phonology, only voiced geminates, but not voiced singletons, devoice to dissimilate from another voiced obstruent. The neutralizability difference follows from a ranking which Japanese speakers created on perceptual grounds: IDENT(voi)$_{\text{Sing}}$ » IDENT(voi)$_{\text{Gem}}$. On the other hand, the trigger of devoicing—OCP(voi)—has no phonetic underpinning because voicing does not have phonetic characteristics that would naturally lead to confusion-based dissimilation (Ohala, Proceedings of Chicago Linguistic Society: Papers from the parasession on language and behaviour, 1981, in: Jones (ed.) Historical linguistics: Problems and perspectives, 1993). OCP(voi) in Modern Japanese originated as a phonetically natural OCP(prenasal) in Old Japanese because the spread out heavy nasalization would lead to perceptual confusion, but it divorced from its phonetic origin when prenasalization became voicing. The interaction of the three constraints in Modern Japanese suggests that phonetic naturalness (the ranking IDENT(voi)$_{\text{Sing}}$ » IDENT(voi)$_{\text{Gem}}$) and unnaturalness (OCP(voi)) co-reside within a single module.

Keywords Phonetic (un)naturalness · Perceptibility · Dissimilation

1 Introduction

This paper addresses the issues surrounding phonetic naturalness in phonology. Many phonological patterns seem to make phonetic sense. For example, many languages disfavor voiced stops, and the dispreference seems rooted in an aerodynamic

S. Kawahara (✉)
Linguistics Department, Rutgers University, The State University of New Jersey,
18 Seminary Place, New Brunswick, NJ 08901-1108, USA
e-mail: kawahara@rci.rutgers.edu

challenge. During a stop closure, intraoral air pressure rises, but the rise would make it difficult to maintain voicing because there needs to be a drop in transglottal air pressure in order to produce voicing (Ohala 1983).

Because many phonological patterns seem to be shaped by phonetic factors, some scholars have proposed that these phonetic factors are encoded in synchronic phonological grammars. In other words, phonetic imperatives—such as minimization of articulatory effort or maximization of perceptual distinctiveness—synchronically shape phonological behaviors (e.g., Boersma 1998; Flemming 1995; Stampe 1973; Steriade 2001).

Other proposals have attributed the naturalness of phonological patterns to phonetically natural diachronic sound changes and kept the synchronic phonology free of phonetics because synchronic phonology can include phonetically unnatural patterns (e.g., Anderson 1981; Bach and Harms 1972; Blevins 2004). The primary argument is that several phonetically natural sound changes can result in phonetically unnatural phonological patterns. For example, Icelandic had palatalization of velar stops before front vowels, a phonetically natural process, given tongue-body coarticulation (Keating and Lahiri 1993). However, Icelandic historically changed *[æ] to [ai], but the diphthong with a back vowel nucleus still triggers fronting (Anderson 1981). Thus the historical diphthongization of [æ] to [ai] yielded a phonetically unnatural pattern, i.e., palatalization triggered by a back vowel. In this view, phonetic naturalness descends only from sound changes, and synchronic phonology is phonetics-free.

I draw on both positions, each of which I believe captures genuine aspects of phonology. I argue that phonetically natural and unnatural patterns can coexist within a single module of a synchronic grammar. Formally, I propose that synchronic phonology is directly constrained by phonetic factors but nonetheless accommodates phonetically unnatural patterns. My proposal rests on a phonological analysis and experimental studies of devoicing of voiced geminates observed in Japanese loanword phonology (Kawahara 2006).

My arguments develop as follows. Section 2 presents patterns of voiced consonants in Japanese phonology and shows that voiced geminates are more devoicable than voiced singletons. In Sect. 3, I analyze the difference in devoicability within Optimality Theory (OT; Prince and Smolensky 2004). In Sect. 4, I show that neutralizability of voicing contrasts depends on the phonetically natural principle of perceptibility. In Sect. 5, I argue that the constraint that triggers devoicing is phonetically unnatural. I summarize the results in Sect. 6, and discuss broader implications of the study. In particular, although phonetically natural and unnatural patterns can coexist within a single grammar, proving that phonological grammars are flexible enough to accommodate both, synchronic phonology nevertheless seems biased toward phonetic naturalness.

2 Data: voiced consonants in Japanese phonology

This section presents the pattern of voiced consonants in Japanese phonology (Kawahara 2006). Native Japanese phonology prohibits voiced geminates; there are no native words like ***wabba*, ***wadda*, or ***wagga*. Potential instances of voiced geminates are resolved by nasalization of the first portion of the geminates. The suffix *-ri* induces gemination of root-final consonants, as in (1a), but when that process would geminate a voiced obstruent, the form surfaces instead with a homorganic nasal-obstruent sequence, as in (1b).

(1) Coda nasalization in Japanese

a. /tapu+μ+ri/	→	[ta**pp**uri]	*[ta**mp**uri]	'a lot of'
b. /zabu+μ+ri/	→	[za**mb**uri]	*[za**bb**uri]	'splashing'

However, voicing in geminates has become contrastive in recent loanwords. In loanwords, coda consonants—including voiced obstruents—which follow a lax vowel in the source language are often borrowed as geminates (Katayama 1998). Some near-minimal pairs of voiced and voiceless geminates in loanwords are given in (2). Henceforth, those words that contain voiced geminates but no other voiced obstruents are referred to as TVDDV words.

(2) Voicing in geminates is contrastive in loanwords

ha**bb**uru	'Hubble'	ka**pp**uru	'couple'
ki**dd**o	'kid'	ki**tt**o	'kit'
e**gg**u	'egg'	ne**kk**u	'neck'

In the TVDDV words shown in (2), voiced geminates do not devoice, suggesting that a voicing contrast is phonemic in geminates. However, Nishimura (2003) has pointed out that voiced geminates can undergo optional devoicing when they co-occur with another voiced obstruent, as illustrated in (3). These words, which contain voiced geminates and additional voiced obstruents, are referred to as DVDDV words in the subsequent discussion.

(3) Voiced geminates devoice when they appear with another voiced obstruent

ge**bb**erusu	~	ge**pp**erusu	'Göbbels'
gu**dd**o	~	gu**tt**o	'good'
do**gg**u	~	do**kk**u	'dog'

Devoicing of voiced geminates in DVDDV words is attributable to the Obligatory Contour Principle (OCP: Leben 1973) on voicing, which is a constraint against two voiced obstruents within the same stem. This constraint is also known as Lyman's Law, and it is visibly active in the prohibition of stems with two voiced obstruents in the native phonology; e.g., *fuda* 'amulet', ***b**uta* 'pig', *****b**u**d**a* (Itô and Mester 1986; Lyman 1894). In loanwords, however, singletons do not devoice—unlike geminates—even when they violate OCP(voi), as illustrated in (4) (henceforth referred to as DVDV words).

(4) Words with two singletons do not undergo devoicing

ba**g**ii	'buggy'	**b**o**g**ii	'bogey'
da**g**u	'Doug'	**d**ai**b**u	'dive'
gi**g**a	'giga (10^9)'	**g**a**b**urieru	'Gabriel'

In short, DVDDV words can undergo devoicing whereas DVDV words cannot: put differently, OCP(voi) can devoice only voiced geminates but not voiced singletons. The next section presents an analysis of why singletons and geminates differ in their devoicablity in Japanese loanword phonology.

One note is in order before proceeding: this paper focuses on the synchronic behavior of a [voi] feature in the sector of the lexicon that can be identified as loanwords by native speakers rather than the process of loanword adaptation. In other words, the analysis concerns what happens *after* speakers borrow new words rather than what happens *when* they borrow these words.

3 An OT analysis

I now turn to an analysis of the patterns summarized in Sect. 2, largely following the analysis presented in Kawahara (2006). To recast the generalization in (2)–(4), a voicing contrast is less stable in geminates than in singletons in that only voiced geminates can devoice under the duress of OCP(voi). To directly implement different neutralizability of singletons and geminates within OT (Prince and Smolensky 2004), we can separate the faithfulness constraint for [voi] into two IDENT(voi) constraints: IDENT(voi)$_{\text{Sing(leton)}}$ prohibits devoicing of singletons and IDENT(voi)$_{\text{Gem(inate)}}$ prohibits devoicing of geminates. In Japanese loanwords, IDENT(voi)$_{\text{Sing}}$ outranks IDENT(voi)$_{\text{Gem}}$. Section 4 justifies the ranking of the two constraints in terms of perceptibility of [voi] in Japanese singletons and geminates.

In addition to splitting IDENT(voi) into two constraints, I argue that the cause of devoicing in DVDDV words is a constraint against a pair of two voiced obstruents within a stem, which is well motivated in the native phonology of Japanese.[1] Following Itô and Mester (1986), I refer to this restriction as OCP(voi).

OCP(voi) and the two separate faithfulness constraints IDENT(voi)$_{\text{Sing}}$ and IDENT(voi)$_{\text{Gem}}$ can account for the patterns presented in (2)–(4). First, since singletons do not devoice due to OCP(voi) in DVDV words, IDENT(voi)$_{\text{Sing}}$ dominates OCP(voi), as shown by the tableau in (5).

(5) IDENT(voi)$_{\text{Sing}}$ » OCP(voi): No devoicing in DVDV words

/bagii/	IDENT(voi)$_{\text{Sing}}$	OCP(voi)
a. → [bagii]		*
b. [bakii]	*!	
c. [pagii]	*!	

In contrast, in DVDDV words, geminates can devoice. Those cases suggest that OCP(voi) can rank above IDENT(voi)$_{\text{Gem}}$, as shown in (6). To allow for the

[1] Since devoicing takes place in words like /deibiddo/ 'David' which contain three underlying voiced consonants, OCP(voi) should assign a violation mark for each pair of voiced consonants within a stem rather than for each stem that contains two voiced obstruents (Tesar 2007).

optionality of devoicing, OCP(voi) and IDENT(voi)$_{Gem}$ can be left unranked with respect to one another (Anttila and Cho 1998).

(6) OCP(voi) » IDENT(voi)$_{Gem}$: Geminates can devoice in DVDDV words

/baggu/	OCP(voi)	IDENT(voi)$_{Gem}$
a. [baggu]	* !	
b. → [bakku]		*

Finally, since voiced geminates surface as voiced unless OCP(voi) is relevant, IDENT(voi)$_{Gem}$ is ranked above the markedness constraint that prohibits voiced obstruent geminates, *VOIOBSGEM, as shown in (7).

(7) IDENT(voi)$_{Gem}$ » *VOIOBSGEM: Geminates do not devoice in TVDDV words

/eggu/	IDENT(voi)$_{Gem}$	*VOIOBSGEM
a. → [eggu]		*
b. [ekku]	*!	

To summarize, the ranking IDENT(voi)$_{Sing}$ » OCP(voi) » IDENT(voi)$_{Gem}$ » *VOIOBSGEM accounts for all of the patterns of [voi] in the loanword phonology of Japanese. The next section justifies positing *two* IDENT(voi) constraints and makes the case for their ranking.

4 Naturalness: phonetically-determined devoicability threshold

4.1 Neutralizability based on perceptibility

Based on the behavior of [voi] in singletons and geminates with respect to OCP(voi), I have argued in Sect. 3 that Japanese loanword phonology exhibits the ranking IDENT(voi)$_{Sing}$ » IDENT(voi)$_{Gem}$. It is worth emphasizing that nothing in the native phonology of Japanese motivates the ranking between IDENT(voi)$_{Sing}$ and IDENT(voi)$_{Gem}$. The native phonology resolves voiced geminates by nasalizing the initial portion of the geminates, not by devoicing (e.g., /zabu+µ+ri/ → [za**mb**uri], *[za**pp**uri], as shown in (1b)). Therefore, prior to loanword adaptation, Japanese speakers have not seen voiced geminates, let alone devoicing of voiced geminates. Given that, why would Japanese speakers devoice only voiced geminates due to OCP(voi) in loanwords?

I argue that the ranking of IDENT(voi)$_{Sing}$ » IDENT(voi)$_{Gem}$ derives from a perceptibility scale of voicing contrasts. I propose that speakers project faithfulness rankings that go beyond what can be inferred from their native phonology. For the case at hand, Japanese speakers know that a voicing contrast is less perceptible in geminates than in singletons and accordingly project the ranking IDENT(voi)$_{\text{Sing}}$ » IDENT(voi)$_{\text{Gem}}$. More generally, speakers exert stronger grammatical pressure against larger perceptual disparities. In OT we can formalize that behavior by making the perceptibility of a phonological alternation correlate with the ranking of a faithfulness constraint that it violates (Steriade 2001); from the perceptibility scale x > y > z, we project the constraint ranking IDENT(x) » IDENT(y) » IDENT(z). To maintain this hypothesis, we must show that a voicing contrast is less perceptible in geminates than in singletons.

4.2 Experimental evidence of unequal perceptibility

Both acoustic and perceptual evidence suggests that a voicing contrast is less perceptible in geminates than in singletons, at least in Japanese.[2] Acoustically, voiced geminates in Japanese are partially devoiced. An experiment reported in Kawahara (2005, 2006) based on the recording of three native speakers shows that voiced geminates in Japanese are partially devoiced whereas voiced singletons are voiced throughout the closure. Figure 1 shows respective closure voicing of singletons and geminates.

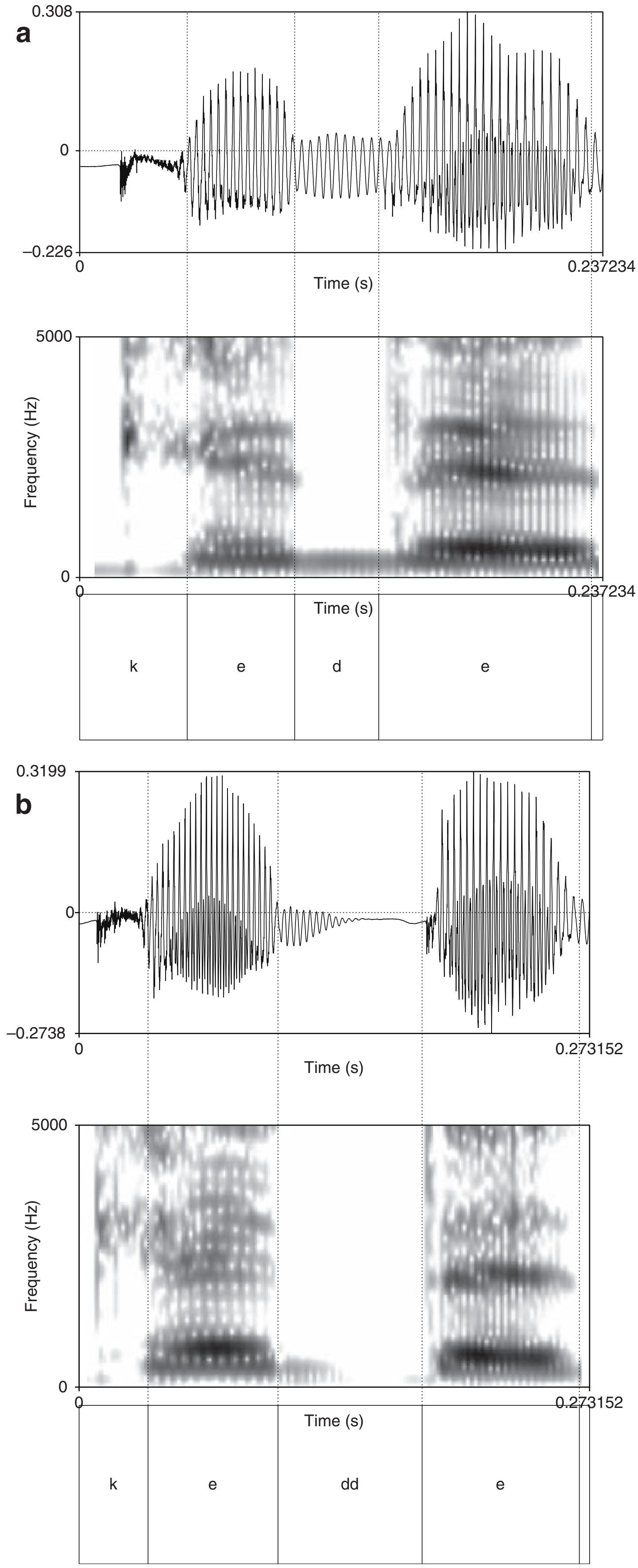
To quantify the general degree of partial devoicing, the proportion of closure voicing duration to closure duration was calculated based on the average of all three speakers. The result shows that while closure voicing is fully maintained in singletons, on average only the first 40% of the entire closure shows voicing in geminates. Considering that closure voicing is one of the most important cues signaling [+voi] (Lisker 1978; Raphael 1981; see Sect. 5.1 below), we expect the partial devoicing to undermine the percept of [+voi] in geminates.

A perceptual experiment reported in Kawahara (2005, 2006) confirms that a voicing contrast is less perceptible in geminates than in singletons. The experiment was a forced-choice identification task with 17 native speakers of Japanese, using naturally produced tokens covered by multi-layered cocktail party noise. The participants identified the voicing quality of [p,t,k,b,d,g,pp,tt,kk,bb,dd,gg], each of which was repeated 72 times. Based on the identification results obtained, I calculated a sensitivity measure (d') for the voicing contrast in singletons and geminates. I used d' rather than a percent correct analysis because d' directly represents the perceptibility of contrasts (Macmillan and Creelman 2005).

The average d' for the voicing contrast in singletons across all of the 17 listeners is 3.79, which significantly differs from zero ($t(16) = 34.15$, $p < .001$). The average d' for a voicing contrast in geminates is .71, which also significantly differs from zero ($t(16) = 11.47$, $p < .001$). These results show that for Japanese listeners, a voicing contrast is perceptible in both singletons and geminates, even under noisy environments. However, the perceptibility of a voicing contrast is much higher for singletons than for geminates; a paired t-test comparing d' for singletons and

[2] To the extent that the low perceptibility of a voicing contrast in geminates is due to partial devoicing, the devoicing in DVDDV words is a case in which language-particular phonetics affect phonological patterns since partial devoicing is specific to Japanese, e.g., it is not observed in Arabic. However, it is unlikely that a voicing contrast can ever be more perceptible in geminates than in singletons because we then incorrectly predict the existence of a language which permits only voiced geminates, not voiced singletons (see Kawahara 2006, Sect. 6.2).

Fig. 1 Waveforms and spectrograms of a singleton [d] (**a**) and geminate [dd] (**b**) (based on Kawahara 2005, 2006)



geminates reveals a significant difference ($t(16) = 27.27$, $p < .001$). The lower perceptibility of a voicing contrast in geminates supports the hypothesis that a phonological contrast more easily neutralizes when it is less perceptible.

The results of the perceptual experiment also reveal that the perceptibility of a voicing contrast in geminates depends on the place of articulation: $d'_{\text{bb-pp}}(.82) > d'_{\text{dd-tt}}(.64) > d'_{\text{gg-kk}}(.15)$. The relation between perceptibility and the place of articulation is at least partially reflected in the likelihood of optional, phonological devoicing of voiced geminates in DVDDV words (shown in (3)). Only one DVDDV word contains /bb/ (*ge**bb**erusu*), so it is difficult to make any conclusive generalizations about /bb/. On the other hand, in Nishimura's (2003) data, the likelihood of optional devoicing for /dd/ and /gg/ is /dd/ = .15 < /gg/ = .24. In other words, the likelihood of devoicing (/dd/ = .15 < /gg/ = .24) negatively correlates with the perceptibility ($d'_{\text{dd-tt}} = .64 > d'_{\text{gg-kk}} = .15$): the more perceptible a voicing contrast is, the less likely it is to be neutralized.

These experimental results suggest that the higher devoicability of geminates is rooted in their low perceptibility. Neutralizability and perceptibility should automatically correlate in a novel loanword phonological pattern if speakers avoid neutralizing a highly perceptible contrast (Steriade 2001). In other words, the ranking IDENT(voi)$_{\text{Sing}}$ » IDENT(voi)$_{\text{Gem}}$ has the phonetically natural basis of a perceptibility scale.

5 Unnaturalness: OCP(voi)[3]

Having justified the ranking of IDENT(voi)$_{\text{Sing}}$ » IDENT(voi)$_{\text{Gem}}$ on perceptual grounds, we now turn to the discussion of the trigger of devoicing, OCP(voi).

5.1 Unnatural status of OCP(voi)

OCP(voi) is a phonetically unnatural constraint because voicing does not have phonetic characteristics that would naturally lead to dissimilation. Ohala (1981, 1993) proposes that dissimilative principles—phonologically expressed as OCP—derive from potential perceptual confusions. Dissimilation takes place between segments that have acoustic correlates that are 'stretched out' or extended due to coarticulation and can therefore be perceived in positions distant from the feature's original host. Some examples of features involving spread-out acoustic correlates are rhoticity, nasalization, and glottalization. Dissimilation can arise when listeners assume that only one segment is specified of that feature because positing one segment alone can explain the spread-out acoustic correlates. Dissimilation would also arise if speakers actively avoid such perceptually confusable configurations (Boersma 1998; Flemming 1995; Liljencrants and Lindblom 1972). Regardless of whether dissimilation is initiated by listeners' innocent hypercorrection or speakers' active avoidance of confusable configurations, perceptual confusions serve as a seed

[3] This section builds on Ohala (1981, 1993). The hypothesis that multiple tokens of voiced consonants do not cause perceptual confusion needs to be tested experimentally, but I leave this task for future research.

for dissimilation, and therefore only features that have stretched-out acoustic correlates are predicted to dissimilate.[4]

If perceptual confusion serves as a seed for a dissimilative requirement, a voicing contrast should not dissimilate because acoustic correlates for a voicing contrast are not stretched out. Ohala (1993, pp. 253–254) explicitly states that grounding dissimilation on potential perceptual confusion predicts

> that [voice] should not be subject to dissimilation. Although one might conclude that there are 'prosodic' or long time-window cues to voicing…in fact there is no evidence supporting such claims…[T]he primary cue to a segment being voiced is the generally robust cue of periodic pulsation in the lower frequencies. This cue operates in a relatively short time-window and does not manifest itself by colouration of adjacent segments; therefore it should not be susceptible to dissimilation.

Some studies have shown that acoustic correlates of a voicing contrast do spread out to some extent, affecting the F0 and F1 of surrounding vowels (Hawkins and Nguyen 2004; van Summers 1988). However, the spread-out acoustic correlates do not seem substantial enough to cause perceptual confusion and hence would not yield dissimilation. First of all, spread-out acoustic correlates of a voicing contrast may not be consistently present. In Japanese, for example, F0 and F1 differences in the steady state of the preceding and following vowels—potential instances of spread cues—are small and subject to large inter-speaker variability (Kawahara 2005).

Second and more importantly, the perceptual impact of spread-out acoustic correlates of a voicing contrast is weak at best. Some evidence shows that closure voicing constitutes the primary cue for a voicing contrast. First, Raphael (1981) found that the absence of closure voicing can cause a strong [−voi] percept in English listeners even when other cues favor a [+voi] percept and concluded that closure voicing is the dominant cue for voicing. Second, Lisker (1978) has shown that consonants with 120 ms closure duration and 40 ms closure voicing are perceived by English listeners as voiceless about 70% of the time even when other cues such as preceding vowel duration favor a [+voi] percept. Third, my perception experiment, discussed above, shows that once closure voicing is diminished in voiced geminates, so is the perceptibility of [+voi]. Finally, some previous studies demonstrate that spread-out correlates of voicing—F0 and F1—visibly affect listeners' categorization only when the dominant cue is ambiguous (Abramson and Lisker 1985; Whalen et al. 1990).

Taken together, these pieces of evidence indicate that the effect of potentially spread-out acoustic correlates of the [voi] contrast is perceptually weak and unreliable,

[4] A phonological characteristic of dissimilative requirements suggests an important caveat for the hypothesis that dissimilation arises from perceptual confusion. Some cases of dissimilation are non-local and apply in a domain as large as words (e.g., Old Japanese; see Unger 1975), but such long-distance misperception is unlikely. Therefore, given a perceptually confusable configuration, speakers may generalize the restriction due to local perceptual confusion to a larger domain. Thus, ''phonetically grounded constraints'' have original phonetic motivations—such as avoidance of perceptually confusable configurations—but also involve abstract generalizations, such as extension of their domains.

and the internal cue—closure voicing—primarily determines the voicing percept. It is thus difficult to argue that multiple tokens of voiced obstruents within a domain have any particular vulnerability to perceptual confusion. Therefore, the presence of OCP(voi) in any phonology is unexpected, given the phonetic grounding of dissimilation.[5]

5.2 A natural origin of OCP(voi) in Japanese

Where does OCP(voi) come from, if not from a phonetic source of dissimilation? The answer lies in the history of Japanese voiced obstruents: voiced obstruents formerly involved a contrast that had spread-out cues. Several pieces of evidence indicate that the voicing contrast in Modern Japanese was a prenasalization contrast in Old Japanese (Unger 1975; Vance 2005). The historical prenasalized status of modern voiced obstruents is evidenced first by the transcriptions of the 17th century missionary Rodriguez in *Arte da Lingoa de Japam* (Rodriguez 1930); some Korean and Chinese transliterations dating to the 14th century also mark voiced obstruents in Japanese with nasality. Second, some conservative dialects (such as Tosa or Tôhoku dialects) still preserve the prenasalization.

Some evidence also shows that the prenasalization in Old Japanese exhibited spread-out acoustic correlates because of coarticulatory nasalization. Rodriguez (1930, p. 637) writes that prenasalization realized itself as partial or full nasalization of the preceding vowel. In fact, prenasalization is cross-linguistically extensive almost to the degree that the whole closure can be nasalized and only the release is oral (Huffman 1993), presumably because nasalization needs extended temporal span for it to be perceived (Whalen and Beddor 1989). Given [^{n}dV^{n}d] sequences, the vowel can then be nasalized right up to the release of the first stop by coarticulation, effectively making that release nasal. Thus the spread-out heavy nasalization makes the prenasalization contrast a plausible candidate for dissimilation because the heavy nasalization can serve as a seed for perceptual confusion (Ohala 1981, 1993).[6,7]

We also have evidence that the OCP affected the phonology of Old Japanese when the contrast involved prenasalization rather than voicing (Unger 1975; Vance

[5] Frisch (2004) proposes that dissimilative restrictions on homorganic consonants found in Arabic and other languages have their functional root in difficulty in serialization. However, the evidence that multiple tokens of non-homorganic voiced obstruents cause a serialization problem is yet to be found. Given the previous findings that dominant cues for a voicing contrast are internal cues and also given the observation that voicing dissimilation can arise only through a diachronic change, it seems reasonable to maintain the hypothesis that voicing dissimilation is phonetically unnatural. Future experiments should address the question of whether two occurrences of voiced obstruents present any challenge to listeners.

[6] Languages that exhibit dissimilation in (pre)nasalization include Muna (Coetzee and Pater 2005), Gurindji (Evans 1995, p. 733), and Takelma (Sapir 1912, pp. 45–46) among others. However, the existence of these patterns per se is not direct evidence that dissimilation in prenasalization is phonetically natural because dissimilation in voicing is arguably phonetically unnatural, but it does exist. To the extent that these patterns spontaneously emerged—as does devoicing of voiced geminates in Japanese loanword phonology—they do provide evidence for phonetic naturalness of dissimilation in nasalization.

[7] Prenasalization may be more vulnerable to perceptual confusion than full nasal segments. Prenasalized segments have an oral release, and thus to cue nasality, spread nasalization in preceding vowels should play a more important role for prenasalized segments than for full nasal segments.

2005). In sum, OCP(voi) in Modern Japanese originated as OCP(prenas) in Old Japanese, which was a phonetically natural constraint. However, when the prenasalization contrast became a voicing contrast, OCP(prenas) was divorced from its original natural phonetic motivation. Configurations with two voiced consonants, since they have no robust spread-out cues, do not present a perceptual confusability problem in the way that two prenasalized consonants do.

5.3 Cross-linguistic look at OCP(voi)

As discussed above, Japanese OCP(voi) originated as OCP(presnas), a phonetically natural constraint. Then what about other cases of voicing dissimilation? Beyond Japanese, two cases of voicing dissimilation are known, but they also originated as a phonetically natural dissimilative requirement on a property other than voicing, or from a natural but non-dissimilative pattern (Ohala 1981, 1993).

First, Dahl's Law in Bantu languages dissimilates [k] into [ɣ] before a voiceless consonant separated by a vowel. However, Bennett (1967, p. 155) compares several languages and dialects which exhibit Dahl's Law and concludes that it should have originated from ''a loss of aspiration by the first of two voiceless and aspirated stops''. In other words, dissimilation of voicelessness originated as the dissimilation of aspiration. Aspiration involves spread-out acoustic correlates like breathy voicing of surrounding vowels, and therefore dissimilation would resolve multiple occurrences of aspirated consonants. Bennett's hypothesis also explains why only velars dissimilate: velars have longer aspiration than consonants at other places of articulation (Maddieson 1997) and are thus most likely to cause perceptual confusion. The dissimilation of aspiration is thus the phonetically natural ancestor of dissimilation of voicing in these Bantu languages.

Second, in Gothic, fricatives dissimilate in voicing from a preceding consonant (Thurneysen's Law: Thurneysen 1897). However, the dissimilation in voicing may have derived from the prosodic-conditioned alternation in Proto-Germanic. A general voicing process of fricatives (Verner's Law: Verner 1875/1967) was blocked by a preceding accented syllable, and since Proto-Germanic had a mobile accent system, this voicing pattern historically resulted in an apparently dissimilative pattern (Flickinger 1981; Ohala 1981; see also Garrett 2007 for a recent summary of previous treatments of Thurneysen's Law as well as a slightly different perspective). In short, the dissimilation in voicing in Gothic derived from a non-dissimilative pattern.

To summarize, all three cases of voicing dissimilation, including that of Japanese, originated as a phonetically natural dissimilative requirement on properties other than voicing, or from a natural but non-dissimilative pattern.[8]

[8] Proto-Indo European turns up in discussions about voicing dissimilation (Meillet 1964), but the consonants in question may have been glottalized rather than voiced (Hopper 1973). If the consonants were voiced rather than glottalized, the system would be an otherwise poorly attested one which has [d] and [g], but lacks [b] (Ohala 1983). Regarding the series as voiceless glottalized consonants results in a more natural system since the lack of voiceless glottalized labials is typologically common. Glottalization involves spread-out cues through creakiness of surrounding vowels, so glottalization would naturally be subject to dissimilation.

6 General discussion and conclusion

In Japanese loanword phonology, only voiced geminates devoice due to OCP(voi). The different devoicability between singletons and geminates has a natural perceptual basis whereas the unnatural principle of OCP(voi) originated from a coincidental historical development.

The phonetically natural ranking IDENT(voi)$_{Sing}$ » IDENT(voi)$_{Gem}$ interacts with the phonetically unnatural OCP(voi) constraint. Therefore, a single grammatical system can have phonetically natural and phonetically unnatural aspects.[9]

Nevertheless, echoing the insight of Stampe (1973) and Hooper (1976, pp. 84–86), I argue that phonetic naturalness and unnaturalness are asymmetric. Natural patterns arise in the absence of positive phonological evidence whereas unnatural patterns do not. In Japanese loanword phonology, devoicability has been determined by a perceptibility scale without any overt phonological evidence. On the other hand, OCP(voi) arose as a coincidental historical development in Japanese as well as in other languages. In fact, all other phonetically unnatural patterns discussed in the literature, beyond OCP restrictions, have developed as a result of sound changes (Anderson 1981; Bach and Harms 1972; Blevins 2004).

Phonetically unnatural constraints—or constraints that do not have phonetic seeds—do not seem to spontaneously emerge as novel phonological patterns but only appear as developments of historical changes. The limited origin of phonetically unnatural patterns, such as OCP(voi), needs to be explained. I propose that speakers—or UG—possess phonetically natural constraints even without overt phonological evidence, but phonetically unnatural constraints are constructed only when confronted with overt evidence (Hayes 1999; see also Stampe 1973 for a similar idea).

The present proposal makes predictions that should be tested in future studies. If the universal set of phonological restrictions is phonetically natural, emergent novel phonological patterns should *always* be phonetically natural because diachronic changes cannot influence them. The prediction has been borne out in a novel devoicing pattern in Japanese loanword phonology, as discussed in Sect. 4 (see also Zuraw 2007 for another pattern). Another domain in which the emergence of phonetic naturalness has been documented is verbal art patterns such as English and Japanese imperfect puns (Fleischhacker 2005; Kawahara and Shinohara 2009) and Japanese rap rhyming (Kawahara 2007). Although these cases suggest that emergent phonological patterns tend to be phonetically natural, the possible phonetic bias of phonological grammars should be tested in wider context.

Acknowledgements Thanks to José Benkí, Kathryn Flack, Edward Flemming, John Kingston, Dan Mash, John McCarthy, Jaye Padgett, Joe Pater, Donca Steriade, Tim Vance, Betsy Wang, John Whitman,

[9] The interaction of these three constraints implies another point. OCP(voi) shows a number of morphophonological characteristics (e.g., its domain is a morphological stem). The fact that OCP(voi) interacts with the phonetically-based ranking, IDENT(voi)$_{Sing}$ » IDENT(voi)$_{Gem}$, suggests that phonetic information can influence morphophonological patterns. This conclusion should be contrasted with a claim made in Lexical Phonology that morphophonological lexical processes should be separated out from more automatic post-lexical rules that resemble phonetic implementation rules (Kiparsky 1985).

an anonymous reviewer, the editors of this volume Keren Rice and Yoonjung Kang, and the audience at ICEAL for their comments. All remaining errors are my own.

References

Abramson, Arthur, and Leigh Lisker. 1985. Relative power of cues: F0 shift versus voice timing. In *Phonetic linguistics: Essays in honor of Peter Ladefoged*, ed. V. Fromkin, 25–33. Orlando: Academic Press.

Anderson, Stephen. 1981. Why phonology isn't 'natural'. *Linguistic Inquiry* 12: 493–539.

Anttila, Arto, and Young-mee Yu Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104: 31–56.

Bach, Emmon, and Robert Harms. 1972. How do languages get crazy rules? In *Linguistic change and generative theory*, ed. R. Stockwell and R. Macaulay, 1–21. Bloomington: Indiana University Press.

Bennett, Patrick. 1967. Dahl's law and Thagicu. *African Language Studies* 8: 127–159.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Boersma, Paul. 1998. *Functional phonology*. The Hague: Holland Academic Graphics.

Coetzee, Andries, and Joe Pater. 2005. Lexically gradient phonotactics in Muna and Optimality Theory. Ms., University of Massachusetts, Amherst.

Evans, Nick. 1995. Current issues in the phonology of Australian languages. In *Handbook of phonological theory*, ed. J.A. Goldsmith, 723–761. Oxford: Blackwell.

Fleischhacker, Heidi. 2005. Similarity in phonology: Evidence from reduplication and loan adaptation. Ph.D. dissertation, UCLA.

Flemming, Edward. 1995. Auditory representations in phonology. Ph.D. dissertation, UCLA.

Flickinger, Daniel. 1981. Dissimilation in Gothic without Thurneysen's Law. *Proceedings of Chicago Linguistic Society* 17: 67–75.

Frisch, Stephan. 2004. Language processing and segmental OCP effects. In *Phonetically-based phonology*, ed. B. Hayes, R. Kirchner, and D. Steriade, 346–371. Cambridge: Cambridge University Press.

Garrett, Andrew. 2007. Thurneysen's Law. A handout from East Coast Indo-European Conference 26, Yale University.

Hawkins, Sarah, and Noel Nguyen. 2004. Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English. *Journal of Phonetics* 32: 199–231.

Hayes, Bruce. 1999. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In *Functionalism and formalism in linguistics*, ed. M. Darnell, E. Moravscik, M. Noonan, F. Newmeyer, and K. Wheatly, 243–285. Amsterdam: John Benjamins.

Hooper, Joan. 1976. *An introduction to Natural Generative Phonology*. New York: Academic Press.

Hopper, Paul. 1973. Glottalized and murmured occlusives in Indo-European. *Glossa* 7: 141–166.

Huffman, Marie. 1993. Phonetic patterns of nasalization and implications for feature specification. In *Phonetics and phonology 5: Nasals, nasalization and the velum*, ed. M. Huffman and R. Krakow, 303–327. San Diego: Academic Press.

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

Katayama, Motoko (1998) Optimality Theory and Japanese loanword phonology. Ph.D. dissertation, UCSC.

Kawahara, Shigeto. 2005. Voicing and geminacy in Japanese: An acoustic and perceptual study. *University of Massachusetts Occasional Papers in Linguistics* 31: 87–120.

Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language* 82: 536–574.

Kawahara, Shigeto. 2007. Half-rhymes in Japanese rap lyrics and knowledge of similarity. *Journal of East Asian Linguistics* 16: 113–144.

Kawahara, Shigeto, and Kazuko Shinohara. 2009. The role of psychoacoustic similarity in Japanese imperfect puns: A corpus study. *Journal of Linguistics* 45.

Keating, Patricia, and Aditi Lahiri. 1993. Fronted velars, palatalized velars, and palatals. *Phonetica* 50: 73–101.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2: 85–138.

Leben, William. 1973. Suprasegmental phonology. Ph.D.dissertation, MIT.

Liljencrants, Johan, and Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48: 839–862.
Lisker, Leigh. 1978. On buzzing the English /b/. *Haskins Laboratories Status Report on Speech Research* SR-55/56: 251–259.
Lyman, Benjamin S. 1894. Change from surd to sonant in Japanese compounds. *Oriental Studies of the Oriental Club of Philadelphia*: 1–17.
Macmillan, Neil, and Douglas Creelman. 2005. *Detection theory: A user's guide* (2nd ed.). Mahwah: Lawrence Erlbaum Associates Publishers.
Maddieson, Ian (1997) Phonetic universals. In *The handbook of phonetics sciences*, ed. W. Hardcastle and J. Laver, 619–639. Oxford: Blackwell.
Meillet, Antoine. 1964. *Introduction à l'étude comparative des langues indo-européennes* [Reprint of 8th Edition]. Alabama: University of Alabama Press.
Nishimura, Kohei. 2003. Lyman's law in loanwords. Ms., University of Tokyo.
Ohala, John. 1981. The listener as a source of sound change. *Proceedings of Chicago Linguistic Society: Papers from the parasession on language and behavior*: 178–203.
Ohala, John. 1983. The origin of sound patterns in vocal tract constraints. In *The production of speech,* ed. P. MacNeilage, 189–216. New York: Springer-Verlag.
Ohala, John. 1993. The phonetics of sound change. In *Historical linguistics: Problems and perspectives,* ed. C. Jones, 237–278. London: Longman Academic.
Prince, Alan, and 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: 126–147.
Rodriguez, Joao. 1930. *Nihon Dai Bunten [Arte da Lingoa de Japam]*. Tokyo: Sanseidoo [Originally written in 1604–1608; translated by Tadao Doi in 1930].
Sapir, Edward. 1912. The Takelma language of southwestern Oregon. In *The handbook of American Indian languages*, ed. F. Boas, 1–296. Washington: Bureau of American Ethnography.
Stampe, David. 1973. A dissertation on natural phonology. Ph.D. dissertation, University of Chicago.
Steriade, Donca. 2001. The phonology of perceptibility effect: The P-map and its consequences for constraint organization. Ms., UCLA.
Tesar, Bruce. 2007. A comparison of lexicographic and linear numeric optimization using violation difference ratio. Ms., Rutgers University.
Thurneysen, Rudolf. 1897. Spirantenwechsel im gotischen. *Indogermanische Forschungen* 8: 208–214.
Unger, Marshall. 1975. Studies in Early Japanese morphophonemics. Ph.D. dissertation, Yale University.
van Summers, W. 1988. F1 structure provides information for final-consonant voicing. *Journal of Acoustical Society of America* 84: 485–492.
Vance, Timothy. 2005. Sequential voicing and Lyman's Law in Old Japanese. In *Polymorphous linguistics: Jim McCawley's legacy*, ed. S. Mufwene, E. Francis, and R. Wheeler, 27–43. Cambridge: MIT Press.
Verner, Kerl. 1875/1967. An exception to the first sound shift. Reprinted in *A reader in nineteenth century historical Indo-European linguistics*, ed. R.P. Lehmann 1967, 132–163. Bloomington: Indiana University Press.
Whalen, Douglas, Arthur Abramson, Leigh Lisker, and Maria Mody. 1990. Gradient effects of fundamental frequency on stop consonant voicing judgments. *Phonetica* 47: 36–49.
Whalen, Douglas, and Patrice Beddor. 1989. Connections between nasality and vowel duration and height: Elucidation of the Eastern Algonquian intrusive nasal. *Language* 65: 457–486.
Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog reduplication. *Language* 83: 277–316.

Kawahara, Shigeto (2011) Japanese loanword devoicing revisited: A rating study. Natural Language and Linguistic Theory 29(3): 705-723.

Nat Lang Linguist Theory (2011) 29:705–723
DOI 10.1007/s11049-011-9131-7

Japanese loanword devoicing revisited: a rating study

Shigeto Kawahara

Received: 19 November 2009 / Accepted: 26 May 2010 / Published online: 9 April 2011

Abstract In Japanese loanword phonology, geminates optionally devoice when there is another voiced obstruent within the same stem, i.e., geminates may optionally devoice when they violate OCP(voice). This devoicing of OCP-violating geminates has received much attention in the recent phonological literature. However, the debates centering around this phenomenon have relied primarily on intuition-based data, and no systematic judgment experiments have been performed. This paper fills that gap. The experiment reported in this paper shows that Japanese speakers do find devoicing of geminates natural when there is another voiced stop within the same word, i.e., when the geminates violate OCP(voice). The experiment moreover finds other interesting aspects of devoicing: (i) the naturalness of devoicing of OCP-violating geminates correlates positively with the lexical frequencies of the words in question, (ii) the naturalness of devoicing of OCP-violating geminates is not significantly affected by place of articulation, (iii) speakers find (context-free) devoicing of geminates more natural than devoicing of OCP-violating singletons, and (iv) speakers find the devoicing of OCP-violating singletons more natural in word-medial position than in word-initial position.

Keywords Laboratory phonology · Phonological judgments · Devoicing · Geminates · OCP(voice) · Lyman's law · Japanese

1 Introduction

1.1 The phenomena

This paper reports a judgment experiment on the devoicing of voiced geminates in Japanese loanword phonology, a topic that has received much attention in the recent

S. Kawahara (✉)
Department of Linguistics and Rutgers University Center for Cognitive Science, Rutgers University, New Brunswick, USA
e-mail: kawahara@rci.rutgers.edu

phonological literature. Japanese native phonology does not permit voiced geminates (Itô and Mester 1995, 1999), but gemination in recent loanwords has brought about voiced geminates in Japanese loanword phonology (e.g. [eggu] 'egg' and [doggu] 'dog'). Itô and Mester (1999) observe that some such geminates are devoiceable whereas others are not; to account for this difference, they treat devoiceable geminates as contained in assimilated foreign items and non-devoiceable geminates as contained in unassimilated foreign items.

Instead of relying on an etymological distinction, Nishimura (2003) has identified a phonological condition which makes devoicing of geminates possible: the presence of another voiced obstruent. Nishimura argues that OCP(voice)$_{stem}$, which prohibits two voiced obstruents within the same stem, is the key to coercing the devoicing of geminates—compare (1) and (2). This argument by Nishimura is attractive since we know independently that OCP(voice) is active in native Japanese phonology (Itô and Mester 1986).[1] One complication, however, is that singletons do not seem to undergo devoicing even when they violate OCP(voice)—compare (1) and (3).

(1) Geminates can optionally devoice if they co-occur with another voiced obstruent[2]
  a. ba**dd**o → ba**tt**o 'bad'
  b. ba**gg**u → ba**kk**u 'bag'
  c. do**gg**u → do**kk**u 'dog'

(2) Geminates do not devoice otherwise
  a. suno**bb**u → *suno**pp**u 'snob'
  b. re**dd**o → *re**tt**o 'red'
  c. e**gg**u → *e**kk**u 'egg'

(3) Singletons do not devoice even when they violate OCP(voice)
  a. gi**b**u → *gi**p**u 'give'
  b. ba**g**u → *ba**k**u 'bug'
  c. da**g**u → *da**k**u 'Doug'

Since Nishimura (2003) identified the patterns in (1)–(3), they have received much attention in the literature (Coetzee and Pater 2011; Crawford 2009; Farris-Trimble 2008; Haraguchi 2006; Hayes 2009; Itô and Mester 2008; Kaneko and Iverson 2009; Kawahara 2006, 2008; McCarthy 2008; Pater 2009; Pycha et al. 2006; Rice 2006; Steriade 2004; Tanaka 2010; Tateishi 2002; Tesar 2007). One debate focuses on the

[1] OCP(voice) in Japanese is also known as Lyman's Law (Lyman 1894). Native Japanese phonology generally does not allow stems with two voiced obstruents. OCP(voice) also blocks Rendaku, voicing of the initial consonant of the second member of a compound, when the second member already contains a voiced obstruent (Itô and Mester 1986; Vance 1980). OCP(voice) in Japanese targets voicing only in obstruents, but not in sonorants (Itô and Mester 1986; Mester and Itô 1989). See Vance (1980) for an experiment on OCP-induced blockage of Rendaku.

[2] An anonymous reviewer pointed out that in some cases these OCP-violating geminates are spelled as voiceless; e.g. *gutto laihu kanpanii* 'Good life company' (http://gootlife.com) and *gutto hikkoshi sentaa* 'Good moving company' (http://a-hikkoshi.com/cooperation/good.html); a google search also reveals that Buddha is often spelled as *butta*.

difference between singletons and geminates—why do Japanese speakers devoice only geminates under the influence of OCP(voice) (Kawahara 2006, 2008; Rice 2006; Steriade 2004)? In answer to this question Kawahara (2006, 2008), for example, has demonstrated that Japanese voiced geminates are phonetically semi-devoiced and that a voicing contrast is therefore perceptually harder to hear in geminates than in singletons; these experimental results arguably show that phonetic perceptibility plays a role in determining phonological neutralizability (Steriade 2008).

Another set of work focuses on the seemingly cumulative behavior of devoicing: neither geminacy nor OCP(voice) alone can coerce devoicing (see (2) and (3)); only when both of the factors are relevant, does devoicing become possible (Coetzee and Pater 2011; Farris-Trimble 2008; Hayes 2009; Pater 2009; Tesar 2007). This cumulative behavior may bear on a general theory of constraint interaction. Such a behavior is impossible to model in a strict ranking-based theory of constraint interaction, such as Optimality Theory (Prince and Smolensky 1993/2004), unless there is a constraint that is violated only when both OCP(voice) and a constraint against voiced geminates are violated (Nishimura 2003), or unless we posit different faithfulness constraints for singletons and geminates (Kawahara 2006, 2008). However, a theory with weighted, rather than ranked, constraints can model the cumulative pattern without positing a complex markedness constraint or differentiated faithfulness constraints (Pater 2009).

Finally, yet another set of work addresses why and how the patterns in (1)–(3) spontaneously emerged in loanword phonology, especially given that native phonology does not allow voiced geminates or OCP-violating singletons. One particular question is the one discussed above: where does the difference between singletons and geminates stem from? Some other work moreover addresses how the emergence of the devoicing of OCP-violating geminates bears on a general theory of loanword adaptation and lexical stratification: the fact that OCP(voice), which is active in native Japanese phonology, plays a role in loanword phonology may shed light on how loanword phonology is related to native phonology (Crawford 2009; Itô and Mester 2008; Tateishi 2002).

In summary, the patterns in (1)–(3) have evoked many theoretical debates. This paper does not attempt to model the intricate patterns of devoicing or resolve the issues that are raised in the literature cited above; interested readers are referred to those works. Rather, the current concern is instead that the data in (1)–(3) is largely based on native authors' intuitions by Nishimura (2003) and Kawahara (2006). Although Nishimura (2003) and Kawahara (2006) back up their intuitions using corpus search and informal native speaker consultation, systematic judgment studies of OCP-violating geminates have not been conducted.

1.2 The need for experimentation

The aim of this paper, therefore, is to validate the generalizations exemplified in (1)–(3), since several important theoretical claims have been made using these patterns. This study is inspired and motivated by an increasing interest in testing the quality of linguistic data using experimental methodology (Berko 1958; Cowart 1997; Dabrowska 2010; Hayes and Londe 2006; Kawahara 2011b; Myers 2009; Nolan 1992; Schütze 1996, among others). Here I briefly summarize why experimentation is necessary/desirable beyond intuition-based data collection (see the work cited and references cited therein for more general and elaborate discussion of the following points).

The first general concern is that some phonological patterns that are used to argue for particular theoretical claims have been shown to be unproductive or non-reproducible in experimental settings (e.g. Alderete and Kochetov 2009; Batchelder 1999; Griner 2001; Jaeger 1983; Ohala 1974; Sanders 2003; Vance 1987). One example is Japanese verbal conjugation patterns; several experiments show that native Japanese speakers do not reproduce alleged phonological alternations in verb conjugations (Batchelder 1999; Griner 2001; Vance 1987). (See Davis and Tsujimura 1991 for a review of alternations in Japanese verbal conjugations and an autosegmental analysis.) More recently, Alderete and Kochetov (2009) show that a case of conflicting directionality of a palatalization feature in Japanese mimetics (Hamano 1986; Mester and Itô 1989) is not productively reproduced by native speakers. These examples highlight the importance of systematic experimentation in order to guarantee the productivity of the phonological patterns under question; otherwise we may run the risk of building a theory based on unproductive linguistic patterns.

The second concern is that of generalizability. When the data are based on the intuition of two authors, we cannot guarantee that their intuitions generalize to the whole population of Japanese speakers. In order to assure that the patterns in (1)–(3) are a general property of Japanese phonology, rather than the phonology of two specific individuals, it is necessary to gather data from a large number of speakers. The third concern is replicability: in an intuition-based approach, we cannot guarantee the replicability of the results, because the procedure of obtaining the intuition-based data is "private", relying on the inner sensation of informants, who are, in the case of Nishimura (2003) and Kawahara (2006) (and often in other cases as well), the authors themselves (see Schütze 1996: 48–52). Because of this problem, we do not have a measure to evaluate the replicability of an informal judgment, unless we follow a rigorous experimental protocol.

The fourth concern of the intuition-based approach is bias (Dabrowska 2010; Gibson and Fedorenko 2010): authors can be biased due to their theoretical commitments. The purely intuition-based approach runs the risk of (unconsciously) skewing and/or oversimplifying the actual data in the process of introspection. To avoid this problem, it would be ideal to obtain data from naive speakers. The final concern is that judgments are often made in terms of a binary, grammatical/ungrammatical choice, as is the case in Nishimura (2003) and Kawahara (2006). However, it is known that grammaticality judgment experiments can reveal more nuanced distinctions among "grammatical" forms or among "ungrammatical forms" (see e.g. Cohn 2006; Cowart 1997; Greenberg and Jenkins 1964; Hayes 2000, 2009; Myers 2009; Pertz and Bever 1975; Pierrehumbert 2001). For these reasons, experimental testing of linguistic data should complement—if not replace—intuition-based data.

1.3 Additional hypotheses tested

These considerations call for verification of the empirical foundations of linguistic theory, beyond an intuition-based approach to linguistics. To that end, I conducted a naturalness rating experiment that tests the validity of the generalizations stated in (1)–(3). In addition, the experiment was designed to address some particular additional aspects of devoicing in (1).

One additional hypothesis that is tested is an informal observation, or an intuition, that Japanese speakers sometimes have about the devoicing of geminates—the more frequent the words are, the more natural it is to devoice OCP-violating geminates. This intuition was shared by native Japanese speakers that I consulted, but has never been substantiated by a systematic study. The effects of lexical frequencies on phonological patterns have been receiving an increasing interest in recent phonological and psycholinguistic studies (see e.g. Bybee 1999, 2001; Coetzee and Kawahara 2011; Coleman and Pierrehumbert 1997; Ernestus and Baayen 2003; Frisch et al. 2000; Hay et al. 2003; Hayes 2009; Hayes and Londe 2006; Zuraw 2009). For example, in Usage-based Phonology (Bybee 1999, 2001), frequency is one major factor that governs and shapes phonological regularity (see also Coleman and Pierrehumbert 1997; Frisch et al. 2000; Hay et al. 2003). Therefore it is important to investigate the extent to which frequency can affect the devoicing of OCP-violating geminates.

Another aspect of the devoicing pattern that is tested in the following experiment is the place effect: According to Nishimura's (2003) data, [gg] is more likely to be devoiced than [dd] (Kawahara 2006), and in relation to this observation, Kawahara (2006) shows that a voicing contrast is less perceptible in [gg] than in [dd]. Kawahara (2006) takes this correlation as additional evidence that the perceptibility of a phonological contrast correlates with the neutralizability of that contrast. A question thus remains as to whether Japanese speakers would indeed show a difference in the devoiceability of OCP-violating geminates based on their place of articulation.

1.4 Why a rating task?

To test these hypotheses, the current study conducted a naturalness rating experiment. Some remarks on why a rating task was specifically chosen are now in order. First, rating judgment tasks with a numerical response scale are known to reveal subtle distinctions of grammaticality beyond the grammatical/ungrammatical dichotomy (see the references above at the end of Sect. 1.2). Second, testing the intuitive correlation between the frequencies of lexical items and the likelihood of devoicing requires us to obtain quantitative measures of grammaticality, comparable to the frequencies of the items under question. In summary, in a naturalness rating study with a larger number of speakers in which we control for the relevant variables, we can gain further insight into the phenomenon.

1.5 A preview of the results

The experiment reported below shows that Japanese speakers do find the devoicing of geminates natural when the geminates violate OCP(voice), supporting the basic intuitions of Nishimura (2003) and Kawahara (2006). Therefore it succeeds in securing the empirical foundations of the theoretical claims reviewed above in Sect. 1.1. However, the experiment reveals other interesting systematic patterns as well: (i) the naturalness ratings of devoicing of OCP-violating geminates positively correlate with the lexical frequencies of the words in question, (ii) the naturalness ratings of the devoicing of voiced geminates are not significantly affected by place of articulation,

(iii) speakers find the (context-free) devoicing of geminates more natural than the devoicing of OCP-violating singletons, and (iv) speakers find the devoicing of OCP-violating singletons more natural in word-medial position than in word-initial position.

1.6 One caveat

Before moving onto the description of the experiment, it needs to be made clear that the patterns in (1)–(3) are a part of loanword phonology rather than the process of loanword adaptation *per se*. In other words, we are interested in how Japanese speakers treat words that they have already borrowed and adapted from other languages (for studies of Japanese loanword adaptation, see Lovins 1973; Kaneko and Iverson 2009; Katayama 1998; Shinohara 2004). On this note, Kaneko and Iverson (2009) conducted a production adaptation study on how Japanese speakers adapt (mostly) nonce English words, and did not find evidence for devoicing of OCP-violating geminates in the process of adaptation. Therefore, if devoicing in (1) happens at all, then it is in the loanword phonology of already-borrowed forms rather than in loanword adaptation.

2 Method

In this experiment, Japanese speakers judged the naturalness of the devoicing of singletons and geminates in various contexts.

2.1 Stimuli

This experiment used real words for two reasons: (i) in order to test the data of Nishimura (2003) and Kawahara (2006) and (ii) to test the effect of lexical frequencies on the naturalness of devoicing. Therefore, the stimuli were taken from the lists of items provided in Kawahara (2006), which itself builds on Nishimura (2003). The design had three conditions: (i) OCP-violating geminates as in (1), (ii) non-OCP-violating geminates as in (2), and (iii) OCP-violating singletons as in (3). Words with only one singleton were not included among the stimuli, because neither Nishimura (2003) nor Kawahara (2006) discuss them.[3] The complete list of the stimuli is provided in Table 1. There is only one word that contains OCP-violating [bb]. The scarcity of this sort of form is due to the fact that in loanword adaptation, [b] tends to resist gemination compared to [d] and [g] (Kaneko and Iverson 2009; Katayama 1998; Shirai 2002).

2.2 Task

The task was a naturalness rating task. In the general instructions, the participants were told that the questionnaire was about the naturalness of devoicing (i.e. that the experiment was about "daku-on" (voiced obstruents)). They were also told that the

[3]For the effect of OCP(voice) on singletons, see Kawahara (2011a, 2011c).

Table 1 The list of stimuli. Based on Kawahara (2006)

OCP-violating geminates		Non-OCP-violating geminates		OCP-violating singletons	
gebberusu	'Göbbels'	webbu	'web'	bagii	'buggy car'
guddo	'good'	sunobbu	'snob'	bobu	'Bob'
beddo	'bed'	habburu	'Hubble'	dagu	'Doug'
doreddo	'dread'	kiddo	'kid'	daiyamondo	'diamond'
deddobooru	'deadball'	reddo	'red'	giga	'giga'
baddo	'bad'	heddo	'head'	gibu	'give'
deibiddo	'David'	suraggaa	'slugger'	bogii	'bogey'
budda	'Buddha'	eggu	'egg'	bagu	'bug'
doggu	'dog'	furaggu	'flag'	daibu	'dive'
baggu	'bag'			doguma	'dogma'
doraggu	'drag'			gaburieru	'Gabriel'
biggu	'big'			gaidansu	'guidance'

experiment is about loanwords. For each question, the participants were first presented with one word from the stimulus list, and they were asked to judge the naturalness of the form that undergoes devoicing (e.g. given [gebberusu], how natural would you find it to pronounce it as [gepperusu]?). They were asked to provide their judgments on a 5-point scale: A. "very natural", B. "somewhat natural", C. "neither natural nor unnatural", D. "somewhat unnatural", and E. "very unnatural".[4] They were then asked to read the stimuli before answering each question, and to base their decision on their auditory impression rather than on orthography. For the stimuli containing OCP-violating singletons (the rightmost column in Table 1), they were asked to judge the naturalness of both devoicing of word-initial singletons and word-internal singletons. These two questions were presented separately.

2.3 Procedure

The test was administered through Sakai (https://sakai.rutgers.edu/portal), a Java-based system that runs online experiments (see Reips 2002 for general discussion of web-based experimentation). The first page of the experimental website showed a consent form for a human subject experiment. Once the participants agreed to participate in the experiment, they were forwarded to a testing site. On each page, Sakai presented one stimulus, and asked how natural the devoiced form of that stimulus was on the 5-point scale described above. The instructions and the options were provided in Japanese orthography; the stimuli and the forms undergoing devoicing were presented in katakana orthography, which is used for loanwords; the instruction sentences and the options were provided in a mixture of kanji and hiragana, following the standard Japanese orthographic convention. The order of the stimuli was randomized by Sakai. At the end of the experiment, they were asked if they were familiar with the

[4] The software used to present the stimuli (see below) did not allow us to present the scale numerically, so numeric conversion was applied later.

devoicing phenomenon of OCP-violating geminates; those who answered positively to this question were excluded from the following analysis.

2.4 Participants

Thirty-eight native speakers of Japanese completed the study. However, two speakers were familiar with the devoicing phenomenon, and therefore were excluded from the following analysis.

2.5 Frequency measures

The frequencies of the items were counted based on a Japanese lexical corpus, Amano and Kondo (2000).[5] The mean log-frequencies of the three conditions were comparable (OCP-violating geminates: 4.57, non-OCP-violating geminates: 4.32, OCP-violating singletons: 4.48).

2.6 Statistics

The responses were first converted to numerical values as follows: "very natural" = 5; "somewhat natural" = 4; "neither natural nor unnatural" = 3; "somewhat unnatural" = 2; "very unnatural" = 1. For statistical analyses, first, a general linear mixed model was run (Baayen et al. 2008; Baayen 2008; Bates 2005) using R (R Development Core Team 1993–2011) with the `lme4` package (Bates et al. 2011). Rating scores were regressed against a model in which OCP, geminacy, place, and frequency were fixed factors and speakers and items were random factors. Since OCP and geminacy were not fully-crossed (there were no non-OCP-violating singletons), the interaction term between them was not coded, although the cumulative effect of OCP and geminacy would show up in their interaction term. Instead this general analysis was followed by specific contrast analyses. The `lm4e` package does not automatically compute p-values because the exact procedure to calculate degrees of freedom has not been discovered. Therefore, the p-values, as well as the 95% confidence intervals of the coefficients of the fixed factors, were calculated by the Markov chain Monte Carlo method using the `pval.fnc()` function of the `languageR` package (Baayen 2009). Finally, the correlations between the naturalness ratings and the lexical frequencies were checked using a Spearman correlation test, again using R.

3 Results

3.1 General results

Figure 1 plots the average naturalness ratings of devoicing in the four conditions: OCP-violating geminates (e.g. /guddo/ → [gutto]); non-OCP-violating geminates

[5] An alternative would be to use frequencies on the internet using search engines such as Google. A problem with this approach is that since Japanese orthography does not insert a space between words, a search for a word automatically includes words that contain the target word; e.g. if we search for the word *doggu* 'dog', the results will for example include *hotto doggu* 'hot dog'.

(e.g. /heddo/ → [hetto]); OCP-violating singletons (initial) (e.g. /bagu/ → [pagu]); OCP-violating singletons (non-initial) (e.g. /bagu/ → [baku]). The y-axis represents the naturalness ratings in the following way: 5 = "very natural", 4 = "somewhat natural", 3 = "neither natural nor unnatural", 2 = "somewhat unnatural", 1 = "very unnatural". The error bars represent 95% confidence intervals based on the variability across all relevant items and speakers and a t-distribution.

Japanese speakers find the devoicing of OCP-violating geminates most natural (the average is 3.89, around "somewhat natural"). Japanese speakers also find the devoicing of non-OCP-violating geminates more natural (the average is 2.40, somewhere between "neither natural nor unnatural" and "somewhat unnatural") than the devoicing of OCP-violating singletons. They also find the devoicing of singletons more natural in medial position (the average is 1.75, near "somewhat unnatural") than in initial position (the average is 1.42, somewhere between "somewhat unnatural" and "very unnatural").

A general linear mixed model showed that geminacy, OCP, place, and frequency all affect naturalness ratings (geminacy: $t = 40.5$, $p < .001$; OCP: $t = 19.58$, $p < .001$; place: $t = 2.78$, $p < .01$; frequency: $t = 4.68$, $p < .001$). A contrast analysis comparing OCP-violating geminates (the first bar in Fig. 1) and non-OCP-violating geminates (the second bar) turned out to be significant ($t = 19.3$, $p < .001$). We also observe a difference between non-OCP-violating geminates (the second bar) and non-initial OCP-violating singletons (the fourth bar) ($t = 22.02$, $p < .001$). (Geminates in Japanese appear only non-initially, and hence this comparison controls for position within words.) Third, speakers found the devoicing of non-initial singletons more natural than the devoicing of initial singletons (the third bar vs. the fourth bar: $t = -6.65$, $p < .001$). Table 2 lists the 95% confidence intervals calculated by the Markov chain Monte Carlo method for all these fixed effect coefficients. None of the intervals overlap with 0.

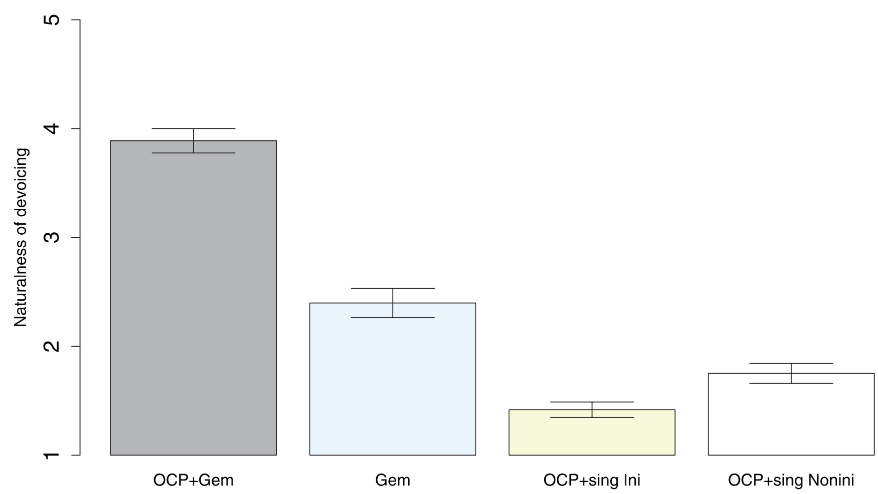


Fig. 1 (Color online) The average naturalness ratings in four conditions: OCP-violating geminates; non-OCP-violating geminates; OCP-violating singletons (initial); OCP-violating singletons (non-initial). The error bars represent 95% confidence intervals, calculated based on variability over each response and a t-distribution

Table 2 The 95% confidence intervals of the coefficients of the linear mixed models. Contrast I = OCP-violating geminates vs. non-OCP-violating geminates; Contrast II = Non-OCP-violating geminates vs. OCP-violating singletons (non-initial); Contrast III = Two positions of singleton consonants

	Lower	Upper
General analysis		
OCP	1.2615	1.5399
Gem	2.1313	2.3496
Place	0.0265	0.1676
Frequency	0.0001	0.0002
Contrast I		
OCP	1.345	1.649
Contrast II		
Gem vs. sing	1.389	1.662
Contrast III		
Initiality	−0.4296	−0.2329

3.2 Place effects

Turning to the effects of place on the devoicing of OCP-violating geminates, Fig. 2 illustrates average ratings in the four conditions. A linear mixed model analysis did not reveal a significant effect of place on the naturalness ratings of OCP-violating geminates ($t = 0.317$, *n.s.*; the average ratings: [bb] = 3.31; [dd] = 4.01; [gg] = 3.82).[6] For non-OCP-violating geminates, backer consonants tended to receive higher ratings, although the effect of place did not reach significance in this condition either ($t = 0.91$, *n.s.*; the average ratings: [b] = 2.28; [d] = 2.44; [g] = 2.48). No consistent place effects were observed in initial singletons ($t = 0.90$, *n.s.*; the average ratings: [bb] = 1.40; [dd] = 1.34; [gg] = 1.52) or in non-initial singletons ($t = 1.28$, *n.s.*; the average ratings: [bb] = 1.50; [dd] = 2.10; [gg] = 1.80).

An anonymous reviewer pointed out that even though the effect of place was not significant in these analyses, labials generally received lower scores than coronals and dorsals. The reviewer further pointed out that the current stimuli contained only one item with geminate [bb] in the OCP-violating geminate condition (see Table 1). The question therefore arises whether the high rating of devoicing of OCP-violating geminates is due to the fact that the OCP-violating geminate condition contained only one item with [bb]. To address this question, a linear-mixed model was rerun excluding all labial stimuli, and still revealed a significant difference between OCP-violating geminates and non-OCP-violating geminates ($t = 16.07$, $p < .001$).

3.3 Lexical frequency effects

Next, Fig. 3 illustrates the correlation between the average ratings and the natural log-frequencies in all four conditions (ln(0) was replaced with 0). Spearman correla-

[6] A question arises as to where the difference between this result and Nishimura's (2003) result comes from. According to Nishimura's (2003) data, [gg] is more likely to devoice than [dd] (Kawahara 2006), whereas in this experiment speakers rated devoicing of [dd] more natural than devoicing of [gg] (although the difference did not reach significance). See Kawahara (2011a)—a follow-up study of the current experiment—for relevant discussion.

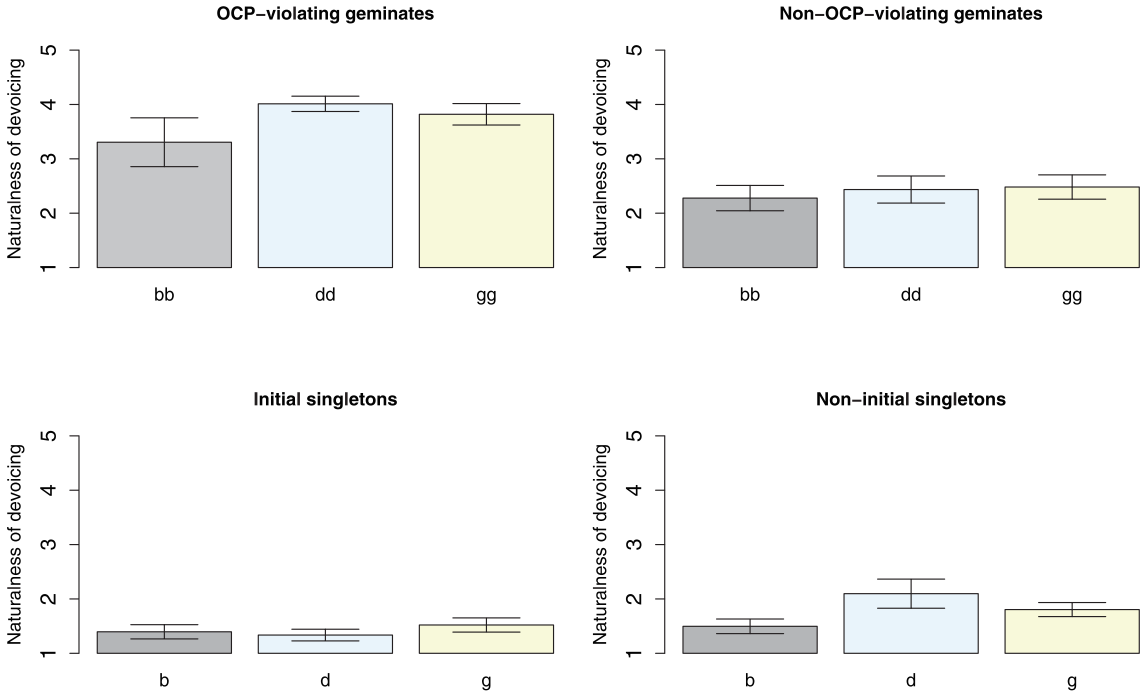


Fig. 2 (Color online) The effects of place on naturalness ratings in the four conditions

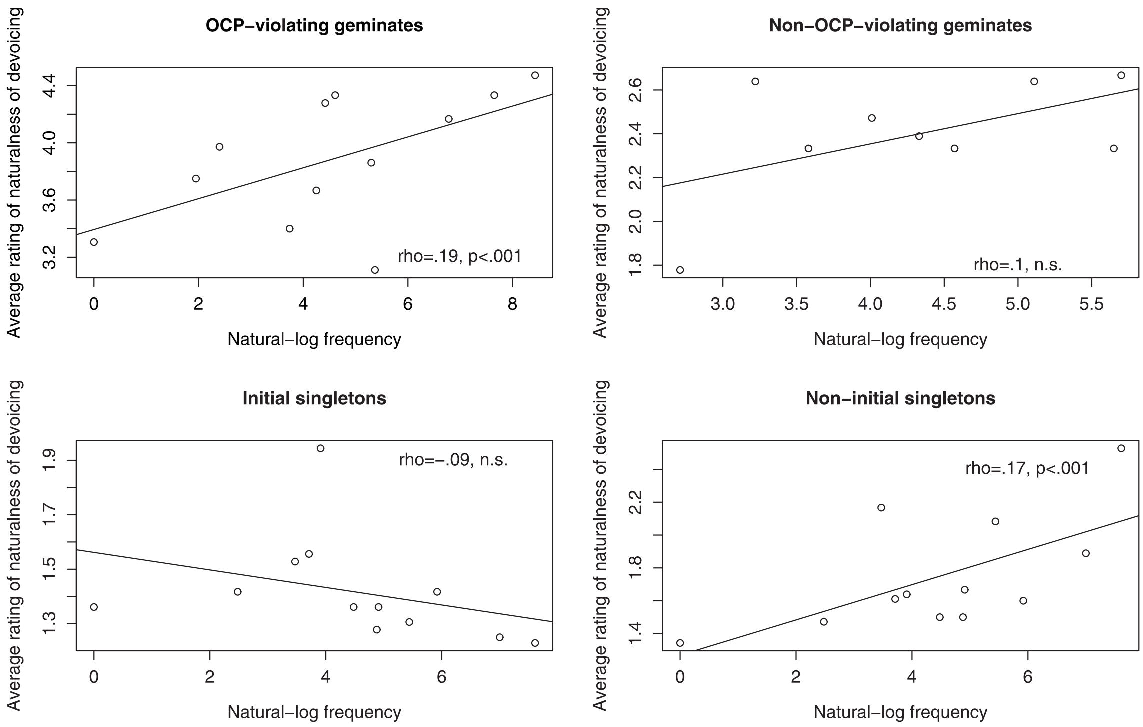


Fig. 3 The correlation between the average naturalness ratings of devoicing and the natural log-frequencies of the items for each condition. X-axies and Y-axies are on different scales in each graph

tion tests reveal a statistically significant correlation between the lexical frequency of words and the corresponding naturalness rating in OCP-violating geminates ($\rho = .19$, $p < .001$) and in non-initial singletons ($\rho = .17$, $p < .001$). In these conditions, high frequency-words show tendency toward higher rating. The correlation was not significant in non-OCP-violating geminates ($\rho = .1$, *n.s.*) and in initial singletons ($\rho = -.09$, *n.s.*). These analyses show that frequency affects naturalness ratings only in some grammatical conditions, including OCP-violating geminates.

3.4 Post-hoc, item-specific analyses

Now I turn to discussion of items that have certain properties.[7] To assess each post-hoc analysis, linear mixed models with speakers and items as random factors were used. To soak up variability as much as possible, other fixed factors were included where possible. Since there were four repeated post-hoc analyses, to avoid inflation of Type I error, the alpha level was adjusted to $.05/4 = .0125$ by Bonferronization.

3.4.1 Non-local effects of OCP

Within items that contain OCP-violating geminates, two items have long-distance triggers: [doreddo] 'dread' and [doraggu] 'drug'. Since dissimilatory force is cross-linguistically known to be stronger between two closer elements (Frisch et al. 2004; Suzuki 1998), one may expect that devoicing of these two words may be considered to be less natural. In fact, Vance (1980) shows in his nonce-word experiment that the blockage of Rendaku due to OCP(voice) (see footnote 1) is more likely when the trigger is closer.

A post-hoc linear mixed model with locality of OCP-violation, together with frequency and place, as fixed independent factors, was run between the two items with long-distance OCP-triggers and the rest of the items. Although we find a trend in which the two items containing long-distance triggers received lower rating (3.43) than the others (3.98), the difference did not reach significance ($t = 1.51$, *n.s.*).

3.4.2 Double triggers

Second, in the OCP-violating geminate condition, there are two items that contain two triggers for OCP(voice): [deddobooru] 'deadball' and [deibiddo] 'David'. As Tesar (2007) points out, devoicing geminates in such words would still entail a configuration that violates OCP(voice): [dettobooru] and [deibitto]. Kawahara (2008) and Tesar (2007) suggest that OCP(voice) can militate against each pair of OCP-violating voiced consonants so that devoicing of geminates would still lessen the violation of OCP(voice). Nevertheless, it may be possible that devoicing of these geminates can be considered less natural than others because devoicing does not resolve OCP(voice) violations completely.

Although the speakers rated the two items with double triggers lower (3.84) than the other items (4.13), a linear-mixed model with the number of triggers (together with frequency and place) as fixed factors did not reveal a significant difference between the two groups ($t = 1.48, n.s.$).

[7] Thanks to an anonymous reviewer who pointed out that these post-hoc analyses may be informative.

3.4.3 *Merger with existing words*

Some items containing OCP-violating geminates merge with some existing lexical items: [baddo] 'bad' vs. [batto] 'bat', [baggu] 'bag' vs. [bakku] 'back', and [doggu] 'dog' vs. [dokku] 'dock'. Some scholars point out that languages may avoid mergers of two existing lexical items (Blevins 2005; Ichimura 2006) (see also general anti-merger constraints within the framework of Dispersion Theory: Flemming 1995; Itô and Mester 2004; Padgett 2002; Urbanczyk 2005).

A linear mixed model with the possibility of merger, place and frequency within the OCP-violating geminate condition did not reveal a significant difference between those that result in merger and those that do not (3.80 vs. 3.91; $t = -0.51$, *n.s.*), although those that do not result in merger showed slightly higher ratings.

3.4.4 *Devoiceability of the following vowel*

Finally, devoiceability of the vowels following the target consonants may affect the naturalness ratings of devoicing for the following reason. Devoicing of consonants may create an environment for high vowel devoicing in Japanese, which takes place between two voiceless consonants and word-finally after voiceless consonants (e.g. Tsuchida 1997). Speakers may disprefer devoicing of voiced consonants that will feed high vowel devoicing, because the devoicing of consonants would yield further unfaithful—or more dissimilar—forms with respect to the original forms.

The stimuli that contain OCP-violating geminates were thus categorized into two groups: one in which devoicing of the target consonants can yield high vowel devoicing and those that cannot. A linear mixed model with place, frequency, and devoiceability of the following vowel shows that the effect of devoiceability of the following vowel did not influence the ratings of OCP-violating geminates (those followed by devoiceable vowels: 3.92 vs. those followed by non-devoiceable vowels: 3.82; $t = -0.56$, *n.s.*).

3.4.5 *Summary of post-hoc analyses*

The first three post-hoc analyses showed interesting trends in directions that we expect from theoretical considerations, although none of the effects turned out to be significant. We should bear in mind, however, that these results are based on post-hoc analyses. Therefore, the devoicing of OCP-violating geminates may provide a testing ground for addressing these theoretical considerations (Kawahara 2011a).

4 Discussion

4.1 Summary

To summarize the results, the current experiment confirms the intuition-based data of Nishimura (2003) and Kawahara (2006) in that Japanese speakers find the devoicing of OCP-violating geminates most natural. In this respect, the current work has succeeded in securing the empirical foundations of the theoretical claims reviewed in Sect. 1.1.[8]

4.2 Gradiency: Beyond a grammatical/ungrammatical dichotomy

In addition to confirming the naturalness of devoicing of OCP-violating geminates, the current experiment makes two findings beyond what Nishimura (2003) and Kawahara (2006) report. First, it found a positive correlation between the naturalness of the devoicing of OCP-violating geminates and the lexical frequencies of the items under question. Interestingly, the correlation was only significant under certain grammatical conditions (OCP-violating geminates and non-initial singletons).

Second, the experiment revealed further distinctions among those devoicing patterns that were judged to be ungrammatical by Nishimura (2003) and Kawahara (2006): Japanese speakers find the devoicing of non-OCP-violating geminates more natural than the devoicing of OCP-violating singletons; they also find the devoicing of singletons more natural in medial position than in initial position.

The fact that Japanese speakers found the devoicing of non-OCP-violating geminates more natural than the devoicing of OCP-violating singletons is interesting, because both structures are illicit in the native phonology and both are licit in the loanword phonology (Itô and Mester 1995, 1999). This difference could arise from the difference in perceptibility of a voicing contrast between singletons and geminates. Kawahara (2006) has demonstrated that a voicing contrast is perceptually less salient in geminates than in singletons; therefore speakers may prefer the devoicing of geminates in general to the devoicing of singletons. Alternatively, one could argue that voiced geminates are more marked than OCP-violating singletons.

The second new finding—that speakers find devoicing of singletons more natural in word-medial initial positions than in word-initial positions—is also interesting, because in modern Japanese voiced stops are allowed in both word-initial and word-internal positions. The dispreference against devoicing of word-initial singletons may have its roots in the psycholinguistic prominence of initial positions (Hawkins and Cutler 1988; Horowitz et al. 1968, 1969; Nooteboom 1981); since word-initial position plays an important role in lexical access, speakers disprefer changing segments in this position (Beckman 1997; Kawahara and Shinohara 2011).[9]

More generally, the results show that, in line with other recent studies, grammatical intuitions are gradient (e.g. Chomsky 1965; Cohn 2006; Coetzee 2008, 2009; Coleman and Pierrehumbert 1997; Fanselow et al. 2006; Frisch et al. 2000, 2004; Greenberg and Jenkins 1964; Hay et al. 2003; Hayes 2000, 2009; Hayes and Londe

[8] One remaining question is whether the judgement pattern reflects the frequencies of devoicing of voiced consonants in Japanese loanword phonology. Unfortunately, Amano and Kondo (2000), being a written database, does not contain many forms that undergo devoicing. A production experiment would be able to fill the gap—see Sect. 4.3.

[9] There are languages which avoid initial voiced stops, such as Old Japanese (Hamano 2000; Vance 1980), as it is articulatorily challenging to initiate voicing in word-initial position (Westbury and Keating 1986). Thus, from an articulatory point of view, one would expect that speakers would prefer devoicing of word-initial voiced stops.

2006; Myers 2009; Pierrehumbert 2001; Schütze 1996; Zuraw 2000) in the following two senses. First, it is not the case that some devoicing patterns are grammatical and some other devoicing patterns are ungrammatical. Each factor (geminacy, OCP, word-position) affects the naturalness of devoicing. Second, lexical frequencies gradiently affect the naturalness of devoicing of geminates, showing that distinctions exist even among OCP-violating geminates.

One general lesson we can draw from the results is that it may be dangerous to rely on an intuition-based dichotomous distinction between "grammatical" and "ungrammatical" processes, as was assumed by Nishimura (2003) and Kawahara (2006). The current results show that our linguistic knowledge provides much more fine-grained distinctions; speakers can tell the difference between three "ungrammatical devoicing patterns" (i.e. devoicing of non-OCP-violating geminates and devoicing of singletons in two positions).

I acknowledge that there is a view that grammar itself is categorical and apparent gradient patterns arise from extra-grammatical factors (Sprouse 2007) (see also Schütze 1996 for extensive discussion). In this view, grammar could consider devoicing of OCP-violating geminates grammatical and the other three types of devoicing ungrammatical (as Nishimura 2003 and Kawahara 2006 did), and some "extra-grammatical factors" could be responsible for yielding the differences between the latter three types of devoicing. For example, these further distinctions could arise from the fact that the scale provided was gradient: the participants were somehow forced to assume that they needed to make a full use of the scale provided. One follow-up experiment that would address the grammar-as-always-binary view is to use a binary forced-choice experiment (for which, see Kawahara 2011c).

4.3 Topics for future experimentation

Finally, this paper opens up (at least) three lines of future research. First, the current experiment used real words for two reasons: (i) in order to test the words that were used by Nishimura (2003) and Kawahara (2006) and (ii) in order to test the effect of lexical frequencies on devoiceability of OCP-violating geminates. However, in order to test the productivity of the devoicing, it would be worthwhile to run a similar experiment with nonce words (Kawahara 2011c).

Second, this experiment used visual stimuli rather than auditory stimuli. However, recall that one explanation for why only geminates can devoice under the duress of OCP(voice) is because a voicing contrast is less perceptible in geminates than in singletons (Kawahara 2006, 2008; Steriade 2004). This explanation relies on a general principle that the less perceptible the phonological change is, the more it is tolerated by speakers (Steriade 2008). An auditory judgment experiment would be better suited for testing this specific hypothesis (Kawahara 2011c).

Third, it would be interesting to see if the patterns obtained in the current experiment can be systematically replicated in a production study in which speakers actually produce the relevant sounds. In this paradigm, one would need to avoid hyperarticulation of the stimuli by the speakers, which often happens in a laboratory setting. Moreover, visual presentation of the stimuli with explicit indication of voicing in orthography may discourage devoicing of the consonants under investigation.

One promising strategy to overcome these potential difficulties is to have speakers engage in a conversation using some key words, which contain the target words of our interest.

4.4 Overall summary

To summarize, many theoretical proposals have been made using the devoicing of OCP-violating geminates in Japanese. This paper succeeded in verifying the empirical foundation of such theoretical claims. However, the current experiment also revealed additional interesting aspects of devoicing patterns in Japanese. The current experiment thus contributes to a growing body of work that shows the importance of experimental verification of linguistic data, beyond—and in tandem with—a traditional approach using intuition-based data.

Acknowledgements Thanks to Ikue Motoyanagi and Kazuko Shinohara for their help in getting many participants for the online test. For comments on previous versions of the paper, I am grateful to Aaron Braver, Maria Gouskova, Kyoko Yamaguchi, the members of Rutgers Experimental Group, including Asya Achimova, Vandana Bajaj, Justyna Grudzinska, Julien Musolino, Marta Suarez, and Kristen Syrett. My particular thanks go to the research assistants of the Rutgers phonetics lab in 2009–2010: Kelly Garvey, Lara Greenberg, Sophia Kao, and Shanna Lichtman. Comments and suggestions from three NLLT reviewers and Junko Itô also improved the content and the style of the present paper. The research is supported by a Research Council Grant from Rutgers University. All remaining errors are mine.

References

Alderete, John, and Alexei Kochetov. 2009. Japanese mimetic palatalization revisited: Implications for conflicting directionality. *Phonology* 26 (3): 369–388.

Amano, Shigeaki, and Tadahisa Kondo. 2000. *NTT database series: Lexical properties of Japanese, 2nd release*. Tokyo: Sanseido.

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

Baayen, Harald R. 2009. LanguageR. R package.

Baayen, Harald R., Doug J. Davidson, and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390–412.

Batchelder, Eleanor Olds. 1999. Rule or rote? Native-speaker knowledge of Japanese verb inflection. In *Proceedings of the second international conference on cognitive science*.

Bates, Douglas M. 2005. Fitting linear mixed models in R. *R News* 5: 27–30.

Bates, Douglas, Martin Maechler, and Ben Bolker. 2011. lme4: Linear mixed-effects models using S4 classes. R package.

Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona vowel harmony. *Phonology* 14 (1): 1–46.

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

Blevins, Juliette. 2005. Understanding antigemination: Natural or unnatural history. In *Linguistic diversity and language theories*, eds. Zygmunt Frajzyngier, David Rood, and Adam Hodges, 203–234. Amsterdam: Benjamins.

Bybee, Joan. 1999. Usage-based phonology. In *Functionalism and formalism in linguistics*, eds. Michael Darnell, Edith Moravcsik, Michael Noonan, Frederick Newmeyer, and Kathleen Wheatly, 211–242. Amsterdam: Benjamins.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.

Coetzee, Andries W. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84 (2): 218–257.

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

Coetzee, Andries W., and Shigeto Kawahara. 2011. Frequency biases in phonological variation. Manuscript. Michigan University and Rutgers University (submitted for publication in *Natural Language & Linguistic Theory*).

Coetzee, Andries W., and Joe Pater. 2011. The place of variation in phonological theory. In *The handbook of phonological theory*, 2nd edn., eds. John A. Goldsmith, Jason Riggle, and Alan Yu. Oxford: Blackwell-Wiley (to appear).

Cohn, Abigail. 2006. Is there gradient phonology? In *Gradience in grammar: Generative perspectives*, eds. Gisbert Fanselow, Caroline Fery, Matthias Schlesewsky, and Ralf Vogel, 25–44. Oxford: Oxford University Press.

Coleman, John, and Janet Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. In *Computational phonology: Third meeting of the ACL special interest group in computational phonology*, 49–56. Somerset: Association for Computational Linguistics.

Cowart, Wayne. 1997. *Experimental syntax: Applying objective methods to sentence judgments*. London: Sage.

Crawford, J. Clifford. 2009. Adaptation and transmission in Japanese loanword phonology. Doctoral dissertation, Cornell University.

Dabrowska, Ewa. 2010. Naive vs. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27 (1): 1–23.

Davis, Stuart, and Natsuko Tsujimura. 1991. An autosegmental account of Japanese verbal conjugation. *Journal of Japanese Linguistics* 13: 117–144.

Ernestus, Mirjam, and Harald R. Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79: 5–38.

Fanselow, Gisbert, Caroline Féry, Ralf Vogel, and Matthias Schlesewsky, eds. 2006. *Gradience in grammar: Generative perspectives*. Oxford: Oxford University Press.

Farris-Trimble, Ashley. 2008. Cumulative faithfulness effects in phonology. Doctoral dissertation, Indiana University.

Flemming, Edward. 1995. Auditory representations in phonology. Doctoral dissertation, University of California, Los Angeles.

Frisch, Stefan, Nathan Large, and David Pisoni. 2000. Perception of wordlikiness: Effects of segment probability and length of the processing of nonwords. *Journal of Memory and Language* 42: 481–496.

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

Gibson, Edward, and Evelina Fedorenko. 2010. Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* 14 (6): 233–234.

Greenberg, Joseph, and James Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20: 157–177.

Griner, Barry. 2001. Productivity of Japanese verb tense inflection: A case study. MA thesis, University of California, Los Angeles.

Hamano, Shoko. 1986. The sound-symbolic system of Japanese. Doctoral dissertation, University of Florida.

Hamano, Shoko. 2000. Voicing of obstruents in Old Japanese: Evidence from the sound-symbolic stratum. *Journal of East Asian Linguistics* 9: 207–215.

Haraguchi, Shosuke. 2006. A theory of voicing. Manuscript. Meikai University (available at http://www.let.leidenuniv.nl/ulcl/faculty/vdweijer/jvoice/haraguchi.pdf).

Hawkins, John, and Anne Cutler. 1988. Psycholinguistic factors in morphological asymmetry. In *Explaining language universals.*, ed. J. A. Hawkins, 280–317. Oxford: Basil Blackwell.

Hay, Jennifer, Janet Pierrehumbert, and Mary Beckman. 2003. Speech perception, well-formedness, and the statistics of the lexicon. In *Papers in laboratory phonology VI: Phonetic interpretation*, eds. John Local, Richard Ogden, and Rosalind Temple, 58–74. Cambridge: Cambridge University Press.

Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. In *Optimality Theory: Phonology, syntax, and acquisition*, eds. Joost Dekkers, Frank Van der Leeuw, and Jeroen Van de Weijer, 88–120. Oxford: Oxford University Press.

Hayes, Bruce. 2009. Embedding grammar in a quantitative framework: Case studies from phonology and metrics. A handout for a minicourse at Brown University.

Hayes, Bruce, and Zsuzsa Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23: 59–104.

Horowitz, Leonard, Margeret White, and Douglas Atwood. 1968. Word fragments as aids to recall: The organization of a word. *Journal of Experimental Psychology* 76 (2): 219–226.
Horowitz, Leonard, Peter Chilian, and Kenneth Dunnigan. 1969. Word fragments and their reintegrative powers. *Journal of Experimental Psychology* 80 (2): 392–394.
Ichimura, Larry. 2006. Anti-homophony blocking and its productivity in transparadigmatic relations. Doctoral dissertation, Boston University.
Itô, Junko, and Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17: 49–73.
Itô, Junko, and Armin Mester. 1995. Japanese phonology. In *The handbook of phonological theory*, ed. John Goldsmith, 817–838. Oxford: Blackwell.
Itô, Junko, and Armin Mester. 1999. The phonological lexicon. In *The handbook of Japanese linguistics*, ed. Natsuko Tsujimura, 62–100. Oxford: Blackwell.
Itô, Junko, and Armin Mester. 2004. Morphological contrast and merger: Ranuki in Japanese. *Journal of Japanese Linguistics* 20: 1–19.
Itô, Junko, and Armin Mester. 2008. Lexical classes in phonology. In *The Oxford handbook of Japanese linguistics*, eds. S. Miyagawa and M. Saito, 84–106. Oxford: Oxford University Press.
Jaeger, Jeri J. 1983. Assessing the psychological status of the vowel shift rule. *Journal of Psycholinguistic Research* 13: 13–36.
Kaneko, Emiko, and 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*, eds. Shunji Inagaki and Makiko Hirakawa. Tokyo: Kuroshio.
Katayama, Motoko. 1998. Optimality Theory and Japanese loanword phonology. Doctoral dissertation, University of California, Santa Cruz.
Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language* 82 (3): 536–574.
Kawahara, Shigeto. 2008. Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* 18 (4): 317–330.
Kawahara, Shigeto. 2011a. Aspects of Japanese loanword devoicing. *Journal of East Asian Linguistics* 20.
Kawahara, Shigeto. 2011b. Experimental approaches in theoretical phonology. In *The Blackwell companion to phonology*, eds. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 2283–2303. Oxford: Blackwell-Wiley.
Kawahara, Shigeto. 2011c. Modes of phonological judgments. Manuscript. Rutgers University. (Talk presented at the University of Pennsylvania, Oct 2010).
Kawahara, Shigeto, and Kazuko Shinohara. 2011. Phonetic and psycholinguistic prominences in pun formation: Experimental evidence for positional faithfulness. In *Japanese/Korean linguistics 18*, eds. W. McClure and M. den Dikken, 177–188. Stanford: CSLI.
Lovins, Julie. 1973. Loanwords and the phonological structure. Doctoral dissertation, University of Chicago.
Lyman, Benjamin S. 1894. Change from surd to sonant in Japanese compounds. *Oriental Studies of the Oriental Club of Philadelphia*.
McCarthy, John J. 2008. *Doing optimality theory*. Oxford: Blackwell-Wiley.
Mester, Armin, and Junko Itô. 1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65: 258–293.
Myers, James. 2009. Syntactic judgment experiments. *Language and Linguistic Compass* 3 (1): 406–423.
Nishimura, Kohei. 2003. Lyman's Law in loanwords. MA thesis, Nagoya University.
Nolan, Francis. 1992. The descriptive role of segments: Evidence from assimilation. In *Papers in laboratory phonology II: Gesture, segment, prosody*, eds. G. R. Docherty and D. R. Ladd, 261–280. Cambridge: Cambridge University Press.
Nooteboom, Sieb. 1981. Lexical retrieval from fragments of spoken words: Beginnings vs. endings. *Journal of Phonetics* 9: 407–424.
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*, eds. J. M. Naderson and Charles Jones, 353–389. New York: Elsevier.
Padgett, Jaye. 2002. Contrast and postvelar fronting in Russian. *Natural Language & Linguistic Theory* 21: 39–87.
Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33: 999–1035.
Pertz, D. L., and T. G. Bever. 1975. Sensitivity to phonological universals in children and adolescents. *Language* 51: 149–162.

Pierrehumbert, Janet B. 2001. Stochastic phonology. *GLOT* 5: 1–13.
Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality theory: Constraint interaction in generative grammar*. Malden and Oxford: Blackwell.
Pycha, Anne, Eurie Shin, and Rhan Shosted. 2006. Directionality of assimilation in consonant clusters: An experimental approach. UC Berkeley Phonology Lab Annual Report.
R Development Core Team. 1993–2011. R: A language and environment for statistical computing. Vienna, Austria. R Foundation for Statistical Computing. Software, available at http://www.R-project.org.
Reips, Ulf-Dietrich. 2002. Standards for internet-based experimenting. *Experimental Psychology* 49 (4): 243–256.
Rice, Keren. 2006. On the patterning of voiced stops in loanwords in Japanese. *Toronto Working Papers in Linguistics* 26: 11–22.
Sanders, Nathan. 2003. Opacity and sound change in the Polish lexicon. Doctoral dissertation, University of California, Santa Cruz.
Schütze, Carlson. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
Shinohara, Shigeko. 2004. Emergence of universal grammar in foreign word adaptations. In *Constraints in phonological acquisition*, eds. René Kager, Joe Pater, and Wim Zonneveld, 292–320. Cambridge: Cambridge University Press.
Shirai, Setsuko. 2002. Gemination in loans from English to Japanese. MA thesis, University of Washington.
Sprouse, Jon. 2007. Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1: 118–129.
Steriade, Donca. 2004. Projecting non-lexical phonology from phonetic knowledge. Talk presented at "Symposium on phonological theory: representations and architectures" (the City University of New York, Feb 20–21st, 2004).
Steriade, Donca. 2008. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The nature of the word*, eds. Kristin Hanson and Sharon Inkelas, 151–179. Cambridge: MIT Press.
Suzuki, Keiichiro. 1998. A typological investigation of dissimilation. Doctoral dissertation, University of Arizona.
Tanaka, Shin-Ichi. 2010. The eurhythmics of segmental melody. *Journal of the Phonetic Society of Japan* 13 (3): 44–52.
Tateishi, Koichi. 2002. Bunpoo no ichibutoshite no goisoo no zehi [Lexical strata as a part of grammar]. *Journal of the Phonetic Society of Japan* 6 (1): 34–43.
Tesar, Bruce. 2007. A comparison of lexicographic and linear numeric optimization using violation difference ratios. Manuscript. Rutgers University.
Tsuchida, Ayako. 1997. Phonetics and phonology of Japanese vowel devoicing. Doctoral dissertation, Cornell University.
Urbanczyk, Suzanne. 2005. Enhancing contrast in reduplication. In *Studies on reduplication*, ed. B. Hurch, 431–454. Mouton: de Gruyter.
Vance, Timothy J. 1980. The psychological status of a constraint on Japanese consonant alternation. *Linguistics* 18: 245–267.
Vance, Timothy J. 1987. *An introduction to Japanese phonology*. New York: SUNY Press.
Westbury, John R., and Patricia Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22: 145–166.
Zuraw, Kie. 2000. Patterned exceptions in phonology. Doctoral dissertation, University of California, Los Angeles.
Zuraw, Kie. 2009. Frequency influences on rule application within and across words. In *Proceedings of Chicago Linguistic Society 43-2*, 283–309. Chicago: Chicago Linguistics Society.

Kawahara, Shigeto (2011) Aspects of Japanese loanword devoicing. Journal of East Asian Linguistics 20(2): 169–194.

J East Asian Linguist (2011) 20:169–194
DOI 10.1007/s10831-011-9074-3

Aspects of Japanese loanword devoicing

Shigeto Kawahara

Received: 19 June 2010 / Accepted: 17 February 2011 / Published online: 12 April 2011

Abstract Nishimura (M.A. thesis, 2003) first pointed out that in Japanese loan words, voiced geminates devoice optionally when they co-occur with another voiced obstruent, i.e., when they violate OCP(voice) (e.g., /baggu/ → [bakku] 'bag'). This devoicing of geminates has been used to make several theoretical claims in the recent phonological literature. However, these claims have so far largely been based on intuition-based data provided by Nishimura (M.A. thesis, 2003) and Kawahara (Language 82(3):536–574, 2006). Kawahara (Nat Lang Linguist Theory, 2011a) addressed this problem by conducting a rating study. The first aim of this study, building on Kawahara (Nat Lang Linguist Theory, 2011a), is to further support the empirical foundation of these theoretical claims by way of a large-scale rating study. The current study shows that (i) the OCP and geminacy each affect naturalness rating of devoicing, and (ii) there is nevertheless something special about the combination of the OCP and geminacy. The second aim is to test an assumption behind the recent literature on this phenomenon. The assumption is that this devoicing pattern is monolithic—i.e., all voiced geminates uniformly undergo devoicing in a certain phonological environment. The current experiment shows that this assumption is too simplistic. In particular it shows (i) speakers rate the devoicing of affricates as natural as that of stops, (ii) speakers find devoicing of items that merge with other lexical items less natural, (iii) speakers rate devoicing as more natural when there are multiple triggers, (iv) speakers find devoicing of [dd] more natural than that of [gg], and (v) speakers find devoicing of more frequent items more natural.

Keywords Japanese · Devoicing · OCP · Geminates · Homophony avoidance · Lexical frequency · Experimental phonology · Phonological judgment

S. Kawahara (✉)
Department of Linguistics, Rutgers University, 18 Seminary Place,
New Brunswick, NJ 08901, USA
e-mail: kawahara@rci.rutgers.edu

1 Introduction

1.1 The phenomenon

Nishimura (2003) first pointed out that in Japanese loanword phonology, voiced geminates optionally devoice when they occur with another voiced obstruent, as in (1).[1] This devoicing of geminates is caused by a restriction against two voiced obstruents within the same stem, which is known as the OCP(voice) (henceforth the OCP) (for the OCP in Japanese, see Itô and Mester 1986, 2003 as well as discussion below). Nishimura (2003) and Kawahara (2006) contrast OCP-violating geminates with voiced consonants in two other contexts: non-OCP-violating voiced geminates and OCP-violating singletons, which are claimed not to undergo devoicing, as in (2)–(3).

(1) Geminates can optionally devoice if they co-occur with another voiced obstruent
    a. ba**dd**o → ba**tt**o 'bad'
    b. dore**dd**o → dore**tt**o 'dread'
    c. deibi**dd**o → deibi**tt**o 'David'
    d. ba**gg**u → ba**kk**u 'bag'
    e. do**gg**u → do**kk**u 'dog'

(2) Geminates do not devoice otherwise
    a. suno**bb**u → *suno**pp**u 'snob'
    b. re**dd**o → *re**tt**o 'red'
    c. e**gg**u → *e**kk**u 'egg'

(3) Singletons do not devoice when they violate the OCP
    a. gi**b**u → *gi**p**u 'give'
    b. ba**g**u → *ba**k**u 'bug'
    c. da**g**u → *da**k**u 'Doug'

Since Nishimura (2003), the patterns in (1)–(3) have received much attention in the theoretical literature (Coetzee and Pater, to appear; Crawford 2009; Farris-Trimble 2008; Haraguchi 2006; Hayes 2009; Itô and Mester 2008; Kaneko and Iverson 2009; Kawahara 2005, 2006, 2008; McCarthy 2008; Nishimura 2006; Pater 2009; Pycha et al., 2006; Rice 2006; Steriade 2004; Tanaka 2010; Tateishi 2002; Tesar 2007). To briefly summarize the debates concerning (1)–(3), one debate addresses how the difference between singletons and geminates arises. Kawahara (2005, 2006) demonstrates through an acoustic experiment that Japanese voiced geminates are phonetically half devoiced. The follow-up perception experiment

[1] In an online adaptation experiment reported in Kaneko and Iverson (2009), Japanese speakers did not show evidence for this OCP-induced devoicing when they adapted new loanwords. Their finding shows that this devoicing occurs in loanword phonology, not in the process of loanword adaptation (Kaneko and Iverson 2009; Kawahara 2006, 2008, 2011a). In other words, devoicing does not happen *when* Japanese speakers borrow these loanwords, but it does *after* they borrow them.

further demonstrates that a voicing contrast is perceptually less salient in geminates than in singletons. Kawahara (2005, 2006, 2008) argues based on these results that this low perceptibility is the source of the higher neutralizability of geminates. Rice (2006) on the other hand argues that a voicing contrast in geminates is more neutralizable than the one in singletons because the former is not contrastive in Japanese native phonology. These studies argue therefore that a difference in phonological neutralizability arises from either a difference in phonetic perceptibility (Steriade 2001/2008) or contrastiveness elsewhere in the phonology (Dresher 2010).

The second debate focuses on the cumulative behavior of devoicing in (1), which is also known as a "gang effect". Nishimura (2003) and Kawahara (2006) claim that neither geminacy nor the OCP alone can cause devoicing, as in (2) and (3), but that when both factors are relevant devoicing can take place. This gang effect may bear on a general debate about how to model constraint interaction (Coetzee and Pater, to appear; Farris-Trimble 2008; Hayes 2009; Nishimura 2003; Pater 2009; Tesar 2007). This sort of gang effect cannot be modeled in a ranking-based theory of constraint interaction, such as Optimality Theory (Prince and Smolensky 1993/2004), without additional mechanisms.[2] However, a theory with weighted constraints can model the cumulative pattern without extra mechanisms (Pater 2009).

The final debate is about how the emergence of the loanword devoicing in (1) bears on a theory of lexical stratification. The OCP is active in native Japanese phonology (Itô and Mester 1986), and the data in (1) show that the OCP produces an emergent phonological pattern in loanword phonology. This connection may shed light on how loanword phonology derives from native phonology (Crawford 2009; Itô and Mester 2008; Tateishi 2002).

In summary, the patterns in (1)–(3) have evoked many theoretical debates. However, one problem identified by Kawahara (2011a) is that these theoretical claims have been primarily based on the intuitions of Nishimura (2003) and Kawahara (2006). Several studies have shown potential pitfalls of an approach relying purely on authors' introspections (see, among others, Alderete and Kochetov 2009; Dabrowska 2010; Gibson and Fedorenko 2010; Griner 2001; Kawahara 2011a; Labov 1975, 1996; Myers 2009; Ohala 1974, 1986; Schütze 1996; Vance 1980 and references cited therein). Since several important theoretical claims have been made based on the Japanese devoicing data, their empirical foundation should ideally be supported by systematic experimentation with a number of theoretically-unbiased native speakers. To address this problem, Kawahara (2011a) conducted a rating experiment with a number of native Japanese speakers. The experiment indeed showed that Japanese speakers generally find devoicing of OCP-violating geminates more natural than devoicing in other environments. Kawahara (2011a) thus succeeds in securing the empirical basis of the claims made based on the patterns in (1)–(3).

However, Kawahara (2011a) found two aspects which go beyond what Nishimura (2003) and Kawahara (2006) report based on their intuitions. First,

[2] Two solutions have been proposed: (i) to posit a complex locally-conjoined (Smolensky 1995) markedness constraint that is violated only when both the OCP and a constraint against voiced geminates are violated (Nishimura 2003, 2006), or (ii) to posit different faithfulness constraints for singletons and geminates (Kawahara 2006, 2008).

Japanese speakers distinguish the grammaticality of two processes that were both judged to be "ungrammatical" by Nishimura (2003) and Kawahara (2006): devoicing of non-OCP violating geminates and devoicing of OCP-violating singletons. Given that there is a three way distinction in acceptability between OCP-violating geminates, non-OCP-violating geminates, and OCP-violating singletons, Kawahara (2011a) contends that it is difficult to draw an objective line between grammatical processes and ungrammatical processes. One of the aims of this paper is to test whether this fine-grained grammatical distinction holds in a wider set of data than considered in Kawahara (2011a).

Second, Kawahara (2011a) found that the devoicing of OCP-violating geminates may not be as monolithic as previous studies have assumed. In other words, the experiment found that there are specific types of words with certain properties—phonological and otherwise—which showed higher/lower ratings than other items. The results thus cast doubt on the assumption that the devoicing of OCP-violating geminates is a uniform, monolithic phonological pattern, as Nishimura (2003) (and subsequent studies) have assumed.[3] Rather, various factors may contribute to the naturalness of devoicing of OCP-violating geminates. This paper therefore reports a follow-up experiment to address questions that are raised—and remained unanswered—in Kawahara (2011a) (and beyond).

1.2 Seven hypotheses tested

The specific hypotheses about Japanese devoicing that this paper addresses are listed in (4).

(4) Hypotheses tested

a. Singletons: Does the OCP affect naturalness of devoicing of singletons? If so, does it uniformly affect singletons and geminates?

b. Affricates: Do Japanese speakers find devoicing of voiced affricates as natural as that of voiced stops?

c. Merger: Do Japanese speakers rate devoicing less natural if devoicing would result in merger with another lexical item?

d. Locality: Does the distance between the trigger and the voiced consonants affect naturalness of devoicing?

e. Multiple trigger: Do the numbers of the triggers affect naturalness of devoicing?

f. Place effects: Does place of articulation affect naturalness of devoicing?

g. Lexical frequency effects: Do the lexical frequencies of the words affect naturalness of devoicing?

First, Kawahara (2011a) found that the OCP does make the devoicing of geminates more natural, but that study did not test whether the OCP makes

[3] This paper is not meant to undermine the value of the discovery by Nishimura (2003). Idealization is a standard methodology in linguistics (and other fields). This paper attempts to reveal a more realistic nature of devoicing in Japanese loanwords. "[I]dealization is only the *starting point* of inquiry, not its goal" (Riemer 2009, p. 626, emphasis in original).

devoicing of singletons more natural. The assumption was that devoicing of OCP-violating singletons was already “bad enough;” i.e., devoicing of singletons would be ungrammatical regardless of whether they violate the OCP or not. However, Kawahara (2011a) found that Japanese speakers distinguish between two “ungrammatical processes” (e.g., devoicing of non-OCP-violating geminates and devoicing of OCP-violating singletons), and therefore, it is of some interest whether Japanese speakers would treat devoicing of singletons differently depending on whether the words under question violate the OCP or not. A related question, to the extent that the OCP does influence the naturalness of devoicing of singletons, is whether the OCP affects singletons and geminates to the same degree.

Second, the status of devoicing of affricates has not been explicitly tested.[4] Neither Nishimura (2003) nor Kawahara (2006) report their intuitions about affricates. Haraguchi (2006) treats some intervocalic affricates as fricatives (which allophonically do alternate in this position; see footnote 4), and judges their devoicing to be “?*”. Nishimura (2006) treats affricates on a par with voiced stops. Kawahara (2011a) did not include affricates in the stimuli. Therefore, whether affricates and stops pattern in the same way or not is an open question.

Third, some items can result in merger with already-existing lexical items ([baggu] ‘bag’ vs. [bakku] ‘back’). Ichimura (2006) argues that (Japanese) speakers avoid such mergers of existing lexical items. More generally, such anti-merger constraints have been proposed by many researchers for various languages (Blevins 2005; Crosswhite 1999; Itô and Mester 2004; Kaplan 2010; Lubowicz 2007; Padgett 2002; Urbanczyk 2005), especially within the context of Optimality-Theoretic (Prince and Smolensky 1993/2004) Dispersion Theory (Flemming 1995). Kawahara (2011a) found a non-significant tendency toward the expected direction, and a follow-up study with more relevant lexical items was claimed to be necessary.

Fourth, there are items in which the trigger of devoicing is distant from the devoicing geminates (e.g. [doreddo] ‘dread’). Dissimilatory force is cross-linguiatically known to be stronger between two closer elements, i.e., with less fewer intervening segments (Frisch et al. 2004; Itô and Mester 2003; Suzuki 1998; Tanaka 2007). In fact, there is evidence within Japanese that locality may matter: In a phenomenon known as Rendaku, the initial consonant of the second member of a compound becomes voiced (e.g. /nise+tanuki/ → [nise-danuki] ‘fake raccoon’); however, if there is already a voiced obstruent in the second member, this voicing is blocked (i.e., Lyman’s Law) (e.g. /nise+tokage/ → [nise-tokage] ‘fake lizard’) (Itô and Mester 1986, 2003; Lyman 1894). Vance (1980) found in a nonce-word experiment that the closer the blocker consonant, the less likely that Rendaku occurs. Therefore, we expect that in words in which the trigger and the geminates are non-local, speakers may find devoicing of OCP-violating geminates less natural. Kawahara (2011a) in fact found such a tendency, although it did not reach statistical significance. However, the stimuli contained only two relevant items ([doreddo] ‘dread’ and

[4] Voiced affricates in Japanese alternate with fricatives intervocalically. Whether this alternation is free-variation or allophonic alternation is controversial (Maekawa 2009, 2010; Vance 1987); for the sake of exposition, I treat them as affricates in this paper. See Maekawa (2009, 2010) for phonetic factors that influence this alternation.

[doraggu] 'drug'), leaving the possibility that that experiment did not have enough statistical power.

Fifth, there are items that contain two triggers of devoicing (e.g. [deibiddo] 'David'). Tesar (2007) points out that in such words, devoicing of geminates would still violate the OCP (e.g., [deibitto]). It is possible that the OCP militates against each pair of voiced obstruents, and therefore devoicing of geminates does ameliorate the violation of the OCP (Kawahara 2008; Tesar 2007). For example, [deibiddo] incurs three violations ([d]–[b], [d]–[dd], [b]–[dd]) whereas [deibitto] incurs only one ([d]–[b]). Nevertheless, it is also possible that speakers may find devoicing of geminates less natural in this kind of word, because of the remaining violation(s) of the OCP. Kawahara (2011a) found a non-statistical tendency toward this direction; however, the stimuli contained only two relevant items ([deibiddo] 'David' and [deddobooru] 'deadball'), again leaving the possibility that the experiment lacked enough statistical power to reveal a real difference.

Sixth, according to the data in Nishimura (2003), [gg] is more likely to devoice than [dd] (Kawahara 2006). In terms of the aerodynamic difficulty that voiced geminates present, this observation makes sense (Hayes and Steriade 2004; Jaeger 1978; Ohala 1983; Ohala and Riordan 1979; Westbury 1979). To maintain voicing, it is necessary to keep intraoral air pressure lower than subglottal air pressure. However, with stop closure, airflow required to maintain voicing increases the intraoral air pressure. The rise in intraoral air pressure occurs more quickly for [gg] than for [dd] because the intraoral cavity behind the constriction is smaller and less flexible to expand for [gg]. Transglottal air pressure drop necessary to sustain voicing should then decease more quickly for [gg] than for [dd] (see the references cited above). On the other hand, Kawahara (2011a) did not find this asymmetry in the judgment patterns; in fact, the experiment found a non-significant tendency toward Japanese speakers rating the devoicing of [dd] more natural than that of [gg]. This follow-up experiment is designed to address this conflict by including more relevant items.

Finally, Kawahara (2011a) found that lexical frequencies of the target items affect the naturalness of devoicing in that devoicing of more frequent words was judged to be more natural. Moreover, this positive correlation was found only in certain grammatical contexts (OCP-violating geminates and word-internal OCP-violating singletons) but not in other contexts (non-OCP-violating geminates and word-initial OCP-violating singletons). However, this limited effect of lexical frequencies on naturalness rating should be taken with caution because some conditions had only nine items in Kawahara (2011a). The current study is thus designed to confirm the limited effect of lexical frequencies with a larger set of data. This question is in part motivated by a growing body of interest in to what extent lexical frequency affects phonological regularity (for different proposals on this issue, see e.g. Boersma and Hayes 2001; Bybee 1999, 2001; Coetzee 2009b; Coetzee and Kawahara 2010; Coleman and Pierrehumbert 1997; Frisch et al. 2000; Hay et al. 2003; Hayes and Londe 2006).

These are the seven questions that the current experiment aims to address. To summarize, Kawahara (2011a) aimed to make a general comparison between OCP-violating geminates and other consonants, and it indeed showed that Japanese speakers treat OCP-violating geminates differently from other consonants. It also found hints of interesting variations within OCP-violating geminates. However, the study was limited in its power because it did not have enough lexical items. The current experiment thus follows up on Kawahara (2011a) to address these remaining questions. The current experiment therefore includes many more lexical items to guarantee sufficient statistical power.

2 Method

2.1 Stimuli

Table 1 lists the entire stimuli. (In this paper I use romanized representation of Japanese except for the affricate [ʤ].) To address the questions reviewed in Sect. 1.2, the stimuli were constructed as follows. The design had four conditions: OCP-violating geminates, non-OCP-violating geminates, OCP-violating singletons, and non-OCP-violating singletons. The current experiment included many relevant lexical items because, for some hypotheses discussed above, Kawahara (2011a) considered only a few relevant items. To that end, first, many words containing OCP-violating geminates were collected, partly based on three previous studies (Kawahara 2011a; Nishimura 2003, 2006), which resulted in 28 items. Within these items, one item contained [bb], 16 items contained [dd], six items contained [gg], and five items contained affricates. There is only one item containing [bb], because [bb] barely appears in Japanese loanwords in the first place; on the other hand, there are many words containing [dd], because [dd] is most commonly found (Kaneko and Iverson 2009; Katayama 1998; Shirai 2002). After collecting items for the OCP-violating gemination condition, for the other three conditions, 28 items were collected with the same numbers of items with matching place of articulation. This resulted in 112 items in total (cf. Kawahara 2011a had 33 items).

The items in which the target consonants come after nasal consonants were avoided because nasals encourage voicing—and may hinder devoicing—in the following consonants (Hayes and Stivers 1995; Pater 1999). Since the OCP-violating geminates generally appear in word-final syllable, the target consonants in the other conditions were placed in word-final syllables as much as possible. In the OCP-violating geminate condition, those that have non-local triggers are [buraddo], [bureddo], [doreddo], [guriddo], [haiburiddo], [madoriddo], [sarabureddo], [doraggu], [bireddʤi], [buriddʤi], and [kemburiddʤi]. Those that have multiple triggers are [bagudaddo], [deibiddo], [sindobaddo], and [debaggu]. Those that would result in merger with other existing lexical loanword items are [baddo] vs. [batto] 'bat', [daddo] vs. [datto] 'DAT', [baggu] vs. [bakku] 'back', and [doggu] vs. [dokku] 'dock'.[5]

[5] Since the instructions of the experiment made it clear that the experiment was about loanwords, merger with non-loanwords was not considered to be merging. For OCP-violating geminates, there were two such items: [budda] and [guddo] could be considered as merging with phrases [but-ta] 'hit (past)' and [gut-to] 'patiently' (the latter with an accentual difference).

Table 1 The list of the stimuli

OCP-violating geminates		Non-OCP-violating geminates		OCP-violating singletons		Non-OCP-violating singletons	
gebberusu	‘Göbbels’	habburu	‘Hubble’	ʤdobu	‘job’	ibu	‘eve’
baddo	‘bad’	hariuddo	‘Hollywood’	baado	‘bird’	haado	‘hard’
bagudaddo	‘Bagdad’	huddo	‘hood’	bado	‘badminton’	haido	‘hide’
budda	‘Buddha’	heddo	‘head’	baraado	‘ballad’	huudo	‘food’
buraddo	‘blood’	kiddo	‘kid’	barikeedo	‘barricade’	kaado	‘card’
bureddo	‘bread’	maddo	‘mad’	bideo	‘video’	koodo	‘cord’
daddo	‘dad’	mesoddo	‘method’	biriyaado	‘billiard’	moodo	‘mode’
deddo	‘dead’	middo	‘mid’	biroodo	‘velvet’	muudo	‘mood’
deibiddo	‘David’	reddo	‘red’	boodaa	‘border’	niido	‘need’
doreddo	‘dread’	rikiddo	‘liquid’	boodo	‘board’	nuudo	‘naked’
goddo	‘god’	roddo	‘rod’	daioodo	‘dyode’	raado	‘lard’
guddo	‘good’	shureddaa	‘shredder’	gaaden	‘garden’	raidaa	‘rider’
guriddo	‘grid’	sureddo	‘thread’	gaado	‘guard’	waado	‘word’
haiburiddo	‘hybrid’	teddo	‘Ted’	gaido	‘guide’	roodo	‘rode’
madoriddo	‘Madrid’	uddo	‘wood’	goorudo	‘gold’	saido	‘side’
sarabureddo	‘thoroughbred’	yunaiteddo	‘United’	gureedo	‘grade’	siido	‘seed’
sindobaddo	‘Sindbad’	piramiddo	‘pyramid’	zoido	(common name)	saado	‘third’
baggu	‘bag’	eggu	‘egg’	bagu	‘bug’	hagu	‘hug’
biggu	‘big’	furaggu	‘flag’	ʤogu	‘jog’	riigu	‘league’
debaggu	‘debug’	furoggu	‘frog’	burogu	‘blog’	magu	‘mug’
doggu	‘dog’	reggu	‘leg’	giga	‘giga’	ragu	‘rag’
doraggu	‘drag’	sumoggu	‘smog’	tagarogu	‘Tagalog’	tagu	‘tag’
badʤi	‘badge’	maridʤi	‘marriage’	dameeʤi	‘damage’	keeʤi	‘cage’
biredʤi	‘village’	hedʤi	‘hedge’	engeeʤi	‘engage’	peeʤi	‘page’
buridʤi	‘bridge’	rodʤi	‘lodge’	gareeʤi	‘garage’	raaʤi	‘large’
ʤadʤi	‘judge’	edʤi	‘edge’	geeʤi	‘gauge’	imeeʤi	‘image’
kemburidʤi	‘Cambridge’	karedʤi	‘college’	ʤooʤi	‘George’	chaaʤi	‘charge’
gudzzu	‘goods’	kidzzu	‘kids’	ʤazu	‘jazz’	ruuzu	‘loose’

2.2 Task

In this experiment Japanese speakers rated the naturalness of devoicing in four grammatical conditions. The instructions explained that the questionnaire was about the naturalness of devoicing in Japanese loanwords. For each question, the participants were presented with one stimulus, and were asked to judge the naturalness of the form that undergoes devoicing of word-internal consonants (e.g., given [baddo], how natural would you find it to pronounce it as [batto]?). For stimuli with two singleton voiced consonants, the devoiced forms were those that devoice the second voiced consonant. The target stimuli as well as their devoiced forms were all presented in Japanese *katakana* orthography. (See Kawahara 2011b for a comparison between orthography-based rating and audio-based rating of the devoicing patterns under discussion.)

Following Kawahara (2011a), the speakers judged the naturalness of devoicing on a 5-point scale: A "very natural," B "somewhat natural," C "neither natural nor unnatural," D "somewhat unnatural," and E "very unnatural." (The software used to present the stimuli (sakai; see below) could not present the scale numerically, so the responses were converted to a numerical scale later). The speakers were asked to read the stimuli before answering each question, and to base their decision on their auditory impression rather than on orthography. The devoiced form of affricates—although they can alternate with fricatives intervocalically (see footnote 4)—were represented as voiceless affricates, rather than voiceless fricatives, following Nishimura (2006).

2.3 Procedure

Sakai (https://sakai.rutgers.edu/portal) was used to run the online experiment (see Reips 2002 for general discussion on online experimentation). The first page of the experimental website presented a consent form for a human subject experiment, and then they were forwarded to a testing site. After that, each page presented one stimulus. The instructions and the options were provided in Japanese orthography. Sakai randomized the order of the stimuli. At the end of the experiment, participants were asked if they knew the devoicing of OCP-violating geminates discussed in Kawahara's work; data from those who answered positively to this question were excluded from the following analysis, in order to prevent the results from being affected by the participants' theoretical orientation.

2.4 Participants

To obtain enough statistical power, the experiment was run until 52 native speakers of Japanese completed the study. Three speakers were familiar with the devoicing phenomenon and therefore were excluded from the following analysis.

2.5 Frequency measures

The frequencies of the stimuli were taken from a Japanese lexical corpus, Amano and Kondo (2000). Following the standards of psycholinguistic studies (Rubin 1976;

Smith and Dixon 1971), the natural log frequencies of these values were calculated (ln(0) was replaced with 0). The mean log-frequencies of the four grammatical conditions were as follows: OCP-violating geminates: 4.39; non-OCP-violating geminates: 4.01; OCP-violating singletons: 4.51; non-OCP-violating singletons: 5.51.

2.6 Statistics

The responses were first converted to numerical values in the following way: "very natural" = 5; "somewhat natural" = 4; "neither natural nor unnatural" = 3; "somewhat unnatural" = 2; "very unnatural" = 1. For statistical analyses, first, a general linear mixed model was run (Baayen et al. 2008; Baayen 2008; Bates 2005; Jaeger 2008) using R (R Development Core Team 1993–2011) with the `lme4` package (Douglas et al. 2011).

The first analysis included all the data points. Rating scores were regressed against a general model in which OCP, geminacy, affricacy (i.e., affricates vs. stops), the possibility of merger, and lexical frequency were fixed factors, and speakers and items were random factors. Two other factors—the number of intervening syllables and the number of triggers—were not included in this general model, because they concern only OCP-violating words. Place was not coded for affricates—hence not included in this general model either—because the place of affricates was largely predictable. This general model included many fixed factors, partly to test whether two grammatical factors—OCP and geminacy—would have effects on naturalness rating beyond the other factors. The general model also included the interaction term between the OCP and geminacy. A significant interaction would mean that there is something special about the combination between the OCP and geminacy, i.e., there is something special about OCP-violating geminates, as Nishimura (2003) and Kawahara (2006) claim.

The second analysis only analyzed OCP-violating geminates, the grammatical condition that is of most interest to us. This model included the possibility of merger, the multiplicity of triggers, the number of intervening syllables, and lexical frequency as fixed factors. Affricacy was not included in this model, because it correlated highly with lexical frequency ($r = 0.37$), and the inclusion of affricacy would have resulted in a multicollinearity problem. More specific hypotheses that were not covered by these general linear mixed models were addressed by simple t-test based contrast analyses. The alpha level was not adjusted because different tests test different, independent pre-planned hypotheses.

The `lme4` package does not automatically compute p-values because the exact procedure to calculate degrees of freedom is not known for linear mixed model analyses. Therefore, they were instead calculated by the Markov chain Monte Carlo method using the `pval.func()` function of the `languageR` package (Baayen 2009). Finally, the correlation between the ratings and the frequencies was tested by a Spearman correlation test using R.

3 Results

3.1 The four grammatical conditions

Figure 1 illustrates the average ratings of the four grammatical conditions (OCP-violating geminates: /baddo/ → [batto], non-OCP-violating geminates: /heddo/ → [hetto], OCP-violating singletons: /baado/ → [baato], and non-OCP-violating singletons: /haado/ → [haato]). Here and throughout, the error bars represent 95% confidence intervals based on the variability across all relevant items and speakers and a *t*-distribution.

Japanese speakers find the devoicing of OCP-violating geminates most natural (average = 3.84), more natural than devoicing of non-OCP-violating geminates (average = 3.24). They also rate the devoicing of non-OCP-violating geminates more natural than that of OCP-violating singletons (the average = 2.39). The comparisons among the first three cases replicate the results of Kawahara (2011a). Moreover, the speakers judged devoicing more natural when singletons violate the OCP (averages = 2.39 vs. 2.08).

Table 2 shows the results of the general linear mixed model. The effects of the OCP and geminacy are both significant, each contributing to the naturalness judgments. Their interaction term is significant, which shows that the OCP affects ratings significantly more in the geminate condition than in the singleton condition (the difference in the geminate conditions = 0.60 (3.84−3.24) vs. the difference in the singleton condition = 0.31 (2.39−2.08)). Affricacy and lexical frequency have an effect on overall naturalness ratings. The possibility of merger does not have a significant impact in this general model.

Table 3 illustrates the linear mixed model analysis only on OCP-violating geminates. In this analysis, multiplicity of trigger and lexical frequency have a significant effect. In the following sections, we look at the specific hypotheses that were reviewed in Sect. 1.2.

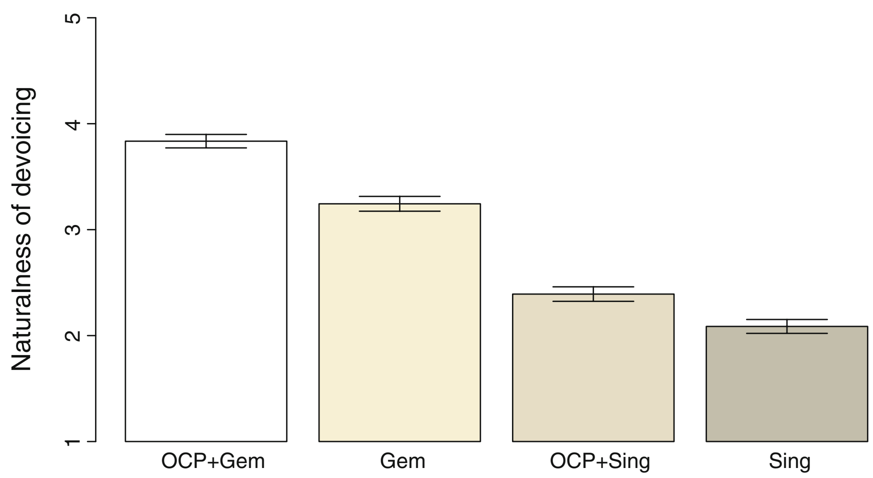


Fig. 1 The average naturalness ratings of OCP-violating geminates, non-OCP-violating geminates, OCP-violating singletons, and non-OCP-violating singletons. The *error bars* represent 95% confidence intervals

Table 2 The results of the genearl linear mixed model

	Coefficients	*t*-value	Significance
OCP	0.324596	7.358	<0.001
Gem	1.191316	27.492	<0.001
OCP:Gem	0.254924	4.240	<0.001
Affricacy	−0.094928	−2.398	<0.05
Merger	−0.022312	−0.558	n.s.
Frequency	0.026240	3.659	<0.001

Table 3 The results of the linear mixed model on OCP-violating geminates

	Coefficients	*t*-value	Significance
Merger	−0.116150	−1.314	n.s.
Multiplicity of trigger	0.233651	2.312	<0.05
Number of intervening syllables	−0.001614	−0.020	n.s.
Frequency	0.122237	8.615	<0.001

3.2 Affricates

Figure 2 compares naturalness ratings between stops and affricates. The difference between stops and affricates is small (stops 2.91 vs. affricates 2.79) but significant overall ($t = -2.398, p < 0.05$). Within OCP-violating geminates, speakers rate the devoicing of affricates slightly more natural than the devoicing of stops. However, the difference is very small (stop 3.83 vs. affricate 3.86) and non-significant according to a post-hoc test comparing the rating scores of OCP-violating geminate stops and those of OCP-violating geminate affricates ($t(1362) = 0.387, \text{n.s.}$). This result is consistent with Nishimura's (2006) treatments of voiced affricates in that Japanese speakers treat devoicing of affricates at least as natural as that of stops (cf. Haraguchi 2006).

3.3 An anti-merger effect

Figure 3 compares the naturalness ratings of items that result in merger and those that do not. Speakers in general disfavor devoicing that would result in merger (averages: 2.79 vs. 2.91), but the overall effect does not reach significance ($t = -0.558$, n.s.). A liner mixed model on OCP-violating geminates does not reveal a significant different either (averages: 3.65 vs. 3.87: $t = -1.314$, n.s.). However, given the non-overlapping error bars in Fig. 3, a post-hoc test was run to compare OCP-violating geminates that would result in merger and those that would not. This analysis revealed a significant difference ($t(1362) = -2.286, p < 0.05$).

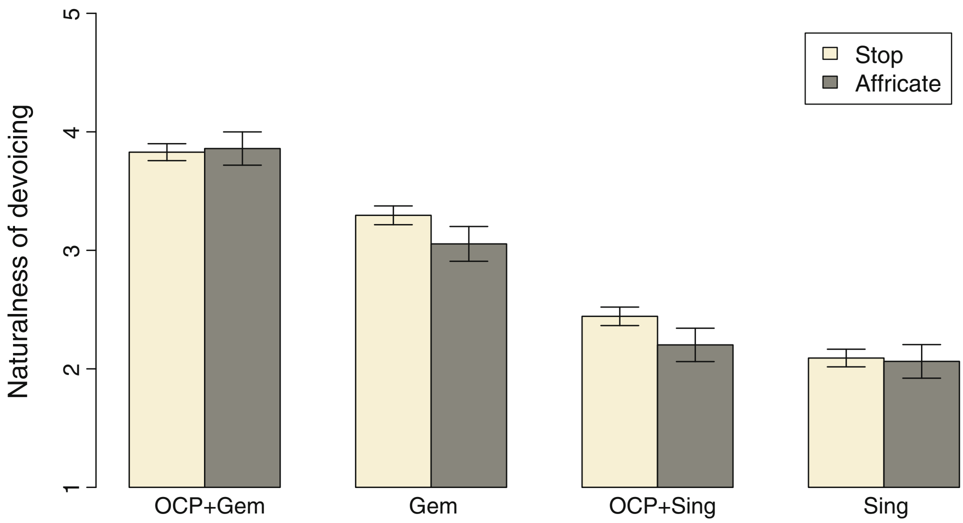


Fig. 2 A comparison between affricates and stops in terms of naturalness ratings

At least within OCP-violating geminates, therefore, speakers find devoicing that results in merger less natural than those that do not.

3.4 Locality

Figure 4 illustrates, for OCP-violating singletons and OCP-violating geminates, the effect of intervening syllables on the naturalness ratings of devoicing; each bar represents cases with a different number of intervening syllables.

For OCP-violating geminates, the number of intervening syllables does not have any effects (3.83 vs. 3.84: $t = -0.020$, n.s.). For OCP-violating singletons, surprisingly, the more intervening syllables exist, the more natural the devoicing is rated. Post-hoc analyses show that devoicing of singletons is rated more natural

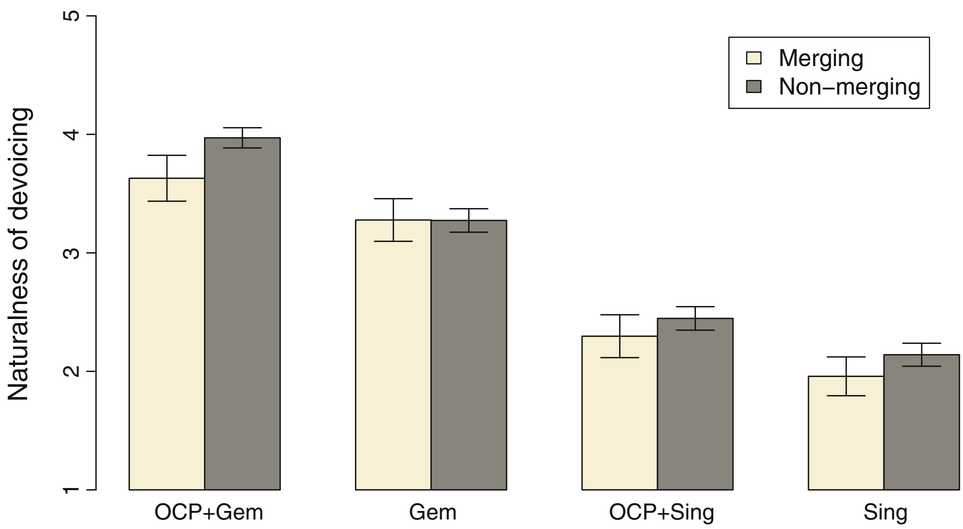


Fig. 3 A comparison between items that would result in merger and those that would not

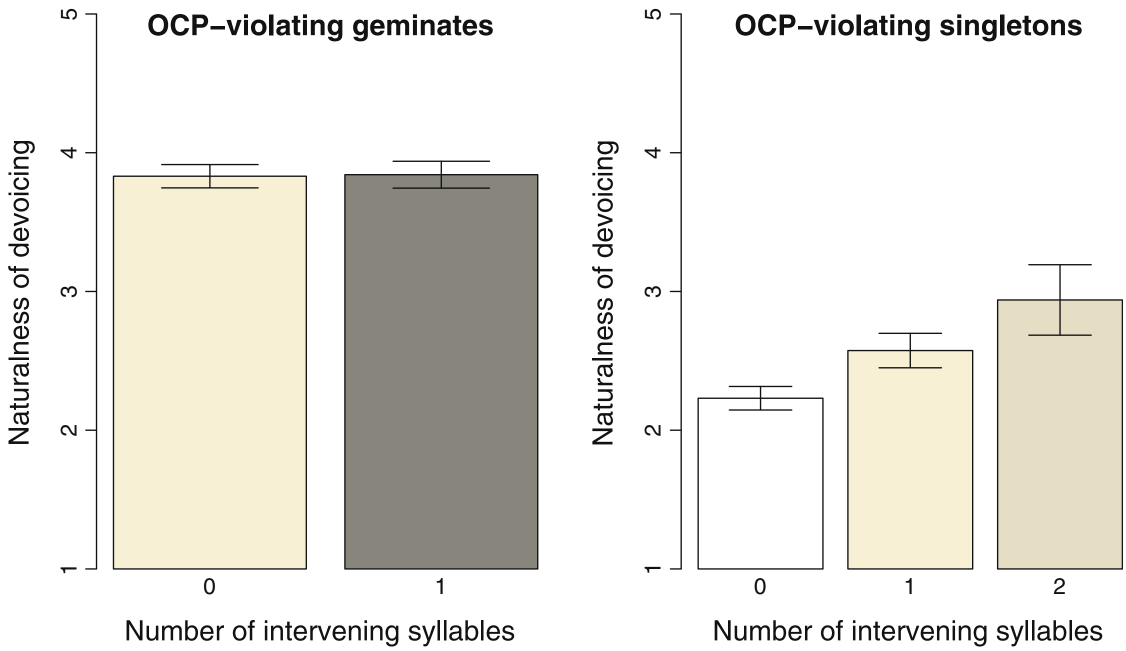


Fig. 4 The average naturalness ratings according to the numbers of intervening syllables. *Left figure* OCP-violating geminates; *right figure* OCP-violating singletons

when there is one intervening syllable than when there is none (2.57 vs. 2.23: $t(1266) = -4.575, p < 0.001$); the comparison between 1 syllable and 2 syllables is also significant (2.57 vs. 2.94: $t(535) = -2.486, p < 0.05$).

3.5 Words with multiple triggers

Figure 5 illustrates, for OCP-violating geminates, the effects of the numbers of trigger on naturalness ratings (in the current stimuli, there are no items for the OCP-violating singleton condition in which there are multiple triggers). Speakers rate devoicing more natural when there are multiple triggers than when there is only one trigger (4.05 vs. 3.80: $t = 2.312, p < 0.05$).

3.6 Place effects

Figure 6 illustrates the effect of place on naturalness ratings for stops. The Japanese speakers rate devoicing of coronal stops more natural than devoicing of dorsal stops, and rate devoicing of labials lower than that of dorsals (labial 2.38, coronal 3.01, dorsal 2.73). Place shows a significant impact on naturalness rating overall ($t = 7.187, p < 0.001$). Post-hoc analyses show that within OCP-violating geminates, speakers rate devoicing of [dd] higher than that of [gg] (3.91 vs. 3.67: $t(1022) = 2.462, p < 0.05$). The difference between [bb] and [gg] is also significant (3.16 vs. 3.67: $t(288) = 2.764$, $p < 0.01$); however, there is only one item containing [bb] in each condition, and therefore it is difficult to make any conclusive generalizations about the labial vs. dorsal comparison.

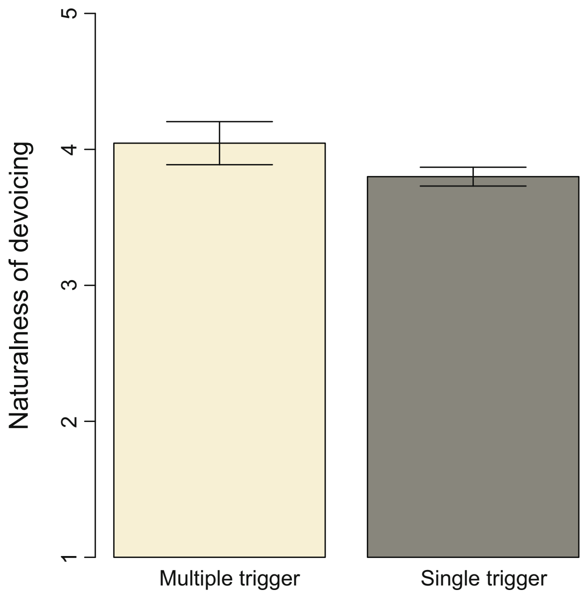


Fig. 5 The effect of the numbers of triggers on naturalness ratings (OCP-violating geminates only)

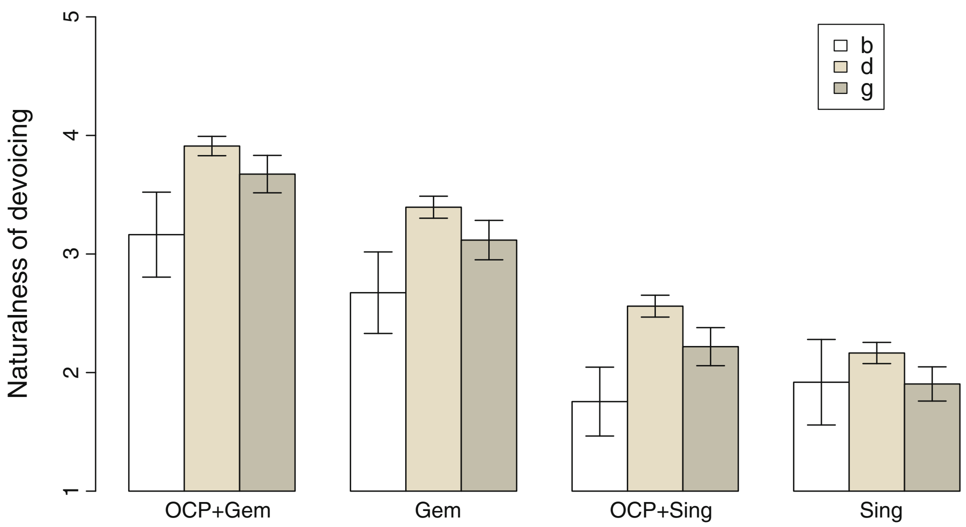


Fig. 6 The average naturalness ratings by place of articulation

3.7 Lexical frequency effects

Finally, the scatterplots in Fig. 7 illustrate the effect of lexical frequencies on naturalness ratings in the four grammatical conditions. Each scatterplot plots, for each lexical item, its average naturalness rating on the *y*-axis against its natural-log frequency on the *x*-axis (the *y*-axis scales are different in different scatterplots). In OCP-violating geminates, naturalness ratings positively correlate with lexical frequencies ($\rho = 0.59, p < 0.001$): Japanese speakers rate devoicing more natural when the word in question is more frequent.[6] On the other hand, lexical fre-

[6] See Coetzee and Kawahara (2010) for an attempt to model this frequency effect.

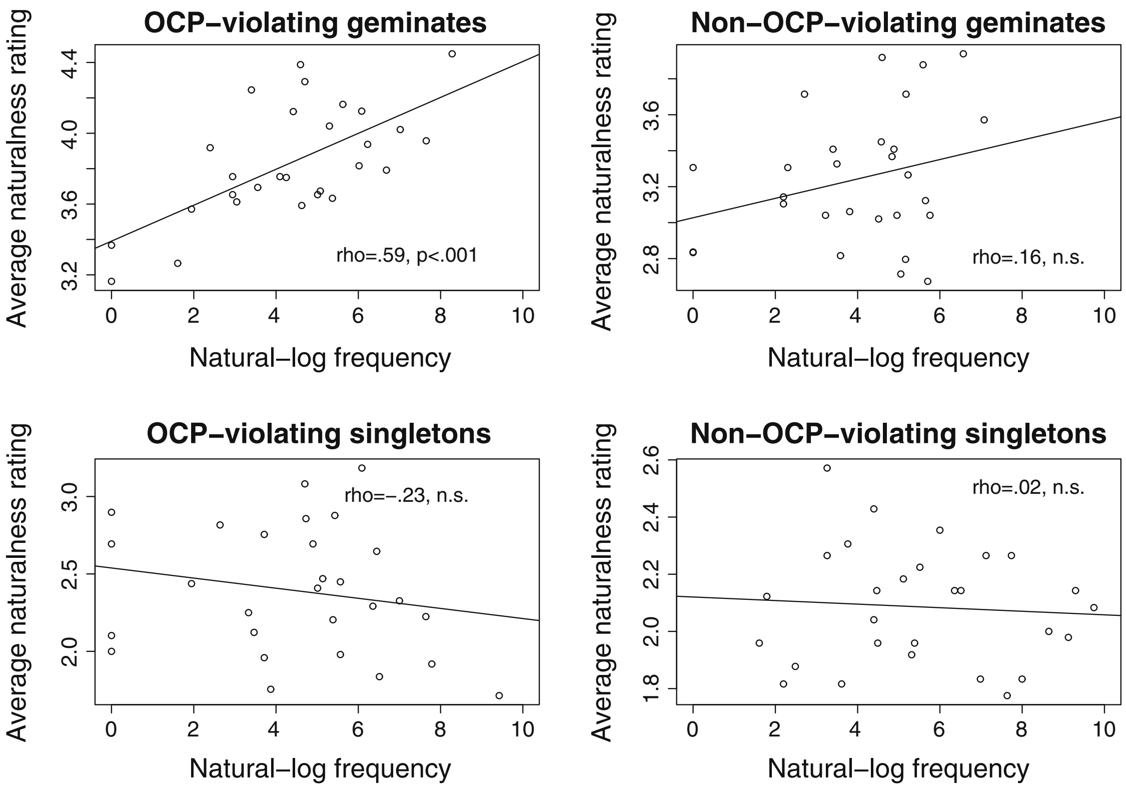


Fig. 7 The effects of lexical frequency on naturalness ratings in the four grammatical conditions: OCP-violating geminates, non-OCP-violating geminates, OCP-violating singletons, non-OCP-violating geminates. Each *dot* represents one lexical item. The *y*-axis scales are different in different sub-figures

quencies do not show any significant effects on naturalness ratings in the other three conditions (non-OCP-violating geminates: $\rho = 0.16$, n.s.; OCP-violating singletons: $\rho = -0.23$. n.s; non-the OCP violating singletons: $\rho = 0.02$, n.s.).

4 Discussion

4.1 The OCP and geminacy

The current rating study first of all has shown that the OCP and geminacy each contribute to the naturalness ratings of devoicing for Japanese speakers. Moreover, the OCP and geminacy interact to a statistically significant degree, indicating that the effect of the OCP is stronger on geminates than on singletons, i.e., the effects of the OCP and geminacy are not additive. This significant interaction is perhaps what underlines the intuition of Nishimura (2003) and Kawahara (2006) that devoicing is only "possible" in OCP-violating geminates (though see Sect. 4.2).

This significant interaction can present a challenge to an analysis of the patterns in (1)–(3) using weighted constraints (Pater 2009). In this analysis, two simple markedness constraints—the OCP and a constraint against voiced geminates—have lower weights than a faithfulness constraint that prevents devoicing. However, the weights of these markedness constraints add up to coerce devoicing of OCP-violating geminates. The analysis derives devoicing of OCP-violating geminates from addition of violations of two lower-weighted constraints. However, the current rating results show that the effects of the OCP and geminacy are not additive.

4.2 Gradiency: beyond a grammatical/ungrammatical dichotomy

The rating patterns of the four grammatical conditions show that grammatical ratings follow a gradient pattern, as Kawahara (2011a) contends. Nishimura (2003) and the following studies have assumed that only devoicing of OCP-violating geminates is grammatical, while context-free devoicing of geminates and OCP-induced devoicing of singletons are ungrammatical, as in (2)–(3). In other words, a grammatical/ungrammatical dichotomy was assumed to exist.

However, the current results show that there is a distinction among "ungrammatical processes:" (i) speakers find context-free devoicing of geminates better than OCP-induced devoicing of singletons, and (ii) speakers find devoicing of singletons more natural when they violate the OCP. Given this result, it is hard to draw a line between "grammatical" devoicing and "ungrammatical" devoicing. These results therefore demonstrate that judgments do not operate as a grammatical/ungrammatical diachotomy, but instead follow a more fine-grained gradient pattern. (For relevant discussion, see among others Albright 2009; Berent et al. 2007; Coetzee 2008, 2009a; Goldrick to appear; Hayes 2000, 2009; Pertz and Bever 1975; Pierrehumbert 2001; see Adli 2010; Chomsky 1965; Myers 2009; Schütze 1996; Sorace and Keller 2005 for a similar observation in syntactic judgments.)[7]

One may argue that the current experiment resulted in a gradient rating pattern because the scale used in the current experiment was gradient. To address this question, Kawahara (2011b) ran a series of follow-up experiments on the same devoicing patterns in Japanese. One experimental variable tested in the study was a mode of judgment, comparing a judgment pattern using a scale, as in the current experiment, and a judgment pattern using a binary forced-choice paradigm. In the latter mode, the participants were asked to judge whether devoicing was possible or impossible in various phonological environments. The results show that the four-way grammatical distinction we observed in the current experiment still holds even with a binary forced-choice task. The results in Kawahara (2011b) thus demonstrate that the four-way grammatical distinction found in the current experiment is not an artifact of an experimental task.

To the extent that acceptability patterns show a four-way distinction, a further question arises: why Japanese speakers make the two distinctions other than the one identified by Nishimura (2003) ((i) context-free devoicing of geminates is better than OCP-induced devoicing of singletons, and (ii) devoicing of singletons is more natural when they violate the OCP). The second case makes sense because Japanese native phonology does avoid words containing OCP-violating singletons (Itô and Mester 1986, 1995, 1999), and this avoidance may have led to the higher naturalness of devoicing of OCP-violating singletons even in loanwords.

[7] An interesting question that arises given this result is how this four-way grammatical judgment distinction is reflected in actual production patterns. A systematic production experiment is called for to address this question.

The first distinction is less straightforward to explain. It may be that, since Japanese voiced geminates are semi-devoiced phonetically and since they are perceptually more similar to voiceless counterparts than singleton voiced consonants are (Kawahara 2005, 2006), Japanese speakers may prefer devoicing of geminates in general to devoicing of singletons. An alternative would be to say that voiced geminates are “more marked” than OCP-violating singletons, although there is no evidence for this postulation in Japanese phonology: They are both prohibited in native phonology, and they are both permitted in loanword phonology (Itô and Mester 1995, 1999).

4.3 A summary of other effects

In addition to revealing the nature of the contributions of the OCP and geminacy on naturalness ratings, the current experiment reveals several aspects of devoicing of OCP-violating geminates: (i) speakers rate the devoicing of affricates as natural as that of stops, (ii) speakers rate devoicing more natural when it does not result in merger with other lexical items, (iii) speakers rate devoicing of voiced geminates more natural when there is more than one trigger, (iv) speakers rate devoicing [dd] more natural than [gg], and (v) speakers rate devoicing of more frequent lexical items more natural.[8]

4.4 Locality of OCP

The current experiment does not find an effect of intervening syllables on the naturalness of devoicing in OCP-violating geminates, although cross-linguistically dissimilatory forces are known to be stronger between two closer elements (Frisch et al. 2004; Itô and Mester 2003; Suzuki 1998; Tanaka 2007). The results may imply, at least in the case of devoicing of OCP-violating geminates in Japanese, that the domain of the OCP is unbounded—as long as two voiced obstruents are within the same domain, an equal amount of force applies (though see below for a potential confound). This conclusion may not come as a surprise, given that blockage of Rendaku due to the OCP occurs across intervening syllables (e.g., /nise+tokage/ → [nisetokage] ‘fake lizard’, Itô and Mester 1986, 2003). This non-local nature of the Japanese OCP constraint, however, is not what is predicted from the results of Vance’s (1980) nonce-word study in which blockage of Rendaku occurred more

[8] Kazu Kurisu and Gunnar Hanson (p.c.) independently asked if accents can affect the naturalness of devoicing. If accented syllables are phonologically strong (Beckman 1997) (although such evidence is scarce in Japanese accented syllables), then speakers may disfavor devoicing in accented syllables. However, since heavy syllables attract accents in Japanese loanwords (Kubozono 2006), in the OCP-violating geminate condition, voiced geminates are generally in the syllable coda of accented syllables; e.g. [bággu] ‘bag’ and [dorággu] ‘drag’. Three exceptions are [déibiddo] ‘David’, [badʤi] ‘badge’, and [bíreddʒi] ‘village’. The averages confirm the expectation that devoicing is less natural in accented syllables: 3.83 (accented syllable) vs. 3.91 (non-accented syllable). However, a more systematic study with auditory stimuli is necessary to address this question more carefully, because some loanwords can be pronounced without an accent when they become familiar, especially in a particular group of people (Inoue, 1998) (e.g., [debaggu] ‘debug’ can be pronounced without an accent when used by computer experts), and there is no way of knowing if/when this deaccenting occurred for particular items for particular speakers.

frequently when the blocker consonant is closer to the consonants that would undergo Rendaku.

Recall next that speakers rate devoicing of singletons more natural when there are intervening syllables. Andries Coetzee (p.c.) pointed out that there could be an assimilative effect in voicing between consonants across vowels (see Hansson 2004; Rose and Walker 2004 for some cases of long-distance voicing assimilation). If that were the case, then it would disfavor devoicing when two voiced consonants appear close to each other. This is an interesting hypothesis that is worth testing in future research, although no evidence is known for this sort of long-distance assimilative effect in Japanese phonology.

Another possibility, raised by Kyoko Yamaguchi (p.c.), is to resort to an idea of positional faithfulness (Beckman 1997). Words with OCP-violating singletons with intervening syllables are, by virtue of having intervening syllables, long (longer than two syllables) (e.g. [baraado] 'ballad', [barikeedo] 'barricade', and [biriyaado] 'billiard'). It may be that in general speakers may allow neutralizing a contrast in late syllables in long words, because preceding syllables provide enough cues for lexical access (as predicted by, for example, a cohort model of lexical access: Marslen-Wilson 1975; Marslen-Wilson and Welsh 1978). Put in phonological terms, later syllables in long words are positionally weak, so that they are more likely to undergo phonological alternations (see Becker and Nevins 2010 for a similar observation about the correlation between word-length and neutralizability in Turkish, and its positional-faithfulness based explanation; see Kawahara and Shinohara 2011 for evidence for the positional strength of initial syllables in Japanese).

4.5 The effect of multiple triggers

The initial hypothesis about words containing multiple triggers was that their devoicing would be rated less natural than that of words that contain only one trigger, because words with multiple triggers would still incur violation(s) of the OCP after they undergo devoicing (Kawahara 2008; Tesar 2007). However, the actual results go in the opposite direction. Devoicing in words with multiple triggers is considered to be more natural.

The higher rating of words with multiple triggers may stem from the fact that there are multiple forces that coerce dissimilation. Suppose the OCP constraint assesses its violation by pairs of two voiced obstruents (Kawahara 2008; Tesar 2007). Then words with two triggers have three violations (e.g., [deibiddo] has three violations: [d]–[b], [d]–[dd], [b]–[dd]), and devoicing the geminates reduces the violations by two (e.g., [deibitto] has one violation: [d]–[b]). On the other hand, words with one trigger have one violation (e.g., [baddo] has one violation: [b]–[dd]) and devoicing would resolve that violation. The former case therefore ends up resolving more of the OCP violations, and hence it could be judged more natural.

This scenario predicts that there can be a case in which dissimilation is coerced only when there are two triggers—Itô and Mester (2003) deny that such a pattern is possible (p. 265, endnote 23) (see also Tanaka 2007; Tesar 2007). However, there are patterns of vowel harmony that are triggered only when there is more than one trigger: e.g., Classical Manchu (Dresher 2010; Walker 2001; Zhang 1996). In

Hungarian, moreover, harmony strength is stronger when there are multiple triggers (Hayes and Londe 2006; Hayes et al. 2009). The result may show that the same tendency holds in dissimilation, although this hypothesis needs to be examined more extensively in light of cross-linguistic patterns of dissimilation.

4.6 Place effects

This study replicated Kawahara (2011a) in that Japanese speakers find the devoicing of [dd] more natural than that of [gg]. This pattern, however, goes counter to the direction that is expected from the aerodynamic challenge imposed on voiced stops (see Sect. 1.2). One possibility is the difference actually comes from the following vowel: In Japanese loanwords, word-final [d] in source words is followed by [o] and [g] is followed by [u] (Katayama 1998). Moreover, word-final high vowels preceded by voiceless consonants are devoiced (Tsuchida 1997). Therefore, devoicing of [gg] word-finally would feed the devoicing of the following [u], whereas devoicing of [d] does not, because the vowel is non-high. As a consequence, Japanese speakers may disfavor devoicing of [gg] compared to devoicing [dd] because devoicing of [gg] would cause a further phonological change. A problem of this explanation is that the previous study by Kawahara (2011a) did not find the effect with devoicability of the following vowel in naturalness rating. Another problem with this explanation, which was raised by an anonymous reviewer, is the fact that the devoicing of [bb] was rated lower than that of [gg], despite the fact that the following vowel is not devoicable in the geminate conditions ([gebberusu] 'Göbbels' and [habburu] 'Hubble'). (Recall however that we should take this pattern with caution because there is only one lexical item with a labial stop per each condition.)

An alternative explanation is that words containing [dd] are more frequent than those that contain [gg] (average log frequencies: 4.72 vs 3.44). However, this frequency-based explanation does not hold for OCP-violating geminates ([dd]: 3.99 vs. [gg]: 4.79). Another potential factor is that [dd] is more frequent as a sound than [gg] in Japanese loanword phonology in general (i.e., it has a high phoneme frequency). Kawahara (2005) finds, based on Amano and Kondo (2000), that [dd] appears 22,896 times (10.04 in log), whereas [gg] appears 1,201 times (7.09 in log). This higher phoneme frequency of [dd] may be a reason for why Japanese speakers judged the devoicing of [dd] to be more natural. To the extent that this explanation is on the right track, the frequencies of sounds in general—rather than the frequencies of the words *per se*—may affect the naturalness ratings, although this hypothesis needs to be tested in future experimentation.

4.7 Lexical frequency effects

Finally, the current study has found that lexical frequencies affect grammatical judgments only in a limited grammatical condition, replicating Kawahara (2011a) with a much larger pool of stimuli. The results show that it is not the case that lexical frequencies govern judgment patterns entirely; only in limited grammatical environments can lexical frequencies exert their effects.

Modeling the limited effect of frequency provides a challenging and interesting task for future research. Jason Shaw (p.c.) raised the following possibility: Since the type frequency of words containing OCP-violating geminates is smaller than the type frequencies of other kinds of words, the devoicing pattern is less regularized for OCP-violating geminates. In other words, in the three environments other than words with OCP-violating consonants, there are sufficient number of words so that their phonological properties are stabilized, resulting in less phonological variation. On the other hand, since words containing OCP-violating geminates are still small in number (this study contained “only” 28 words with OCP-violating geminates), its phonological property as a group is yet to be stabilized, leaving room for lexical frequencies to affect the phonological variation (see Albright and Hayes 2003; Bybee 2001; Bybee and Pardo 1981; Pierrehumbert 2001 for a relevant discussion). This hypothesis provides an interesting line of approach for modeling the limited influence of lexical frequencies on the naturalness judgement pattern, but verifying the hypothesis is left for future research.

4.8 Overall summary

The current experiment shows that the OCP and geminacy each affect naturalness rating of devoicing in Japanese. It moreover reveals a significant interaction between these two factors, suggesting that the OCP affects rating more in geminate consonants, supporting the original observation of Nishimura (2003). In addition, this study reveals that various factors—the possibility of merger, place of articulation, and the lexical frequencies of the words—affect naturalness ratings of devoicing. The current study shows that devoicing of OCP-violating geminates is not as monolithic as previous phonological studies have assumed, and in this regard, it highlights the importance of an experimental approach to theoretical linguistics.

One final remaining question that arises given the current results is which factors are due to grammar and which factors are due to task effects. It seems safe to conclude that the influence of the OCP, geminacy, affricacy, and place, are due to grammatical factors, because these are notions intrinsic to phonology (i.e., any phonological system must encode these notions in the grammar). Some other cases are not so clear-cut. The anti-merger effect, for example, may be due to the fact that Japanese speakers are consciously avoiding lexical mergers, not necessarily that this effect is encoded in the grammar. The length effect we observed in Sect. 4.4 can also be interpreted as speakers “consciously not caring” about neutralizing a contrast in late syllables in long words, without this effect being grammaticalized. Teasing apart grammatical effects from non-grammatical ones is not always an easy task (e.g., Berent et al. 2007; Featherson 2005; Goldrick to appear; Schütze 1996), and the decision often depends on the entire architectural structure of grammar. Addressing this question is thus an important issue in future linguistic research.

Acknowledgements My thanks first of all go to the participants of this experiment, without whom this paper would not have existed. Particularly, I am grateful to Toshio Matsuura for encouraging his students to participate in the experiment. I am also grateful to Kazuko Shinohara for her help in constructing the stimuli and to Sophia Kao for her help in setting up the online experiment. For their comments on earlier versions of this paper, I would like to thank the following people: Michael Becker, Andries Coetzee, Bruce Hayes, Kazu Kurisu, Shanna Lichtman, Toshio Matsuura, Jason Shaw, Kyoko Yamaguchi, three anonymous reviewers for *JEAL*, and the audiences at the University of British Columbia, especially Bryan Gick and Gunnar Hanson. All remaining errors are mine.

References

Adli, Aria. 2010. Constraint cumulativity and gradience: Wh-scrambling in Persian. *Lingua* 120 (9): 2259–2294.

Albright, Adam. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology* 26 (1): 9–41.

Albright, Adam, and Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90: 119–161.

Alderete, John, and Alexei Kochetov. 2009. Japanese mimetic palatalization revisited: Implications for conflicting directionality. *Phonology* 26 (3): 369–388.

Amano, Shigeaki, and Tadahisa Kondo. 2000. *NTT database series: Lexical properties of Japanese, 2nd release*. Tokyo: Sanseido.

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

Baayen, Harald R. 2009. LanguageR. R package. http://cran.r-project.org/web/packages/languageR/index.html.

Baayen, Harald R., Doug J. Davidson, and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390–412.

Bates, Douglas M. 2005. Fitting linear mixed models in R. *R News* 5: 27–30.

Bates, Douglas, Maechler, Martin, and Bolker Ben. 2011. Linear mixed-effects models using S4 classes. R Package. http://cran.r-project.org/web/packages/lme4/index.html.

Becker, Michael, and Andrew Nevins. 2010. Initial-syllable faithfulness as the best model of wordsize effects in alternations. Paper presented at the 84th meeting of LSA, Baltimore.

Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona vowel harmony. *Phonology* 14 (1): 1–46.

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

Blevins, Juliette. 2005. Understanding antigemination: Natural or unnatural history. In *Linguistic diversity and language theories*, ed. Zygmunt Frajzyngier, David Rood, and Adam Hodges, 203–234. Amsterdam: Benjamins.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45–86.

Bybee, Joan. 1999. Usage-based phonology. In Functionalism and formalism in linguistics, ed. Michael Darnell, Edith Moravscik, Michael Noonan, Frederick Newmeyer, and Kathleen Wheatly, 211–242. Amsterdam: Benjamins.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan, and Elly Pardo. 1981. On lexical and morphological conditioning of rules: A nonce-probe experiment with Spanish verbs. *Linguistics* 19: 937–968.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.

Coetzee, Andries W. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84 (2): 218–257.

Coetzee, Andries W. 2009a. Grammar is both categorical and gradient. In *Phonological Argumentation: Essays on evidence and motivation*, ed. Steve Parker, 9–42. London: Equinox.

Coetzee, Andries W. 2009b. Phonological variation and lexical frequency. In *Proceedings of North East Linguistic Society 38,* ed. Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman, 189–202. Amherst: GLSA Publications.

Coetzee, Andries W., and Shigeto Kawahara. 2010. Frequency and other biases in phonological variation. Ms., Michigan University and Rutgers University (submitted for publication in Natural Language and Linguistic Theory).

Coetzee, Andries W., and Joe Pater. to appear. The place of variation in phonological theory. In *The handbook of phonological theory,* 2nd edn, ed. John A. Goldsmith, Jason Riggle, and Alan Yu. Oxford: Blackwell-Wiley.

Coleman, John, and Janet Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. In *Computational phonology: Third meeting of the ACL special interest group in computational phonology*, 49–56. Somerset: Association for Computational Linguistics.

Crawford, J. Clifford. 2009. Adaptation and transmission in Japanese loanword phonology. Ph.D. dissertation, Cornell University.

Crosswhite, Katherine. 1999. Intra-paradigmatic homophony avoidance in dialects of Slavic. In *UCLA working papers in linguistics: Papers in phonology 2*, ed. Matthew K. Gordon, 48–67.

Dabrowska, Ewa. 2010. Naive vs. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27 (1): 1–23.

Dresher, B. Elan. 2010. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Farris-Trimble, Ashley. 2008. Cumulative faithfulness effects in phonology. Ph.D. dissertation, Indiana University.

Featherson, Sam. 2005. Universals and grammaticality: Wh-constraints in German and English. *Linguistics* 43: 667–711.

Flemming, Edward. 1995. Auditory representations in phonology. Ph.D. dissertation, University of California Los Angeles.

Frisch, Stefan, Nathan Large, and David Pisoni. 2000. Perception of wordlikiness: Effects of segment probability and length of the processing of nonwords. *Journal of Memory and Language* 42: 481–496.

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

Gibson, Edward, and Evelina Fedorenko. 2010. Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* 14 (6): 233–234.

Goldrick, Matthew. to appear. Utilizing psychological realism to advance phonological theory. In *The handbook of phonological theory II*, ed. John A. Goldsmith, Jason Riggle, and Alan Yu. Oxford: Blackwell-Wiley.

Griner, Barry. 2001. Productivity of Japanese verb tense inflection: A case study. MA thesis, University of California, Los Angeles.

Hansson, Gunnar Olafur. 2004. Long-distance voicing agreement: An evolutionary perspective. In *Proceedings of Berkeley Linguistic Society 30*, ed. M. Fleischer Ettlinger and M. Park-Doob, 130–141. Berkeley: Berkeley Linguistic Society.

Haraguchi, Shosuke. 2006. A theory of voicing. ms., Meikai University. http://www.let.leidenuniv.nl/ulcl/faculty/vdweijer/jvoice/haraguchi.pdf.

Hay, Jennifer, Janet Pierrehumbert, and Mary Beckman. 2003. Speech perception, wellformedness, and the statistics of the lexicon. In *Papers in laboratory phonology VI: Phonetic interpretation*, ed. John Local, Richard Ogden, and Rosalind Temple, 58–74. Cambridge: Cambridge University Press.

Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. In *Optimality Theory: Phonology, syntax, and acquisition*, ed. Joost Dekkers, Frank Van der Leeuw, and Jeroen Van de Weijer, 88-120. Oxford: Oxford University Press.

Hayes, Bruce. 2009. Embedding grammar in a quantitative framework: Case studies from phonology and metrics. Handout for a minicourse at Brown University.

Hayes, Bruce, and Zsuzsa Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23: 59–104.

Hayes, Bruce, and Donca Steriade. 2004. Introduction: The phonetic bases of phonological markedness. In *Phonetically based phonology*, ed. B. Hayes, R. Kirchner, and D. Steriade, 1–33. Cambridge: Cambridge University Press.

Hayes, Bruce, and Tanya Stivers. 1995. Postnasal voicing. ms., University of California, Los Angeles. http://www.linguistics.ucla.edu/people/hayes/phonet/ncphonet.pdf.

Hayes, Bruce, Kie Zuraw, Péter Siptár, and Zsuzsa Londe. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85 (4): 822–863.

Ichimura, Larry. 2006. Anti-homophony blocking and its productivity in transparadigmatic relations. Ph.D. dissertation, Boston University.

Inoue, Humio. 1998. *Nihongo wocchingu [Watching Japanese]*. Tokyo: Iwanami.

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

Itô, Junko, and Armin Mester. 1995. Japanese phonology. In *The handbook of phonological theory*, ed. John Goldsmith, 817–838. Oxford: Blackwell.

Itô, Junko, and Armin Mester. 1999. The phonological lexicon. In *The handbook of Japanese linguistics*, ed. Natsuko Tsujimura, 62–100. Oxford: Blackwell.

Itô, Junko, and Armin Mester. 2003. *Japanese morphophonemics*. Cambridge: MIT Press.
Itô, Junko, and Armin Mester. 2004. Morphological contrast and merger: Ranuki in Japanese. *Journal of Japanese Linguistics* 20: 1–19.
Itô, Junko, and Armin Mester. 2008. Lexical classes in phonology. In *The Oxford handbook of Japanese linguistics*, ed. S. Miyagawa and M. Saito, 84–106. Oxford: Oxford University Press.
Jaeger, Florian T. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59: 434–446.
Jaeger, Jeri J. 1978. Speech aerodynamics and phonological universals. In *Proceedings of Berkeley Linguistic Society 4*, ed. Jeri Jaeger, Anthony Woodbury, Farrel Ackerman, Christine Chiavello, Orin Gensler, John Kingston, Eve Sweetner, Henry Chompson, and Kenneth Whistler, 311–325. Berkeley: Berkeley Linguistics Society.
Kaneko, Emiko, and 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*, ed. Shunji Inagaki and Makiko Hirakawa. Tokyo: Kuroshio Publications.
Kaplan, Abby. 2010. How much homophony is normal? Paper presented at CUNY Conference on the Word in Phonology, CUNY.
Katayama, Motoko. 1998. Optimality Theory and Japanese loanword phonology. Ph.D. dissertation, University of California, Santa Cruz.
Kawahara, Shigeto. 2005. Voicing and geminacy in Japanese: An acoustic and perceptual study. In *University of Massachusetts occasional papers in linguistics 31*, ed. Kathryn Flack and Shigeto Kawahara, 87–120. Amherst: GLSA.
Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language* 82 (3): 536–574.
Kawahara, Shigeto. 2008. Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* 18 (4): 317–330.
Kawahara, Shigeto. 2011a. Japanese loanword devoicing revisited: A rating study. *Natural language and Linguistic Theory* 29.
Kawahara, Shigeto. 2011b. Modes of phonological judgments. ms., Rutgers University.
Kawahara, Shigeto, and Kazuko Shinohara. 2011. Phonetic and psycholinguistic prominences in pun formation: Experimental evidence for positional faithfulness. In *Japanese/Korean linguistics 18*, ed. W. McClure and M. den Dikken, 177–188. Stanford: CSLI.
Kubozono, Haruo. 2006. Where does loanword prosody come from?: A case study of Japanese loanword accent. *Lingua* 116: 1140–1170.
Labov, William. 1975. What is a linguistic fact? In Empirical foundations of linguistic theory: The scope of American linguistics, ed. Robert Austerlitz, 77–113. Lisse: The Peter de Ridder Press.
Labov, William. 1996. When intuitions fail. In *Papers from the 32nd regional meeting of Chicago Linguistic Society: Papers from the parasession on theory and data in linguistics*, ed. Lisa McNair, Kora Singer, Lise, Dolbrin, and Michelle Aucon, 77–106. Chicago: Chicago Linguistics Society.
Lubowicz, Anna. 2007. Paradigmatic contrast in Polish. *Journal of Slavic linguistics* 2: 229–262.
Lyman, Benjamin S. 1894. Change from surd to sonant in Japanese compounds. *Oriental Studies of the Oriental Club of Philadelphia*, 1–17.
Maekawa, Kikuo. 2009. Nihongo za-gyo-no choo-on yooshiki-no hen-i-ni tsuite [On the articulatory variability of /z/ in Japanese]. *Dai 23kai Nihon Onsei Gakkai Zenkoku Taikai Yokoushuu*, 169–174.
Maekawa, Kikuo. 2010. Coarticulatory reinterpretation of allophonic variation: Corpus-based analysis of /z/ in spontaneous Japanese. *Journal of Phonetics* 38 (3): 360–374.
Marslen-Wilson, William. 1975. Sentence perception as an interactive parallel process. *Science* 189: 226-228.
Marslen-Wilson, William, and Alan Welsh. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10: 29–63.
McCarthy, John J. 2008. *Doing Optimality Theory*. Oxford: Blackwell-Wiley.
Myers, James. 2009. Syntactic judgment experiments. *Language and Linguistic Compass* 3 (1): 406–423.
Nishimura, Kohei. 2003. Lyman's Law in loanwords. M.A. thesis, Nagoya University.
Nishimura, Kohei. 2006. Lyman's Law in loanwords. *Phonological Studies [Onin Kenkyuu]* 9: 83–90.
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*, ed. J.M. Naderson and Charles Jones, 353–389. New York: Elsevier.
Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In *The production of speech*, ed. Peter MacNeilage, 189–216. New York: Springer.

Ohala, John J. 1986. Consumer's guide to evidence in phonology. *Phonology* 3: 3–26.

Ohala, John J., and C. J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. In *Speech communication papers*, ed. J.J. Wolf and D.H. Klatt, 89–92. New York: Acoustical Society of America.

Padgett, Jaye. 2002. Contrast and postvelar fronting in Russian. *Natural Language and Linguistic Theory* 21: 39–87.

Pater, Joe. 1999. Austronesian nasal substitution and other NC effects. In *The prosody-morphology interface*, ed. Rene Kager, Harry van der Hulst, and Wim Zonneveld, 310–343. Cambridge: Cambridge University Press.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33: 999–1035.

Pertz, D.L., and T.G. Bever. 1975. Sensitivity to phonological universals in children and adolescents. *Language* 51: 149–162.

Pierrehumbert, Janet B. 2001. Stochastic phonology. *GLOT* 5: 1–13.

Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden: Blackwell.

Pycha, Anne, Eurie Shin, and Rhan Shosted. 2006. Directionality of assimilation in consonant clusters: An experimental approach. *UC Berkeley Phonology Lab Annual Report*.

R Development Core Team. 1993–2011. R: *A language and environment for statistical computing*. Vienna, Austria. R Foundation for Statistical Computing. Software. http://www.R-project.org.

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

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.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80 (3): 475–532.

Rubin, David. 1976. Frequency of occurrence as a psychophysical continuum: Weber's fraction, Ekman's fraction, range effects, and the phi-gamma hypothesis. *Perception & Psychophysics* 20: 327–330.

Schütze, Carlson. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Shirai, Setsuko. 2002. Gemination in loans from English to Japanese. M.A. thesis, University of Washington.

Smith, Roger, and Theodore Dixon. 1971. Frequency and judged familiarity of meaningful words. *Journal of Experimental Psychology* 88: 279–281.

Smolensky, Paul. 1995. On the internal structure of the constraint component CON of UG. Paper presented at the University of California, Los Angeles.

Sorace, Antonella, and Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115 (11): 1497–1524.

Steriade, Donca. 2001/2008. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The nature of the word*, ed. Kristin Hanson and Sharon Inkelas, 151–179. Cambridge: MIT Press (Originally circulated in 2001).

Steriade, Donca. 2004. Projecting non-lexical phonology from phonetic knowledge. Paper presented at Symposium on phonological theory: Representations and architectures, the City University of New York.

Suzuki, Keiichiro. 1998. A typological investigation of dissimilation. Ph.D. dissertation, University of Arizona.

Tanaka, Shin-Ichi. 2007. On the nature of typology of dissimilation. *English Linguistics* 24 (1): 279–313.

Tanaka, Shin-Ichi. 2010. The eurhythmics of segmental melody. *Journal of the Phonetic Society of Japan* 13 (3): 44–52.

Tateishi, Koichi. 2002. Bunpoo no ichibutoshite no goisoo no zehi [Lexical strata as a part of grammar]. *Journal of the Phonetic Society of Japan* 6 (1): 34–43.

Tesar, Bruce. 2007. A comparison of lexicographic and linear numeric optimization using violation difference ratios. ms., Rutgers University.

Tsuchida, Ayako. 1997. Phonetics and phonology of Japanese vowel devoicing. Ph.D. dissertation, Cornell University.

Urbanczyk, Suzanne. 2005. Enhancing contrast in reduplication. In *Studies on reduplication*, ed. B. Hurch, 431–454. Mouton: de Gruyter.

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

Vance, Timothy J. 1987. *An introduction to Japanese phonology*. New York: SUNY Press.
Walker, Rachel. 2001. Round licensing, harmony, and bisyllabic triggers in Altaic. Natural Language and Linguistic Theory 19: 827–878.
Westbury, John R. 1979. Aspects of the temporal control of voicing in consonant clusters in English. Ph.D. dissertation, University of Texas, Austin.
Zhang, Xi. 1996. Vowel systems of the Manchu-Tungus languages of China. Ph.D. dissertation, University of Toronto.

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

DE GRUYTER MOUTON
DOI 10.1515/ling-2013-0050 — Linguistics 2013; 51(6): 1271–1299

Shigeto Kawahara

Testing Japanese loanword devoicing: Addressing task effects

Abstract: In the loanword phonology of Japanese, voiced obstruent geminates ([bb, dd, gg]) have been claimed to devoice when they co-occur with another voiced obstruent within the same morpheme (e.g., /beddo/ → [betto] 'bed'). This devoicing pattern has contributed much to address a number of theoretical issues in the recent phonological literature. However, the relevant data have been primarily based on intuition-based data provided by Nishimura (2003) and Kawahara (2006). Kawahara (2011a, 2011b) addressed this issue by conducting rating studies using naive native speakers of Japanese. The results generally supported the intuition-based data by Nishimura (2003) and Kawahara (2006). However, the rating studies also revealed several aspects of the devoicing pattern that go beyond the intuition-based data as well.

The current study further investigates the devoicing pattern by varying several task variables. In particular, this paper builds on Kawahara (2011a, 2011b) by adding (i) nonce word stimuli, (ii) a binary yes/no experiment, and (iii) auditory stimuli. The results show that (i) nonce words and real words behave similarly, but nonce words nevertheless show less variability across different grammatical conditions than real words; (ii) the binary yes/no experiment shows results similar to those of the scale-based experiment; and (iii) while auditory stimuli yield results comparable with those of orthographic stimuli, they also show an exaggerated effect of a phonetic implementation pattern. Overall, this paper uses Japanese as a case study, and finds some task effects in phonological judgment experiments. It is hoped that this paper stimulates further experimental research on phonological judgments of other phenomena in Japanese as well as in other languages.

Keywords: Japanese, geminates, devoicing, experimental phonology, task effects

Shigeto Kawahara: Institute of Cultural and Linguistic Studies, Keio University, 2–15–45 Mita, Minato-ku, Tokyo, JAPAN 108-8345. E-mail: kawahara@icl.keio.ac.jp

1 Introduction

1.1 The phenomenon

This paper is about devoicing of obstruents in the loanword phonology of Japanese. It has been known that voiced obstruent geminates ([bb, dd, gg]) in Japanese loanwords can be devoiced (Itô and Mester 1995, 1999; Quakenbusch 1989; Vance 1987), but exactly when such devoicing occurs remained unclear. For example, Itô and Mester (1999) argued that some items can undergo devoicing while other items cannot, and considered the first type of words as “assimilated foreign items” and the second, non-devoicing type as “unassimilated foreign items”. Instead of relying on a (more or less) arbitrary etymological distinction, Nishimura (2003) proposes a *phonological* characterization of this distinction, claiming that voiced obstruent geminates optionally devoice when they co-occur with another voiced obstruent within the same stem, as exemplified by the data in (1). He further claims that this devoicing is due to a restriction against having two voiced obstruents within the same stem. In Japanese phonology, this restriction has long been known as Lyman’s Law (Kawahara 2012b; Lyman 1894; Vance 2007), and has been formalized as OCP[voice] (Obligatory Contour Principle: Leben 1973; henceforth simply the OCP) (Itô and Mester 1986, 1998, 2003). In other words, devoicing is possible in (1) whereas it is impossible in non-OCP-violating voiced geminates, as shown in (2). Moreover, Nishimura (2003) argues that devoicing is also impossible in OCP-violating singletons, as in (3).

(1) Voiced obstruent geminates optionally devoice if they co-occur with another voiced obstruent; i.e., when they violate OCP[voice].

be**dd**o	→	be**tt**o	‘bed’
ba**gg**u	→	ba**kk**u	‘bag’
bi**gg**u	→	bi**kk**u	‘big’

(2) Voiced obstruent geminates do not devoice if they do not violate OCP[voice].

suno**bb**u	→	suno**bb**u	*suno**pp**u	‘snob’
he**dd**o	→	he**dd**o	*he**tt**o	‘head’
re**gg**u	→	re**gg**u	*re**kk**u	‘leg’

(3) Voiced singletons do not devoice even when they violate OCP[voice].

da**b**u	→	da**b**u	*da**p**u	‘Dove’
do**g**uma	→	do**g**uma	*do**k**uma	‘dogma’
da**g**u	→	da**g**u	*da**k**u	‘Doug’

The patterns in (1)–(3) have attracted much attention in the recent phonological literature. It is beyond the scope of this paper to settle these debates; however, to briefly summarize, the devoicing pattern triggered three major theoretical debates: (i) how to explain the difference between singletons (= the data in (3)) and geminates (= the data in (1)) (Kawahara 2006, 2008; Rice 2006; Steriade 2004); (ii) how to capture the cumulative markedness requirement of devoicing in (1) (Farris-Trimble 2008; Nishimura 2003; Pater 2009, forthcoming; Tesar 2007); and (iii) how the spontaneous emergence of loanword devoicing in (1) bears on the theory of lexical stratification – a theory of how loanword phonology is related to native phonology (Crawford 2009; Itô and Mester 2003, 2008; Tateishi 2002). See Kawahara (2011a) and Kawahara (2012a) for recent summaries (the former in English and the latter in Japanese).

In short, the Japanese loanword devoicing pattern has contributed much to several theoretical debates in recent years. However, Kawahara (2011b) raises one issue: the Japanese loanword devoicing data are primarily based on the intuitions of two linguists, namely, Nishimura (2003) and Kawahara (2006); i.e., the grammaticality judgments in (1)–(3) primarily come from the authors themselves.[1] Many studies have raised concerns about research exclusively relying on authors' own introspections (e.g., Dąbrowska 2010; Gibson and Fedorenko 2010; Griner 2001; Labov 1996; Myers 2009; Ohala 1986; Schütze 1996). To address this problem, Kawahara (2011b) conducted a rating experiment with 38 native Japanese speakers who did not know about the devoicing pattern. The experiment indeed showed that Japanese speakers generally judge devoicing of OCP-violating geminates as more natural than devoicing of non-OCP-violating geminates or devoicing of OCP-violating singletons. In this regard, Kawahara (2011b) succeeded in supporting the empirical basis of the claims made about the patterns in (1)–(3). Kawahara (2011a) reports a follow-up experiment using a larger set of stimuli with 49 naive native speakers, which again supported the idea that devoicing of OCP-violating geminates is the most natural environment for native speakers of Japanese.

1.2 The current study

There are some remaining questions, however. First, both Kawahara (2011a) and Kawahara (2011b) used only real words. In the case of Japanese loanword

1 See Nishimura (2003), Kawahara and Sano (2013) and Sano and Kawahara (forthcoming) for some evidence based on corpus data.

devoicing, it is of some interest to investigate whether the results obtained for real words generalize to nonce words. An often-used test on phonological productivity is a wug-test (Berko 1958), in which participants are asked to inflect nonce words. Some previous wug-tests have failed to replicate phonological patterns that apply to real words, in which case it is often concluded that the alleged phonological patterns are not productive; i.e., they are lexicalized (Griner 2001; Ohala 1974; Sanders 2003). (See also Shademan (2007) for some related discussion.) If the phonological pattern under discussion is not productive with nonce words, the pattern should probably not be used for phonological argumentation.

In fact, there is an example from Japanese phonology whose productivity has been questioned by way of experiments using nonce words. Several phonological changes occurring in Japanese verbal paradigms (Davis and Tsujimura 1991; Tsujimura 1996) were not replicated in nonce word experimentation (Batchelder 1999; Griner 2001; Vance 1987). In short, there is no guarantee that we can generalize the patterns of real words to nonce words, and it is vital to test the productivity of the phenomenon under question using nonce words. Experiment III in Kawahara (2012b) addressed this question, although in that paper, the comparison between real words and nonce words was not the main focus. The Section 2 of this paper therefore reports that experiment in more detail to address this issue of whether the previous results can be replicated with nonce words.

The first aim of this paper is to therefore re-examine Experiment III of Kawahara (2012) to mainly address the question of whether the results obtained in Kawahara (2011a) and Kawahara (2011b) can be replicated with nonce words. This report also allows us to compare the results of that experiment with two other experiments reported in the paper.

The second aim of this paper is to test the gradiency of judgment patterns found in the previous experiments. Kawahara (2011a, 2011b) found that Japanese speakers distinguish the naturalness of two processes that were both judged to be "ungrammatical" by Nishimura (2003) and Kawahara (2006), with the devoicing of non-OCP-violating geminates (= the examples in (2)) rated as more natural than the devoicing of OCP-violating singletons (= (3)). One may wonder whether this gradient effect was due to a task effect; the reason being that Kawahara (2011a, 2011b) uses a gradient scale. Testing this issue is in part motivated by the debate concerning the gradient nature of phonological judgments. It is known that grammatical judgments show distinctions beyond a simple, binary "grammatical" vs. "ungrammatical" dichotomy, especially in experimental settings (see e.g., Albright 2009; Coetzee 2008; Coleman and Pierrehumbert 1997; Daland et al. 2011; Dankovičová et al. 1998; Goldrick 2011; Greenberg and Jenkins 1964; Hayes 2000; Hayes and Wilson 2008; Pertz and Bever 1975; Pierrehumbert 2001; Shademan 2007 for phonological/phonotactic judgments; Chomsky 1965; Myers

2009; Schütze 1996; Sorace and Keller 2005 for syntactic judgments). However, one may contend the idea that we obtain gradient results in experimental settings because these experiments use scales. Therefore, the second aim of this paper is to test whether the gradient results that Kawahara (2011a, 2011b) obtained can be replicated using a binary yes/no task. Some previous studies (Bader and Mäussler 2010; Coleman and Pierrehumbert 1997; Dankovičová et al. 1998; Frisch et al. 2004) raised similar issues and found gradient results using a binary yes/no format. The current study thus builds on them and aims to address the gradient nature of phonological judgments in the case of Japanese loanword devoicing.

Finally, the third issue that this paper addresses is that Kawahara (2011a, 2011b) used visual, orthographic stimuli, although the instructions in these studies encouraged the participants to read the stimuli in their heads and use an auditory impression to make judgments. While many judgment experiments in linguistics are run with orthography, it is worth running the same experiment with auditory stimuli for a few reasons. First, one explanation for why voiced geminates, but not voiced singletons, can devoice is because a phonological voicing contrast is auditorily less perceptible in geminates than in singletons (Kawahara 2006, 2008). An auditory judgment experiment would help to address this specific hypothesis. Second, it would be interesting to investigate whether the results of Kawahara (2011a, 2011b) can be replicated with auditory stimuli, because phonology is concerned with sounds. Replicating the Japanese devoicing pattern with auditory stimuli is therefore the third aim of this paper.

To summarize, there are three issues that this paper aims to address: (i) the judgment patterns on devoicing as revealed by nonce words, (ii) the effect of using a binary yes/no format, and (iii) the effect of using auditory stimuli. This paper reexamines Kawahara (2012b) and reports two additional experiments in order to address these three issues. More generally, by varying experimental variables, the current project aims to further examine the empirical basis of the theoretical debates reviewed in Section 1.1, beyond Kawahara (2011a, 2011b).

Before reporting the actual experiments, a few remarks are in order. First, the experiments reported in this paper are judgment experiments for a phonological process, i.e., devoicing. The task is for native speakers to judge the naturalness or possibility of a phonological pattern, or in other words, a pairing between one form and another form (i.e., in this case, a phonological form and its optional variant). This task therefore differs from phonotactic wellformedness judgment tasks in which speakers judge the wellformedness of surface forms only (e.g., Bailey and Hahn 2001; Coetzee 2008; Coleman and Pierrehumbert 1997; Daland et al. 2011; Dankovičová et al. 1998; Greenberg and Jenkins 1964; Shademan 2007). Second, this paper offers a case study in Japanese of such a phonological judgment study. Although its scope is thus limited, it is hoped that this paper will stimulate further studies on different phonological phenomena in different languages (including Japanese).

2 Kawahara (2012b) Experiment III and beyond

This experiment, briefly reported as Experiment III in Kawahara (2012b), is an orthography-based rating experiment. This section reexamines this experiment in detail, since some details and analyses of this experiment were omitted from Kawahara 2012b (Kawahara 2012b was written after the current paper),[2] and also since the other two experiments in the current paper crucially build on this experiment and I will make many cross-experimental comparisons in what follows. In particular, this section reexamines the experiment from the perspectives that are discussed in the introduction, mainly with the focus of comparing real words and nonce words.

This section thus addresses three issues: (i) to replicate Kawahara (2011a, 2011b); and, more importantly, (ii) to test whether the results obtained with real words in the previous studies generalize to nonce words; and finally, (iii) to compare the patterns of real words and nonce words.

2.1 Method

2.1.1 Stimuli

The stimuli consisted of four grammatical conditions: (i) OCP-violating geminates, (ii) non-OCP-violating geminates, (iii) OCP-violating singletons, and (iv) non-OCP-violating singletons, as summarized in (4), each with a representative example. In this design, two factors – OCP and GEM – were fully crossed. This paper uses CAPITAL LETTERS to represent variable names.

2 In Kawahara (2012b), together with two other experiments on rendaku, this experiment is also reported (in much less detail) to show the activity of the OCP – or Lyman's Law – in both loanwords and nonce words. That paper does not extensively discuss the effect of geminacy, its interaction with OCP, or on the comparison between real words and nonce words. Neither does that paper report the issue of gradiency at all; i.e., the analyses presented in Figures 2 and 3 below are new to this paper. Figure 1 is reproduced with permission from Elsevier.

(4) The four grammatical conditions
   a. OCP-violating geminates (e.g., [ba**gg**u])
   b. non-OCP-violating geminates (e.g., [e**gg**u])
   c. OCP-violating singletons (e.g., [da**g**u])
   d. non-OCP-violating singletons (e.g., [ma**g**u])

The experiment had 9 items per each condition. All the stimulus items were disyllabic, and all the target consonants were word-internal (since all lexical geminates in Japanese appear word-internally: Kawahara (forthcoming)). The stimulus set was constructed in the following way: first, real disyllabic words containing OCP-violating geminates were chosen; this case has the least number of existing items in the Japanese lexicon. This selection process resulted in 9 items. Among those 9 items, 6 items contained [dd] followed by epenthetic [o], and the remaining 3 items contained [gg] followed by epenthetic [u]. No stimuli with [bb] were found; in fact, no disyllabic words with OCP-violating [bb] exist which is not unexpected given that [bb] is very rare in Japanese loanwords (Katayama 1998). Then the items for the other three conditions were selected, consisting of 6 items for [d(d)] and 3 items for [g(g)], as listed in Table 1. Across all conditions, the number of items for each place of articulation was controlled for. Short vowels were used before geminates and singleton [g]. Long vowels and diphthongs had to be used before singleton [d], because disyllabic loanwords with an initial short vowel almost always have a geminate [dd], and not a singleton [d]. This pattern is due to a productive gemination process in loanword adaptation (e.g., [baddo] 'bad'; see e.g., Katayama 1998; Kubozono et al. 2009).[3] All of the stimuli have a pitch accent on the initial syllable, which is phonetically realized as a HL falling F0 contour.

Table 1: The list of the stimuli that are real words.

OCP-Gem	Gem	OCP-Sing	Sing
baddo 'bad'	heddo 'head'	bado 'badminton'	muudo 'mood'
beddo 'bed'	reddo 'red'	gaido 'guide'	waido 'wide'
daddo 'dad'	uddo 'wood'	zoido common name	haido 'hide'
deddo 'dead'	kiddo 'kid'	boodo 'board'	roodo 'road'
guddo 'good'	maddo 'mad'	gaado 'guard'	riido 'lead'
goddo 'god'	roddo 'rod'	baado 'bird'	huudo 'food'
baggu 'bag'	eggu 'egg'	dagu 'Doug'	hagu 'hug'
biggu 'big'	reggu 'leg'	bagu 'bug'	magu 'mag'
doggu 'dog'	taggu 'tag'	jogu 'jog'	ragu 'rag'

3 [bado] is a truncated form of [badominton].

Table 2: The list of the stimuli that are nonce words.

OCP-Gem	Gem	OCP-Sing	Sing
buddo	keddo	budo	hudo
boddo	koddo	dado	rado
doddo	ruddo	dodo	rudo
geddo	yuddo	dedo	rido
gaddo	taddo	gado	yudo
giddo	kuddo	gudo	wado
boggu	uggu	degu	hegu
gaggu	oggu	dogu	negu
goggu	naggu	gegu	mugu

The nonce word stimuli are listed in Table 2. These stimuli had the same phonological structures as the real word stimuli, except that all nonce word stimuli had short initial vowels, including those nonce words that contain a singleton [d], which can also have a short vowel.

2.1.2 Task

In this experiment Japanese speakers were asked to rate the naturalness of devoicing in the four grammatical conditions. The instructions stated that the questionnaire was about the naturalness of devoicing in Japanese loanwords. Using the same format as Kawahara (2011a, 2011b), for each question, the participants were presented with one stimulus item and asked to judge the naturalness of the form that undergoes devoicing of word-internal consonants (e.g., given [baddo], how natural would you find it to pronounce it as [batto]?). The instructions and stimuli were presented in Japanese orthography. The *katakana* orthography was used for the stimuli (for both [baddo] and [batto] in the example above), for both real words and nonce words, since *katakana* is conventionally used for loanwords and nonce words in standard Japanese orthography (Labrune 2012). Although the test was based on orthography, the participants were asked to read each stimulus in their heads, and make judgments based on their auditory impression rather than on the orthography.

Following Kawahara (2011a, 2011b), in this experiment, the speakers judged the naturalness of devoicing using a 5-point scale, as follows: A. “very natural”, B. “somewhat natural”, C. “neither natural nor unnatural”, D. “somewhat unnatural”, and E. “very unnatural”. Since the software that ran the experiment (see Section 2.1.3) could not present the scale numerically, the responses were later converted to a numerical scale.

The main session was blocked into two parts. The first block presented all the real word stimuli, followed by a break sign. The second block presented all the nonce word stimuli. The entire experiment was structured in this way because it was assumed that making judgments about real words would be easier than making judgments about nonce words for the participants, allowing the participants to first gain familiarity with the task before giving judgments for nonce words.[4]

2.1.3 Procedure

Sakai, an online system which runs questionnaires, was used to run the current online experiment. An advantage of this internet-based methodology is the fact that it is easy to get a large number of participants. This advantage is particularly important when the researcher does not reside in an area where there are many local speakers of the target language. A potential disadvantage is that the researcher cannot control the environment in which the participants take the experiment, although Sprouse (2011) shows that linguistic judgment data gathered with this sort of method are comparable with the data gathered in the laboratory (see Reips (2002) and Sprouse (2011) for further, general discussion about online experimentation in psychology and linguistics).

The experimental website first presented a consent form and the instructions of the experiment. Then the main session started, with one trial presented per page. The order of the stimuli within each block was randomized.

2.1.4 Participants

Thirty-three native speakers of Japanese, who were mainly students at a Japanese university, participated in the experiment. One speaker reported that they are familiar with the devoicing pattern, and hence his/her data were excluded from the following analysis.

4 Kawahara (2010) reports an experiment that addresses the question of how this organization may have affected the results.

2.1.5 Statistics

The responses were first converted to numerical values as follows: A. "very natural" = 5; B. "somewhat natural" = 4; C. "neither natural nor unnatural" = 3; D. "somewhat unnatural" = 2; E. "very unnatural" = 1. For statistical analyses, a general linear mixed model was run (Baayen et al. 2008; Baayen 2008) using R (R Development Core Team, 1993–2013) with the lme4 package (Bates et al. 2011). The fixed factors were OCP and GEM.[5] The p-values were calculated by the Markov chain Monte Carlo method using the languageR package (Baayen 2009).

2.2 Results

Figure 1 illustrates average rating scores. In real words, the average naturalness ratings showed the following order: OCP-violating geminates (4.23) > non-OCP-violating geminates (3.29) > OCP-violating singletons (2.69) > non-OCP-violating singletons (2.21). Simply put, devoicing of OCP-violating geminates was rated as more natural than the devoicing of non-OCP-violating geminates which, in turn, was rated as more natural than the devoicing of OCP-violating singletons, which was rated as more natural than devoicing of non-OCP-violating singletons, replicating the results of previous studies (Kawahara, 2011a, 2011b). Statistically, for real words, all factors are significant: OCP ($t = 5.29$, $p < .001$), GEM ($t = 11.81$, $p < .001$), and the interaction between OCP and GEM ($t = 2.68$, $p < .01$). The significance of the main effects shows that OCP and GEM each affect naturalness ratings on devoicing, and the significant interaction term indicates that the effect of OCP is bigger on the geminate pair (4.23 – 3.29 = 0.94) than on the singleton pair (2.69 – 2.21 = 0.48).

For nonce words, the order of the naturalness ratings is the same as the real word condition: OCP-violating geminates (3.64) > non-OCP-violating geminates (3.41) > OCP-violating singletons (3.06) > non-OCP-violating singletons (2.81). The statistical analysis shows that both OCP ($t = 2.56$, $p < .05$) and GEM ($t = 6.44$, $p < .001$) are significant, but, unlike the result of real words, their interaction is

5 To make the interpretation of the statistical analyses simpler, this model left out the effect of lexical usage frequencies on naturalness ratings. See Coetzee and Kawahara (2013), Kawahara (2011a) and Kawahara and Sano (2013) for discussion and also modeling of lexical frequency effects in the Japanese loanword devoicing pattern. Also, to avoid interpreting complex interaction terms, the difference between real words and nonce words was not coded in this model. The targeted comparison between real words and nonce words is provided in the discussion section (Section 2.3).

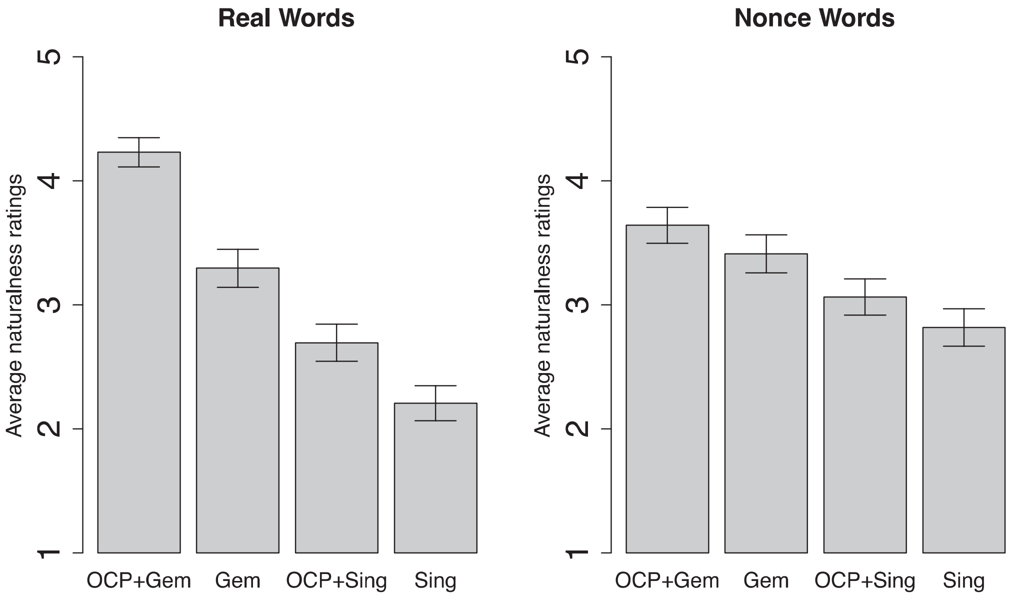


Fig. 1: The average naturalness ratings in the orthography-based rating experiment. Error bars represent 95% confidence intervals. (Reprinted with permission from Elsevier)

not ($t = 0.06$, *n.s.*). For nonce words, the effect of OCP on naturalness ratings is comparable between the singleton condition (3.64 – 3.41 = 0.23) and the geminate condition (3.06 – 2.81 = 0.25).

2.3 Discussion

2.3.1 Real words vs. nonce words

First, we observe the same order of the four grammatical conditions across real words and nonce words. This order also matches with the results of two previous studies using real words (Kawahara 2011a, 2011b). In this sense, the current experiment has shown that the results of the previous studies that use real words generalize to nonce words. Most importantly, even in nonce words, OCP-violating geminates received the highest naturalness ratings, supporting the original observation by Nishimura (2003). The current experiment thus contributes further empirical support for theoretical claims made about the Japanese loanword devoicing pattern (see Section 1.1).

At the same time, we observe a difference between real words and nonce words: in nonce words, there is less variability in naturalness ratings across the four grammatical conditions than in real words. In other words, devoicing in nonce words showed less variation in naturalness ratings across the four conditions than devoicing in real words did. The condition rated as having the most natural devoicing pattern is OCP-violating geminates; devoicing in OCP-violating geminates is judged to be less natural in nonce words than in real words. The least natural devoicing pattern is non-OCP-violating singletons; this condition is judged to be more natural in nonce words than in real words.

To statistically assess this difference between real words and nonce words, for each speaker, the standard deviations across all tokens were calculated separately for real words and nonce words. These standard deviations were then compared between the two conditions using a non-parametric within-subject Wilcoxon test. This analysis shows that the average standard deviations are 1.30 for the real words and 1.03 for the nonce words, and that the difference is significant ($p < .001$).

This reduction of variability across the four grammatical conditions in nonce words could be responsible for the absence of a significant interaction between OCP and GEM in nonce words; there may not be a space left for OCP-violating geminates to have naturalness ratings that are high enough to yield a significant interaction between OCP and GEM.

A question arises as to where the difference between real words and nonce words comes from. Presumably the participants have encountered real instances of devoicing in real words, which would make them more confident about what would happen to each target word. On the other hand, the participants have not seen nonce words before, and therefore they may feel less committed about making extreme judgments in general; i.e., they are reluctant to use endpoints of judgment scales. Despite this difference between real words and nonce words, as discussed, we observe the same ordering between the four grammatical conditions in real words and nonce words.

2.3.2 Gradiency

Second, the current study found gradient grammatical distinctions among the four grammatical conditions, just like the two previous studies (Kawahara, 2011a, 2011b). It does not seem possible to divide the judgment patterns simply into the "grammatical" category and the "ungrammatical" category. In this sense, the current results agree with the previous studies in finding distinctions that go beyond what Nishimura (2003) first proposed.[6]

6 Kawahara (2011a, 2011b) speculates about why Japanese speakers find the devoicing of non-OCP-violating geminates more natural than that of OCP-violating singletons. Beyond the specu-

One question that arises is whether this four-way distinction is due to a non-homogeneous speech community. That is, one could argue that the response from each speaker is always binary which follows a "grammatical" vs. "ungrammatical" dichotomy, but averaging over the responses from different speakers results in a gradient pattern. This hypothesis predicts that distributions of responses are at two extremes, because people should consistently rate each devoicing pattern either as completely natural (i.e., grammatical = 5 in rating) or completely unnatural (i.e., ungrammatical = 1 in rating). In this view, the differences between the four grammatical conditions arise from the difference in the number of speakers who assign grammatical status (= 5 in rating) to each condition. To examine this prediction, Figures 2 and 3 provide histograms that show the

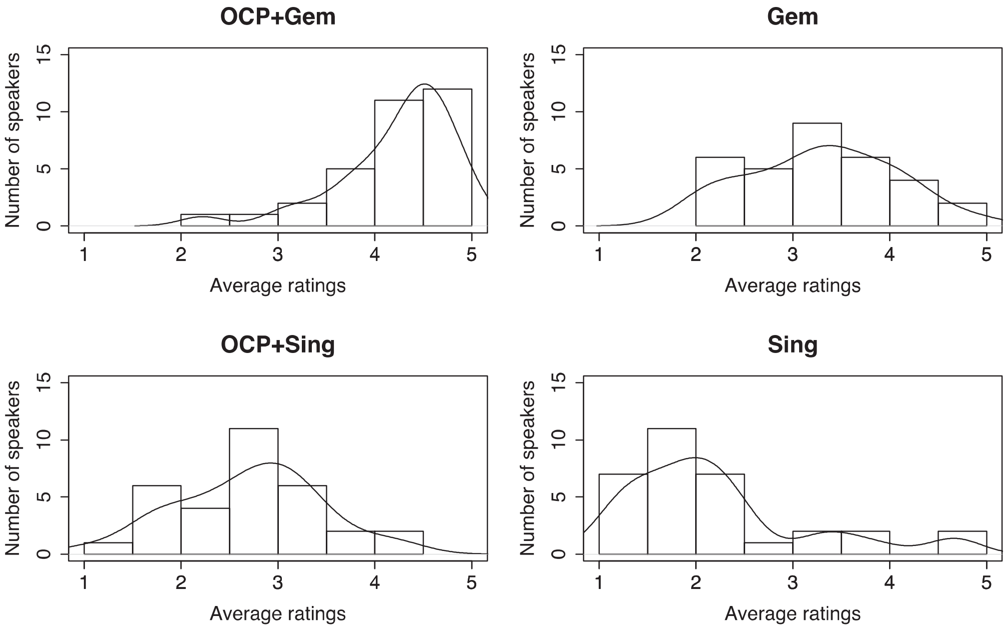


Fig. 2: A histogram of naturalness ratings (number of speakers) for real words.

lation presented there, yet another possibility is that a constraint against voiced geminates is a phonetically natural one (Ohala 1983), whereas OCP[voice] in Japanese is not (Kawahara 2008; Ohala 1981). In fact, there is evidence that children acquiring Japanese show a stage in which they apparently do not show the effect of OCP[voice] (Fukuda and Fukuda 1994), implying that this constraint may have to be learned rather than being innate (Kawahara 2008). Given the characteristics of OCP[voice] in Japanese, the speakers may have found the *phonetically* natural devoicing (= geminate devoicing) more *grammatically* natural. This possibility was brought to my attention by Armin Mester (p.c. August 2011).

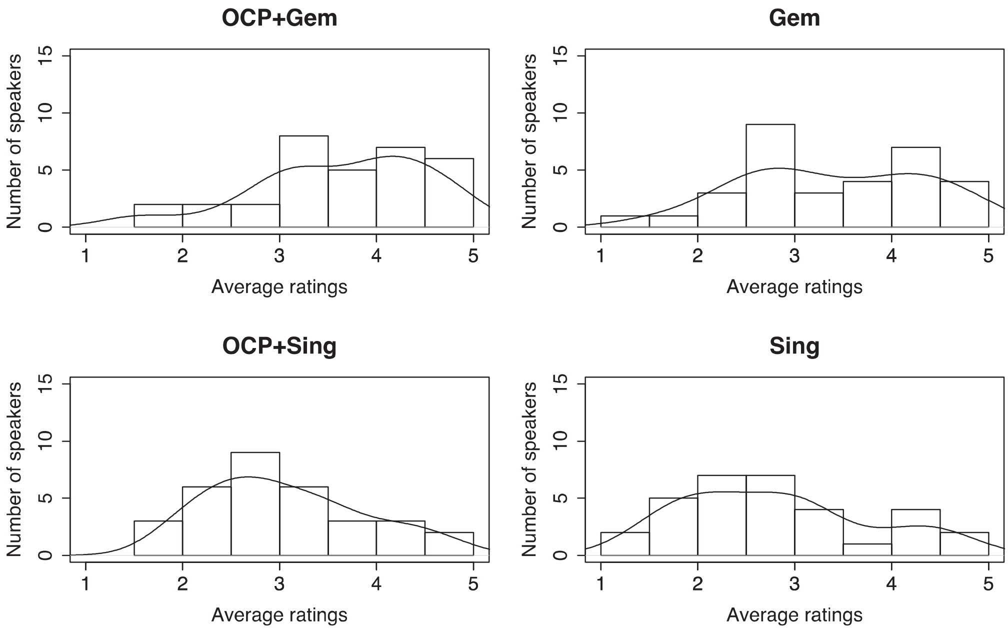


Fig. 3: A histogram of naturalness ratings (number of speakers) for nonce words.

distributions of average scores for each speaker in each grammatical condition. We observe that, contra the hypothesis, there are many speakers who show intermediate average scores in each grammatical condition.

An alternative to the hypothesis we examine in Figures 2 and 3 is to say that items within each grammatical condition showed a binary grammatical vs. ungrammatical pattern, but averaging over non-homogeneous set of items resulted in a gradient pattern. To check this possibility, Figures 4 and 5 illustrate the distributions of average naturalness ratings for each individual item. The hypothesis predicts that average scores for each item distribute at the two extreme ends, around grammatical (= 5 in rating) and ungrammatical (= 1 in rating). This prediction, however, is not supported by the actual data in Figures 4 and 5.

In summary, gradiency does not come from averaging over a non-homogeneous speech community or a non-homogeneous set of test items. It seems safe to conclude that the naturalness patterns in the Japanese devoicing case show a gradient distinction, which goes beyond the "grammatical" vs. "ungrammatical" dichotomy (Albright, 2009; Coetzee 2008; Coleman and Pierrehumbert 1997; Daland et al. 2011; Dankovičová et al. 1998; Goldrick 2011; Greenberg and Jenkins 1964; Hayes 2000; Hayes and Wilson 2008; Pertz and Bever 1975; Pierrehumbert, 2001; Shademan 2007).

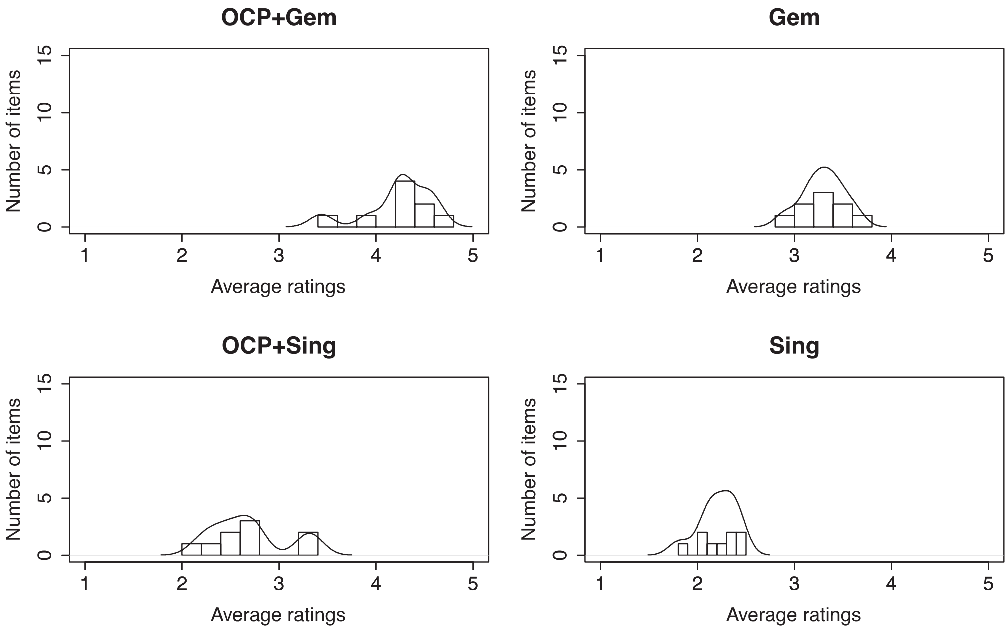


Fig. 4: A histogram of naturalness ratings (number of items) for real words.

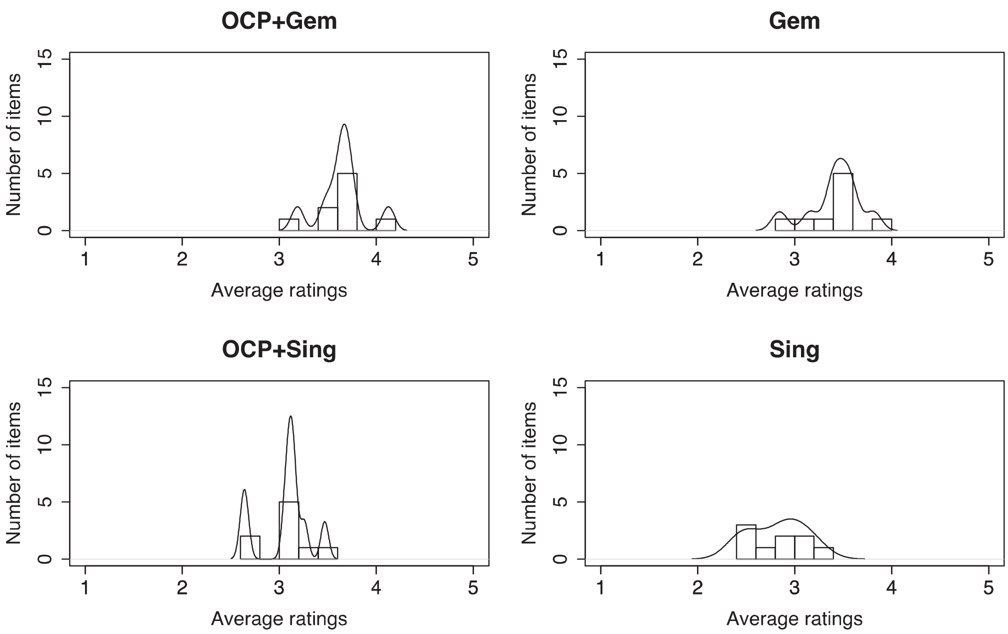


Fig. 5: A histogram of naturalness ratings (number of items) for nonce words.

3 Experiment I: Orthography-based yes/no experiment

Building on Kawahara (2012b) and to a lesser extent on Kawahara (2011a) and Kawahara (2011b), Experiment I is an orthography-based experiment that uses a binary yes/no format, rather than a scale-based rating one. The primary aim of this experiment is to address whether the gradient effect we observed in the previous three studies (Kawahara 2011a, 2011b, 2012b) can be replicated using a binary yes/no format. In these studies, given a 5-point scale, the participants may have felt obliged to use intermediate points (Schütze 2011). To avoid this task effect, the current experiment used a binary yes/no format.

3.1 Method

Experiment I is similar to Experiment III of Kawahara (2012b) reviewed and discussed in Section 2, but it instead asked native speakers whether devoicing in each of the four grammatical conditions is possible or not using a binary yes/no format. Experiment I used the same set of stimuli as Kawahara (2012b). Thirty-seven native speakers of Japanese, again mainly university students in Japan, participated in this experiment. There was no overlap between the participants of Kawahara (2012b) and those of Experiment I. No participants reported that they were familiar with the theoretical issues surrounding the devoicing phenomenon. Since the responses were binary, a logistic linear mixed model was used to analyze the results (Jaeger 2008; Quené and van den Berg 2008).

3.2 Results

Figure 6 illustrates the average ratios of DEVOICING POSSIBLE responses – the numbers of items participants chose DEVOICING POSSIBLE divided by the total number of items – of each condition, both for real words and nonce words. The ratio followed the same hierarchy as the rating experiment for both real words and nonce words: OCP-violating geminates (0.90) > non-OCP-violating geminates (0.62) > OCP-violating singletons (0.34) > non-OCP-violating singletons (0.22) for real words, and OCP-violating geminates (0.76) > non-OCP-violating geminates (0.62) > OCP-violating singletons (0.40) > non-OCP-violating singletons (0.33) for nonce words.

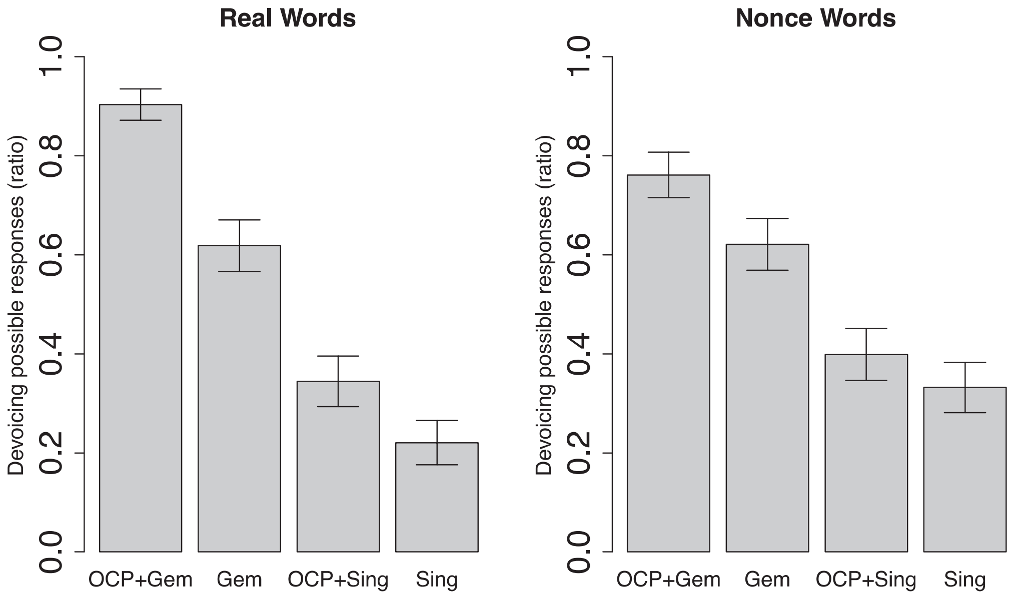


Fig. 6: Average DEVOICING POSSIBLE response ratios in the orthography-based yes/no test. Error bars represent 95% confidence intervals.

A logistic linear mixed model run on real words shows that OCP ($z = 4.17$, $p < .001$), GEM ($z = 11.09$, $p < .001$), and their interaction ($z = 3.67$, $p < .01$) are all significant. OCP and GEM each increase the possibility of devoicing. The significant interaction shows that the effect of OCP is bigger on the geminate pair (0.28 increase in ratio (0.90 – 0.62)) than on the singleton pair (0.12 increase in ratio (0.34 – 0.22)).

For nonce words, OCP ($z = 2.17$, $p < .05$) and GEM ($z = 8.56$, $p < .001$) are significant, but their interaction is not ($z = 1.65$, *n.s.*). There is some difference in the effect of OCP between the geminate pair (0.76 – 0.62 = 0.14) and the singleton pair (0.40 – 0.33 = 0.07), but this difference did not reach statistical significance.

3.3 Discussion

3.3.1 The rating experiment vs. the yes/no experiment

First of all, the rating experiment (Kawahara 2012b) and the binary yes/no experiment (the current experiment) yielded the same ordering between the four grammatical conditions. The results further support Nishimura's (2003) original observation, since naive Japanese speakers find devoicing of OCP-violating geminates possible more frequently than devoicing of non-OCP-violating geminates.

The results extend beyond Kawahara (2011a), Kawahara (2011b), and Kawahara (2012b) by showing this pattern with a yes/no format.

Second, even when the speakers made binary yes/no judgments, we observe a four-way grammatical distinction. This result shows that the gradient pattern obtained in Kawahara (2012b) was not due to the fact that the participants used a scale for their judgments (see Coleman and Pierrehumbert 1997; Dankovičová et al. 1998; Frisch et al. 2004 for similar results in wellformedness/word-likeliness judgment tasks). The phonological judgment pattern, at least in the case of Japanese devoicing, shows a gradient distinction that goes beyond a "grammatical" vs. "ungrammatical" dichotomy, regardless of whether we use a scale-based task or a binary yes/no task as the experimental format.

One may argue that this four-way grammatical distinction had arisen from averaging over a non-homogeneous speech community or a non-homogenous set of items. To address this possibility, analyses similar to those reported in Figures 2–5 were run for the current experiment, and these analyses showed that the four-way grammatical distinction did not arise from averaging over a non-homogeneous speech community or a non-homogeneous set of items.

3.3.2 Real words vs. nonce words

As with Kawahara (2012b), we again observe a reduction of variability across the four grammatical conditions in nonce words. As observed in Figure 6, OCP-violating geminates show fewer DEVOICING POSSIBLE responses in nonce words than in real words, and non-OCP violating singletons show more DEVOICING POSSIBLE responses in nonce words than in real words. To assess this decrease in variability in nonce words with respect to real words, standard deviations across the four grammatical conditions in the number of DEVOICING POSSIBLE responses for each condition were calculated. The average standard deviations in the numbers of DEVOICING POSSIBLE responses were 3.04 for the real word condition and 2.36 for the nonce word condition, and the difference is significant according to a within-subject Wilcoxon test ($p < .001$). Speakers make less consistent, less committed responses to each grammatical condition in nonce words than in real words, which results in less variability across the four grammatical conditions in nonce words.

4 Experiment II: Audio-based yes/no experiment

The second experiment is an audio-based experiment that used a yes/no format. The primary purpose of the experiment is to investigate whether the results of the previous orthography-based experiments (the previous two experiments as well as those reported in Kawahara 2011a, 2011b) can be replicated with auditory stimuli.

4.1 Method

4.1.1 Stimuli

Experiment II used the same set of stimuli as the previous experiments (Kawahara 2012b and Experiment I above). To obtain the auditory stimuli, a female native speaker of Japanese, who was naive to the purpose of this paper, pronounced all the stimuli (both faithful renditions of the stimuli (e.g., [doggu]) and forms undergoing devoicing (e.g., [dokku])) seven times in a sound-attenuated booth. She was asked to read all the stimuli with a pitch accent on the initial syllable, i.e., with HL tonal contour.

Her speech was recorded with an AT4040 Cardioid Capacitor microphone with a pop filter and amplified through an ART TubeMP microphone pre-amplifier (JVC RX 554V), digitized at a 44K sampling rate. From the seven repetitions, tokens that have phonetic deviance – such as heavy creakiness or unusual F0 contours – were first excluded. Among those that did not have such problems, one token was chosen for each test item. To equalize the amplitudes of the stimuli, the peak amplitude of all stimuli was modified to 0.8 by Praat (Boersma and Weenink 1999–2013). Then the files were converted to mp3 files and embedded in a Sakai test. In her pronunciation, as expected, voiced geminates were semi-devoiced phonetically (Kawahara 2006; see also Hirose and Ashby 2007 and Matsuura 2012). As illustrated in the right panel of Figure 7, voicing during closure ceases at an early phase of the constriction interval. (However, see Kawahara 2006 for evidence that this phonetic semi-devoicing does not itself result in neutralization of a phonological voicing contrast in geminates.)

4.1.2 Participants and procedure

Experiment II was a judgment experiment using a yes/no format; the participants were presented with an original form and a form that undergoes the devoicing in audio formats, and were asked if the second form was a possible pronunciation of the original form. Twenty-five speakers participated in this experiment. The experiments were run in a quiet room at a Japanese university, using headphones. Other aspects of the experiment were identical to the previous two experiments,

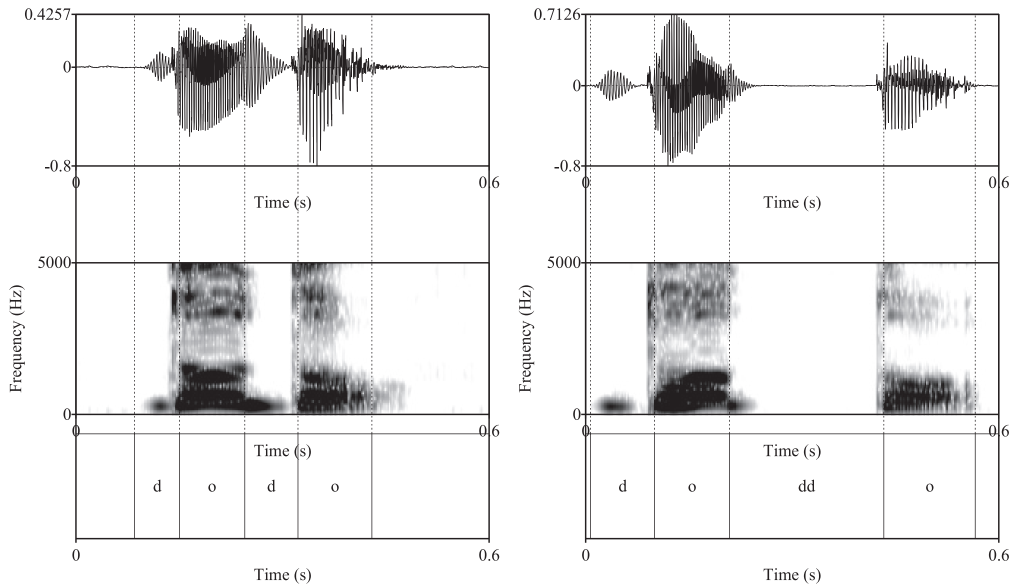


Fig. 7: A comparison of a singleton [d] and a geminate [dd] in the current stimuli.

except that the experimenter sat with the participants. As with Experiment I, within each trial, the participants were presented with an original form (e.g., [doggu] 'dog') and the form that undergoes devoicing (e.g., [dokku]). They were asked whether the second form is a possible pronunciation of the original form or not. No orthographic representations of the stimuli were given – the participants only saw play buttons. Since the two stimuli were presented as two separate play buttons, there was no fixed inter-stimulus interval. Participants were allowed to listen to the stimuli as many times as they liked.

4.2 Results

Figure 8 illustrates the results of Experiment II. The real words show the by-now familiar order: OCP-violating geminates (0.87) > non-OCP-violating geminates (0.68) > OCP-violating singletons (0.17) > non-OCP-violating singletons (0.12). For real words, GEM ($z = 11.12$, $p < .001$) is significant, and OCP is not ($z = 1.42$, *n.s.*). However, the interaction is significant ($z = 2.18$, $p < .05$), reflecting the fact that OCP has a more tangible effect on the geminate pair than on the singleton pair. Within the geminate pair, OCP is significant ($z = 4.94$, $p < .001$).

The nonce words show non-significant reversals within the geminate and the singleton pairs: non-OCP-violating geminates (0.87) > OCP-violating geminates (0.84) > non-OCP-violating singletons (0.36) > OCP-violating singletons (0.35).

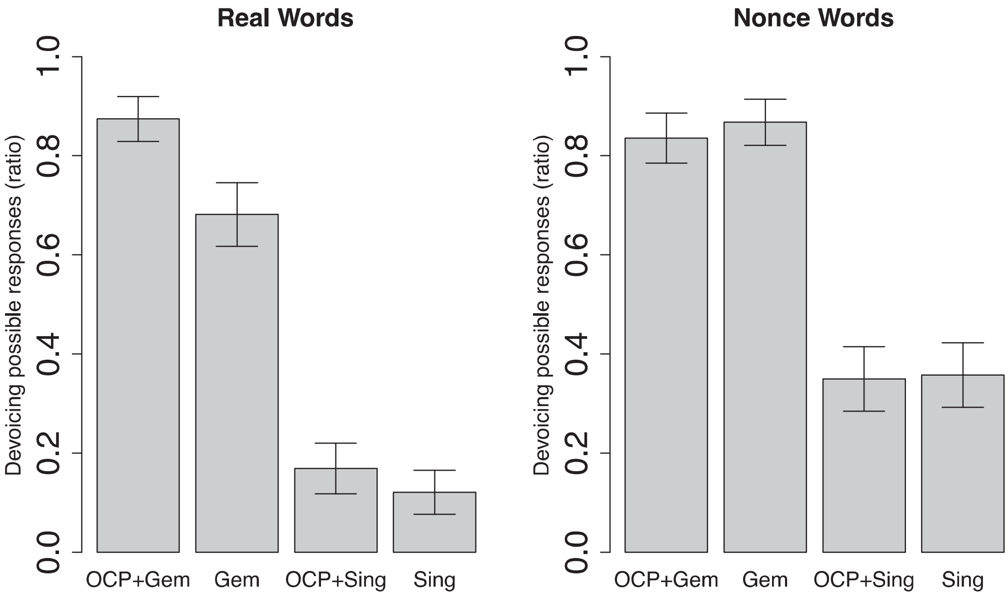


Fig. 8: Average DEVOICING POSSIBLE response ratios in the audio-based yes/no test.

The statistical test shows that only GEM ($z = 10.78$, $p < .001$) is significant, but not OCP ($z = -0.12$, *n.s.*) or the interaction ($z = -0.76$, *n.s.*). The reversal is not significant in the geminate pair ($z = -1.15$, *n.s.*) nor in the singleton pair ($z = -0.13$, *n.s.*).

4.3 Discussion

4.3.1 Orthography stimuli vs. auditory stimuli

The ordering between the four grammatical conditions in real words in Experiment II is identical to that observed in Experiments I and the previous studies (Kawahara 2011a, 2011b, 2012b). At least in the real word condition, the experiment with auditory stimuli yielded results similar to those in the orthography-based tests. In nonce words, the difference due to the OCP disappeared in both the singleton pair and the geminate pair.

One noticeable difference between auditory stimuli and orthographic stimuli is that the effect of GEM is larger in the current audio-based experiment than in the orthography-based experiment (Experiment I). The average difference between the geminate conditions and the singleton conditions in the number of DEVOICING POSSIBLE responses is 14.43 in Experiment I and 20.17 in Experiment II. To assess this difference statistically, a between-subject Wilcoxon test was run and it showed a significant effect ($p < .001$). The magnified effect of GEM in the auditory condition may be responsible for the lack of effect of OCP in nonce words, since the participants' attention was directed to the difference due to GEM more in the audio-based experiment. The difference due to OCP was diminished in nonce words, since the variability between the four conditions was reduced in general in nonce words (see below in Section 4.3.2).

The reason for this magnified effect of GEM in Experiment II perhaps lies in the phonetic semi-devoicing in Japanese voiced geminates. As we observe in Figure 7, Japanese voiced geminates are phonetically semi-devoiced. Therefore, the participants heard renditions of voiced geminates that were already close to voiceless counterparts. On the other hand, voiced singleton stops were fully voiced, which sound more different from their voiceless counterparts. This difference in the perceptibility of the [voice] contrasts was demonstrated in the perception experiment reported in Kawahara (2006). Therefore, the effect of a particular phonetic implementation pattern – semi-devoicing in this case – is likely to have affected the possibility of devoicing in the current experiment. The current result thus accords well with Kawahara's (2006) hypothesis that the higher voicing neutralizability of geminates may have its roots in the phonetic semi-devoicing of voiced geminates in Japanese.

4.3.2 Reduction of variability in nonce words

Again, similar to the previous two experiments, differences in naturalness ratings across the four different conditions are reduced in nonce words. Average standard deviations in the numbers of DEVOICING POSSIBLE responses are 3.54 for the real words and 2.77 for the nonce words ($p < .001$).

5 General discussion

5.1 Summary

To summarize, we started with three questions regarding the judgment patterns of devoicing in Japanese: (i) the similarity and the difference between real words and nonce words, (ii) the difference between scale-based judgments and yes/no judgments, and (iii) the difference between orthographic stimuli and auditory stimuli. The findings are that, throughout all the experiments, nonce words and real words generally show similar patterns, but nonce words show less variability across the four grammatical conditions than real words. The comparison between

Kawahara (2012b) and Experiment I shows that experiments using a scale-based rating and those using a binary yes/no format show very similar results. The comparison between Experiment I and Kawahara (2012b) on the one hand and Experiment II on the other shows that auditory stimuli and orthographic stimuli yield comparable results, especially in real words. However, the effect of a particular phonetic implementation – semi-devoicing in Japanese voiced geminates – is exaggerated in the audio-based experiment.

5.2 Supporting the intuition-based data

Concerning the status of OCP-violating geminates, which were treated as special by Nishimura (2003) and Kawahara (2006), all the experiments, except for the nonce word condition in Experiment II, showed that OCP-violating geminates received highest naturalness scores, or were judged to be most likely to devoice. In the current experiments, this status of OCP-violating geminates is thus shown to hold even under different modes of phonological judgments, including nonce words. In this regard, the experiments further support the intuition-based data provided by Nishimura (2003) and Kawahara (2006). Therefore, expanding on Kawahara (2011a, 2011b) by testing various modes of phonological judgment, the current experiments contribute to further secure the empirical bases of the debates that were based on Japanese loanword devoicing phenomena, briefly reviewed in Section 1.1. In other words, we can perhaps conclude that the use of intuition-based data by Nishimura (2003) and Kawahara (2006) was reliable.

More generally, the current results are in line with the body of recent experimental work by Sprouse and his colleagues (Sprouse and Almeida 2010; Sprouse et al. 2011; Sprouse and Almeida 2011, 2012) showing that intuition-based data used in generative syntax are generally reliable given that they are replicated by experiments using naive native speakers. I do not wish to imply that experimental verification of linguistic data is hence not necessary. Given some cases that cannot be replicated by experiments (recall the discussion in Section 1.1), we should continue to experimentally verify the quality of the phonological data that we use in building phonological theories and, indeed as a result we may discover finer-grained distinctions, as was the case here.

5.3 Beyond the intuition-based data

While the experimental results generally agree with the introspection-based data by Nishimura (2003) and Kawahara (2006), the experiments have also demonstrated that both the naturalness hierarchy (Kawahara 2012b) and devoiceability hierarchy (Experiments I and II) show a distinction that goes beyond a binary “grammatical” vs “ungrammatical” distinction. This gradient pattern is observed even when the participants use a binary yes/no method (see also Bader and Mäussler 2010; Coleman and Pierrehumbert 1997; Dankovičová et al. 1998; Frisch et al. 2004 for similar results). The current experiments thus show that gradient judgment patterns do not necessarily arise because many experiments in the past have used a rating scale; i.e., that it is not a task effect (c.f. Gorman (forthcoming) and Schütze (2011)). In this sense, experimentation can reveal subtle aspects of our linguistic knowledge which can be missed by an approach that is exclusively based on intuition. Therefore, experimental approaches to phonological patterns can complement – but not replace – a more-traditional approach to phonology.

5.4 Where does gradience come from?

The current experiments show that Japanese speakers’ judgments on devoicing is generally gradient, even when a yes/no format is used. One question that arises is where this gradience comes from. Even given this result, one could still hold that grammar is dichotomous, and that it is performance that is gradient (e.g., Sprouse 2007). However, recall that generally OCP and GEM both contribute to the naturalness/possibility of devoicing, and these two forces are most likely grammatical. A remaining question therefore is to identify where the gradience comes from – whether it be the grammar or performance – and if it is performance, how the two grammatical factors can derive such gradiency in performance. See Gorman (forthcoming) for recent related discussion.

5.5 Conclusion

To conclude, the three experiments generally replicated the results of the previous studies on Japanese loanword devoicing (Kawahara 2011a, 2011b) with different experimental settings. However, they also revealed interesting differences between certain conditions (for example, the difference between real words and nonce words). Although this paper used Japanese loanword devoicing as a case study, and thus its contribution is limited in its scope, it is hoped that further experimentation will reveal how systematic these differences are across different phonological phenomena and across different languages. To the extent that they are different, further theoretical research should address how to model such differences.

Acknowledgments: The experiments reported in this paper were first reported in a longer manuscript circulated in 2010 (Kawahara 2010). For their insightful comments on this paper and Kawahara (2010), I am grateful to Osamu Fujimura, Kazu Kurisu, Julien Musolino, Jeremy Perkins, Jason Shaw, Mariko Sugahara, Kristen Syrett, Kyoko Yamaguchi, four anonymous reviewers, and the audience at the colloquium talks at the University of Pennsylvania and Johns Hopkins University, and the participants of Japan Phonology Forum 2011. Sophia Kao and Melanie Pangilinan helped me with the stimulus preparation for Experiment II. Finally, I would like to thank Hope McManus for her edits at the final stage of the manuscript preparation. This research is partly supported by Research Council grant to the author from Rutgers University. All remaining errors are mine.

References

Albright, Adam. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology* 26. 9–41.

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

Baayen, R. Harald. 2009. LanguageR: R package. http://cran.r-project.org/web/packages/language (accessed 5 September 2013)

Baayen, R. Harald., Doug. J. Davidson & Douglas. M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412.

Bader, Markus & Jana Mäussler. 2010. Toward a model of grammaticality judgments. *Journal of Linguistics* 46. 273–330.

Bailey, Todd & Ulrike Hahn. 2001. Determinants of wordlikeliness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44. 568–591.

Batchelder, Eleanor Olds. 1999. Rule or rote? Native-speaker knowledge of Japanese verb inflection. *Proceedings of the Second International Conference on Cognitive Science* and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference (ICCS/JCSS99). 141–146.

Bates, Douglas, Martin Maechler & Ben Bolker. 2011. lme4: Linear mixed-effects models using S4 classes. R package. http://cran.r-project.org/web/packages/lme4/index.html (accessed 5 September 2013)

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

Boersma, Paul & David Weenink. 1999–2013. Praat: Doing phonetics by computer. http://www.praat.org/ (accessed 5 September 2013)

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Coetzee, Andries W. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84. 218–257.

Coetzee, Andries W. & Shigeto Kawahara. 2013. Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 30. 47–89.
Coleman, John & Janet Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. In *Computational phonology: Third meeting of the ACL special interest group in computational phonology*, 49–56. Somerset: Association for Computational Linguistics.
Crawford, J. Clifford. 2009. *Adaptation and transmission in Japanese loanword phonology*. Ithaca, NY: Cornell University dissertation.
Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis & Ingrid Norrmann. 2011. Explaining sonority projection effects. *Phonology* 28. 197–234.
Dankovičová, Jana, Paula West, John Coleman & Andrew Slater. 1998. Phonotatic grammaticality is gradient. Poster presented at LabPhon 6, University of York, 4 July.
Davis, Stuart & Natsuko Tsujimura. 1991. An autosegmental account of Japanese verbal conjugation. *Journal of Japanese Linguistics* 13. 117–144.
Dąbrowska, Ewa. 2010. Naive vs. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27. 1–23.
Farris-Trimble, Ashley. 2008. *Cumulative faithfulness effects in phonology*. Bloomington, IN: Indiana University dissertation.
Frisch, Stefan, Nathan Large & David Pisoni. 2004. Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language* 42. 481–496.
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.
Gibson, Edward & Evelina Fedorenko. 2010. Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* 14. 233–234.
Goldrick, Matthew. 2011. Utilizing psychological realism to advance phonological theory. In John A. Goldsmith, Jason Riggle & Alan Yu (eds.) *The handbook of phonological theory,* 2nd edn., 631–660. Oxford: Blackwell-Wiley.
Gorman, Kyle. forthcoming. Gradient and categorical aspects of wordlikeness judgments. *Proceedings of the 43rd annual meeting of North East Linguistic Society*.
Greenberg, Joseph & James Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20. 157–177.
Griner, Barry. 2001. *Productivity of Japanese verb tense inflection: A case study*. Los Angeles, CA: University of California Los Angeles MA Thesis.
Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*, 88–120. Oxford: Oxford University Press.
Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic inquiry* 39. 379–440.
Hirose, Aki & Michael Ashby. 2007. An acoustic study of devoicing of the geminate obstruents in Japanese. *Proceedings of ICPhS* XVI. 909–912.
Itô, Junko & Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17. 49–73.
Itô, Junko & Armin Mester. 1995. Japanese phonology. In John Goldsmith (ed.), *The handbook of phonological theory*, 817–838. Oxford: Blackwell.
Itô, Junko & Armin Mester. 1998. Markedness and word structure: OCP effects in Japanese. Ms. University of California Santa Cruz.

Itô, Junko & Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 62–100. Oxford: Blackwell.
Itô, Junko & Armin Mester. 2003. *Japanese morphophonemics*. Cambridge, MA: MIT Press.
Itô, Junko & Armin Mester. 2008. Lexical classes in phonology. In Shigeru Miyagawa & Mamoru Saito (eds.), *The Oxford handbook of Japanese linguistics*, 84–106. Oxford: Oxford University Press.
Jaeger, Florian T. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59. 434–446.
Katayama, Motoko. 1998. *Optimality Theory and Japanese loanword phonology.* Santa Cruz, CA: University of California Santa Cruz dissertation.
Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82. 536–574.
Kawahara, Shigeto. 2008. Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* 17. 317–330.
Kawahara, Shigeto. 2010. Modes of phonological judgments. Ms. Rutgers University.
Kawahara, Shigeto. 2011a. Aspects of Japanese loanword devoicing. *Journal of East Asian Linguistics* 20. 169–194.
Kawahara, Shigeto. 2011b. Japanese loanword devoicing revisited: A rating study. *Natural language and Linguistic Theory* 29. 705–723.
Kawahara, Shigeto. 2012a. Gairaigo yuusei sokuon no museika: Rironteki kouken [Japanese loanword devoicing: A review]. In Phonological Society of Japan (ed.), *On'in kenkyu* [Phonological studies] 15, 93–104. Tokyo: Kaitakusha.
Kawahara, Shigeto. 2012b. Lyman's Law is active in loanwords and nonce words: Evidence from naturalness judgment experiments. *Lingua* 122. 1193–1206.
Kawahara, Shigeto. forthcoming. The phonetics of obstruent geminates, *sokuon*. In Haruo Kubozono (ed.), *The handbooks of Japanese language and linguistics: Phonetics and phonology*. Berlin & Boston: De Gruyter Mouton.
Kawahara, Shigeto & Shinichiro Sano. 2013. A corpus-based study of geminate devoicing in Japanese: Linguistic factors. Language Sciences 40: 300–307.
Kubozono, Haruo, Junko Itô & Armin Mester. 2009. Consonant gemination in Japanese loanword phonology. In Linguistic Society of Korea (ed.), *Current issues in unity and diversity of languages: Collection of papers selected from the 18th International Congress of Linguists*, 953–973. Seoul: Dongam Publishing Co.
Labov, William. 1996. When intuitions fail. In Lisa McNair, Kora Singer, Lise Dolbrin, & Michelle Aucon (eds.), *Proceedings of Chicago Linguistic Society 32: Papers from the parasession on theory and data in linguistics*, 77–106. Chicago: Chicago Linguistics Society.
Labrune, Laurence. 2012. *The phonology of Japanese*. Oxford: Oxford University Press.
Leben, Will. 1973. *Suprasegmental phonology*. Cambridge, MA: MIT dissertation.
Lyman, Benjamin S. 1894. Change from surd to sonant in Japanese compounds. *Oriental Studies of the Oriental Club of Philadelphia*.
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]. Paper presented at the Phonetic Society of Japan, September 29th.
Myers, James. 2009. Syntactic judgment experiments. *Language and Linguistic Compass* 3. 406–423.
Nishimura, Kohei. 2003. *Lyman's Law in loanwords*. Nagoya: Nagoya University MA thesis.

Ohala, John J. 1974. Experimental historical phonology. In John M. Naderson & Charles Jones (eds.), *Historical linguistics II: Theory and description in phonology. Proceedings of the first international linguistic conference on historical linguistics*, 353–389. New York: Elsevier.

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

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter MacNeilage (ed), *The production of speech*, 189–216. New York: Springer.

Ohala, John J. 1986. Consumer's guide to evidence in phonology. *Phonology* 3. 3–26.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33. 999–1035.

Pater, Joe. forthcoming. Universal grammar with weighted constraints. In John J. McCarthy & Joe Pater (eds.), *Harmonic grammar and harmonic serialism.* London: Equinox.

Pertz, Doris L. & Thomas G. Bever. 1975. Sensitivity to phonological universals in children and adolescents. *Language* 51. 149–162.

Pierrehumbert, Janet B. 2001. Stochastic phonology. *GLOT* 5. 1–13.

Quakenbusch, Hiroko. 1989. Gairaigo-no sokuonka-ni tsuite [On geminaton in loanwords]. *Ryuugakusei Nihongo Kyouiku-ni Kansuru Rironteki Jissenteki Kenkyuu* [Theoretical and practical research on Japanese language education for foreign students]. 1–7.

Quené, Hugo & Huub van den Berg. 2008. Examples of mixed effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59. 413–425.

R Development Core Team. 1993–2013. *R: A language and environment for statistical computing*. Vienna, Austria. R Foundation for Statistical Computing. Software, available at http://www.R-project.org.

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

Rice, Keren. 2006. On the patterning of voiced stops in loanwords in Japanese. *Toronto Working Papers in Linguistics* 26. 11–22.

Sanders, Nathan. 2003. *Opacity and sound change in the Polish lexicon*. Santa Cruz, CA: University of California Santa Cruz dissertation.

Sano, Shinichiro & Shigeto Kawahara. forthcoming. A corpus-based study of geminate devoicing in Japanese: The role of the OCP and external factors. Gengo Kenkyuu [The Journal of the Linguistic Society of Japan].

Schütze, Carlson. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Schütze, Carlson. 2011. Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Reviews: Cognitive Science* 2. 206–211.

Shademan, Shabhame. 2007. *Grammar and analogy in phonotactic well-formedness judgments*. Los Angeles, CA: University of California Los Angeles dissertation.

Sorace, Antonella & Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115. 1497–1524.

Sprouse, Jon. 2007. Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1. 123–134.

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

Sprouse, Jon & Diogo Almeida. 2010. A quantitative defense of linguistic methodology. Ms. University of California, Irvine.

Sprouse, Jon & Diogo Almeida. 2011. Power in acceptability judgment experiments. Ms. University of California, Irvine.
Sprouse, Jon & Diogo Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics* 48. 609–652.
Sprouse, Jon, Carlson Schütze & Diogo Almeida. 2011. Assessing the reliability of journal data in syntax: Linguistic Inquiry 2001–2010. Ms. University of California, Irvine, University of California Los Angeles & Michigan State University.
Steriade, Donca. 2004. Projecting non-lexical phonology from phonetic knowledge. Paper presented at the Symposium on Phonological Theory: Representations and Architectures (City University of New York, February 20–21).
Tateishi, Koichi. 2002. Bunpou-no ichibutoshite-no goisou-no zehi [Lexical strata as a part of grammar]. *Onsei Kenkyuu* [Journal of the Phonetic Society of Japan] 6. 34–43.
Tesar, Bruce. 2007. A comparison of lexicographic and linear numeric optimization using violation difference ratios. Ms. Rutgers University.
Tsujimura, Natsuko. 1996. *An introduction to Japanese linguistics* (Blackwell Textbooks in Linguistics 10). Cambridge, MA & Oxford: Blackwell Publishers.
Vance, Timothy J. 1987. *An introduction to Japanese phonology*. New York: SUNY Press.
Vance, Timothy J. 2007. Have we learned anything about *rendaku* that Lyman didn't already know? In Bjarke Frellesvig, Masayoshi Shibatani & John Carles Smith (eds.), *Current issues in the history and structure of Japanese*, 153–170. Tokyo: Kuroshio.

Kawahara, Shigeto and Shin-ichiroo Sano (2013) A corpus-based study of geminate devoicing in Japanese: Linguistic factors. Language Sciences 40: 300-307.

Language Sciences 40 (2013) 300–307

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

A corpus-based study of geminate devoicing in Japanese: linguistic factors

Shigeto Kawahara [a,*], Shin-ichiro Sano [b]

[a] Keio University, Japan
[b] Tokyo University of Agriculture and Technology, Japan

A R T I C L E I N F O

Article history:
Received 1 April 2013
Received in revised form 10 June 2013
Accepted 5 July 2013
Available online xxxx

Keywords:
Japanese
OCP
Lyman's Law
Geminate
Devoicing
Variation

A B S T R A C T

Ever since Nishimura (2003, 2006) pointed out that voiced geminates can optionally devoice in Japanese when they co-occur with another voiced obstruent (e.g. /baddo/ → [batto] 'bad'), the pattern has been analyzed within a number of theoretical frameworks and studied in several experimental studies. However, there are only a few studies on actual production data. Moreover, most of the previous studies have generally assumed that this pattern is a linguistically monolithic phenomenon—all OCP-violating geminates would have equal probability of devoicing. By studying the Corpus of Spontaneous Japanese (Kokuritsu-Kokugo-Kenkyuujo, 2008), we show that many linguistic factors affect the probability of the devoicing of voiced geminates: in addition to OCP (voice), we find effects of the location of the trigger with respect to the target, the number of triggers, place of articulation, and lexical usage frequency. All of these observed patterns accord well with phonetic considerations and/or cross-linguistic tendencies. We conclude that geminate devoicing in Japanese phonology is not a linguistically monolithic phenomenon, because the probability of devoicing is affected by several linguistic factors. We suggest that future analyses of this phenomenon should take into account the factors that are identified in this project.

1. Introduction

In the native phonology of Japanese, voiced obstruent geminates are not allowed; however, recent borrowings from English and other languages result in voiced obstruent geminates due to a word-final gemination process (Irwin, 2011; Itô and Mester, 1995, 1999; Kubozono et al., 2008; Vance, 1987).[1] Nishimura (2003, 2006) pointed out that such voiced geminates can optionally devoice when they occur with another voiced obstruent in Japanese loanwords, as in (1). Ever since then, the pattern has been analyzed within a number of theoretical frameworks and studied in several experimental studies. Furthermore, Nishimura (2003, 2006) claims that this devoicing is caused by a well-known restriction in native Japanese phonology that does not allow two voiced obstruents within the same morpheme (also known as Lyman's Law, or OCP (voice),[2] henceforth the OCP), as shown by the (purportedly) ungrammatical devoicing in (2).

* Corresponding author. Address: Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan. Tel.: +81 03 5427 1455.
E-mail address: kawahara@icl.keio.ac.jp (S. Kawahara).

[1] Emphatic gemination can create voiced obstruent geminates in native words, as in /sugoi/ → [suggoi] 'super' and /hidoi/ → [hiddoi] 'awful'. This emphatic gemination process is generally non-structure-preserving in that it creates kinds of geminates that are not allowed to make lexical contrasts (Kawahara, 2001; Vance, 1987).

[2] OCP (voice) in Japanese only targets voicing in obstruents, not in sonorants. The formulation of Lyman's Law as the OCP effect on a [(+)voice] feature is originally due to Itô and Mester (1986). Their analysis is further developed in their subsequent work (Itô and Mester, 1996, 1998, 2003; Mester and Itô, 1989). See also Vance (2007) on the discussion of Lyman's original article (Lyman, 1894) on this restriction, and Ihara et al. (2009), Kawahara (2012), and Vance (1980) for the evidence for the synchronic psychological reality of Lyman's Law from the perspective of Rendaku, a process of voicing of the initial consonant of the second member in compounding. See also Vance (in press) for a comprehensive recent overview of related issues.

0388-0001/$ - see front matter
http://dx.doi.org/10.1016/j.langsci.2013.07.001

(1) OCP-violating geminates can optionally devoice
  a. be**dd**o → be**tt**o 'bed'
  b. ba**gg**u → ba**kk**u 'bag'
  c. bi**gg**u → bi**kk**u 'big'

(2) Non-OCP-violating geminates do not devoice
  a. suno**bb**u → *suno**pp**u 'snob'
  b. he**dd**o → *he**tt**o 'head'
  c. re**gg**u → *re**kk**u 'leg'

Several researchers have analyzed this devoicing pattern using various theoretical mechanisms, including local conjunction (McCarthy, 2008; Nishimura, 2003, 2006), phonetically-based phonology (Kawahara, 2006, 2008; Steriade, 2004), the theory of markedness and contrast (Rice, 2006), Harmonic Phonology (Farris-Trimble, 2008; Pater, 2009, in press), Noisy Harmonic Phonology (Coetzee and Kawahara, 2013; Coetzee and Pater, 2011), and Maximum Entropy Grammar (Coetzee and Pater, 2011). In this sense, this pattern has contributed much to many recent debates in phonological theory (see Kawahara, 2011a for a more extensive review).

Furthermore, Kawahara (2011a,b, 2012, 2013) has studied the devoicing pattern in several naturalness judgment experiments. The results generally support the intuition-based data in (1)–(2) in that naive native Japanese speakers find the devoicing of OCP-violating geminates more natural than the devoicing of non-OCP-violating geminates. However, all of these studies also found that Japanese speakers do not find devoicing of non-OCP-violating geminates, as in (2), to be entirely unnatural, contrary to the intuition reported by Nishimura (2003, 2006) and Kawahara (2006).

Our first aim of this study is thus to examine the role of the OCP in the geminate devoicing pattern in actual production data, using a large-scale speech corpus (our companion paper, Sano and Kawahara (2013), offers a similar attempt, but the current paper goes a few steps further in this respect; see below). Although this pattern has been studied from a variety of theoretical and experimental perspectives, actual production patterns have been generally understudied. Nishimura (2003, 2006), in the appendices to his papers, reports some analyses of the devoicing pattern using the older versions of the Corpus of Spontaneous Japanese (the CSJ, see below), and shows the effect of the OCP on devoicing. The first aim of this study is to replicate this result with an updated version of the CSJ.[3] In addition, we go one step further and examine how other linguistic factors related to OCP, such as the number of triggers and the location of the trigger with respect to the target, affect the probability of devoicing.

The second question addressed in this study is whether the devoicing of geminates is a linguistically monolithic phenomenon—that is, whether all OCP-violating geminates are equally likely to devoice, regardless of other factors that may possibly affect devoicing. Most previous theoretical studies have assumed that the geminate devoicing pattern is a monolithic phonological phenomenon, i.e., it is as simple as "geminates optionally devoice when they violate the OCP." However, Kawahara (2011a) found that various linguistic factors affect the naturalness judgments of the devoicing of geminates, as judged by a number of naive native speakers of Japanese. Moreover, our companion paper (Sano and Kawahara, 2013), through an extensive corpus study of the CSJ, found that various linguistic-external factors, such as gender, age, and education level, affect the devoicability of geminates in actual production.[4] These studies (Kawahara, 2011a; Sano and Kawahara, 2013) raise the possibility that geminate devoicing in Japanese loanwords is not a monolithic phenomenon, because various factors affect the probability—or naturalness—of devoicing. This study thus aims to test the hypothesis that the devoicing of voiced geminates in Japanese is affected by several linguistic factors, using the production data provided by the CSJ.

In summary, the current study uses the CSJ to test how various linguistic factors affect the devoicability of geminates in Japanese. We in particular examine those factors that are known to affect devoicing and the effects of OCP cross-linguistically.[5] To provide a preview of our results, we will observe that many linguistic factors do indeed affect the devoicing probability of voiced geminates: the OCP, the number of triggers, the distance between the trigger and the geminate, place of articulation, and lexical usage frequency. Furthermore, all of these observed patterns accord well with cross-linguistic tendencies and/or phonetic considerations. We conclude that geminate devoicing in Japanese phonology is not a linguistically monolithic phenomenon, being affected by several linguistic factors, and that future theoretical analyses of this phenomenon should take into account the factors that are identified in this project.

2. Method

To investigate how the OCP and other linguistic factors affect the probability of devoicing, we searched through the Corpus of Spontaneous Japanese, version 2 (the CSJ; Kokuritsu-Kokugo-Kenkyuujo, 2008; Maekawa, 2003, 2004; Maekawa et al.,

[3] Nishimura (2003) used the monitor—trial—version, and Nishimura (2006) used the first version (p.c. Kohei Nishimura, Aug. 2012). For the updated information of the 2nd version, which the current study has used, see http://www.ninjal.ac.jp/csj/data/previous/2nd/02revision/.

[4] We report the effects of linguistic factors and those of extra-linguistic factors separately in a different paper (Sano and Kawahara, 2013) for the sake of exposition. In doing so, we also follow the claim that linguistic factors and extra-linguistic factors do not interact (Labov, 1982; Sankoff and Labov, 1979; Weiner and Labov, 1983).

[5] See Sano (2013) for an examination of many other factors as well as the diachronic development of this devoicing pattern.

2000; Sano and Hibiya, 2012—for the evaluation of the CSJ's reliability as a spoken corpus, see Maekawa, 2003 in particular). This is one of the largest and most reliable corpora in spoken Japanese, based on 662 h of speech with 7.5 million words, produced by a total of 1417 speakers. This corpus is large in size and comes with a rich annotation system. Another characteristic of this database is that it provides both underlying forms and surface forms (coded in the forms of *hatsuon-kei*). Using this feature, we retrieved a set of words with voiced geminates, and studied how the words are actually pronounced using the phonetic transcription provided by the CSJ.[6]

We first extracted words with underlying voiced obstruent geminates ($N = 1666$). As the focus of this study was on devoicing, we excluded tokens in which the voiced geminates went through changes other than devoicing, such as degemination. This process resulted in 1617 tokens (97%).[7] Among those tokens, 472 of them showed devoicing (29%). We then tested how the OCP and various other linguistic factors affected the probability of devoicing. Any token containing another voiced obstruent within 6 preceding or following moras was taken to violate the OCP.[8] To calculate the probability of devoicing for each condition, we calculated the proportion of relevant tokens which showed devoicing (=devoiced tokens/total relevant tokens).

3. Results and discussion

3.1. The general role of OCP

Fig. 1 shows the probability of devoicing in the OCP-violating condition (e.g. *beddo* 'bed') and the non-OCP-violating condition (e.g. *shureddaa* 'shredder'). It demonstrates that devoicing is more likely for OCP-violating geminates (438/1099 = 40%) than for non-OCP-violating geminates (26/518 = 5%) ($\chi^2(1) = 207.1$, $p < .001$).[9] These data support the intuition by Nishimura (2003, 2006) and Kawahara (2006) that the OCP is a crucial factor in inducing devoicing, and also replicate the earlier corpus studies reported in the appendices of Nishimura (2003, 2006) which used older versions of the CSJ. However, we also note that devoicing is not entirely impossible for non-OCP-violating geminates (it happens about 5% of the time), which may be the basis of the judgment patterns found in Kawahara (2011a,b, 2012, 2013).

3.2. The distance effect

Having established that the OCP plays a crucial role in triggering devoicing in the previous section, Fig. 2 examines the distance effect; i.e., the distance between the trigger and the potential undergoer of devoicing. In some cases, the trigger and the undergoer are in adjacent moras, as in *beddo* 'bed'[10]; on the other hand, they can be separated by intervening moras, as in *neebaahuddo* 'neighborhood'. The distance effect is interesting to examine, because it is known cross-linguistically that dissimilatory force is stronger between closer segments (Frisch, 2004; Frisch et al., 2004; Ihara et al., 2009; Itô and Mester, 2003; Kawahara et al., 2006; Odden, 1994; Pulleyblank, 2002; Suzuki, 1998; Vance, 1980; Tanaka, 2007).

Against this background, Fig. 2—together with Table 1—lays out the probabilities of devoicing for each number of intervening moras. We observe a general trend in which the closer the trigger is to the target, the more likely devoicing is to occur ($\chi^2(5) = 191.1$, $p < .001$). The effect of the OCP is strongest when the trigger is in the adjacent mora, and disappears when the trigger is more than 5 moras away.

This observation accords well with the cross-linguistic observation that the OCP functions more strongly between closer segments (Frisch, 2004; Frisch et al., 2004; Ihara et al., 2009; Itô and Mester, 2003; Kawahara et al., 2006; Odden, 1994; Pulleyblank, 2002; Suzuki, 1998; Vance, 1980; Tanaka, 2007). For example, in Arabic consonant co-occurrence restrictions, constraints against a pair of homorganic consonants are stronger between syllable-adjacent consonants than non-syllable-adjacent consonants (Frisch, 2004; Frisch et al., 2004).[11] This result shows that the Japanese OCP (voice) constraint shares the same characteristic as the OCP effects found in other languages.

---

[6] As made explicit in some previous work (Kawahara, 2006, 2008, 2011b; Kaneko and Iverson, 2009), devoicing occurs in loanword phonology, rather than in the process of loanword adaptation (the latter process referring to borrowing of words into Japanese from other languages). Therefore, we assume that the underlying forms are those forms that are already borrowed into Japanese, not those of source language; e.g. for 'dog', it is [doggu] rather than [dɔg].

[7] Devoicing accounted for 92% of the repair strategies that happened to voiced geminates; the next most frequent repair strategy was degemination (7%); other strategies happened very rarely. See Sano (2013) for details.

[8] We will observe in Fig. 2 that most of the OCP-induced devoicing was caused by a local trigger at the moraic level, but we included non-local triggers as well in this study to see how the distance between the trigger and the target affects devoicability. See Fig. 2, which shows that the OCP's effects virtually disappear with more than 5 intervening moras.

[9] This result is also reported in our companion paper (Sano and Kawahara, 2013).

[10] We assume, following the majority of the literature, that the basic prosodic unit in Japanese is a (C) V mora. In Japanese, vowels and any coda consonants (coda nasals or coda part of geminates) count as one mora, and onsets count as being in the same mora as their nucleus. For example, *beddo* is divided into *be-d-do* in terms of moras in Japanese. See Kubozono (1999), Labrune (2012a,b), Poser (1990), Vance (1987, 2008) and many references cited therein for extensive discussion on the status of mora as a basic prosodic unit in Japanese prosodic phonology.

[11] This observation also accords well with another observation about the OCP-driven blockage of Rendaku in Japanese. The OCP blocks a morphophonemic alternation which voices the initial consonant of compounds (Rendaku) (Itô and Mester, 1986, 2003; Kawahara, 2012; Kubozono, 2005; Vance, 2007), and in some nonce-word experiments, this blockage effect was found to be stronger when the blocker is closer to the potential undergoer of Rendaku (Ihara et al., 2009; Vance, 1980) (though see Kawahara, 2012 who did not find this distance effect in naturalness judgment experiments with contemporary speakers).

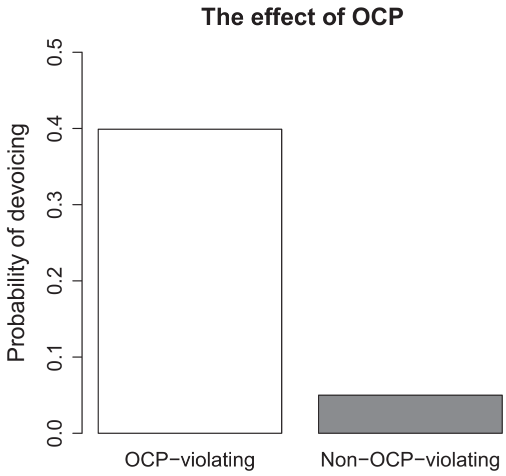


Fig. 1. The effect of OCP on devoicing.

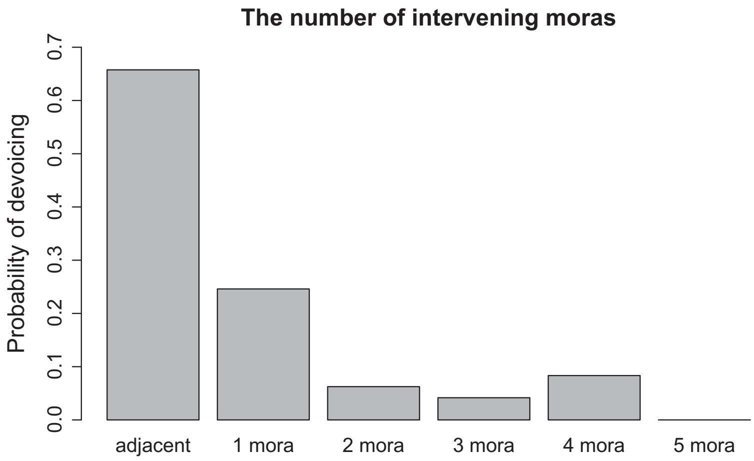


Fig. 2. The number of intervening moras and the probability of devoicing. The more moras that intervene between the trigger and the target, the less likely devoicing is to occur.

Table 1
The number of intervening moras and the probability of devoicing (actual numbers).

	Total	Devoiced	%
Adjacent	476	313	65.8
1 mora	126	31	24.6
2 moras	112	7	.6
3 moras	24	1	.4
4 moras	12	1	.8
5 moras	4	0	0

3.3. Number of triggers

Next we examine the effect of the numbers of triggers; in some case, there is only one trigger for OCP-induced devoicing, as in *beddo* ‘bed’, but there can be two (e.g. *debaggu* ‘debug’) or three (e.g. *bagudaddo* ‘Bagdad’). Kawahara (2011a), in his naturalness experiment, found that the number of triggers do indeed impact the naturalness of devoicing; see also Tesar (2007) for related discussion on the effect of multiple triggers from a theoretical perspective.

Fig. 3 shows the effect of the number of triggers on the devoicability of OCP-violating geminates. It shows that the more voiced obstruents there are, the more likely it is that devoicing occurs: one trigger (353/754 = 46.8%), two triggers (56/102 = 54.9%), and three triggers (23/29 = 79%) ($\chi^2(2) = 13.7$, $p < .01$). This result is compatible with the naturalness judgement patterns obtained in Kawahara (2011a). This case instantiates a pattern in which dissimilation is coerced more strongly with more trigger consonants.

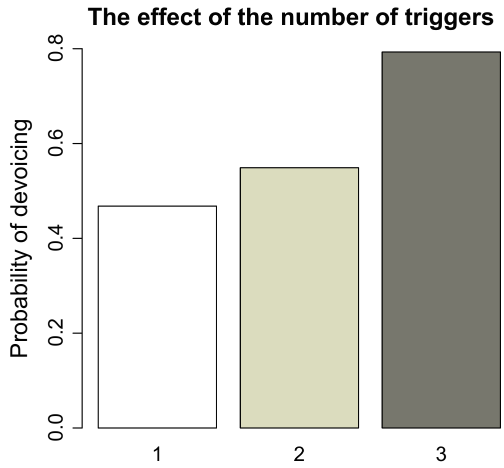


Fig. 3. The effect of the number of triggers. The more triggers exist, the more likely that geminates devoice.

This case shows that more triggers can make a variable process more likely to apply. A question arises whether there can be a categorical pattern of dissimilation which occurs only when there are more than one trigger. There are patterns of vowel harmony that are triggered only when there is more than one trigger: e.g. Classical Manchu (Dresher, 2010; Walker, 2001; Zhang, 1996) and Oroqen (Walker, 2001). The current result shows that the same pattern may hold in dissimilation—dissimilative force is stronger with multiple triggers (Kawahara, 2011a). However, Itô and Mester (2003) argue that there should not be dissimilative pattern which is triggered only by multiple triggers. Given the current results, it thus remains to be seen whether there can be a dissimilation pattern which can be triggered only by multiple triggers.

3.4. The effect of place of articulation

We now look at factors not related to the OCP, starting with the effect of place of articulation of the voiced geminates. Place of articulation is interesting to examine, because cross-linguistically it is known that backer places of articulation suffer from more articulatory difficulty in maintaining voicing for aerodynamic reasons (Hayes, 1999; Hayes and Steriade, 2004; Jaeger, 1978; Kawahara, 2006; Kingston, 2007; Ohala, 1983; Ohala and Riordan, 1979; Westbury, 1979; Westbury and Keating, 1986).

Fig. 4 shows the probabilities of devoicing according to place of articulation: [bb], [dd], and [gg]. We observe that as place of articulation goes further back in the mouth, the more likely the geminate to be devoiced: [bb] (1/18 = 5.6%), [dd] (319/1100 = 29%), [gg] (115/230 = 50%) ($\chi^2(2) = 44.3$, $p < .001$). We do acknowledge that the number of tokens for [bb] is rather small, but nevertheless we maintain that this effect of place articulation is compatible with what we expect from the aerodynamics of voiced stops, as we discuss below.[12]

This effect of place of articulation is compatible with a well-known aerodynamic consideration. Intraoral air pressure goes up with stop closure, but this rise in pressure makes it hard to maintain the transglottal air pressure drop that is necessary to produce voicing. The further back the place of articulation, the worse the problem, because the intraoral cavity behind the closure is smaller and more difficult to expand. As a result, the further back the place of articulation, the harder it is to maintain voicing (see Hayes, 1999; Hayes and Steriade, 2004; Jaeger, 1978; Kawahara, 2006; Kingston, 2007; Ohala, 1983; Ohala and Riordan, 1979; Westbury, 1979; Westbury and Keating, 1986). Therefore, the devoicability hierarchy follows the order that is predicted from an aerodynamic point of view.

3.5. The effect of lexical frequency

Finally, Fig. 5 plots the effect of lexical usage frequency on the *x*-axis against the probability of devoicing on the *y*-axis for OCP-violating geminates. The figure shows a trend for a positive correlation between the two factors. Statistically, the Pearson correlation coefficient *r* between the log-transformed lexical frequency and devoicing probability is 0.49. This result matches well with what Kawahara (2011a,b) found, although, unfortunately, the correlation coefficient did not reach significance due to small N, because there are not many words containing OCP-violating geminates (type-wise). There is one outlier whose frequency is relatively high (90) and the devoicing probability is relatively low (0.01) (*neebaahuddo* 'neighborhood'). If we exclude this item as an outlier, the *r* Pearson coefficient becomes as high as 0.69 and is statistically significant ($p < .01$).

[12] This figure sets aside [zz] for two reasons. First, we should not treat voiced stops on par with voiced fricatives, because they involve different—though related—types of aerodynamic problems (Ohala, 1983; Lindblom and Maddieson, 1988). Second, in Japanese [zz] variably alternates with [ddz], and the variation pattern is governed by some complicating phonetic factors (Maekawa, 2010). Therefore, it does not seem safe to directly compare voiced stops and voiced fricative. See Sano (2013) for some discussion on the behavior of [zz].

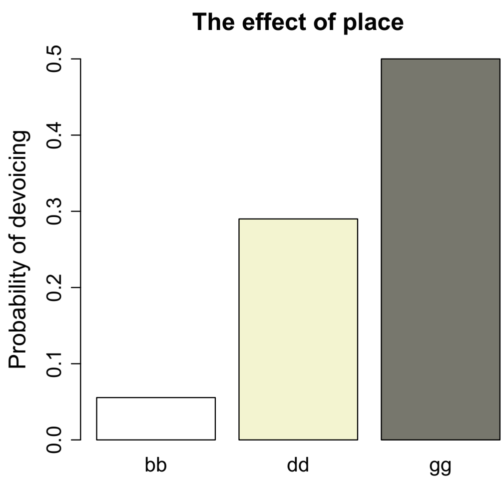


Fig. 4. The effect of place of articulation on devoicing. The backer the place of articulation, the more likely that geminates are devoiced.

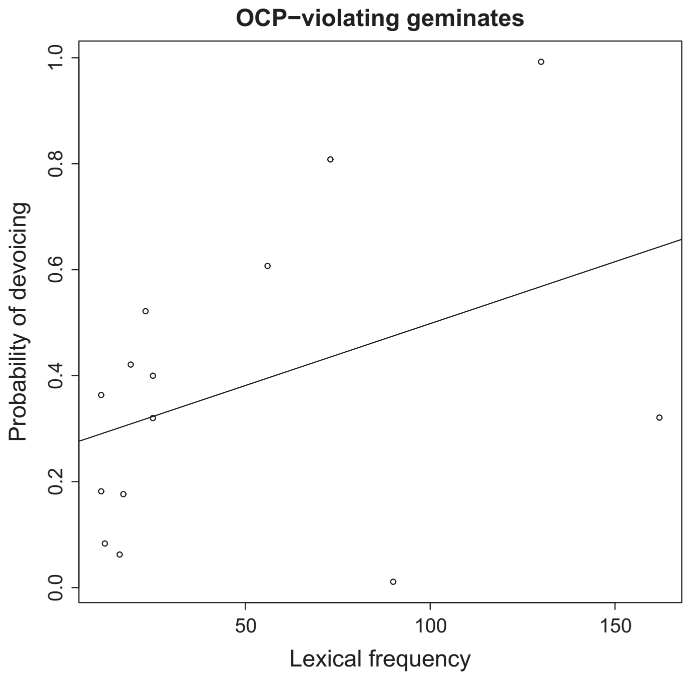


Fig. 5. The effect of lexical frequency. There is a trend toward a positive correlation relationship between lexical usage frequencies and devoicability probability.

We acknowledge, however, that the number of items under discussion is rather small, and this correlation should thus be interpreted with caution. However, we would like to point out that the correlation between lexical usage frequency and devoicing probability is compatible with Kawahara's (2011a) claim that the lexical usage frequency and devoicability stand in positive correlation, at least in naturalness judgment patterns. The result also accords well with the general, cross-linguistic observation that items with higher lexical frequency undergo phonological processes with higher probability (see Bybee, 2001, 2002; Coetzee, 2009; Coetzee and Kawahara, 2013; Goeman, 1999; Lacoste, 2008; Phillips, 2006, among others); for example, the English t/d-deletion rule is more likely to apply to more frequent items (see Coetzee and Kawahara, 2013 and references cited therein).

The claim that lexical usage frequency and probability of devoicing positively correlate can and should be validated with more solid statistical data, once Japanese speakers start using many more borrowings that contain OCP-violating geminates.

4. Conclusion

In conclusion, by studying the Corpus of Spontaneous Japanese, we have found that several linguistic factors affect the devoicing probability of voiced geminates: the OCP, the location of the trigger with respect to the geminate, the number of triggers, place of articulation, and lexical usage frequency (the last of which may be a trend). These patterns are all in accord with phonetic considerations and/or cross-linguistic tendencies.

Overall the current results show that the devoicing of geminates in Japanese loanword phonology is not a monolithic phenomenon, as assumed in the previous theoretical analyses (except for Kawahara, 2011a). Our companion paper (Sano and Kawahara, 2013) shows that external factors, such as age and gender too, affect the probabilities of devoicing. It is hoped that further theoretical analyses of the devoicing pattern will incorporate several factors, both internal and external, in their analyses.

Acknowledgements

The first author is mainly responsible for the write-up and theoretical discussion, and the second author is mainly responsible for the corpus search and data analysis. There is a companion paper to the current paper (Sano and Kawahara, 2013), which examines the effects of extra-linguistic/external factors, such as gender, speech style, educational level, and age. We are grateful to the audience at GLOW in Asia VI, ICPP 2013 and RUMMIT 2013 as well as Aaron Braver, Hiroaki Kato, Yosuke Igarashi, Jeremy Perkins, and a number of anonymous reviewers for their comments. Remaining errors are ours.

References

Bybee, Joan, 2001. Phonology and Language Use. Cambridge University Press, Cambridge.
Bybee, Joan, 2002. Word frequency and context of use in lexical diffusion of phonetically conditioned sound change. Language Variation and Change 14, 261–290.
Coetzee, Andries W., 2009. Phonological variation and lexical frequency. In: Schardl, Anisa, Walkow, Martin, Abdurrahman, Muhammad (Eds.), Proceedings of North East Linguistic Society, vol. 38. GLSA Publications, pp. 189–202.
Coetzee, Andries W., Kawahara, Shigeto, 2013. Frequency biases in phonological variation. Natural Language and Linguistic Theory 30 (1), 47–89.
Coetzee, Andries W., Pater, Joe, 2011. The place of variation in phonological theory. In: Goldsmith, John A., Riggle, Jason, Yu, Alan (Eds.), The Handbook of Phonological Theory, second ed. Blackwell-Wiley, Oxford, pp. 401–431.
Dresher, B. Elan, 2010. The Contrastive Hierarchy in Phonology. Cambridge University Press, Cambridge.
Farris-Trimble, Ashley, 2008. Cumulative faithfulness effects in phonology. Doctoral Dissertation, Indiana University.
Frisch, Stephan, 2004. Language processing and segmental OCP effects. In: Hayes, Bruce, Kirchner, Robert, Steriade, Donca (Eds.), Phonetically Based Phonology. Cambridge University Press, Cambridge, pp. 346–371.
Frisch, Stephan, Pierrehumbert, Janet, Broe, Michael, 2004. Similarity avoidance and the OCP. Natural Language and Linguistic Theory 22, 179–228.
Goeman, Ton, 1999. T-deletie in Nederlandse dialecten. Kwantitatiewe analyse van structurele, ruimtelijke en temporele variate. Doctoral Dissertation, Vrije Universiteit.
Hayes, Bruce, 1999. Phonetically-driven phonology: the role of Optimality Theory and inductive grounding. In: Darnell, Michael, Moravscik, Edith, Noonan, Michael, Newmeyer, Frederick, Wheatly, Kathleen (Eds.), Functionalism and Formalism in Linguistics, General Papers, vol. 1. John Benjamins, Amsterdam, pp. 243–285.
Hayes, Bruce, Steriade, Donca, 2004. Introduction: the phonetic bases of phonological markedness. In: Hayes, Bruce, Kirchner, Robert, Steriade, Donca (Eds.), Phonetically Based Phonology. Cambridge University Press, Cambridge, pp. 1–33.
Ihara, Mutsuko, Tamaoka, Katsuo, Murata, Tadao, 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. Dongam Publishing Co., Republic of Korea.
Irwin, Mark, 2011. Loanwords in Japanese. John Benjamins, Amsterdam, Philadelphia.
Itô, Junko, Mester, Armin, 1986. The phonology of voicing in Japanese: theoretical consequences for morphological accessibility. Linguistic Inquiry 17, 49–73.
Itô, Junko, Mester, Armin, 1995. Japanese phonology. In: Goldsmith, John (Ed.), The Handbook of Phonological Theory. Blackwell, Oxford, pp. 817–838.
Itô, Junko, Mester, Armin, 1996. Rendaku I: Constraint Conjunction and the OCP. Ms. University of California, Santa Cruz (ROA 144).
Itô, Junko, Mester, Armin, 1998. Markedness and Word Structure: OCP Effects in Japanese. Ms. University of California, Santa Cruz.
Itô, Junko, Mester, Armin, 1999. The phonological lexicon. In: Tsujimura, Natsuko (Ed.), The Handbook of Japanese Linguistics. Blackwell, Oxford, pp. 62–100.
Itô, Junko, Mester, Armin, 2003. Japanese Morphophonemics. MIT Press, Cambridge.
Jaeger, Jeri J., 1978. Speech aerodynamics and phonological universals. In: Jaeger, Jeri, Woodbury, Anthony, Ackerman, Farrel, Chiavello, Christine, Gensler, Orin, Kingston, John, Sweetner, Eve, Chompson, Henry, Whistler, Kenneth (Eds.), Proceedings of Berkeley Linguistic Society, vol. 4. Berkeley Linguistics Society, Berkeley, pp. 311–325.
Kaneko, Emiko, Iverson, Gregory, 2009. Phonetic and other factors in Japanese on-line adaptation of English final consonants. In: Inagaki, Shunji, Hirakawa, Makiko (Eds.), Studies in Language Sciences 8: Papers from the Eighth Annual Conference of the Japanese Society for Language Science. Kuroshio Publications, Tokyo, pp. 179–195.
Kawahara, Shigeto, 2001. Similarity among variants: output-variant correspondence. B.A. 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, 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).
Kawahara, Shigeto, Ono, Hajime, Sudo, Kiyoshi, 2006. Consonant co-occurrence restrictions in Yamato Japanese. In: Vance, T., Jones, K. (Eds.), Japanese/Korean Linguistics, vol. 14. CSLI, Stanford, pp. 27–38.
Kingston, John, 2007. The phonetics–phonology interface. In: de Lacy, Paul (Ed.), The Cambridge Handbook of Phonology. Cambridge University Press, Cambridge, pp. 401–434.
Kokuritsu-Kokugo-Kenkyuujo, 2008. Nihongo hanashikotoba koopasu [The Corpus of Spontaneous Japanese], Ver. 2. Database.
Kubozono, Haruo, 1999. Mora and syllable. In: Tsujimura, Natsuko (Ed.), The Handbook of Japanese Linguistics. Blackwell, Oxford, pp. 31–61.
Kubozono, Haruo, 2005. Rendaku: its domain and linguistic conditions. In: van derWeijer, Jeroen, Nanjo, Kensuke, Nishihara, Tetsuo (Eds.), Voicing in Japanese. Mouton De Gruyter, Berlin, New York, pp. 5–24.
Kubozono, Haruo, Itô, Junko, Mester, Armin, 2008. Consonant gemination in Japanese loanword phonology. Current Issues in Unity and Diversity of Languages. Collection of Papers Selected from the 18th International Congress of Linguists. The Linguistic Society of Korea. Dongam Publishing Co., Republic of Korea. pp. 953–973.

Labov, William, 1982. Building on empirical foundations. In: Lehmann, Winfred P., Malkiel, Yakov (Eds.), Perspectives on Historical Linguistics. John Benjamins, pp. 17–92.
Labrune, Laurence, 2012a. The Phonology of Japanese. Oxford University Press, Oxford.
Labrune, Laurence, 2012b. Questioning the universality of the syllable: evidence from Japanese. Phonology 29 (1), 113–152.
Lacoste, Véronique, 2008. Learning the sounds of Standard Jamaican English: variationist, phonological and pedagogical perspectives on 7-year-old children's classroom speech. Doctoral Dissertation, University of Essex.
Lindblom, Björn, Maddieson, Ian, 1988. Phonetic universals in consonant systems. In: Hyman, L., Li, C. (Eds.), Language Speech and Mind. Routledge, London, pp. 62–80.
Lyman, Benjamin S., 1894. Change from surd to sonant in Japanese compounds. Oriental Studies of the Oriental Club of Philadelphia.
Maekawa, Kikuo, 2003. Corpus of Spontaneous Japanese: its design and evaluation. In: Proceedings of ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003).
Maekawa, Kikuo, 2004. Nihongo hanashikotoba koopasu-no gaiyoo (an overview of the Corpus of Spontaneous Japanese). Nihongo Kagaku 15, 111–133.
Maekawa, Kikuo, 2010. Coarticulatory reintrepretation of allophonic variation: corpus-based analysis of /z/ in spontaneous Japanese. Journal of Phonetics 38 (3), 360–374.
Maekawa, Kikuo, Koiso, Hanae, Furui, Sadaoki, Isahara, Hitoshi, 2000. Spontaneous speech corpus of Japanese. In: Proceedings of the Second International Conference of Language Resources and Evaluation.
McCarthy, John J., 2008. Doing Optimality Theory. Blackwell-Wiley, Oxford.
Mester, Armin, Itô, Junko, 1989. Feature predictability and underspecification: palatal prosody in Japanese mimetics. Language 65, 258–293.
Nishimura, Kohei, 2003. Lyman's Law in loanwords. M.A. Thesis, Nagoya University.
Nishimura, Kohei, 2006. Lyman's Law in loanwords. Onin Kenkyuu (Phonological Studies) 9, 83–90.
Odden, David, 1994. Adjacency parameters in phonology. Language 70, 289–330.
Ohala, John J., 1983. The origin of sound patterns in vocal tract constraints. In: MacNeilage, Peter (Ed.), The Production of Speech. Springer, New York, pp. 189–216.
Ohala, John J., Riordan, C.J., 1979. Passive vocal tract enlargement during voiced stops. In: Wolf, J.J., Klatt, D.H. (Eds.), Speech Communication Papers. Acoustical Society of America, New York, pp. 89–92.
Pater, Joe, 2009. Weighted constraints in generative linguistics. Cognitive Science 33, 999–1035.
Pater, Joe, in press. Universal grammar with weighted constraints. In: McCarthy, John J., Pater, Joe (Eds.), Harmonic Grammar and Harmonic Serialism. Equinox, London.
Phillips, Betty S., 2006. Word Frequency and Lexical Diffusion. Palgrave Macmillan, New York.
Poser, William, 1990. Evidence for foot structure in Japanese. Language 66, 78–105.
Pulleyblank, Douglas, 2002. Harmony drivers: no disagreement allowed. Proceedings of Berkeley Linguistic Society, vol. 28. Berkeley Linguistics Society, Berkeley, pp. 249–267.
Rice, Keren, 2006. On the patterning of voiced stops in loanwords in Japanese. Toronto Working Papers in Linguistics 26, 11–22.
Sankoff, Gillian, Labov, William, 1979. On the use of variable rules. Language in Society 8, 189–222.
Sano, Shin-ichiro, 2013. Patterns in avoidance of marked segmental configurations in Japanese loanword phonology. In: Proceedings of GLOW in Asia 2012: Main Session, pp. 245–260.
Sano, Shin-ichiro, Hibiya, Junko, 2012. Nihongo hanasikotoba koopasu-o tukau (Usage of the Corpus of Spontaneous Japanese). In: Hibiya, Junko (Ed.), Hajimete manabu syakaigengogaku (Introduction to Sociolinguistics). Minerva Shobo, Tokyo, pp. 80–99.
Sano, Shin-ichiro, Kawahara, Shigeto, 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).
Steriade, Donca, 2004. Projecting non-lexical phonology from phonetic knowledge. Talk presented at "Symposium on Phonological Theory: Representations and Architectures" (The City University of New York, February 20th–21st, 2004).
Suzuki, Keiichiro, 1998. A typological investigation of dissimilation. Doctoral Dissertation, University of Arizona.
Tanaka, Shin-Ichi, 2007. On the nature of typology of dissimilation. English Linguistics 24 (1), 279–313.
Tesar, Bruce, 2007. A comparison of lexicographic and linear numeric optimization using violation difference ratios. Ms. Rutgers University.
Vance, Timothy, in press. Rendaku. In: Kubozono, Haruo (Ed.), The Handbook of Japanese Language and Linguistics: Phonetics and Phonology. Mouton De Gruyter.
Vance, Timothy J., 1980. The psychological status of a constraint on Japanese consonant alternation. Linguistics 18, 245–267.
Vance, Timothy J., 1987. An Introduction to Japanese Phonology. SUNY Press, New York.
Vance, Timothy J., 2007. Have we learned anything about *rendaku* that Lyman didn't already know? In: Frellesvig, Bjarke, Shibatani, Masayoshi, Carles Smith, John (Eds.), Current Issues in the History and Structure of Japanese. Kuroshio, Tokyo, pp. 153–170.
Vance, Timothy J., 2008. The Sounds of Japanese. Cambridge University Press, Cambridge.
Walker, Rachel, 2001. Round licensing, harmony, and bisyllabic triggers in Altaic. Natural Language and Linguistic Theory 19, 827–878.
Weiner, E. Judith, Labov, William, 1983. Constraints on the agentless passive. Journal of Linguistics 19, 29–58.
Westbury, John R., 1979. Aspects of the temporal control of voicing in consonant clusters in English. Doctoral Dissertation, University of Texas, Austin.
Westbury, John R., Keating, Patricia, 1986. On the naturalness of stop consonant voicing. Journal of Linguistics 22, 145–166.
Zhang, Xi, 1996. Vowel systems of the Manchu-Tungus languages of China. Doctoral Dissertation, University of Toronto.

Sano, Shin-ichiroo and Shigeto Kawahara (2013) A corpus-based study of geminate devoicing in Japanese: The role of the OCP and external factors. Journal of Linguistic Society of Japan (Gengo Kenkyuu) 144: 103-118.

言語研究（*Gengo Kenkyu*）**144**: 103–118（2013）

【Forum】

A Corpus-Based Study of Geminate Devoicing in Japanese: The Role of the OCP and External Factors

Shin-ichiro Sano — *Okayama Prefectural University*
Shigeto Kawahara — *Keio University*

Abstract: Nishimura (2003, 2006) pointed out that in Japanese loanwords, voiced obstruent geminates can optionally devoice when they co-occur with another voiced obstruent (e.g. /doggu/ → [dokku] 'dog'). This devoicing pattern has been analyzed within a number of theoretical frameworks, and has thereby contributed to address several theoretical issues. The pattern, moreover, has been studied in several experimental, judgment-based studies. However, there are only a few studies on actual production data. Furthermore, all of the previous studies have generally assumed that the devoicing pattern under question is a sociolinguistically monolithic phenomenon. This paper addresses these two issues. By studying the Corpus of Spontaneous Japanese (Kokuritsu-Kokugo-Kenkyuujo 2008), we first confirm the previous claim that the OCP makes devoicing of geminates more likely in actual production data. Moreover, the results also reveal that many external, sociolinguistic factors affect the applicability of devoicing. The overall results thus contribute to the deeper understanding of the phenomenon by revealing various hitherto unnoticed factors that affect the applicability of devoicing.*

Key words: geminate, devoicing, the OCP, the CSJ, sociolinguistic factors

1. Introduction

Nishimura (2003, 2006) points out that in Japanese loanwords, voiced obstruent geminates can optionally devoice when they co-occur with another voiced obstruent, as exemplified by the data in (1). He further points out that this devoicing is due to a restriction against two voiced obstruents, a restriction which can be formalized as the OCP (voice) (Itô and Mester 1986, 2003), also known as Lyman's Law (Lyman 1894; Vance 2007). In other words, geminates do not devoice unless

* We would like to thank three anonymous reviewers for their comments on previous versions of this paper. Portions of this paper were presented at Keio University in February 2013 and the University of Massachusetts, Amherst in April 2013, and we received helpful suggestions from the audiences at these occasions. Finally, thanks are due to Nat Dresher, Chris Kish, Jess Trombetta, and Donna Erickson for proofreading the paper. The usual disclaimer applies. This work is partly supported by JSPS KAKENHI Grant #25770157, and #25280482.

they occur with another voiced obstruent; i.e. unless they violate the OCP, as shown in (2). Moreover, Nishimura (2003, 2006) points out that devoicing is also impossible in OCP-violating singletons, as in (3).

(1) OCP-violating geminates can optionally devoice.
 a. be**dd**o → be**tt**o 'bed'
 b. ba**gg**u → ba**kk**u 'bag'
 c. bi**gg**u → bi**kk**u 'big'
(2) Non-OCP-violating geminates do not devoice.
 a. suno**bb**u → *suno**pp**u 'snob'
 b. he**dd**o → *he**tt**o 'head'
 c. re**gg**u → *re**kk**u 'leg'
(3) Singletons do not devoice, even when they violate the OCP.
 a. gi**b**u → *gi**p**u 'give'
 b. da**g**u → *da**k**u 'Doug'
 c. ba**g**u → *ba**k**u 'bug'

This paradigm, illustrated in (1)–(3), has been analyzed using various theoretical mechanisms: e.g. local conjunction (McCarthy 2008: 219–220; Nishimura 2003, 2006); phonetically-based phonology (Kawahara 2006, 2008; Steriade 2004); the theory of markedness and contrast (Rice 2006); Harmonic Phonology (Farris-Trimble 2008: 22–28; Pater 2009, to appear); Noisy Harmonic Phonology (Coetzee and Kawahara 2013; Coetzee and Pater 2011); and Maximum Entropy Grammar (Coetzee and Pater 2011).[1] This devoicing phenomenon is also discussed to address the issue of how loanword phonology and native phonology are related to one another (Crawford 2009; Itô and Mester 2008; Tateishi 2002). In summary, this pattern has contributed to many recent debates in phonological theory (see Kawahara 2011a and Kawahara 2012a for more extensive reviews, the former in English and the latter in Japanese).

Moreover, the Japanese loanword devoicing pattern has been studied in a number of naturalness judgment experiments (Kawahara 2011a, 2011b, 2012b, 2013b), partly because the pattern has played a non-trivial role in the recent phonological literature, and therefore its empirical foundations needed to be confirmed. In these studies, naive native speakers of Japanese judged the naturalness of devoicing in various contexts, including (1)–(3). The results generally corroborate the intuition-based data in (1)–(3) in that native Japanese speakers find the devoicing of OCP-violating geminates most natural. However, all of these studies also found that devoicing of non-OCP-violating geminates, as in (2), is judged to be not completely unnatural, despite the intuition by Nishimura (2003, 2006) and Kawahara (2006) to the contrary.

[1] Space limitation does not permit us to go into the details of these theoretical analyses. Kawahara (2012a) and Kawahara (2013a) present a summary of most of these analyses (the former in Japanese and the latter in English).

As much as this pattern has been studied from a variety of perspectives, both theoretical and experimental, there is one aspect of this phenomenon that needs to be studied more extensively. That is to study the data of actual production patterns: most of the theoretical analyses are based on the intuitions of Nishimura (2003, 2006) and Kawahara (2006). Although some experimental work more or less confirmed the intuitions (Kawahara 2011a, 2011b, 2012b, 2013b), they do not nevertheless answer the question of whether devoicing happens in actual production or not, and if so, how. Nishimura (2003, 2006), in the appendices to his papers, reports some analyses of the devoicing patterns using the Corpus of Spontaneous Japanese (the CSJ, see below), and shows the effect of the OCP on devoicing. The first aim of this study is to replicate this result with an updated version of the CSJ.[2]

Another question arises from our interests in the sociolinguistic aspects of devoicing pattern. The theoretical and experimental studies reviewed above assume that devoicing is sociolinguistically monolithic, abstracting away from external (extra-linguistic, sociolinguistic) factors.[3] A corpus-based study may show that this assumption may be too simplistic. We address these questions by using the Corpus of Spontaneous Japanese (Kokuritsu-Kokugo-Kenkyuujo 2008; Maekawa et al. 2000; Maekawa 2003, 2004; Sano and Hibiya 2012).[4]

To summarize the current questions:

(4) a. Do geminates devoice when they violate the OCP in the actual production patterns?
b. Can geminates ever devoice without violating the OCP?
c. The previous work has assumed that this devoicing pattern is sociolinguistically monolithic. Is this assumption true?

This study focused on how voiced geminates—but not voiced singletons—behave, because previous judgment studies (Kawahara 2011a, 2011b, 2012b, 2013b) show that devoicing is judged to be more likely for geminates than for singletons, and that devoicing of singletons is judged to be very unnatural. We therefore leave the study of singletons as a topic for future study.

[2] Two notes are in order. First, Nishimura (2003) used the monitor (trial) version, and Nishimura (2006) used the first version (p.c. Kohei Nishimura, Aug. 2012).

Second, we have also investigated the effects of some linguistic factors that are related to the OCP, but we report those results in a companion paper (Kawahara and Sano 2013), and the current paper focuses on external factors. We report linguistic and external factors in two separate papers for the sake of exposition, but in doing so, we also follow the claim that linguistic factors and external factors do not interact (Labov 1982; Sankoff and Labov 1979; Weiner and Labov 1983).

[3] One exception is the effect of lexical usage frequencies on the naturalness judgment of devoicing, investigated by Kawahara (2011a) and modeled by Coetzee and Kawahara (2013). The effect of lexical usage frequencies on actual production patterns in the CSJ is reported in Kawahara and Sano (2013).

[4] For other studies using the CSJ in a similar spirit, see for example, Sano (2008, 2011, 2012). See also http://www.ninjal.ac.jp/csj/.

To provide a brief preview of the results, we first confirm that devoicing is more likely when geminates violate the OCP, replicating Nishimura (2003, 2006); however, devoicing does nevertheless occur even when geminates do not violate the OCP, albeit only infrequently. We also find that some external factors affect the applicability of geminate devoicing in non-trivial ways. Moreover, the ways in which external factors affect the likelihood of devoicing are compatible with the previous observations on patterns of sociolinguistic variation and change.

The rest of this paper is structured as follows: Section 2 presents a detailed description of our data collection method. Section 3 looks at the effect of the OCP and other sociolinguistic factors on devoicing. Section 4 presents a multiple logistic regression analysis with all the factors included in one model. The final section offers brief concluding remarks.

2. Method

To investigate whether and how the OCP and external factors affect the devoicing of geminates in Japanese loanwords in actual production, we conducted an exhaustive search of the Corpus of Spontaneous Japanese, version 2 (the CSJ) (Kokuritsu-Kokugo-Kenkyuujo 2008; Maekawa 2003, 2004; Maekawa et al. 2000). This large-scale corpus is based on 662 hours of speech with 7.5 million words, produced by a total of 1,417 speakers. In addition to its large size, rich annotation system, and ease of searchability, another important characteristic of this database, which is particularly relevant for the purpose of our study, is that it provides both underlying forms and surface forms in terms of hatsuonkei ("pronounced forms"). This system allows us to retrieve a set of words with particular phonological characteristics based on underlying forms and study how the words are actually pronounced using the phonetic transcription provided.

We first extracted words with underlying voiced geminate obstruents from the CSJ (N=1,666), and then excluded tokens in which the voiced geminates underwent some changes other than devoicing, such as degemination and complete deletion, since our focus was on devoicing. This elimination process resulted in 1,617 data points (i.e. 97% of the data remained; see Sano 2013 for details). Among those, 464 tokens showed devoicing (28.7%). We then tested how the OCP and various sociolinguistic factors affect the probability of devoicing. We extracted all the tokens that fit each condition (=n), and counted how many of them are devoiced (=m). We then calculated the percentages of devoicing over n (i.e. 100 * (m/n)). The OCP is defined as containing another voiced obstruent within six preceding or following moras;[5] we deployed sociolinguistic factors that are encoded in the CSJ.

[5] It turned out that most of the OCP-induced devoicing was caused by a trigger in adjacent moras, but we included non-local triggers as well in this study. In Kawahara and Sano (2013), we study in detail how linguistic factors affect the applicability of devoicing. Those linguistic factors include the locality between the trigger and the geminate, and the effect of the number of triggers, among others. Our current study focuses on external factors (and the general role of the OCP).

3. Results and Discussion

3.1. Devoicing and the OCP

Figure 1 illustrates the likelihood of devoicing of voiced geminates for the OCP-violating condition (e.g. [beddo]) and the non-OCP-violating condition (e.g. [heddo]). We observe that devoicing happens about 40% of the time in the OCP-violating condition (=438/1099), while devoicing happens only about 5% of the time in the non-OCP violating condition (=26/518) ($\chi^2(1) = 207.1, p < .001$).

This production-based data supports the intuitions of Nishimura (2003, 2006) and Kawahara (2006) that the OCP is a crucial factor in inducing devoicing (Kawahara and Sano 2013). The current results replicate the corpus studies reported in the appendices of Nishimura (2003, 2006) which used older versions of the CSJ; the current results also support the experimental results that the OCP makes devoicing of geminates more natural (Kawahara 2011a, 2011b, 2012b, 2013b).

It is not the case, however, that devoicing is completely impossible for non-OCP-violating geminates, as it does happen about 5% of the time. The fact that context-free devoicing of geminates is not impossible (though unlikely) may be the basis of the judgment patterns in Kawahara (2011a, 2011b, 2012b, 2013b).

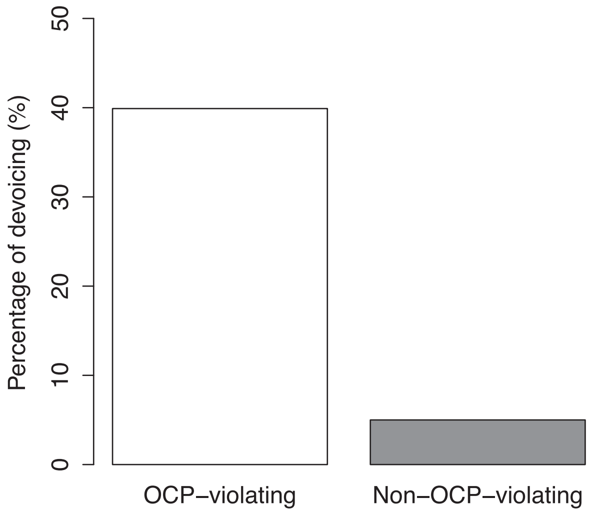


Figure 1. The effect of OCP on devoicing

3.2. Age

Now we move on to the effect of sociolinguistic factors. The following discussion treats the faithful rendition (voiced version) as "norm" and "standard", and the devoiced rendition as "innovative" and "vernacular", as the faithful forms are older forms, reflecting the pronunciations of the donor languages more accurately. For example, a standard dictionary like Shimmura (2008) lists [doggu], not [dokku], as its lexical entry for the word 'dog'.[6]

[6] As an anonymous reviewer points out, for some forms at least, the devoiced renditions

Figure 2 shows the correlation between the speakers' birth years (categorized with 10 year increments)[7] and the likelihood of devoicing, together with Table 1, which provides the actual numbers of occurrences of devoicing. We observe a positive correlation between these two parameters in that younger speakers tend to devoice more often. Although the correlation did not reach statistical significance due to the small N ($N = 6$; $p = .18$), the non-parametric correlation measure ρ is reasonably high (=.66).[8] This observation is compatible with the general sociolinguistic observation that younger speakers tend to prefer innovative, vernacular pronunciations (Chambers 2002; Labov 1966, 1972, 1994, 2001b; Romaine 1984).

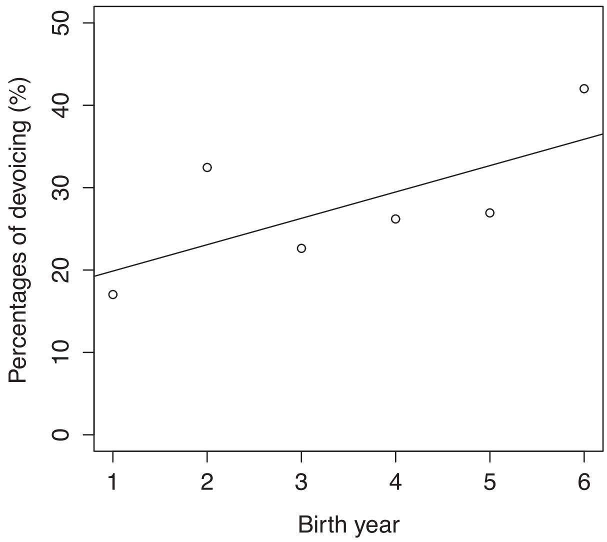


Figure 2. The effect of birth years by 10 year increments (Category 1=1925–1934; 2=1935–44; 3=1945–54; 4=1955–64; 5=1965–74; 6=1975–84)

Table 1. The effect of birth years by 10 year increments (actual numbers)

	1	2	3	4	5	6
n	8/47	37/114	43/190	115/439	153/568	108/257
%	17%	32.5%	22.6%	26.2%	26.9%	42%

may be the "standard", but we cannot think of a good way to objectively determine which lexical items take the voiced rendition as the standard form and which forms take the voiceless rendition as the standard form. Our method at least provides an objective way to define what is standard. This assumption should of course be examined more carefully in future studies.

[7] The CSJ provides birth-years in 5-year increments. Since there were some gaps when we analyzed the data with 5-year increments, we used 10-year increments.

[8] Since the birth-year categories are an ordered pseudo-numerical variable, and since we are interested in the linear correlation between birth years and devoicing percentages, we deployed a non-parametric correlation analysis.

3.3. Gender

Figure 3 shows the effect of gender on devoicing, which shows that female speakers show devoicing more often (207/503=41.2%) than male speakers (257/1114=23.1%) ($\chi^2(1) = 54.5$, $p < .001$). This observation makes sense from a diachronic point of view—female speakers are known to initiate sound changes (Eckert 1989; Labov 1966, 1990; Trudgill 1972; Romaine 2003). To the extent that a variable phonological pattern can be considered as an on-going diachronic change (Weinreich et al. 1968), the results are thus compatible with the previous observation in the sociolinguistic literature.[9]

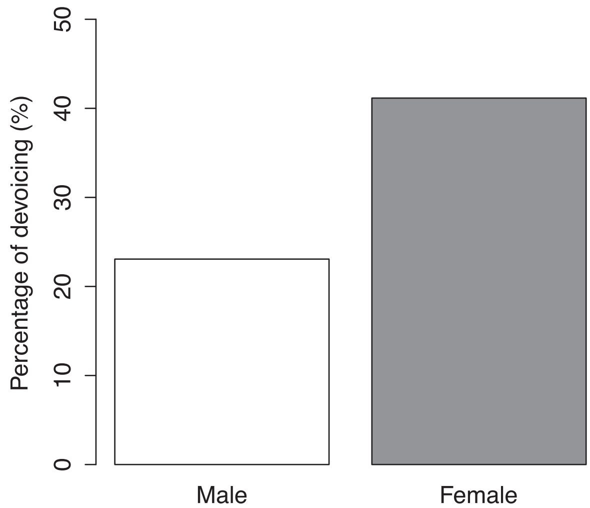


Figure 3. The effect of gender

3.4. Speech style: APS vs. SPS

Figure 4 shows the effect of speech style on the probability of devoicing. APS (for "Academic Presentation Speech") is live recording of academic presentations in various academic societies, whereas SPS (for "Simulated Public Speaking") consists of general speeches by laypeople on everyday topics. APS is characterized by a formal speaking style, whereas SPS is characterized by a casual and informal style.[10]

[9] An anonymous reviewer pointed out that female speakers may prefer to use prestigious forms (Trudgill 1972), and to the extent that voiced forms are more prestigious (because they are "standard"), the pattern in Figure 3 could be a contradiction to this claim. However, as Labov (1990: 215) argues, "[i]n change from below, women are most often the innovators." The devoicing case at hand is "in change from below" because it is a systematic change, which is more often observed in casual speech style (Section 3.4). We therefore believe that this gender effect is not an anomaly, if we follow Labov's observation about the effect of gender on sound changes.

[10] We have also checked the effect of the 5-point scale rating on speech formality provided by the CSJ. The analysis revealed a similar effect. Since the results are similar, we only report the APS/SPS distinction.

The result of this analysis shows that devoicing is more likely in SPS (297/714=41.6%) than in APS (155/872=17.8%) ($\chi^2(1) = 108.2$, $p < .001$), and this trend makes sense given the previous observation that speakers tend to use standard forms more often in formal speech styles than in casual speech styles (Labov 1963, 1966, 2001a).

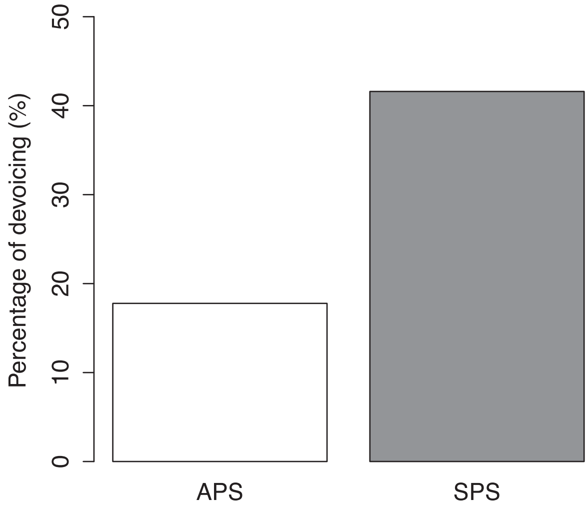


Figure 4. The effect of speech style (APS=Academic Presentation Speech; SPS= Simulated Public Speaking)

3.5. Self-confidence in public speech

Figure 5 shows the effect of self-confidence about public speaking, in which speakers who consider themselves as "not confident" show a high probability of devoicing. There is a clear division between the leftmost category and the other three categories in Figure 5—the actual numbers are provided in Table 2 ($\chi^2(4) = 42.7$, $p < .001$). To assess this observation statistically, post-hoc tests with a Bonferroni correction ($\alpha = .05/6 = .008$) were run, which show that the first condition is statistically different from the other three (1st vs. 2nd: $\chi^2(1) = 34.6$, $p < .001$; 1st vs. 3rd: $\chi^2(1) = 32.9$, $p < .001$; 1st vs. 4th: $\chi^2(1) = 10.0$, $p < .008$), but that no differences among the last three conditions are significant (2nd vs. 3rd: $\chi^2(1) = 0.1$, *n.s.*; 2nd vs. 4th: $\chi^2(1) = 0.1$, *n.s.*; 3rd vs. 4th: $\chi^2(1) = 0.2$, *n.s.*).[11]

[11] These post-hoc statistical analyses on all the possible comparisons were prompted by an anonymous reviewer.

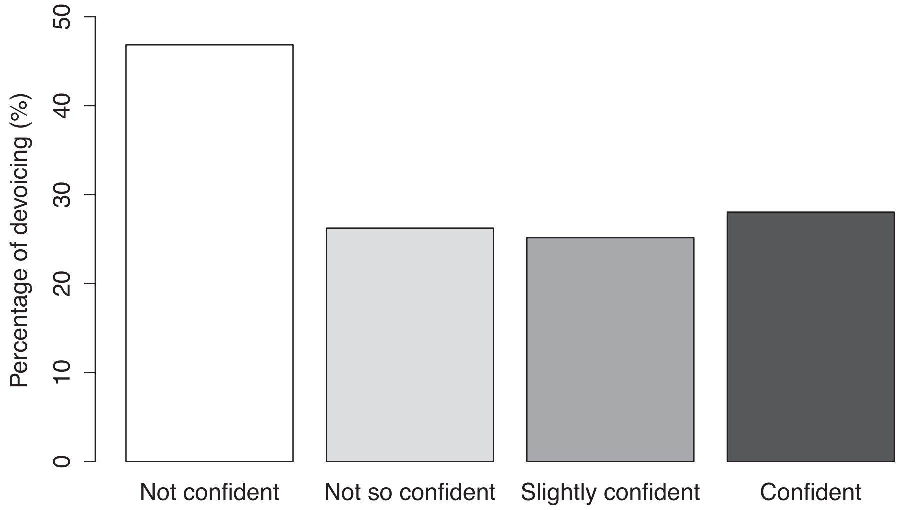


Figure 5. The effect of self-confidence about public speaking

Table 2. The effect of confidence level on devoicability (actual numbers)

	Not confident	Not so confident	Slightly confident	Confident
n	111/237	196/747	119/473	30/107
%	46.8%	26.2%	25.2%	28%

3.6. Educational background

Figure 6 illustrates the effect of educational background of the speakers.[12] The figure shows that the higher the educational background of the speakers, the less likely that they show devoicing (no higher education: 99/209=47.4%; undergraduate: 237/697=34%; graduate: 128/709=18.1%) ($\chi^2(2) = 84.4$, $p < .001$). Post-hoc multiple comparisons with Bonferroni correction ($\alpha = .05/3 = .013$) show that all the differences are significant (no higher education vs. undergraduate: $\chi^2(1) = 11.7$, $p < .001$; no higher education vs. graduate: $\chi^2(1) = 73.0$, $p < .001$; undergraduate vs. graduate: $\chi^2(1) = 45.7$, $p < .001$).

This correlation holds most likely because people with higher education are more likely to know that geminates are voiced in the donor languages, and tend to prefer to use forms that are faithful to the source language. The correlation is also compatible with the sociolinguistic observation that people with higher social class are more likely to use standard forms, while people in lower social classes are more likely to use vernacular forms (Hibiya 1995; Labov 1963, 1972, 1994, 2001b; Trudgill 1974).

12 In other sociolinguistic studies, social class is more often used as a predictor variable (Labov 1966, 2001b). We use educational background as a replacement, as social class is not a standard classification in the Japanese society and hence is not encoded in the CSJ.

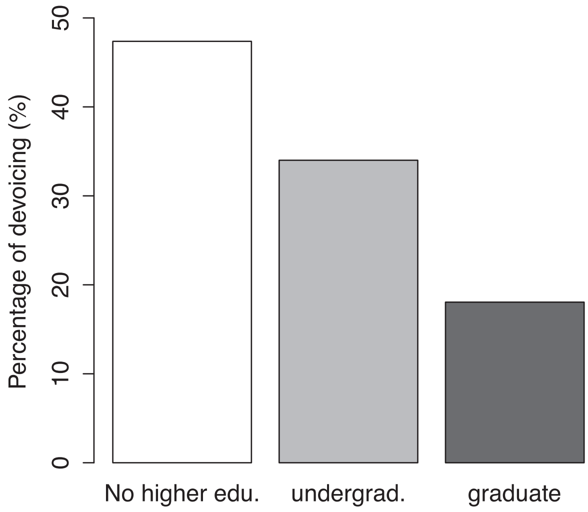


Figure 6. The effect of educational background

3.7. Previous experiences in public speech

Finally, Figure 7, together with Table 3, shows the effect of previous public speaking experiences on devoicing.[13] We observe a general trend in which the more experiences the speakers have, the less likely that devoicing occurs—there is a general decline among the first three conditions, and especially the difference between the second condition and the third condition is apparent. To assess this observation, multiple-comparisons with a Bonferroni correction (α = .05/10 = .005) were run, which show that the differences between the first and the other conditions are significant or marginal (1st vs. 2nd: $\chi^2(1) = 3.3$, $p = .067$; 1st vs. 3rd: $\chi^2(1) = 52.7$, $p < .001$; 1st and 4th: $\chi^2(1) = 28.6$, $p < .001$; 1st vs. 5th: $\chi^2(1) = 70.1$, $p < .001$), that the differences between second and the last three conditions are significant (2nd vs. 3rd: $\chi^2(1) = 22.8$, $p < .001$; 2nd vs. 4th: $\chi^2(1) = 13.1$, $p < .001$; 2nd vs. 5th: $\chi^2(1) = 28.1$, $p < .001$), but that the other differences are minimal (3rd vs. 4th: $\chi^2(1) = 0.00$, *n.s.*; 3rd vs. 5th: $\chi^2(1) = 0.00$, *n.s.*; 4th vs. 5th: χ^2 (1) = 0.00, *n.s.*).

This effect of previous experiences makes sense from the previous observations in Figure 6. Those without much higher education are less likely to have previous public speaking experiences.[14] Therefore, they are more likely to be unaware of—or do not pay careful attention to—the original pronunciations of voiced geminates.

[13] We deployed the categorization coding for the number of public speaking experiences from the CSJ.

[14] There is, in fact, a fairly high correlation between those two factors in our data ($r = 0.57$). See Section 4 for a multiple logistic regression analysis, which shows that even when these two factors are encoded in the same model, they both have a significant impact on devoicing.

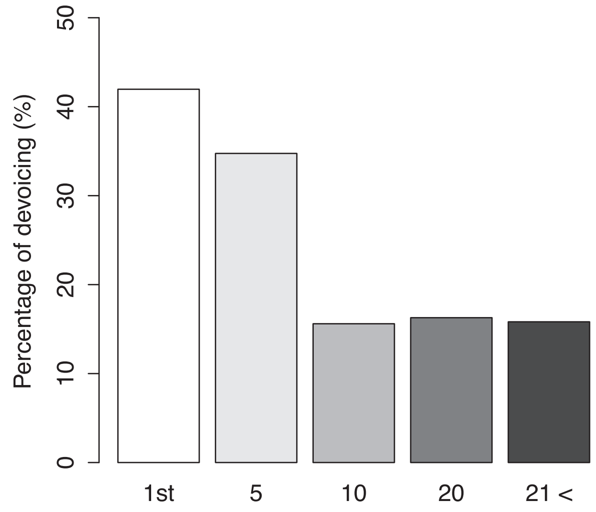


Figure 7. The effect of previous public speaking experiences

Table 3. The effect of previous public speaking experiences (actual numbers)

	1st	5	10	20	21<
n	240/572	82/236	39/250	21/129	59/373
%	42%	34.7%	15.6%	16.3%	15.8%

The last two observations thus converge on one conclusion: geminate devoicing may be to some degree under a conscious control—those who are likely to know the pronunciations of the donor language tend to keep the voicing value of the donor language. As discussed in the previous section, those people may prefer to use forms that reflect the original pronunciations more accurately.

3.8. Other external factors

Other external factors that we investigated, which showed no correlations with the probability of devoicing, include the following: speed (speech rate), spontaneity of the speech, and articulatory clarity.

4. A Logistic Regression Analysis

Although we have seen that several external factors affect the devoicing likelihood of geminates, one may be concerned that some factors are correlated with others, and effects of some factors arise from that correlation. To address this issue, a multiple logistic regression analysis was run with the following model.[15] The

[15] A logistic regression analysis was run because the dependent variable is a binary opposition. We did not encode interaction terms, because encoding interaction terms among all seven factors would make the interpretation of the results extremely hard. We did not use VARBRUL method (Cedergren and Sankoff 1974), because logistic regression is more widely used methodology.

dependent variable was whether a particular token was devoiced or not. The independent variables were the OCP, age, gender, style, self-confidence, educational background, and previous experiences. The self-confidence and previous experiences were recoded because, as we have seen before, their effects were non-linear. Self-confidence was recoded as a binary opposition between "not confident" vs. "everything else", and previous experiences were recoded as a ternary opposition between "1", "5" and "more". The analysis was run by R (R Development Core Team 1993–2013), using the **glm** function.

The result of the logistic regression analysis is shown in Table 4. Almost all the factors still have a significant effect in this multiple regression model, even when all the factors are tested at once. A few exceptions are that educational background is only marginally significant, and that style did not have any significant effect.

Table 4. The result of the logistic regression

	Estimate	StEr	z	p
OCP	2.48477	0.22106	11.240	$< .001$
Age	0.21390	0.05509	3.883	$< .001$
Gender	–0.32577	0.15187	–2.145	$< .05$
Style	–0.22752	0.23053	–0.987	*n.s.*
Self-confidence	–0.62682	0.17163	–3.652	$< .001$
Educational background	–0.27088	0.14070	–1.925	$= 0.054$
Previous experience	–0.31134	0.09724	–3.202	$< .01$

To understand why style did not have a statistical impact in this model, we calculated the correlation matrix of the variables used in this model. The result shows that style had a high correlation with educational background ($r = 0.76$) and previous speech experience ($r = 0.68$). It is likely that those who provide APS (Academic Presentation Speech) have high academic background and many previous speech experiences. The lack of significant effect of style in this regression model may be because the variability is subsumed by these two factors.

Except for this factor, however, the impacts of other extralinguistic factors are reliable in the multiple regression model (the effect of educational background was only marginally significant, whose variability may be partially subsumed by previous experiences and other factors). The analysis shows that even when all these factors are encoded in the same model, most of them each have a reliably significant impact on devoicability.

5. General Discussion

This paper investigated the patterns of geminate devoicing in Japanese loanwords using the large-scale corpus, the Corpus of Spontaneous Japanese. A thorough search of this corpus shows (i) that the OCP plays a crucial role in inducing geminate devoicing, (ii) that devoicing does occur even without the presence of another voiced obstruent, although it is much less likely, and (iii), most importantly for the current purpose, that various external, sociolinguistic factors affect the likelihood of devoicing. The ways in which external factors affect devoicing generally make sense given the findings of the previous sociolinguistic literature: the variation in the devoicing patterns follows the general patterns of language variation and change.

Moreover, we observe some evidence that geminate devoicing may be consciously controllable in the sense that those who are likely to have knowledge of the donor languages may attempt to refrain from devoicing. The overall patterns therefore show that grammar-driven (or markedness-driven) devoicing can be suppressed by conscious control, as reflected in variation patterns affected by various sociolinguistic factors.

In conclusion, our corpus study has confirmed some of the previous observations about geminate devoicing in Japanese loanwords (Nishimura 2003, 2006 *et seq.*), but also has found that many external factors affect the applicability of devoicing. Our results thus contribute to the deeper understanding of the phenomenon by revealing various hitherto unnoticed factors that affect the likelihood of devoicing. We hasten to add, however, that our aim in conducting this project is in no way in conflict with the previous studies of the devoicing phenomenon, which did not consider external, sociolinguistic factors. Instead, it is hoped that further research will seek for a deeper understanding of the phenomenon by considering—and modeling—how both linguistic and external factors shape devoicing patterns.

References

Cedergren, Henrietta J. and David Sankoff (1974) Variable rules: Performance as a statistical reflection of competence. *Language* 50: 333–355.

Chambers, J.K. (2002) Patterns of variation including change. In: J.K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.) *The handbook of language variation and change*, 349–372. Oxford: Blackwell.

Coetzee, Andries W. and Shigeto Kawahara (2013) Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 30: 47–89.

Coetzee, Andries W. and Joe Pater (2011) The place of variation in phonological theory. In: John A. Goldsmith, Jason Riggle and Alan Yu (eds.) *The handbook of phonological theory, second edition*, 401–431. Oxford: Blackwell-Wiley.

Crawford, J. Clifford (2009) Adaptation and transmission in Japanese loanword phonology. Doctoral dissertation, Cornell University.

Eckert, Penelope (1989) The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1: 245–267.

Farris-Trimble, Ashley (2008) Cumulative faithfulness effects in phonology. Doctoral dissertation, Indiana University.

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

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

Itô, Junko and Armin Mester (2003) *Japanese morphophonemics*. Cambridge: MIT Press.

Itô, Junko and Armin Mester (2008) Lexical classes in phonology. In: Shigeru Miyagawa and Mamoru Saito (eds.) *The Oxford handbook of Japanese linguistics*, 84–106. Oxford: Oxford Univeristy Press.

Kawahara, Shigeto (2006) A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82: 536–574.

Kawahara, Shigeto (2008) Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* 17: 317–330.

Kawahara, Shigeto (2011a) Aspects of Japanese loanword devoicing. *Journal of East Asian Linguistics* 20: 169–194.

Kawahara, Shigeto (2011b) Japanese loanword devoicing revisited: A rating study. *Natural Language and Linguistic Theory* 29: 705–723.

Kawahara, Shigeto (2012a) Gairaigo yuusei sokuon no museika: Rironteki kooken. [Japanese loanword devoicing: A review]. In: Phonological Society of Japan (ed.) *On'in kenkyu* [Phonological studies] 15, 93–104. Tokyo: Kaitakusha.

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

Kawahara, Shigeto (2013a) Developments in laboratory phonology: The case of Japanese devoicing. Talk presented at Keio University, February 27th, 2013.

Kawahara, Shigeto (2013b) Testing Japanese loanword devoicing: Addressing task effects. *Linguistics* 51: 1271–1299.

Kawahara, Shigeto and Shin-ichiro Sano (2013) A corpus-based study of geminate devoicing in Japanese: Linguistic factors. *Language Sciences* 40: 300–307.

Kokuritsu-Kokugo-Kenkyuujo (2008) Nihongo hanashikotoba koopasu. [The Corpus of Spontaneous Japanese]. Ver. 2. Database.

Labov, William (1963) The social motivation of a sound change. *Word* 19: 273–309.

Labov, William (1966) *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, William (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William (1982) Building on empirical foundations. In: Winfred P. Lehmann and Yakov Malkiel (eds.) *Perspectives on historical linguistics*, 17–92. Amsterdam: John Benjamins.

Labov, William (1990) The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2: 205–254.

Labov, William (1994) *Principles of linguistic change, volume 1: Internal factors*. Oxford: Blackwell.

Labov, William (2001a) The anatomy of style-shifting. In: Penelope Eckert and John R. Rickford (eds.) *Style and sociolinguistic variation*, 85–108. Cambridge: Cambridge University Press.

Labov, William (2001b) *Principles of linguistic change, volume 2: Social factors*. Oxford: Blackwell Publishers.

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

Maekawa, Kikuo (2003) Corpus of Spontaneous Japanese: Its design and evaluation. *Proceedings of ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003)*, 7–12.

Maekawa, Kikuo (2004) Nihongo hanashikotoba koopasu-no gaiyoo. [An overview of the Corpus of Spontaneous Japanese]. *Nihongo Kagaku* 15: 111–133.

Maekawa, Kikuo, Hanae Koiso, Sadaoki Furui and Hitoshi Isahara (2000) Spontaneous

Speech Corpus of Japanese. *Proceedings of the Second International Conference of Language Resources and Evaluation*, 947–952.

McCarthy, John J. (2008) *Doing Optimality Theory*. Oxford: Blackwell-Wiley.

Nishimura, Kohei (2003) Lyman's Law in loanwords. MA thesis, Nagoya University.

Nishimura, Kohei (2006) Lyman's Law in loanwords. In: Phonological Society of Japan (ed.) *On'in kenkyu* [Phonological studies] 9, 83–90. Tokyo: Kaitakusha.

Pater, Joe (2009) Weighted constraints in generative linguistics. *Cognitive Science* 33: 999–1035.

Pater, Joe (to appear) Universal grammar with weighted constraints. In: John J. McCarthy and Joe Pater (eds.) *Harmonic Grammar and Harmonic Serialism*. London: Equinox.

R Development Core Team (1993–2013) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Software, available at http://www.Rproject. org.

Rice, Keren (2006) On the patterning of voiced stops in loanwords in Japanese. *Toronto Working Papers in Linguistics* 26: 11–22.

Romaine, Suzanne (1984) *The language of child and adolescents*. Basil: Blackwell.

Romaine, Suzanne (2003) Variation in language and gender. In: Janet Holmes and Miriam Meyerhoff (eds.) *The handbook of language and gender*, 98–118. Oxford: Blackwell.

Sankoff, Gillian and William Labov (1979) On the use of variable rules. *Language in Society* 8: 189–222.

Sano, Shin-ichiro (2008) Nihongo hanashikotoba koopasu-ni arawareru sairekotoba-ni kansuru suuryooteki bunseki [A quantitative analysis of sa-insertion in Corpus of Spontaneous Japanese]. *Gengo Kenkyu* [Journal of the Linguistic Society of Japan] 133: 77–106.

Sano, Shin-ichiro (2011) Real-time demonstration of the interaction among internal and external factors in language change: A corpus study. *Gengo Kenkyu* [Journal of the Linguistic Society of Japan] 139: 1–27.

Sano, Shin-ichiro (2012) Predicting ongoing global paradigm optimization in voice in Japanese: Corpus-based OT analysis. In: Jaehoon Choi, Alan Houge, Jessamyn Schertz, Jeff Punske, Deniz Tat and Alex Trueman (eds.) *Proceedings of the 29th West Coast Conference on Formal Linguistics*, 398–406. Somerville: Cascadilla Press.

Sano, Shin-ichiro (2013) Patterns in avoidance of marked segmental configurations in Japanese loanword phonology. *Proceedings of GLOW in Asia 2012: Main session*, 245–260.

Sano, Shin-ichiro and Junko Hibiya (2012) Nihongo hanashikotoba koopasu-o tsukau. [Usage of the Corpus of Spontaneous Japanese]. In: Junko Hibiya (ed.) *Hajimete manabu shakaigengogaku* [Introduction to sociolinguistics], 80–99. Tokyo: Minerva Shobo.

Shimmura, Izuru (ed.) (2008) *Koojien.* Sixth edition. Tokyo: Iwanami Shoten.

Steriade, Donca (2004) Projecting non-lexical phonology from phonetic knowledge. Talk presented at "Symposium on Phonological Theory: Representations and Architectures" (The City University of New York, February 20th-21st, 2004).

Tateishi, Koichi (2002) Bunpoo-no ichibutoshite-no goisoo-no zehi. [Lexical strata as a part of grammar]. *Onsei Kenkyuu* [Journal of the Phonetic Society of Japan] 6(1): 34–43.

Trudgill, Pater (1972) Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1: 179–195.

Trudgill, Pater (1974) *The social differentiation of English in Norwich.* Cambridge: Cambridge University Press.

Vance, Timothy J. (2007) Have we learned anything about *rendaku* that Lyman didn't already know? In: Bjarke Frellesvig, Masayoshi Shibatani and John Carles Smith (eds.) *Current issues in the history and structure of Japanese*, 153–170. Tokyo: Kurosio.
Weiner, E. Judith and William Labov (1983) Constraints on the agentless passive. *Journal of Linguistics* 19: 29–58.
Weinreich, Uriel, William Labov and Marvin Herzog (1968) Empirical foundations for a theory of language change. In: Winfred P. Lehmann and Yakov Malkiel (eds.) *Directions for historical linguistics: A symposium*, 95–195. Austin: University of Texas Press.

Author's contact information:
Okayama Prefectural University
111 Kuboki, Soja-shi
Okayama 719-1197, Japan
e-mail: s-sano@dgn.oka-pu.ac.jp

[Received 10 August 2012;
Accepted 3 September 2013]

【要　旨】

コーパスを用いた日本語有声促音の無声化に関する研究
――必異原理と言語外的要因の役割――

佐野真一郎　　川原　繁人
岡山県立大学　　慶應義塾大学

Nishimura（2003，2006）は，日本語の借用語における有声促音が他の有声阻害音と共起する場合，無声化し得ると指摘している。この無声化のパターンについては，これまで理論・実験の観点から多くの分析があり，理論的諸問題の解決に貢献してきている。しかしながら，自然発話のデータを基にした研究はほとんど例がなく，社会言語学的な要因も仮定されていない。これらの背景を踏まえ本稿にて『日本語話し言葉コーパス』を用いて検証を行った結果，以下の２点が確認された。まず先行研究において確認されている，必異原理が有声促音の無声化を促進する効果が自然発話データにおいても確認された。次に，多くの言語外的・社会言語学的要因が無声化の適用・不適用に影響を与えているということが確認された。本稿における取り組みにより，これまでの研究で注目されることのなかった無声化のパターンを統御する潜在的要因が新たに明らかとなった。

Coetzee, Andries and Shigeto Kawahara (2013) Frequency biases in phonological variation. Natural Language and Linguistic Theory 31(1): 47-89.

Nat Lang Linguist Theory
DOI 10.1007/s11049-012-9179-z

Frequency biases in phonological variation

Andries W. Coetzee · Shigeto Kawahara

Received: 24 August 2010 / Accepted: 23 May 2011

Abstract In the past two decades, variation has received a lot of attention in mainstream generative phonology, and several different models have been developed to account for variable phonological phenomena. However, all existing generative models of phonological variation account for the overall rate at which some process applies in a corpus, and therefore implicitly assume that all words are affected equally by a variable process. In this paper, we show that this is not the case. Many variable phenomena are more likely to apply to frequent than to infrequent words. A model that accounts perfectly for the overall rate of application of some variable process therefore does not necessarily account very well for the actual application of the process to individual words. We illustrate this with two examples, English *t/d*-deletion and Japanese geminate devoicing. We then augment one existing generative model (noisy Harmonic Grammar) to incorporate the contribution of usage frequency to the application of variable processes. In this model, the influence of frequency is incorporated by scaling the weights of faithfulness constraints up or down for words of different frequencies. This augmented model accounts significantly better for variation than existing generative models.

Keywords Variation · Usage frequency · Harmonic Grammar · *t/d*-deletion · Japanese geminate devoicing

Electronic supplementary material The online version of this article (doi:10.1007/s11049-012-9179-z) contains supplementary material, which is available to authorized users.

A.W. Coetzee (✉)
Department of Linguistics, University of Michigan, 440 Lorch Hall, 611 Tappan Street, Ann Arbor, MI 48109, USA
e-mail: coetzee@umich.edu

S. Kawahara
Department of Linguistics, Rutgers University, 18 Seminary Pl., New Brunswick, NJ 08901-1108, USA
e-mail: kawahara@rci.rutgers.edu

Published online: 22 September 2012

1 Introduction

1.1 The changing prospects of variation

Although the existence of phonological variation has been acknowledged since the early years of generative phonology (Postal 1966:185, 1968:14–15), variation received relatively little attention in mainstream generative phonology during the first 25 years of the history of this field. To the extent that variation was acknowledged, it was usually relegated to the late stages of phonology or to phonetic implementation, and was hence not considered a part of the core of phonological grammar. In Lexical Phonology, for instance, it was assumed that lexical rules apply obligatorily while "postlexical rules can be optional and subject to variation" (Kaisse and Shaw 1985:6; see also Kiparsky 1985:86).

This low valuation of variation in mainstream generative phonology contrasts with how it was viewed in the Labovian variationist tradition. This research tradition, spearheaded by Labov's work in the late 1960's (Labov 1966, 1969, etc.), developed concurrently with mainstream generative phonology, but had little impact on this field. In this approach, variation is central to grammar rather than an accidental property that applies only on the edges of grammar. In fact, Labov (2004:6) claims that variation is "the central problem of linguistics".

In the past 15 years, the prospects of variation in generative phonology have changed dramatically. It now occupies a central place in the study of phonology, and to some extent dictates the architecture of phonological grammar. A clear indication of this change is how variation has been treated in handbooks of phonological theory. The first edition of the Blackwell *Handbook of Phonological Theory* (Goldsmith 1995), which reflects the situation in generative phonology at the beginning of the 1990's, does not even contain the word "variation" in its subject index. In contrast, every handbook since contains a chapter dedicated to variation (Anttila 2002b, 2007; Coetzee 2012; Coetzee and Pater 2011; Guy 2011). Similarly, several articles on variation have appeared in theoretical, generatively-oriented journals over the past decade (Anttila 2002a, 2006; Anttila et al. 2008; Boersma and Hayes 2001; Coetzee 2006; etc.).

This same period has seen the development of several versions of current generative phonological grammar intended to deal with variation. These models have all been developed in some version of a constraint-based grammar, be that classic discrete Optimality Theory (OT) (Anttila 1997, 2002a, 2006, 2007; Anttila et al. 2008; Bane 2011, to appear; Coetzee 2004, 2006, 2009c; Kiparsky 1993; Reynolds 1994), stochastic OT (Boersma 1997; Boersma and Hayes 2001), or noisy Harmonic Grammar (HG) (Coetzee 2009a; Coetzee and Pater 2011; Jesney 2007).[1]

In fact, variation has become so important that the ability of a grammatical model to account for variation is now often used as one of the measures of the model's sufficiency. Anttila (2002b:211) claims that an adequate theory of phonology should account for the "locus of variation" (where variation is observed and where it is not),

[1] Noisy HG was first implemented by Paul Boersma in *Praat* (Boersma and Weenink 2009) as early as 2006.

and the “degrees of variation” (the frequency of different variants). Using these two criteria as a measure of success, most of the models mentioned above have been very successful. All of these models have formal mechanisms that can account for the locus of variation. With the exception of Coetzee’s 2004/2006 model, these models also predict the degrees of variation. In fact, they have all been shown to be relatively successful in modeling the frequency with which different variants are observed for a range of variable phenomena.

In spite of the obvious progress that has been made in accounting for phonological variation, much work still remains. All of the existing generative models mentioned above are purely grammatical models that do not incorporate the influence of non-grammatical factors on variation. Decades of research in variationist sociolinguistics and more recent investigations of large speech corpora, however, have shown that variation is influenced by many factors in addition to grammar. In this paper, we take the next logical step in accounting for phonological variation by developing an extension of one of the existing generative models of phonological variation (noisy HG) that allows both grammatical and non-grammatical factors to impact the pattern of observed variation.

1.2 Non-grammatical influences on variation

One of the persistent results of the variationist research tradition is that variation is influenced, in addition to grammatical factors, by many non-grammatical factors. In fact, reviewing this tradition, Bayley (2002:118) identified “the principle of multiple causes” as one of the four core principles of this tradition. These non-grammatical factors include speech genre (word lists, informal conversations, read speech, etc.), discourse situation, age, sex or educational background of the speaker, etc.

Although mainstream generative phonology has adopted the variationist tradition’s higher valuation of variation over the past decade, mainstream approaches have focused nearly exclusively on the grammatical factors that impact variation. Existing generative models make no formal allowance for the influence of other factors. Yet, the variationist tradition has established that phonological variation is influenced by many factors in addition to grammar. The next step, then, is to augment generative models so that they can account for both the grammatical and non-grammatical factors that influence variation. This idea is not original to us. Boersma and Hayes (2001) mention this explicitly with regard to their stochastic OT model of variation, and suggest a way in which their model could be augmented to incorporate some non-grammatical aspects of variation. This paper follows up in more detail on their suggestion (although we will develop our model in noisy HG rather than their stochastic OT).

1.3 Usage frequency as a non-grammatical influence on variation

As mentioned above, many non-grammatical factors influence the application of variable phonological processes. In this paper, we focus on usage frequency—i.e., the observation that some variable processes apply at different rates to words that differ in frequency. Our selection of usage frequency is one of convenience: since frequency is already quantitative, it is straightforward to incorporate it into a quantitative model of variation. We also acknowledge that usage frequency would not be considered external to the grammar in all grammatical models. In fact, in several recent models of grammar, grammar can be described as structured memory encoding of frequency—see the usage-based and exemplar models of grammar, for instance (Bybee 2001, 2006, 2007; Gahl and Yu 2006, and papers therein; Pierrehumbert 2001; etc.). In the generative tradition, however, usage frequency is not encoded in the grammar—generative models do not treat two words differently merely because they differ in their usage frequencies. In this paper, we subscribe to the standard generative assumption, and we will hence treat usage frequency as external to the grammar. See also Sect. 5.1 for further discussion.

Some variable phonological processes (typically reduction or simplification processes, though see Sect. 5.2) are more likely to affect words with higher than lower usage frequency. For example, Bybee reports that the schwa in frequent *memory* is more likely to delete than the schwa in the nearly identical, but infrequent, *mammary* (Hooper 1976; see also Bybee 2000:68).

This correlation between frequency and simplification processes is widespread and has been reported for many different phonological processes. For instance, the variable deletion of word-final *t/d* from consonant clusters in English is more likely to apply to frequent than infrequent words—i.e., more deletion from frequent *just* than infrequent *jest* (Bybee 2000:69–70, 2002; Coetzee 2009a:272–273, 2009c; Lacoste 2008:187–207). The same process also applies in Dutch, where the correlation between frequency and the probability of deletion also holds (Goeman 1999:182; Phillips 2006:65). (See Sect. 3 for a more detailed discussion of *t/d*-deletion.) A similar correlation of usage frequency and variation has also been illustrated for flapping in American English (Patterson and Connine 2001), word-medial *t*-deletion in English (Raymond et al. 2006), word-final *s*-lenition in Spanish (File-Muriel 2010), *l*-vocalization in American English (Lin et al. 2011), and for geminate devoicing in Japanese loans (on which there is more in Sect. 4; see Kawahara 2011a, 2011b). See Phillips (2006) for a recent review of many more similar examples.

A model of variation that incorporates only grammatical influences on variation cannot capture the influence of factors like usage frequency. As a concrete illustration, we include Fig. 1, which represents the rate of *t/d*-deletion from word-final clusters in English for a selection of words from the *Buckeye Corpus* (Pitt et al. 2007), plotted against the log frequency of the words, as measured in CELEX (Baayen et al. 1995).[2] (See Sect. 3.1 on the details of how these data were extracted from the *Buckeye Corpus*.) The three panels show the rate of deletion before consonant-initial words (*west bank*), vowel-initial words (*west end*), and before pause (*west.*). The broken horizontal lines show the overall deletion rate in each context—i.e., deletion rate based on token counts. In existing grammatical models of variation, these are the variation

[2]Throughout this paper, all logarithmic transformations use a base of 10. For instance, the word *and* has a CELEX frequency of 514,946, and hence a log frequency of $\log_{10}(514{,}946) = 5.71$. In the Buckeye Corpus, *and* appears in pre-vocalic context 3,273 times, and in 2,966 of these occurrences its final /d/ was deleted. In this context, *and* therefore shows a deletion rate of 90.6 %. In the middle panel of Fig. 1, the data point that appears in the upper right-hand corner of the graph therefore corresponds to *and* at a log frequency of 5.71 on the x-axis, and at a deletion rate of 90.6 % on the y-axis.

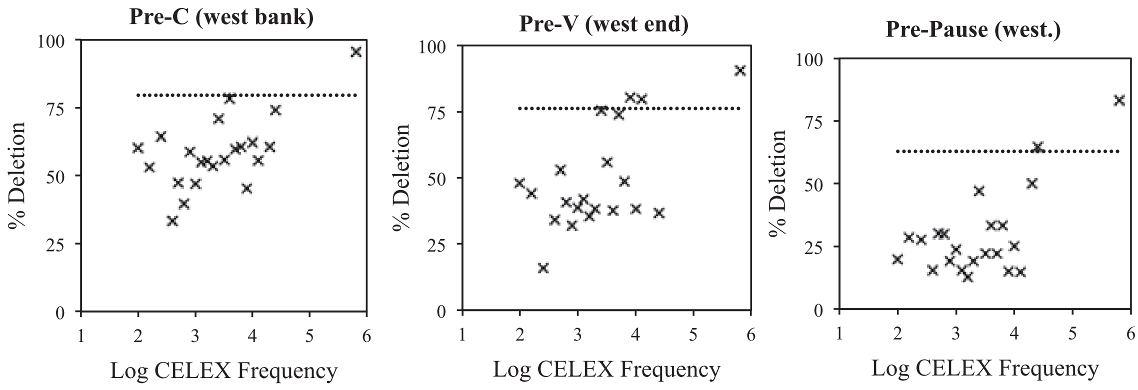


Fig. 1 Relation between deletion rate and frequency in the *Buckeye Corpus*

frequencies that are modeled. These rates capture the difference between the different grammatical contexts (most deletion pre-consonantally, then pre-vocalically, and least deletion pre-pausally). However the actual, observed rates deviate quite drastically from the overall rates, especially for words of lower frequency. To account not only for the grammatical influences on variation, but also for the influence of usage frequency, existing grammatical models would need to be augmented in some way. In the rest of this paper, we augment one of the existing grammatical models of variation (noisy HG). We add an extra parameter, incorporating usage frequency into the noisy HG model of variation, and show that this augmented model accounts significantly better for the deletion rates of words with different usage frequencies.

Although we treat frequency as if it is a standalone property of a word, it is actually only one subpart of the larger concept of predictability. A word's predictability depends on many factors in addition to its frequency, as has been documented by many studies in speech processing and production. A word is, for instance, primed by other words to which it is semantically (McNamara 2005; etc.) or phonologically (Goldinger et al. 1992; etc.) related, or by repetition (Versace and Nevers 2003; etc.). On the other hand, a word is inhibited (i.e., becomes less predictable) if it inhabits a dense lexical neighborhood (Luce and Pisoni 1986; Vitevitch and Luce 1998, 1999; etc.). Many studies have documented that factors such as these influence speech production, with the general result being that less predictable words (words that are inhibited or less strongly primed) tend to be produced more slowly, and with more effort or clarity (Baese-Berk and Goldrick 2009; Bell et al. 2009; Gahl 2008; Jurafsky et al. 2001; Scarborough 2004, 2010; etc.). Similar results have also been reported in the literature on "Uniform Information Density" (Frank and Jaeger 2008; Jaeger 2010; etc.), which shows that speakers have a tendency to spread out information equally across an utterance. Since more predictable words carry less information, speakers tend to reduce these words. Ultimately, it would be necessary to determine an overall measure of the predictability of a word that includes contributions from all of these aspects. Our focus on usage frequency is only an initial step.

2 Noisy Harmonic Grammar with weight scaling

We develop our model in a noisy version of Harmonic Grammar (Pater 2009; Smolensky and Legendre 2006). HG is a constraint-based theory that is closely related to OT (Prince and Smolensky 1993, 2004) and is, in fact, an historical predecessor of OT (Goldsmith 1993; Legendre et al. 1990). The main difference between HG and OT is that HG works with weighted rather than ranked constraints. Noisy HG is a stochastic implementation of HG, similar to the noisy implementation of OT, known as stochastic OT (Boersma 1997; Boersma and Hayes 2001). Noisy HG and stochastic OT are closely related; we could have developed our model in this paper just as successfully in stochastic OT rather than noisy HG (see Coetzee and Pater 2011 for evidence that noisy HG and stochastic OT account for most variable phenomena equally well). In the rest of this section, we first show how noisy HG accounts for variation, and then how we will augment this model to incorporate the influence of frequency on variation.

2.1 Noisy Harmonic Grammar

HG, like OT, is a constraint-based theory of grammar. The main difference between HG and OT is that OT relies on constraint ranking, and HG on constraint weighting. This difference is illustrated in the tableaux in (2) using the familiar OT constraints in (1). These tableaux represent the grammar of a language that does not allow tautosyllabic consonant clusters, and that repairs such clusters via deletion. In the HG tableau, *w*(CON) stands for the weight of constraint CON.

(1) MAX — Assign one violation mark for every segment in the input that lacks a correspondent in the output (no deletion). (McCarthy and Prince 1995:371)

DEP — Assign one violation mark for every segment in the output that lacks a correspondent in the input (no epenthesis). (McCarthy and Prince 1995:371)

*COMPLEX — Assign one violation for every tautosyllabic consonant cluster. (Prince and Smolensky 1993:96)

(2) a. Optimality Theory: DEP ≫ *COMPLEX ≫ MAX

/lʊst/	DEP	*COMPLEX	MAX
lʊst		*!	
☞ lʊs			*
lʊs.ti	*!		

b. Harmonic Grammar: *w*(DEP) > *w*(*COMPLEX) > *w*(MAX)

/lʊst/	5 DEP	1.5 *COMPLEX	1 MAX	H
lʊst		–1		–1.5
☞ lʊs			–1	–1
lʊs.ti	–1			–5

In HG, each constraint is weighted, and these weights are indicated with Arabic numerals above the constraint names in HG tableaux.[3] Constraint violations are marked with negative whole numbers rather than asterisks. A *harmony score* H is calculated for every candidate, using the formula in (3)—i.e., by taking the product of the weight of each constraint and the violation index of the candidate, and summing these products. These H-scores are indicated in the last column of the tableau. The H-score of the first candidate in (2b), for instance, is calculated as follows: The weight of DEP (= 5) is multiplied by the violation index of the candidate in terms of DEP (zero, since this candidate does not violate DEP). The weight of *COMPLEX (= 1.5) is then multiplied with the violation index of the candidate for *COMPLEX (−1), giving −1.5. Similarly, the weight of MAX (= 1) is multiplied with the violation index of the candidate (zero again). Finally, these products are summed, giving an H-score of −1.5 for this candidate. Since H-scores are negative, the candidate with the H-score closest to zero wins.

(3) $$\mathsf{H}(cand) = \sum_{i=1}^{n} w_i C_i(cand)$$

Where w_i is the weight of constraint C_i, and $C_i(cand)$ is the number of times that candidate *cand* violates C_i, expressed as a negative integer.

The version of HG illustrated above is not noisy HG, and cannot generate variation—given these constraints and weights, the grammar will always map /lʊst/ onto [lʊs]. However, HG has an implementation known as "noisy HG" that can generate variable outputs (Coetzee 2009a; Coetzee and Pater 2011; Jesney 2007). Noisy HG is closely related to stochastic OT (Boersma 1997; Boersma and Hayes 2001). In stochastic OT, constraint ranking is along a continuous scale, rather than a discrete scale as in classic OT. Every time that the grammar is used, the ranking of each constraint is perturbed by a negative or positive noise value (randomly selected from a normal distribution with a mean of zero). Because of this noisy evaluation, the relative ranking between two constraints can differ from one occasion to the next, resulting in variation. Noisy HG shares with stochastic OT this noisy evaluation procedure. The only difference is that, in noisy HG, the weights of constraints rather than their rankings are perturbed by random noise. If the weights of two conflicting constraints are close enough, the noisy evaluation can result in their relative weights flipping around between evaluation occasions, potentially causing variation.

In (4), the HG tableau from (2) is repeated, this time with noise. In these tableaux, *w* stands for the weight of a constraint and *nz* for the noise added to a constraint at the specific evaluation occasion. The effective weight of constraints (the sum of *w* and *nz*) is given in parentheses after the constraint names. In the first tableau, the weight of *COMPLEX is adjusted down by the addition of noise at −0.4, and the weight of MAX is adjusted upward by a positive noise value of 0.2. The effect is that violation of *COMPLEX is now less serious than the violation of MAX, so that the

[3] In noisy HG, the weights of the constraints are determined by a gradual learning algorithm, closely related to the learning algorithm developed by Boersma and Hayes (2001) for their stochastic OT model. For more on this, see Sect. 3.2.2.

faithful candidate has the highest H-score, and is selected as the output. In the second tableau, the weight of *COMPLEX is adjusted upward and that of MAX downward, so that the deletion candidate has the highest H-score and is selected as the output. These tableaux show how the same grammar (the same constraints with the same weights) can select different outputs on different evaluation occasions because of the addition of noise to the evaluation. An updated version of the formula used to calculate H-scores that include noise is given in (5).

(4) a. Faithful candidate optimal

	w	*nz*	*w*	*nz*	*w*	*nz*	
	5	–0.7	1.5	–0.4	1	0.2	
/lʊst/	DEP (4.3)		*COMPLEX (1.1)		MAX (1.2)		H
☞ lʊst			–1				–1.1
lʊs					–1		–1.2
lʊs.ti	–1						–4.3

b. Deletion candidate optimal

	w	*nz*	*w*	*nz*	*w*	*nz*	
	5	–0.8	1.5	0.1	1	–0.1	
/lʊst/	DEP (4.2)		*COMPLEX (1.6)		MAX (0.9)		H
lʊst			–1				–1.6
☞ lʊs					–1		–0.9
lʊs.ti	–1						–4.2

(5) $$\mathsf{H}(cand) = \sum_{i=1}^{n}(w_i + nz_i)C_i(cand)$$

Where w_i is the weight of constraint C_i, nz_i the noise associated with constraint C_i at this evaluation occasion, and $C_i(cand)$ is the number of times that *cand* violates C_i, expressed as a negative integer.

Several authors have shown have shown that this model of phonological variation can account for a variety of variable phenomena (Coetzee 2009a; Coetzee and Pater 2011; Jesney 2007). Coetzee and Pater (2011), in particular, show that it performs at least as well as stochastic OT. This model, however, still treats all words exactly the same. There is no place in the formula in (5) where any factor such as usage frequency can impact the H-score of a candidate. In the next section, we augment this model to allow for factors such as usage frequency to impact the H-score of a candidate.

2.2 Weight scaling

We need a model that can account for the fact that more frequent words are more likely to be treated unfaithfully. This correlation can be captured by scaling the weight of faithfulness constraints down for frequent words and up for infrequent words. Violating a faithfulness constraint will then contribute less to the H-score of a frequent word, resulting in unfaithfulness being more likely, while it will contribute more to the H-score of an infrequent word, resulting in faithfulness being more likely.

There are precedents for this idea in the literature. Van Oostendorp (1997) and Itô and Mester (2001), for instance, suggested that the higher likelihood of faithfulness in more formal speech registers can be captured by ranking faithfulness constraints higher in formal speech situations—an idea that echoes the concept of "carefulness weights" in Lindblom's Hyper- and Hypoarticulation theory of speech production (Lindblom 1990). Boersma and Hayes (2001: Appendix C) similarly suggest scaling the ranking values of constraints to account for different rates of unfaithfulness observed with different speech registers.

By adding such weight scaling to the model, two words that differ in usage frequency may be evaluated differently in the same grammatical context. Continuing with the example from the previous section, assume that /lʊst/ and /nʊst/ differ in frequency such that /lʊst/ is frequent and /nʊst/ infrequent. For the sake of the illustration, assume that /lʊst/ will be associated with a weight scaling factor of -1, and /nʊst/ with a factor of $+1$. The weight of faithfulness constraints will be scaled down by one unit in the evaluation of /lʊst/, and up by one unit in the evaluation of /nʊst/. The tableaux in (6) show how this addition of scaling factors affects the evaluation of these words. In these tableaux, the same grammatical settings (the same constraint weights and noise values) are used. All that differs is the scaling factors associated with the faithfulness constraints (marked by *sf* in the tableaux). The result is that frequent /lʊst/ is mapped onto its unfaithful candidate [lʊs], while infrequent /nʊst/ is mapped onto its faithful candidate [nʊst]. An updated version of the H-score formula that incorporates the scaling factor is given in (7).

(6) a. Evaluating frequent /lʊst/, with $sf = -1$

	w	*nz*	*sf*	*w*	*nz*	*w*	*nz*	*sf*	
	5	0.7	–1	1.5	0.1	1	0.2	–1	
/lʊst/	DEP (4.7)			*COMPLEX (1.6)		MAX (0.2)			H
lʊst				–1					–1.6
☞ lʊs						–1			–0.2
lʊs.ti	–1								–4.7

b. Evaluating infrequent /nʊst/, with $sf = +1$

	w	*nz*	*sf*	*w*	*nz*	*w*	*nz*	*sf*	
	5	0.7	1	1.5	0.1	1	0.2	1	
/nʊst/	DEP (6.7)			*COMPLEX (1.6)		MAX (2.2)			H
☞ nʊst				–1					–1.6
nʊs						–1			–2.2
nʊs.ti	–1								–6.7

(7) $$\mathsf{H}(cand) = \sum_{i=1}^{n}(w_i + nz_i)M_i(cand) + \sum_{j=1}^{m}(w_j + nz_j + sf)F_j(cand)$$

Where M_i is the i-th markedness constraint, w_i the weight associated with M_i, nz_i the noise associated with M_i at this evaluation occasion, and $M_i(cand)$ the number of times that *cand* violates M_i (expressed as a negative integer); and where F_j is the j-th faithfulness constraint, w_j the weight associated with F_j, nz_j the noise associated with F_j at this evaluation occasion, and $F_j(cand)$ the number of times that *cand* violates F_i (expressed as a negative integer); and where *sf* is the scaling factor associated with the specific word being evaluated.

In this model, only faithfulness constraints have scaling factors. The same effect could also be achieved by scaling markedness weights, or even by scaling the weights of both markedness and faithfulness constraints. In fact, Boersma and Hayes (Boersma and Hayes 2001: Appendix C) propose scaling the ranking values of both markedness and faithfulness constraints to incorporate style effects into their stochastic OT model. Although there are subtle differences in the variation patterns predicted by these different options, any of these options could have accounted equally well for the data that we discuss in this paper. We return to this issue briefly in Sect. 5.2, but leave the question of the difference between these options for future research.

2.3 A linking function between frequency and scaling factors

The final part of our model is a linking function between frequency and scaling factors: Given a word of some frequency, what is the scaling factor that should be used in evaluating this word? This problem could be approached from two different directions. One possibility is that the mapping between frequency and scaling factors has to be learned on a language-by-language basis. The language learner will then have to take note of how words that are equivalent in their phonological properties but differ in frequency are treated differently by the grammar. From this information, he/she will deduce a function that best maps from frequency to scaling factors. Since the linking function is then determined on a language-particular basis, we would not necessarily expect to see universal tendencies in how frequency maps to scaling factors. See Coetzee (2009a) for an implementation of this kind of approach.

A different possibility is that there is some universal linking function that applies similarly to all languages. The expectation would then be that frequency has the same basic influence in all languages. Given the large amount of evidence that frequency has the same basic influence in all languages (More frequent words are more likely to undergo reduction processes—see the references above in Sect. 1.3 and the discussion in Sect. 5.2 below.), we pursue the second option—that is, that the same basic linking function applies in all languages. In this paper, we illustrate how such a universal mechanism accounts well for two different variable phenomena in two unrelated languages (*t/d*-deletion in English, and geminate devoicing in Japanese).

We propose that every word is associated with a distribution function, whose shape is determined by the frequency of the word. These functions are modeled as instantiations of the *beta* distribution (Gupta and Nadarajah 2004), and the scaling factor associated with a word is read off its distribution function.[4] The formula of the *beta* distribution is given in (8). In addition to its argument x, the distribution has three parameters. ρ specifies the range of the function as spanning from $-\rho$ to ρ. α and β are shape parameters that determine the skewness of the distribution. When $\alpha = \beta$,

[4]See later in this section on why we use the *beta* distribution rather than a more well-known distribution such as the normal distribution.

the distribution is symmetric around zero. When $\alpha > \beta$, it is left-skewed, and when $\alpha < \beta$, it is right-skewed. Additionally, the larger the difference between α and β, the more severe the skewness of the distribution is.

(8) $$f(x, \alpha, \beta, \rho) = \rho \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 x^{\alpha-1}(1-x)dx}$$

Frequent words must have a negative, and infrequent words a positive scaling factor. But what counts as "frequent" or "infrequent"? A reference frequency has to be established such that words that appear more frequently than this reference frequency will be treated as frequent, and words that appear less frequently will be treated as infrequent. There are several ways in which such a reference point can be established. The average or median frequency of all the words in the lexicon could be used, for instance. We explored several different options, and settled on the one that resulted in the best fit of our model to the data. Specifically, the reference frequency is set in such a way that (at least) half of the tokens in the corpus are being treated as frequent, and (at most) half as infrequent. The exact way in which we determine the reference frequency is stated in (9).

(9) Let N be the total number of tokens in the corpus.

i. Order the words in the corpus in terms of frequency.
ii. Determine the point on this ordering such that at least $N/2$ of all the tokens are above this point.
iii. Determine the log frequency of the word just above this point, and the word just below this point.
iv. Let the reference frequency be halfway between these two log frequencies.

We illustrate how this algorithm works with an example. In Sect. 3, we work with a corpus of *t/d*-deletion examples, extracted from the *Buckeye Corpus* (Pitt et al. 2007). The corpus contains 16,460 tokens. Ordering the tokens according to their CELEX frequencies (Baayen et al. 1995), the word *and* occupies the topmost position. It also accounts for more than half of the tokens in the corpus (*and* appears 8,827 times in our corpus). The reference point is halfway between the log CELEX frequency of *and* and the log CELEX frequency of *just*, the next most frequent word in our corpus. For reasons that we explain in Sect. 3, we grouped words together into larger log groups. *Just* was placed into the 4.4 log frequency group, while *and* went into the 5.8 log frequency group. The midpoint between these two is 5.1, and this value serves as the reference point in our modeling of the data in our *t/d*-deletion corpus.

Having established the reference frequency, the values of the shape parameters (α and β) of the *beta* distribution associated with each word, as well as the scaling factor associated with each word, can now be determined. Specifically, we propose that α is set equal to the log reference frequency, and β to the log frequency of the specific word. The α-parameter therefore represents the reference frequency (i.e., neither frequent nor infrequent). The β-parameter represents the frequency of a specific word. For a word that appears less often than the reference frequency (so that $\alpha > \beta$), the distribution will be left-skewed and hence have a positive mode—see the distribution for *interrupt* in Fig. 2. We propose that the mode is used as the frequency scaling factor associated with a specific word. For a word that appears less frequently than the reference frequency, the scaling factor will therefore be positive. The weight of faithfulness constraints will be scaled up in the evaluation of such a word, so that the word will resist an unfaithful mapping more strongly. On the other hand, for a word that appears more often than the reference frequency (so that $\alpha < \beta$), the distribution will be right-skewed, and the mode thus negative—see the function for *and* in Fig. 2. The scaling factor of such a frequent word (the mode of the *beta* distribution) will be negative, diminishing the contribution of faithfulness constraints in evaluating the word, resulting in a higher likelihood of an unfaithful mapping. The table in (10) summarizes the effect of the values of α and β on the skewness of the *beta* distribution, and the effect that this has on the mode of the distribution (and the scaling factors in the model that we propose here).

(10) Determining the values of α, β, and the scaling factor associated with each word

Shape parameters				
Reference frequency		Frequency of specific word	Skewness	Mode (= scaling factor)
α	<	β	Right	negative
α	=	β	Symmetrical	zero
α	>	β	Left	positive

The last parameter to set is the range parameter ρ. ρ does not influence the shape of the *beta* distribution, but only its range. In particular, it specifies the minimum and maximum value of the function on the x-axis: The higher the value of ρ, the higher the absolute value of the mode. The higher ρ is, the higher the scaling factors will be. And the higher the scaling factors, the more influence the frequency of words can have on their evaluation. ρ therefore determines how much frequency is allowed to influence how the grammar functions. We propose that the value of ρ be fit to the data—i.e., for every corpus, the value of ρ that results in the best fit between the model and the data is used.[5]

In (11), we give examples of the parameter values and the modes for three words from our *t/d*-deletion corpus. *And* is used as an example of a frequent word. *And*'s distribution function is right-skewed, so that the mode of this function, and hence *and*'s scaling factor, is negative. *Interrupt* and *weekend* both appear less frequently than the reference frequency, and both serve as examples of infrequent words. Their distributions are left-skewed, so that their modes are positive, and the scaling factors associated with these two words are also positive. Although both *interrupt* and *weekend* are infrequent, they differ in frequency. *Interrupt* has a CELEX log frequency of

[5]We also leave open the possibility that the value of ρ can vary across different speech styles. A larger value for ρ results in a larger range for the *beta* distribution, and hence in modes that deviate more from zero. Since the mode of the *beta* distribution is used as the scaling factor in the evaluation of some word, a larger ρ (and hence more extreme mode and scaling factor) will increase the influence that frequency can have on the determination of H-scores. It is therefore possible that the value of ρ may fluctuate to account for speech situations in which frequency has a bigger or smaller impact. We do not explore this possibility further in this paper, however.

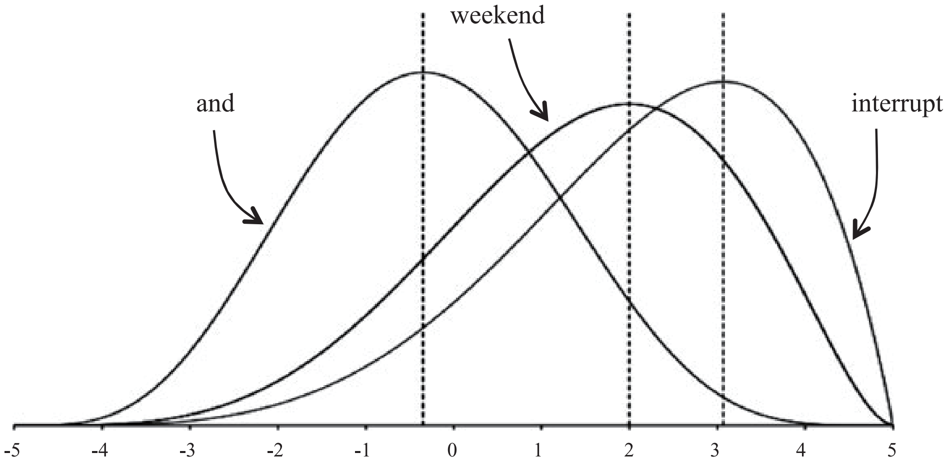


Fig. 2 *Beta* distributions for words from (11) with $\rho = 5$. Vertical broken lines mark the modes for the distributions, and hence the scaling factors associated with these words

1.98 and *weekend* has one of 2.76. In the distribution function associated with *interrupt*, the difference between the values of α and β is hence larger than in *weekend* ($\alpha = 5.1$, $\beta = 1.98$ vs. $\alpha = 5.1$, $\beta = 2.76$). We include both of these words to show that the larger the difference between α and β, the more skewed the distribution, and hence the more extreme the mode of the distribution. The more the frequency of a word (represented by β) differs from the reference frequency (represented by α), the more its scaling factor will differ from zero. Faithfulness will hence be scaled down more for more frequent words, and up more for less frequent words. The table also gives the modes for these distributions at three different values of ρ. Note how a change in ρ influences only the absolute value of the modes, and not their signs. In Fig. 2 we show the shape of the distribution functions for these tokens when $\rho = 5$ (the value that we use for ρ in Sect. 3).[6]

(11) Examples of scaling factors in the *t/d*-deletion corpus (see Sect. 3.3)

Word	Shape parameters		Skew	Mode (scaling factor)		
	α (reference frequency)	β (log frequency of word)		ρ = 1	ρ = 5	ρ = 10
and	5.1	5.71	Right	–0.07	–0.35	–0.70
weekend	5.1	2.76	Left	0.40	2.00	4.00
interrupt	5.1	1.98	Left	0.61	3.07	6.14

In principle, scaling factors could be deduced from a more well-known distribution such as the normal distribution. Our selection of the *beta* rather than the normal distribution is motivated by the fact that the *beta* distribution has a finite range (specified by ρ), while the normal distribution has an infinite range. The finite range of the *beta* distribution places an absolute limit on the influence that non-grammatical factors such as frequency can have via weight scaling. If scaling factors were taken from the

[6]An Excel file for the calculation of the *beta* distribution's mode under different settings of the three parameters is available from http://www.quantitativeskills.com/sisa/rojo/distribs.htm. In this file, the range parameter ρ is represented by A and B, with $A = -\rho$ and $B = \rho$. The shape parameter α is represented by p, and β by q.

normal distribution with its infinite range, there would be no principled limit on the extent to which non-grammatical factors could influence the application of variation. See Sect. 5.1 for more detailed discussion.

3 English *t/d*-deletion

Word-final *t/d* variably deletes from consonant clusters in English, so that a word like *west* can be pronounced as [wɛst] or [wɛs]. This deletion process has been described in detail for countless dialects of English (see Coetzee 2004: Chap. 5 for a review), and even for languages other than English (on Dutch, see Goeman 1999; Goeman and van Reenen 1985; Schouten 1982, 1984). Since this process has been studied so extensively, the factors (both grammatical and non-grammatical) that influence its application are reasonably well understood. We begin this section by first reviewing some of the grammatical and non-grammatical factors that are known to influence this process, focusing on those aspects for which we will provide an account. We then develop a purely grammatical account in the noisy HG framework. Once the grammatical account has been established, we augment it to account for the influence of usage frequency according to the method described above in Sects. 2.2 and 2.3.

3.1 Grammatical and non-grammatical influences

We first review evidence that this process is influenced by the same kinds of grammatical considerations as those that influence "ordinary" non-variable phonological rules. Echoing an idea that has been present throughout the variationist research tradition for nearly four decades, Anttila (1997:44) takes this fact to be a motivation for expecting phonological grammar to account for at least part of variation: "... if variation preferences are based on phonological variables, then it seems reasonable to expect phonology to make sense of them."

In a summary of the grammatical factors that influence *t/d*-deletion, Labov (1989) includes the following: (i) *Stress*: *t/d* is more likely to delete from an unstressed syllable (*cúbis<u>t</u>*) than a stressed syllable (*insís<u>t</u>*); (ii) *Cluster size*: Deletion is more likely from tri-consonantal (*tanked* [tæŋk<u>t</u>]) than from bi-consonantal clusters (*tacked* [tæk<u>t</u>]); (iii) *Similarity to preceding segment*: Deletion is more likely after consonants that share more features with *t/d* than consonants that share fewer features—there is more deletion from *kissed*, where [s] shares place (coronal) and sonorancy (non-sonorant) with the following [t] than from *seemed*, where [m] shares no major features with the following [d]; (iv) *Morphology*: *t/d* that functions as the past tense suffix of a regular past tense verb (*missed*) is less likely to delete than *t/d* that functions as the past tense suffix in a semi-weak verb (*kept*), which is less likely to delete than *t/d* that is part of a morphological root (*mist*).

Another grammatical factor that influences *t/d*-deletion is the context that follows the word-final *t/d*. We use this factor as an example of a grammatical factor in the rest of this section, and will therefore discuss it in more detail. In every dialect of English for which *t/d*-deletion has been studied, it has been found that deletion is most likely if the next word begins with a consonant (*west bank*). Dialects diverge on whether a following vowel-initial word (*west end*) or a pause (*west.*) results in more deletion. The table in (12) contains a sample of the data available on the influence of the following context.[7] The data on all but Columbus English are taken from the literature, with references given in footnote 8.

(12) Percent *t/d*-deletion in different English dialects in pre-consonantal, pre-vocalic, and pre-pausal contexts.[8]

Relative deletion rate		Pre-C *west bank*	Pre-Pause *west.*	Pre-V *west end*
Pre-C > Pre-Pause > Pre-V	AAVE (Washington, DC)	76	73	29
	Jamaican English	85	71	63
	Tejano English	62	46	25
	Trinidadian English	81	31	21
Pre-C > Pre-V > Pre-Pause	Chicano English	62	37	45
	Columbus English	80	63	76

The data on Columbus English were extracted from the *Buckeye Corpus* (Pitt et al. 2007). This is a corpus of conversational speech collected from 40 lifelong residents of Columbus, Ohio. All of the speech was both orthographically and phonetically transcribed. In order to compile a list of words from the corpus to which *t/d*-deletion could apply, we extracted all words that end orthographically in -*Ct* or -*Cd* (where *C* stands for any consonant). Since *t/d* that corresponds to the past tense suffix is consistently treated differently (see discussion above), and since our focus is on the influence of the phonological context, we excluded words with this suffix. The principle by which we selected tokens from the corpus already excluded past tense forms that end orthographically in -*ed*. We manually removed the semi-weak past tense forms, such as *kept*. We also removed a few other classes of words. First, due to the difficulty of determining whether word-final *t/d* has been realized before a word that starts with [t] or [d], we removed all such tokens from the list. Secondly, we removed words that end orthographically in -*rt*/-*rd* or -*lt*/-*ld*. These tokens showed unexpectedly low deletion rates in the corpus. In these tokens, *r* and *l* were often phonetically realized as coloring on the preceding vowel rather than as a separate consonant, so that -*rt*/-*rd* and -*lt*/-*ld* words often do not actually end in consonant clusters phonologically (Guy and Boberg 1997). Lastly, we removed words such as *thought* and *could*, that end orthographically but not phonologically in -*Ct*/-*Cd*. This whole procedure left a list of 16,460 tokens, representing 459 different words. The phonetic transcription in the corpus for each of the token words was then consulted, and each token was coded as either "*t/d* deleted" or "*t/d* retained".[9] Each token was also classified as

[7]These data are simplified with regard to the pre-consonantal context. Labov (1989) and Guy (1991), among others, show that *t/d*-deletion rates are different before consonants of different types. We follow the practice in the vast majority of the *t/d*-deletion literature of lumping all of the consonants together.

[8]Sources: AAVE (Fasold 1972), Jamaican (Patrick 1992), Tejano (Bayley 1995), Trinidad (Kang 1994), Chicano (Santa Ana 1991).

[9]A token was coded as "*t/d* deleted" if no segment was transcribed for the underlying *t/d*. In the *Buckeye Corpus*, underlying *t/d* was transcribed with several different surface realizations, including faithful realizations [t] or [d], glottalized realizations [$\text{t}^{ʔ}$] or [$\text{d}^{ʔ}$], flap [ɾ], etc. All tokens transcribed with one of these

pre-consonantal, pre-vocalic, or pre-pausal based on the context in which the token appeared in the corpus.[10,11]

Several non-grammatical factors that influence the application of *t/d*-deletion have also been documented in the variationist literature. For example, biographical factors, such as the age, sex, or ethnicity of the speaker, have been shown to influence application of the process. Additionally, speech register also influences the deletion rate, with less formal registers associated with higher deletion rates. Browman and Goldstein (1990), for instance, found little evidence of *t/d*-deletion in the reading of a word list, but they did find evidence for the process in a more casual conversational speech style. Mitterer and Ernestus (2006) studied the analogous process in Dutch in two speech corpora. One corpus consisted of read speech (literally, novels read on tape for the blind)—i.e., a rather formal speech register. The other corpus consisted of recordings of casual speech. They found evidence of deletion in both corpora, but at very different rates (8 % for the read speech vs. 45 % for the casual speech).

The non-grammatical factor on which we focus is usage frequency, and we therefore report on it in more detail. As we already showed in Sect. 1.3, phonological processes such as *t/d*-deletion usually apply at higher rates to words of higher frequency—i.e., there is more deletion from frequent *just* than from phonologically similar but infrequent *jest*. Bybee (2000:69–70), for instance, analyzes Santa Ana's 1991 corpus of Chicano English, and finds a deletion rate of 54.4 % in high frequency words compared to 34.4 % for low frequency words.[12] Phillips (2006:65) shows that frequency has the same influence in the analogous process in Dutch.

In order to investigate the influence of frequency on *t/d*-deletion in the *Buckeye Corpus*, we determined the frequency of each of the words that we selected from this corpus in CELEX (Baayen et al. 1995), and then transformed these counts by

realizations were coded as "*t/d* retained". Since the corpus contains no articulatory data, deletion is defined here as the absence of any acoustic evidence of *t/d*. An actually articulated *t/d* might not have any acoustic realization when it is articulated before a labial consonant. If the labial closure of the following consonant is made before the release of the *t/d*, the potential acoustic effect of the coronal release is masked by the labial closure, and hence becomes inaudible (Browman and Goldstein 1990). The actual articulatory *t/d*-deletion rate before consonants may therefore be somewhat lower than the acoustic rate reported here. As a check of the potential influence that this acoustic masking could have on our data, we counted the number of tokens in our pre-consonantal category followed by labial and non-labial consonants. We found that more than 80 % of the pre-consonantal tokens appear before non-labial consonants.

[10] The coding conventions in the *Buckeye Corpus* do not actually include a category for pauses. We coded as pre-pausal the following tokens: (i) tokens where the corpus indicates that silence followed an utterance; (ii) tokens where the corpus indicates that an utterance was followed by the interviewer speaking, and where it was clear from the context that the interviewer did not interrupt the interviewee mid-utterance; (iii) utterances followed by some kind of non-speech vocalization noise, and where the context made it clear that this vocalization noise did not occur mid-utterance.

[11] The corpus of *t/d*-words that we used is available as "supplementary material" on the Springer link for this article, or from the first author upon request.

[12] Bybee (2001) and Jurafsky et al. (2001:252–255) show that mere lexical usage frequency does not capture the full influence of frequency. Just as important, and in some instances maybe even more important, is frequency of use within a specific syntagmatic context. That is, the [t] in *best* may delete more often from a more frequent phrase such as *best friend* than from a less frequent phrase such as *best fruit*. Although an adequate account of phonological variation will ultimately have to incorporate all relevant types of frequency influences (and all other relevant influences), we will focus only on lexical usage frequency in this article.

taking their logarithms (with base 10).[13,14] Because the *Buckeye Corpus* is relatively small, words with a low CELEX frequency count appear infrequently in the corpus. (In fact, several words appear only once.) It is consequently not possible to calculate reliable deletion rates for individual words, and we therefore divided the words into frequency bins before calculating deletion rates (cf. also Bybee 2000:69–70; Lacoste 2008:188–189). Most of the frequency bins spanned 0.1 intervals on the log-transformed frequency values. If some bin contained fewer than 50 tokens, we combined it with one of its adjacent bins so that a few bins spanned a wider range than 0.1. In total, 23 frequency bins were created ranging in log-transformed frequency from (0 to 2.0) up to (5.7 to 5.8).[15] The deletion rate in each of the three contexts (pre-vowel, pre-consonant, pre-pause) was then calculated for each frequency bin. This procedure gives a data set where deletion rates in each of the contexts can be plotted against frequency to look for a correlation, as in Fig. 3. This figure shows a positive correlation between frequency and deletion rate in all three contexts. In fact, the correlation is significant in all three contexts (Pre-C: $r^2 = 0.46$, $p < 0.01$; Pre-V: $r^2 = 0.39$, $p < 0.01$; Pre-Pause: $r^2 = 0.43$, $p < 0.01$).[16]

In the next section, we first develop an account for the influence of the following phonological context on *t/d*-deletion in Columbus English, as given in (12). In doing this, we will abstract away from the influence of usage frequency, shown in Fig. 3.

---

[13] Since log of zero is undefined, a constant of one was added to all frequencies before they were log-transformed.

[14] One could raise some concerns about using CELEX to measure usage frequency. First, CELEX is a British corpus, and usage frequency may differ between CELEX and the American speakers included in the *Buckeye Corpus*. Second, although CELEX includes some spoken sources, the majority of the frequency counts in CELEX come from written texts. Usage frequency may be different between spoken and written language.

A possibly more accurate measure of the usage frequency of words for the speakers who contributed to the *Buckeye Corpus* would be the *Buckeye Corpus* itself—i.e., just counting the frequency with which each token appears in the corpus. However, since the *Buckeye Corpus* is comparatively small, it does not differentiate well between words with low usage frequencies—many words appear only once in the corpus. Facing the same problem with regard to the *Buckeye Corpus* and CELEX, Raymond et al. (2006) showed that CELEX and *Buckeye* frequencies are highly correlated ($r = 0.82$). In fact, using CELEX for frequency counts, even when dealing with American English, is standard practice in the field (Albright 2009; Coetzee 2005, 2008). We therefore follow the standard practice, using CELEX for frequency counts in our study.

[15] The decision to use 23 frequency bins is to some extent arbitrary. A finer-grained division into more bins could potentially give a more detailed picture of how usage frequency interacts with deletion. However, relying on more bins also results in some bins containing too few data points to reliably calculate deletion rates. There is a trade-off between the reliability of the deletion rate for each frequency bin and the fine-grainedness with which the frequency range is sampled. We decided to use bins that contain at least 50 tokens each, resulting in the 23 bins used here.

[16] On each of the three graphs, there is one data point with an extremely high log frequency, just below 6. This data point corresponds to the word *and*, which accounts for more than half of all the tokens in our corpus. If this data point is removed, the positive correlation between frequency and deletion rate remains, even if it is less strong (Pre-C: $r^2 = 0.21$, $p < 0.05$; Pre-V: $r^2 = 0.22$, $p < 0.05$; Pre-Pause: $r^2 = 0.14$, $p < 0.11$). Due to the fact that extremely high frequency words such as *and* show much higher deletion rates, these words are often excluded from the data sets used in variationist sociolinguistic studies of *t/d*-deletion (Patrick 1992:172). By including frequency as a factor in our model, we do not have to exclude frequent words. Their seemingly anomalous behavior is no longer anomalous, but rather expected given the model that we develop.

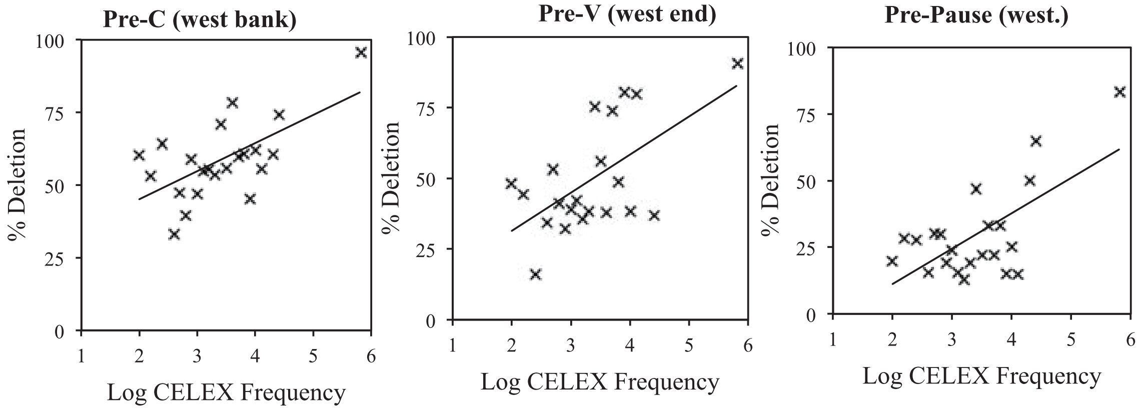


Fig. 3 The relation between frequency and deletion rate in Columbus English in Pre-C, Pre-V and Pre–Pause contexts. The x-axis represents log-transformed CELEX frequencies. Deletion rate is plotted on the y-axis

Once this grammatical account is in place, we will augment it to incorporate the influence of frequency.

3.2 A grammatical account

In this section we develop a noisy HG account for the overall deletion rates observed in Columbus English, as shown in the table in (12). For similar accounts of the other data from this table, see Coetzee and Pater (2011).

3.2.1 Constraints

The constraints that we use are given in (13). The two contextual faithfulness constraints are in the spirit of Steriade's "licensing by cue" constraints—i.e., they protect segments from deletion in contexts where the cues for their perception are saliently licensed (Steriade 1999, 2001; Côté 2004).

(13)	*CT]$_{\text{Word}}$	Assign one violation mark for every word that ends in the sequence [-Ct] or [-Cd].[17]
	MAX	Assign one violation mark for every input segment lacking an output correspondent (no deletion). (McCarthy and Prince 1995:371)
	MAX-PRE-V	Assign one violation mark for each segment that appears in pre-vocalic context in the input, and that does not have a correspondent in the output (no deletion before a vowel). (Côté 2004:22)

[17] This constraint is a special version of the more general *COMPLEX, which applies only to a subclass of consonant clusters, and only when these clusters appear in word-final position. As it stands, the constraint is too specific. For instance, deletion of [p] from words like *ramp*, *wisp*, etc., and deletion of [k] from words like *whisk*, *task*, etc. are also observed. To account for these deletions, the constraint should probably be generalized so that it penalizes all [...C+stop] sequences. However, the literature contains virtually no information on the deletion of [p] and [k], probably because there are so few [...Cp] and [...Ck] words in English. For this reason, we assume the more specific constraint here. See Coetzee (2004: Chap. 5) for an exploration of a more general constraint.

MAX-PRE-PAUSE	Assign one violation mark for each segment that appears in pre-pausal context in the input, and that does not have a correspondent in the output (no deletion before a pause).

Steriade proposes that a segment is protected by special faithfulness constraints in contexts where its perceptual cues are robustly licensed. The consonant release burst can cue both place (Lahiri et al. 1984; Stevens and Blumstein 1978) and manner information (Stevens and Keyser 1989). The formant transitions into a following vowel also carry information about both place (Martínez-Celdrán and Villalba 1995; Eek and Meister 1995; Fowler 1994; Fruchter and Sussman 1997; Kewley-Port 1983; Kewley-Port et al. 1983; Nearey and Shammas 1987; Stevens and Blumstein 1978; Sussman et al. 1991; etc.) and manner (Diehl and Walsh 1989; Walsh and Diehl 1991). To motivate the existence of the positional versions of MAX, it is therefore necessary to show that release bursts and formant transitions are more robustly licensed in pre-vocalic and pre-pausal position than in pre-consonantal position.

In pre-consonantal position, the likelihood of a consonantal release being realized is relatively small. Zsiga (2000:78) reports a release rate of as low as 18 % in this context for English (see also Browman and Goldstein 1990). Except when the following consonant is a sonorant, there is also no opportunity for the realization of formant transitions, and even into a following sonorant, robust transitions are less likely than into a following vowel. Pre-consonantal position is hence the context in which *t/d* is least well cued, so that there is no special faithfulness constraint that protects against deletion specifically in this context.

In pre-pausal position, formant transitions into a following segment cannot be realized. However, it is possible to release stops in this position—Byrd found that 57 % of alveolar stops were released in the TIMIT corpus (Byrd 1992:37). There is also evidence that utterance-final released consonants are perceived more accurately than unreleased consonants (Malécot 1958). In pre-vocalic position, both formant transitions and releases can be realized. Only one of the cues can therefore be realized pre-pausally while both cues can be realized pre-vocalically. On the other hand, the pre-vocalic cues can only be realized across a word boundary. The crossing of the word boundary may result in a penalty for cue robustness in pre-vocalic position. The listener may, for instance, incorrectly perceive the *t/d* as the first segment of the following word rather than the last segment of the preceding word. As such, the additional acoustic cue available in this context would not necessarily result in easier perception and lexical access for the listener. A question is whether there is a universal difference in cue robustness between pre-pausal and pre-vocalic contexts. In Steriade's "licensing by cue" model of faithfulness, constraints protecting inherently more robust sponsoring contexts universally rank higher than constraints protecting less robust sponsoring contexts. If there is an inherent robustness difference between pre-vocalic and pre-pausal contexts, the two positional versions of MAX will therefore be in a universally fixed ranking.

Exactly how the ranking between cue-licensing constraints is established is still an unresolved topic. These rankings could be hard-wired into Universal Grammar or they could emerge during acquisition, influenced by misperception on the side of the language learner (Boersma 2008). In English dialects where pre-pausal *t/d* is seldom released, the child acquiring the grammar will more often not perceive *t/d* in this position, even if his/her parents actually produced *t/d* in this context. Such a learning situation might lead to the lower ranking of the constraint MAX-PRE-PAUSE in the grammar of such a child. On the other hand, a child acquiring a dialect where pre-pausal stops are more often released may actually perceive *t/d* more often in this context, resulting in a higher ranking of MAX-PRE-PAUSE in the grammar of such a child. The rankings could therefore result from the concrete experience of the language learner as a listener. This is also in agreement with Kawahara's claims that rankings between cue-based faithfulness constraints are based on the actual perceptibility of contrasts in different contexts (Kawahara 2006). On the other hand, Moreton (2008, 2010) has shown that some typological tendencies may result from hard-coding of rankings into UG rather than from experience with actual perceptibility.

Given that this issue is still unresolved, we will not take a stance here on how exactly the ranking between cue-based faithfulness constraints comes about. We do note that, given the data reported in (12), it is necessary to allow MAX-PRE-PAUSE and MAX-PRE-V to rank differently in the grammars of different dialects/languages in order to account for the difference between dialects that show more deletion in pre-pausal position and those that show more deletion in pre-vocalic position.

3.2.2 The learning simulation and results

The constraint weights for Columbus English were determined by running a learning simulation with *Praat*'s noisy HG learning algorithm (Boersma and Weenink 2009). For details on this learning algorithm, see Boersma and Pater (2008) and Coetzee and Pater (2008). In creating an input file for the algorithm, we assumed that each of the contexts (pre-consonantal, pre-vocalic, pre-pausal) appears 100 times. Deletion was represented in the 100 tokens in each context proportional to the overall deletion rates from (12)—i.e., in pre-consonantal context, 80 tokens were coded as pronounced with deletion and 20 with a final *t/d*, in pre-pausal context 63 with deletion and 37 with retention, and in pre-vocalic context 76 with deletion and 24 with retention.[18] We based the learning input file on the overall deletion rate, following the tradition in the literature. The account that we develop here will therefore not take into account the contribution of the usage frequency of individual words. In the next section, we will augment our account by implementing weight scaling. In running the learning simulation, we set the "decision strategy" to "Linear OT" (*Praat*'s implementation of the noisy HG learning algorithm). All other settings were kept at *Praat*'s defaults.[19]

[18] The *Praat* input file is available as "supplementary material" on the Springer link for this article, or from the first author upon request.

[19] In particular, the following settings were used: (i) The initial weights of all constraints were set to 100. Changing the initial weights may influence the speed of learning, but as long as sufficient learning time is allowed, it will not influence the final grammar that is learned; (ii) An evaluation noise of 2.0 was used. Changing the evaluation noise may influence the absolute difference in weight between constraints, but will not influence the eventual performance of the grammar; (iii) The initial plasticity was set to 1.0, with 4 decrements of 0.1 in plasticity at every 100,000 replications. As explained by Boersma and Hayes (2001) with regard to their GLA for stochastic OT, starting out with a higher initial plasticity results in faster initial

Once the grammar had been learned, *Praat*'s "To output Distributions" function was used to test the predicted output of the grammar.[20]

The constraint weights that were learned are given in (14). Before this grammar is used to evaluate output candidates, noise is added to the constraint weights. In the noisy HG implementation in *Praat*, this noise is randomly selected from a normal distribution with a mean of zero. Under the default *Praat* setting, the standard deviation of the distribution is 2.[21]

(14)	*CT]$_{\text{Word}}$	101.16
	Max	98.84
	MAX-PRE-V	−1.51
	MAX-PRE-PAUSE	0.96

In (15), we show the output patterns generated by a grammar with the weights in (14). As expected, there is a close match between the observed deletion rates (on which the learning input file was based), and the deletion rates predicted by the grammar. As has been shown before, noisy HG can replicate variation rates extremely well (Coetzee 2009a; Coetzee and Pater 2011; Jesney 2007). However, as we had shown earlier, words of different frequencies are subject to deletion at very different rates. Since high frequency words contribute more to the overall deletion rate, the deletion rate predicted by the grammar learned in this section, based on the overall deletion rate in the corpus, is relatively close to the deletion rates observed for high frequency words. Low frequency words, on the other hand, show deletion rates that are considerably lower than this overall deletion rate. In the next section, we augment this grammar to take into account the difference between words of different frequencies.

(15)

Context	Observed deletion	Expected deletion
Pre-C	80 %	79.6 %
Pre-V	76 %	76.2 %
Pre-Pause	63 %	62.8 %

3.3 Incorporating the frequency bias through weight scaling

In order to apply weight scaling, the scaling factors for words of different frequencies need to be determined, and to do that, the values of the parameters (α, β, and ρ) of

learning. Decreasing plasticity later in learning results in more accurate frequency matching of the learning input. An equally good grammar could be learned by starting out with a small plasticity, but more learning time might be required.

[20] For this production-oriented simulation, we also used *Praat*'s default settings: (i) An evaluation noise of 2.0 was used—the same value used during the learning simulation; (ii) Each input type (pre-consonantal, pre-vocalic and pre-pausal) was submitted to the grammar 100,000 times, and the frequency with which each output candidate (deletion or retention) was selected was tallied.

[21] If the sum of a constraint's weight and the noise added to this weight at a particular evaluation occasion is less than zero, *Praat* resets it to zero during evaluation. This adjustment prevents a candidate from being rewarded in its H-score for violating a constraint—a negative constraint weight multiplied by the negative integer used to mark constraint violation would have increased the H-score.

the *beta* distribution associated with words of different frequencies need to be determined. We start by showing how the values of α and β are determined. As explained in Sect. 2.3, the value of α is set to the logarithm of the reference frequency—i.e., that frequency that divides the words into the frequent and infrequent sets. Following the procedure illustrated in (9) in Sect. 2.3, the log reference frequency, α, for our Columbus English *t/d*-deletion corpus was determined to be 5.1. For all words, the value of α is hence set to 5.1. The value of β is set to the log frequency of the bin to which the word belongs. For the word *and*, for instance, α is set to 5.1, and β to the log bin to which *and* belongs, namely 5.8.

As shown in Sect. 2.3, ρ only influences the size of the scaling factors and not their signs. Its role is to determine how much influence usage frequency (via weight scaling) can have on the functioning of the grammar. We propose that the value of ρ is determined by fitting the model to the data. This value therefore has to be determined separately for each language (represented by some corpus). To determine the value of ρ that results in the best fit to our data, we ran multiple simulations, keeping the values of α and β constant while increasing the value of ρ by whole number steps from 1 upwards. We then compared the weight scaled models with the baseline model without weight scaling in terms of their mean square errors relative to the observed deletion rates. The improvement of the weight scaled grammars at different integer values of ρ could then be compared, and the value of ρ could be selected where the improvement reaches its maximum.[22]

(16) Scaling factors for words of different frequencies, at different values of ρ

			ρ					
		Baseline	3	4	5	6	7	
Frequency bins	2.0	0.0	1.82	2.43	3.04	3.65	4.26	Scaling factor
	2.6	0.0	1.32	1.76	2.20	2.63	3.07	
	3.0	0.0	1.03	1.38	1.72	2.06	2.41	
	3.5	0.0	0.73	0.97	1.21	1.45	1.69	
	4.0	0.0	0.47	0.62	0.78	0.93	1.09	
	4.4	0.0	0.28	0.37	0.47	0.56	0.65	
	5.8	0.0	–0.24	–0.32	–0.40	–0.47	–0.55	

In (16), we give the scaling factors associated with words belonging to different frequency bins in our corpus at different values of ρ. As the frequency increases (top to bottom), the scaling factors decrease, corresponding to the fact that faithfulness constraints play a less important role in the evaluation of more frequent words. For the most frequent frequency bin (5.8), the scaling factor is negative, since for words in this bin α (the reference value, 5.1) is smaller than β (the log frequency of the bin, 5.8), resulting in a right-skewed *beta* distribution with a negative mode. As the value of ρ increases (from left to right), the absolute values of all the scaling factors increase, even though their signs do not change. This correlation corresponds to the fact that frequency has a larger influence (via the scaling factors) at larger val-

[22] Using whole number increments for ρ is motivated by practical considerations. If smaller increments were used, it is possible that a slightly better fit could be achieved.

ues for ρ. The "baseline" column represents the basic grammar without frequency scaling.

Once the scaling factors for words of different frequencies at different values of ρ have been determined, weight scaling can be implemented formally. We use the scenario with $\rho = 5$ as an example. The same procedure is followed for all other values of ρ. The scaling factors listed in (16) represent the amount with which the weight of each faithfulness constraint has to be increased or decreased in the evaluation of words with a specific usage frequency. For instance, when evaluating a word with a usage frequency of 2.0 the weight of all faithfulness constraints has to be increased by 3.04. When evaluating a word with a frequency of 5.8, the weight of all faithfulness constraints has to be decreased by 0.40, etc. In (17), we show the weight scaled grammars for different frequency bins when $\rho = 5$. To get these grammars we added the scaling factors from (16) to the faithfulness constraint weights of the baseline model from (14). Once these weight scaled grammars were determined, we manually edited the *Praat* grammar file for the baseline model that was learned in Sect. 4.2 above. Specifically, we created separate grammar files for each of the different frequency bins by changing the weights of the faithfulness constraints according to the scaling factor for each of the frequency bins, as reflected in (17). Once different grammar files for each frequency bin have been created, we again used *Praat*'s "To output Distributions" function to determine the deletion frequency predicted by each of these frequency scaled grammars.

(17) Frequency scaled grammars at $\rho = 5$

		Scaling factor	*CT]$_{\text{Word}}$	MAX	MAX-PRE-V	MAX-PRE-PAUSE
Baseline		0	101.16	98.84	–1.51	0.96
Frequency bin	2.0	3.04	101.16	101.88	1.53	4
	2.6	2.20	101.16	101.04	0.69	3.16
	3.0	1.72	101.16	100.56	0.21	2.68
	3.5	1.21	101.16	100.05	–0.3	2.17
	4.0	0.78	101.16	99.62	–0.73	1.74
	4.4	0.47	101.16	99.31	–1.04	1.43
	5.8	–0.40	101.16	98.44	–1.91	0.56

In (18) we show the deletion rates in pre-consonantal position predicted for a selection of frequency bins, at the different values of ρ from (16). Since frequency has no influence in the baseline grammar, the same deletion rate is expected for all frequency bins. For all of the other values of ρ, deletion rates increase as frequency increases (top to bottom), given that the scaling factors decrease as frequency increases. Lower scaling factors imply lower effective weights for faithfulness constraints, and hence higher rates of unfaithfulness. For all but frequency bin 5.8, deletion rates decrease as the value of ρ increases (left to right). These frequency bins represent words that appear less often than the reference frequency, and as shown in (16), these bins are therefore associated with positive scaling factors. Also shown in (16) is that the values of the scaling factors increase with ρ. At higher values of ρ, the faithfulness constraints will hence have higher effective weights, and therefore exert more influence on the selection of the output, with the resulting higher rates of faithfulness. Frequency bin 5.8 is the only bin with a frequency higher than the reference frequency of 5.1. As shown in (16), the scaling factors associated with this bin are hence negative, and decrease as ρ increases. As a result, for this frequency bin, deletion rates increase as ρ increases. The contribution of ρ to the model should now be clear. Higher values of ρ result in an increased contribution of frequency to the selection of the output. If a word is frequent and therefore has a higher than overall deletion rate, its deletion rate will be even higher at higher values of ρ. On the other hand, if a word is infrequent and therefore has a lower than overall deletion rate, its deletion rate will be even lower at higher values of ρ.

(18) Predicted deletion rates (%) in pre-consonantal context at different values of ρ

			ρ					
		Baseline	3	4	5	6	7	
Frequency bins	2.0	79.4	56.9	48.4	39.7	32.0	24.5	Predicted % deletion
	2.6	79.4	63.7	58.0	51.4	45.9	39.5	
	3.0	79.4	67.7	63.0	57.9	53.6	48.9	
	3.5	79.4	71.1	68.3	64.4	62.0	58.6	
	4.0	79.4	74.5	72.8	70.3	68.5	67.1	
	4.4	79.4	76.5	75.6	74.1	73.2	72.2	
	5.8	79.4	81.6	82.4	82.9	83.9	84.2	

The table in (19) compares the performance of the model at different values for ρ in terms of mean square errors.[23] For each value of ρ, we also give the percentage of improvement of the model relative to the baseline model. The performance of the model steadily increases up to a value of 5 for ρ, after which it starts declining again. Based on this, we set the value of ρ for the Columbus English *t/d*-deletion corpus at 5. Figure 4 shows the performance of the baseline model relative to a frequency scaled model with $\rho = 5$. The broken line represents the baseline model, and the solid line the frequency scaled model. The scaled model predicts a higher than overall deletion rate for words in frequency bin 5.8, and lower than overall deletion rates for other frequency bins. This figure also shows that the frequency scaled model fits the data better than the baseline model. In fact, as shown in (19), it improves on the baseline by nearly 80 %.

[23] Mean square error is calculated according to the formula $\sum_{i=1}^{n}(P_i - O_i)^2$, where P_i is the value predicted for observation i, and O_i the observed value for observation i. This value is an overall index of the deviation between the model prediction and the actually observed data. Improvement relative to the baseline model is calculated by first determining the difference in mean square error between the baseline and the model being evaluated—this difference represents the improvement of the new model relative to the baseline in terms of mean square error. This difference is then converted into an improvement percentage. For instance, to determine the improvement of a model with $\rho = 5$ relative to the baseline in (19), we first determine the difference in mean square error between the two models (i.e., $1009.7 - 208.2 = 801.5$). We then convert this to a percentage (i.e. $801.5/1009.7 \times 100 = 79.4$ %).

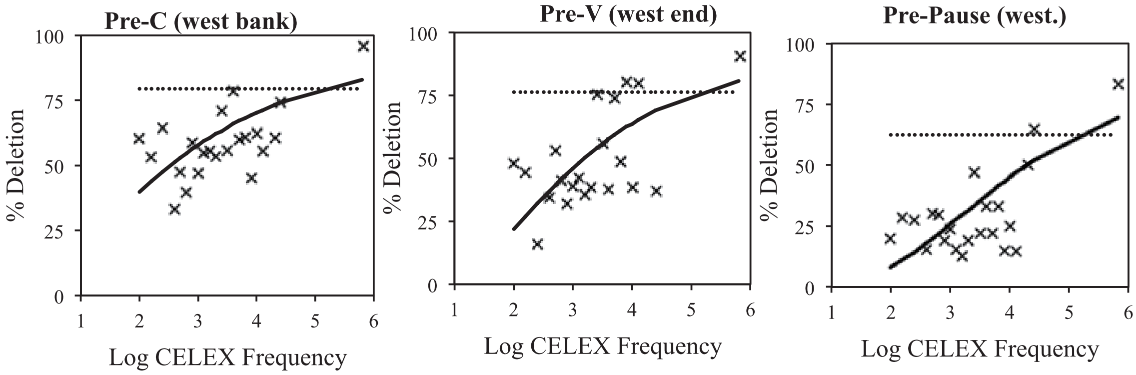


Fig. 4 Observed and predicted *t/d*-deletion rates in Columbus English. The *broken line* indicates the predictions based on the baseline, unscaled HG. The *solid line* shows the predictions based on the frequency weighted HG with a ρ-value of 5

(19) Mean square errors and percentage of improvement relative to the baseline, unscaled grammar at different values of ρ

		ρ				
	Baseline	3	4	5	6	7
Mean Square Error	1,009.7	354.0	280.4	208.2	218.8	425.9
Improvement Percentage Relative to Baseline	0 %	64.9 %	72.2 %	79.4 %	78.3 %	57.8 %

The fact that the scaled model fits the data better is not surprising—the scaled model incorporates one more parameter (frequency) than the baseline model, and given that frequency significantly impacts application of *t/d*-deletion, it is to be expected that a model with this additional parameter will fit the data better. To determine whether this improvement of 80 % is sufficient to warrant the additional complexity we used the Akaike Information Criterion (AIC; Akaike 1973, 1983). Roughly speaking, AIC is an estimate of the amount of information lost when using a specific model relative to the true model. A smaller AIC value associated with a model therefore indicates that the model more closely approximates the true model. To calculate the AIC for the baseline and scaled models, we use the partial AIC derivation (Burnham and Anderson 2004:268–269). The formula used is given in (20) where *MSE* is the mean square error associated with a model and k is the number of parameters used in the model. A model that fits the data better will have a smaller *MSE* and hence, all else being equal, a smaller (or better) AIC. On the other hand, the larger the number of parameters included in a model, the larger k will be. All else being equal, a model with more parameters will therefore have a higher (or less good) AIC than a model with fewer parameters. AIC therefore rewards a model for a better fit with the data (lower *MSE*), but penalizes a model for including more parameters (higher k), so that AIC gives a measure of the tradeoff between model complexity and model fit. The value of n is the number of observations in the dataset being modeled.

(20) $\mathrm{AIC} = n \log_e(MSE) + 2k$

In calculating AIC for the baseline and frequency scaled models, we assume that each of the constraints in our HG grammar counts as one parameter. The baseline model therefore has 4 parameters. The frequency scaled model has one additional parameter (i.e., 5) due to the addition of the frequency scaling factor to this model. Using the *MSE*s for the respective models reported in (19) above, the formula in (20), and setting $n = 65$ (since there are 65 total data points in the corpus), the AICs for the two models can be calculated: $\text{AIC}_{\text{Baseline}} = 457.6$, $\text{AIC}_{\text{Scaled}} = 357.0$. As Burnham and Anderson (2004:271) note, a model with an AIC that is more than 10 units larger than the best model has "essentially no support". Given that the frequency scaled model has an AIC that is 100 units smaller than the baseline model, we can hence conclude that the additional complexity of the scaled model is well warranted by the better fit that this model achieves relative to the baseline model.

4 Geminate devoicing in borrowings in Japanese

4.1 The data

In this section, we present another case study to show the generality of the model that we developed above. Although Japanese native phonology does not tolerate voiced geminates, these sounds have been introduced into Japanese via borrowings. Due to Japanese coda restrictions (Itô 1988), closed syllables are frequently borrowed with an epenthetic vowel. Additionally, when the coda consonant in a borrowed word is preceded by a lax vowel, the consonant is often geminated (Katayama 1998). When the coda consonant is also a voiced obstruent, the combination of these processes results in a voiced geminate. In words that contain another voiced obstruent, the geminate optionally devoices, as in the examples in (21) (all examples from Kawahara 2006:538).

(21)	guddo	∼	gutto	'good'
	beddo	∼	betto	'bed'
	deibiddo	∼	deibitto	'David'
	doggu	∼	dokku	'dog'
	baggu	∼	bakku	'bag'
	doraggu	∼	dorakku	'drug'
	biggu	∼	bikku	'big'

This optional devoicing in loanwords has received a lot of attention in recent years so that the factors that condition its application are now well understood. We refer the reader to the literature for a discussion of these factors (Crawford 2009; Kaneko and Iverson 2009; Kawahara 2005, 2006, 2008, 2011a, 2011b; Nishimura 2003, 2006; Tanaka 2009 and references cited there). Our focus here will be on how this process is influenced by usage frequency. In two recent studies, Kawahara has found a strong positive correlation between geminate devoicing and word frequency (Kawahara 2011a, 2011b). We will develop an account of the results of Kawahara (2011a) here. We summarize the most important aspects of his results below, and refer the reader to the original paper for more details on the design of the experiment.

Kawahara presented 52 native Japanese speakers with 28 loanwords like those in (21) with the task of rating the naturalness of a pronunciation in which the voiced geminate has been devoiced. Participants indicated their responses on a 5-point scale, with [5] corresponding to "very natural", and [1] to "very unnatural". The raw usage frequency of each loan word token was taken from the Amano and Kondo Japanese lexical corpus (Amano and Kondo 2000), and log-transformed. Figure 5 plots the average naturalness rating that each token received against its log-transformed frequency. Performing a linear regression on these data confirms that log frequency and naturalness are positively correlated ($r^2 = 0.43$, $p < 0.01$).

The best way to collect data on devoicing rates in actual speech production would be to investigate the prevalence of devoicing in a large, phonetically transcribed corpus of spoken Japanese—similar to how we investigated the prevalence of *t/d*-deletion in the *Buckeye Corpus* above. Unfortunately, no such corpus exists for Japanese that is large enough to contain enough examples of loanwords. A second option would be to conduct a production experiment, designed to collect data on loanwords. Participants in such experiments usually use a rather formal speech style in which optional processes, such as geminate devoicing, are often inhibited. We therefore work under the assumption that naturalness ratings such as those in Kawahara (2011a) originate in the same grammar that governs speech production, and that these naturalness ratings therefore also reflect the frequency with which devoicing will apply to the loanwords in actual speech. Even if this is accepted, it is still necessary to convert the 5-point naturalness scale to devoicing rate in some manner. Little is known about how naturalness ratings are related to production patterns (though see Kempen and Harbusch 2008 for some ideas involving syntactic data), and we therefore explored several different options for transforming the naturalness ratings of Kawahara (2011a) to devoicing rates. In all of the transformations that we explored, the positive correlation between frequency and rate of devoicing was preserved. We report here on only one of these transformations, a simple linear transformation.[24] This is the transformation on which our model had the best performance.

In order to transform the natural ratings to devoicing rates, we made the assumption that a rating of [5] corresponds to a token that is always produced with devoicing, a rating of [4] to a token that is produced with devoicing four-fifths of the time (i.e., with 80 % devoicing), etc. The formula used to transform the naturalness ratings is given in (22). Figure 6 plots the deletion rate under this transformation against the log frequency of the tokens. As this figure shows, the correlation between frequency and devoicing is preserved under this transformation ($r^2 = 0.43$, $p < 0.01$).

[24] Specifically, in addition to the linear transformation defined in (22), we also used an exponential and sigmoid transformation. The formulas used in these two transformations are given below. Under both of these transformations, the positive correlation between frequency of devoicing and usage frequency is preserved: exponential: $r^2 = 0.34$, $p < 0.01$; sigmoid: $r^2 = 0.41$, $p < 0.01$.

Let r be the average naturalness rating that some token t received, and $devoice(t)$ the rate of devoicing in token t. Let $norm^r$ be the standardized value of r. Then:

$$\text{Exponential transformation} \quad devoice(t) = \left(\frac{e^r}{e^5}\right)(100)$$

$$\text{Sigmoid transformation} \quad devoice(t) = \left(\frac{1}{1 + r^{-norm^r}}\right)(100)$$

Fig. 5 The relation between frequency and devoicing in Kawahara (2011a). The *x-axis* represents log-transformed frequencies from Amano and Kondo (2000). The naturalness rating of devoicing is plotted on the *y-axis*. The *line* indicates the best-fit linear regression line

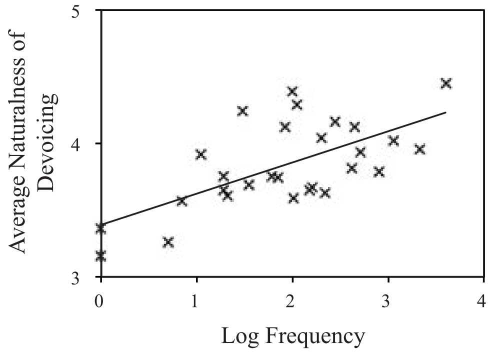
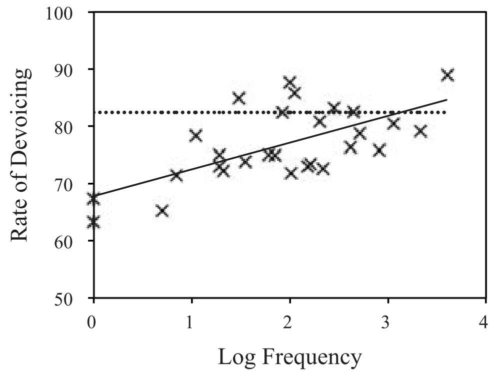


Fig. 6 The relationship between frequency and rate of geminate devoicing under a simple linear transformation of the natural ratings from Kawahara (2011a). The *solid line* represents the result of a linear regression. The *broken line* represents the overall devoicing rate



(22) Let r be the average naturalness rating that some token t received, and $devoice(t)$ the rate of devoicing in token t. Then:

$$devoice(t) = \left(\frac{r}{5}\right)(100)$$

To determine the overall devoicing rate under this transformation, we created a corpus assuming that each loanword appears in the corpus with its frequency in Amano and Kondo (2000). The loanword /budda/ 'Buddha', for instance, has a frequency of 99 in Amano and Kondo, and /budda/ was hence represented 99 times in our corpus. Each token was represented with devoicing according to the transformation given in (22). Devoicing in /budda/ received an average rating of 4.39. Performing the transformation on this score results in a devoicing rate of 87.8 %, and this percentage of the 99 occurrences of /budda/ in the corpus was hence represented with devoicing (i.e., 87 tokens with and 12 without devoicing). The same was done for all loanwords in the corpus. The overall devoicing rate in the corpus was then calculated to be 82.4 %. This overall rate is marked with a broken line in Fig. 6. As with the overall rate of *t/d*-deletion in the *Buckeye Corpus* (see Fig. 1), the overall rate of devoicing is closer to the rate observed for the more frequent words.

In the rest of this section, we develop an account for this transformed corpus. As with *t/d*-deletion, we first develop a purely grammatical model based on the overall devoicing rate in the corpus. We then augment this model with weight scaling according to the method described above in Sects. 2.2 and 2.3.

4.2 A grammatical account

We rely on the three constraints in (23)—see Nishimura (2003), Kawahara (2006) and Pater (2009) for analyses using slightly different constraints. As with *t/d*-deletion, we used the noisy HG learning algorithm in *Praat* to learn the weights associated with these constraints. The learning input file contained 100 tokens, with the proportion of tokens represented with devoicing determined by the overall rate of devoicing in the corpus (i.e., 82 out of the 100 tokens).[25] The learning file was submitted to *Praat*'s learning algorithm, using all of the default settings in *Praat*. The constraint weights that were learned are given in (24). Once the grammar had been learned, the "To output Distributions" function in *Praat* was used to determine the predicted rate of devoicing for the learned grammar. This returned an expected devoicing of 82.2 %, which very closely matches the observed deletion rate of 82.4 % in our corpus. However, as before, this grammar produces devoicing at the overall devoicing rate in the corpus, and treats all words of all frequencies the same. In the next section, we augment this account to incorporate the contribution of usage frequency to the rate of devoicing.

(23)	*GEMINATE	Assign one violation mark for every consonant linked to two timing slots.
	*VOICEDOBS	Assign one violation mark for every voiced obstruent.
	IDENT[voice]	Assign one violation mark for every output segment that has a different specification for the feature [voice] than its input correspondent.

(24)	*GEMINATE	100.0
	*VOICEDOBS	101.3
	IDENT[voice]	98.7

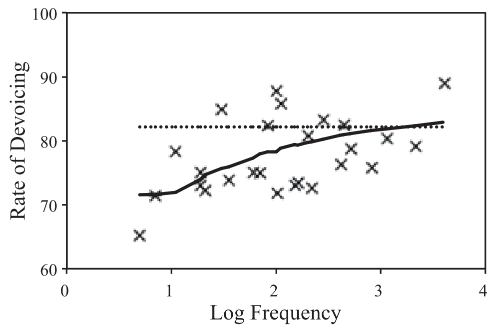
4.3 Incorporating the frequency bias through weight scaling

We incorporate the contribution of usage frequency into the model developed in the previous section in the same way as we did for *t/d*-deletion in Sect. 3.3. What is required is to scale the weight of the faithfulness constraint IDENT[voice] up for infrequent words so that they are more likely to be treated faithfully, and conversely to scale the weight of IDENT[voice] down for frequent words. First, we determined the reference point between frequent and infrequent words according to the method described in (9). In total, our corpus contains 11,000 tokens. The two most frequent words account for over half of the 11,000 tokens (/bagudaddo/ 'Baghdad', frequency: 3951;[26] /baggu/ 'bag', frequency: 2103). The reference point is hence halfway between the log frequency of /baggu/ (3.32) and the log frequency of the next most

[25] The learning input file is available as "supplementary material" on the Springer link for this article, or from the first author upon request.

[26] The high frequency of /bagudaddo/ in Amano and Kondo (2000) is a result of their frequency counts being taken from a corpus of newspapers including the time after the American invasion of Iraq. Although it is not clear that /bagudaddo/ will still have such a high frequency for the average Japanese speaker,

Fig. 7 Observed and predicted devoicing rates. The *broken line* indicates the prediction based on the basic, unscaled HG. The *solid line* shows the predictions based on the frequency weighted HG



frequent word, /bajji/ 'badge' (3.05),[27] or 3.19. With this reference value in hand, the *beta* distribution associated with each word can now be determined. For all words, the value of α is the reference log frequency of 3.19, and the value of β is the log frequency of the specific word. The value of the range parameter ρ is set to maximize the fit of the model's predictions with the data being modeled exactly as it was done for *t/d*-deletion above in Sect. 3.3. For the corpus with which we are working here, this value for ρ was found to be 1.

Once the value of ρ for a corpus has been determined, the scaling factor associated with each word can be determined. The weight of the faithfulness constraint can then be scaled according to this scaling factor for each word, and the predicted rate of devoicing can be determined for individual words using the "To output Distributions" function in *Praat*. Figure 7 shows how the baseline, unscaled HG model compares with the frequency scaled model. The broken line plots the prediction of the baseline model, and the solid line the prediction of the scaled model. This figure clearly shows that the scaled model fits the data better. This is confirmed by the mean square errors (*MSE*) of each of the models. The *MSE* of the baseline model is 52.7, and that of the scaled model 24.5, so that the scaled model represents a 53.5 % improvement over the baseline model.[28] As with the *t/d*-deletion account above, this improved fit is to be expected, given that the scaled model contains an extra parameter (frequency) that is known to be relevant. In order to determine whether the additional complexity of the scaled model is warranted by the increase in fit, we calculated AIC values for the baseline and scaled models, as we did above for *t/d*-deletion. The AIC value for the baseline model was found to be 109.1, and that for the scaled model was found to be 91.2. Since the scaled model has an AIC that is more than 10 units smaller than the baseline model, we conclude with confidence that there is sufficient support for the additional complexity of the scaled model.

we opted not to adjust its frequency for the purposes of this paper. The participants in Kawahara's experiment were mostly university students who were probably familiar with this event, so that /bagudaddo/ would have had a high frequency for them. The fact that /bagudaddo/ pronounced with devoicing, i.e., as [bagudatto], received a high naturalness rating in Kawahara (2011a) suggests that this might be correct.

[27] Following standard conventions in the literature on Japanese phonology, we use /j/ here for the affricate /dʒ/.

[28] As explained in footnote 24, we also explored an exponential and sigmoid transformation of the naturalness ratings. Frequency scaled models for corpora based on these transformations also performed better than baseline models, although the improvement was slightly less good than what we found for the linear transformation reported in the text. Improvement of the frequency scaled model over the baseline model was as follows: exponential transformation = 49.0 %; sigmoid transformation = 42.1 %.

5 Discussion

The model of phonological variation that we developed above incorporates the effects of usage frequency into a generative phonological grammar. Two case studies have shown that this model performs better than a model based on grammar alone. In this section, we discuss some general properties of our model, as well as some still unresolved and underexplored issues.

5.1 Grammar dominance

Although the model that we propose in this paper allows non-grammatical factors such as usage frequency to influence phonological variation, it is a grammar dominant model. Grammar sets the limits of what patterns of variation are possible, and all that the frequency can do is to determine how variation is realized within these limits. The dominance of grammar realizes itself in both universal terms and in the grammars of individual languages.

First consider the universal aspects of grammar dominance. In HG (as in OT), Universal Grammar is represented in the constraint set. Classic OT (Prince and Smolensky 1993, 2004) assumes that the constraint set is universal, so that the grammar of every language contains exactly the same constraints. From this assumption it follows that there are certain logically possible grammatical constraints that do not exist, and if some constraint does not exist then some logically possible grammatical patterns cannot be expressed. For example, in our analysis of *t/d*-deletion, we proposed positional MAX constraints for pre-vocalic and pre-pausal position, but argued that no such positional constraint exists for pre-consonantal position. If this is a true restriction on the constraint set, deletion in pre-consonantal context will always violate only a subset (MAX) of the faithfulness constraints violated by deletion in pre-vocalic (MAX, MAX-PRE-V) or pre-pausal (MAX, MAX-PRE-PAUSE) position. In (25), we show the consequences that this stringency relationship has for the H-score of deletion candidates in the different contexts. The H-score of deletion in pre-consonantal position will always be higher than that of deletion in the other two contexts. This effect cannot be overridden by weight scaling in our model, since we assume that all faithfulness constraints are scaled by the same factor (i.e., the scaling factor is not indexed to a particular faithfulness constraint). In any language, for a word of any frequency, deletion will always be most likely in pre-consonantal position. All that frequency can do is to increase or decrease the likelihood of deletion in all three contexts, but it will do so to the same extent in all three contexts.[29]

[29] Since a process cannot apply at a rate of higher than 100 %, this statement has to be qualified. Imagine a grammar where pre-consonantal context has a base deletion rate of 80 % and pre-pausal context of 50 %. Deletion in pre-consonantal position can be increased by at most 20 % by the contribution of scaling factors. The same holds for scaling factors that reduce the application of a simplification process and the floor of application, 0 %.

(25)

	Max	Max-Pre-V	Max-Pre-Pause	H
/wɛst bæŋk/ → [wɛs_ bæŋk]	–1			–*w*(Max)
/wɛst ɛnd/ → [wɛs_ ɛnd]	–1	–1		–*w*(Max) – *w*(Max-Pre-V)
/wɛst/ → [wɛs_]	–1		–1	–*w*(Max) – *w*(Max-Pre-Pause)

A similar point can be made with regard to geminate devoicing in Japanese. In our analysis, we assumed a markedness constraint that penalizes voiced obstruents, but no constraint that penalizes voiceless obstruents. If no constraint against voiceless obstruents exists, a language that has context-free voicing of obstruents (whether as a categorical or variable process) is impossible. It does not matter how frequent a word is: Since this process is ruled out by the grammar, it is predicted never to be observed.

Grammar also takes precedence over usage frequency at the level of individual languages. In the grammar developed for Columbus English above, the weight of Max-Pre-V (-1.51) is lower than that of Max-Pre-Pause (0.96), corresponding to the fact that this dialect of English shows more deletion in pre-vocalic than pre-pausal position. Since the weights of all faithfulness constraints are scaled by the same amount, the relative difference in the effective weights of Max-Pre-V and Max-Pre-Pause will be preserved under all scaling conditions. No matter how frequent a specific word is, on average a pre-vocalic deletion candidate will have a higher H-score than a pre-pausal deletion candidate. The grammar of Columbus English dictates that deletion is more likely in pre-vocalic context, and frequency cannot override this.

This dominance of grammar depends on the assumption that at a given instance of using the grammar (evaluation of a specific word, at a specific instance) the weights of *all* faithfulness constraints are scaled by the same amount. If weight scaling could variably affect different faithfulness constraints, the dominance of grammar could be lost. In this regard, our proposal diverges from the related proposal made by Boersma and Hayes (2001: Appendix C). Their model is developed in stochastic OT, and they therefore assume constraint ranking rather than weighting. They propose that the ranking values of some constraints can be changed in different speech situations. But crucially, they propose that some constraints can be ranked higher, others lower, and that constraint rankings do not have to be changed by the same amount. As a consequence, their model does not have the property of grammar dominance.

The dominance of grammar is also not a property of other models of phonological variation. In some implementations of usage-based models (Bybee 2001, 2006, 2007; etc.), or exemplar models (Gahl and Yu 2006 and papers therein; Pierrehumbert 2001; etc.), no formal distinction is made between grammatical and non-grammatical factors. In fact, in describing usage-based grammar, Bybee first defines the usage-based conceptualization of grammar as "the cognitive organization of one's experience with language" (Bybee 2006:711). Later on the same page she describes how this organization is done as follows: "… the general cognitive capabilities of the human brain, which allow it to categorize and sort for identity, similarity, and difference, go to work on the language events a person encounters, categorizing and entering in memory these experiences." Grammar is the result of cognitive organization achieved with general cognitive abilities, not with grammar or language specific abilities. Exactly the same cognitive abilities that organize our experience with social interactions and with our physical environment organize our experience with language. No formal distinction is made between how language and other aspects of our experience are processed or stored in the mind. If a child acquiring a language were to be exposed to a set of experiences where deletion happens to be observed more often in pre-vocalic than pre-consonantal context, the general abilities of the mind to classify would notice this pattern, and codify this as the grammar. This view of grammar is fundamentally different from the type of approach that we advocate above. Under our approach, there are language specific cognitive capacities (Universal Grammar represented in the constraint set, as well as the principles for how constraints interact via their weights). Language is processed according to these principles and not with general cognitive capabilities. This places a limit on the types of grammars that can be learned. As we showed above, the assumptions about Universal Grammar under which we operate imply that no grammar that produces more deletion in pre-vocalic than pre-consonantal context is possible.

More research is necessary to determine to what extent certain types of grammars are truly impossible. A long tradition of typological research has established strong universal patterns across languages, a result that could be interpreted as favoring a system that includes a strong Universal Grammar. Recent research in artificial grammar learning has also shown that linguistic patterns that counter such universal trends are either unlearnable or at least not easily learnable (Carpenter 2006, 2010; Coetzee 2009b; Moreton 2008; Pater and Tessier 2006). On the other hand, there are also unambiguous examples of languages with grammars that counter universal trends (Coetzee and Pretorius 2010; Hyman 2001), showing that it should be possible for language learners to acquire grammars that do not fit neatly into the limits of Universal Grammar. Along similar lines, Bybee (2002:275) shows that in one dialect of English some words, under some circumstances, show more word-final *t/d*-deletion in pre-vocalic than pre-consonantal context. With conflicting data from the current literature it is impossible to choose definitively between a model with grammar dominance and a model in which grammar is afforded no special place. However, given that the evidence for strong universal tendencies is currently more copious than evidence for linguistic systems that counter these tendencies, we opt for the more restrictive model where Universal Grammar places limits on possible languages.

5.2 What processes are influenced by frequency?

In the model that we developed above, only the weights of faithfulness constraints are affected by frequency. From this restriction it follows that all and only those phonological processes that violate some faithfulness constraint will be affected by frequency scaling. In this paper, we have focused on two such processes—consonant cluster simplification and geminate devoicing. In both of these processes, it is the relative weight of some faithfulness constraint(s) (MAX/MAX-PRE-V/MAX-PRE-PAUSE or IDENT[voice]) and some markedness constraint(s) (*CT]$_{\text{Word}}$ or *VOICEDOBS/*GEMINATE) that determines whether the process applies. Since weight scaling affects the weights of the faithfulness constraints, it affects the relative weights of faithfulness and markedness constraints, and hence the likelihood that these processes will apply.

The processes on which we focused in this paper are both examples of simplification or reductive processes—i.e., the form that has undergone the process is in some sense articulatorily simpler or more reduced than the input. There is ample evidence from the literature that such reductive processes are indeed subject to the influence of frequency as predicted by the model that we developed above. In Sect. 1.3, we provided references for word-final obstruent deletion, unstressed vowel deletion, obstruent devoicing, and *l*-vocalization as examples.

However, the application of augmentation processes also depends on the relative weights of markedness and faithfulness constraints. In a language that avoids tautosyllabic consonant clusters via epenthesis, for instance, the application of epenthesis (arguably not a reductive process) depends on the relative weights of the anti-cluster markedness constraint *COMPLEX and the anti-epenthesis faithfulness constraint DEP. In a language in which such a process applies variably, the model developed above would predict that epenthesis will be observed more often in more frequent words than in less frequent words. Although there are examples in the literature that discuss such variable augmentation processes (see Auger 2001 on variable epenthesis in Vimeu Picard; Nevins 2007 on variable epenthesis in Brazilian Portuguese), we do not know of any example where the application of these processes is discussed in relation to usage frequency. If indeed variable augmentation processes are affected by frequency in the same way as variable reductive processes, it would be additional evidence for the model that we developed above. On the other hand, if augmentation processes are not affected by frequency in the same manner, the model would need to be revised in some way in order to differentiate between augmentation and reduction processes.

Given that only the weights of faithfulness constraints are affected by frequency scaling in the model developed above, variable phonological phenomena that do not depend on faithfulness constraints should not be affected by frequency in the same way. Under the assumption that there are no faithfulness constraints for prosodic structures (McCarthy 2003: Sect. 6), variable prosodification is not expected to be sensitive to frequency. As an example, consider Hammond's analysis of variable stress placement in Walmatjari (Hammond 1994; see also Anttila 2002b). In Walmatjari, tri-syllabic words are either stressed on the initial or the second syllable so that the underlying form /kaŋani/ 'carried' can be realized as [káŋani] or [kaŋáni]. Neither surface form violates any faithfulness constraints. The selection between the candidates is hence done by markedness constraints alone—in Anttila's account, by the constraints TROCHEE, FTBIN and *LAPSE (Anttila 2002b). Since only faithfulness constraints are sensitive to weight scaling, and since faithfulness constraints are irrelevant in the choice between these two variants, this choice cannot be influenced by frequency in the model developed above. We do not know of any literature that discusses such variable phenomena in relation to usage frequency, and we therefore cannot determine whether this prediction is borne out by actual data. If, in fact, processes such as these are also sensitive to frequency, the model developed above will need to be augmented in some way to account for this.

There is another set of variable phenomena that are known to be sensitive to usage frequency, but that are not accounted for in the model that we developed above. Morphological regularization (analogical leveling) is less likely to apply to more frequent words. As an example, Bybee (1985:119–120; also Hooper 1976) shows that regularization of the English past tense is more likely to apply to infrequent words than to frequent words—a regular past tense form for infrequent *weep* (*weeped* instead of *wept*) is more likely than for frequent *keep* (*keeped* instead of *kept*). See also Phillips (1984, 2001) for more similar examples. Processes such as these are governed by the relations between morphologically related words, and hence by output-output correspondence constraints (Benua 2000) rather than by regular faithfulness constraints. Although our model cannot account for the role of frequency in these types of phenomena, the model could be extended in a straightforward manner to do so. An infrequent word (such as *weep*) is more likely to have a uniform paradigm. This implies that the OO-correspondence constraints that are responsible for enforcing paradigm uniformity should have higher weights in the evaluation of infrequent words than in the evaluation of frequent words. In the same way that we scale the weight of faithfulness constraints up for infrequent words, the weights of OO-correspondence constraints can be scaled up for infrequent words. However, we leave full development of this option for future research.

5.3 Modeling acquisition

In Sects. 3.2 and 4.2, we illustrated how a variable grammar can be learned using the noisy HG learning algorithm implemented in *Praat*. We also showed how this model can be augmented to account for the influence of usage frequency on variation. Two more questions need to be considered in this regard: (i) What predictions does this approach make with regard to the acquisition of variable phonological processes, and (ii) do these predictions correlate with how variable processes are acquired in reality? Although both of these questions are worth considering, we also want to make explicit that our goal in this paper is not to model the actual acquisition process of variable phenomena, but rather to show what a grammatical model would look like that can account for the variation observed in speech, and to show that such a grammatical model is in principle learnable. The goal of learnability theory is not to model how language is actually acquired, but to show whether a specific grammar is learnable from a given set of data (Pullum 2003:432.) Although we consider the possible implications of our model for acquisition, we do not believe that the value of our model crucially depends on how well it models actual acquisition processes.

We first want to set aside two simplifying assumptions that we made, and that do not constitute claims about actual acquisition. We assume that the learner has access to the correct underlying form of the words encountered. In the English *t/d*-deletion case, for instance, upon hearing an utterance like [wɛs bæŋk] for ‘west bank’, we assume that the learner knows that the underlying form of the first word in the utterance is really /wɛst/. This assumption is part of all of the main learning algorithms used in phonology (Boersma and Hayes 2001:51; Tesar and Smolensky 1998:237). The learning discussed here is hence learning at a later stage of acquisition, after underlying forms have already been acquired. For a development of the formal mechanisms involved in learning underlying forms in a constraint-based grammar, see Tesar and Smolensky (1996:40–44) and especially Merchant and Tesar (2005) and Tesar (2006).

The second simplifying assumption has to do with the role of usage frequency during grammar learning. We modeled the grammar learning stage above as if usage frequency of individual words plays no role during grammar learning, since we augmented the grammar with weight scaling only after the grammar has been learned (see also Hayes and Londe 2006 for a similar two-stage approach to learning). Another option that should be explored is one where usage frequency is incorporated into the grammar learning stage itself.

The noisy HG learning algorithm implemented in *Praat* is an error-driven learning algorithm. The basic steps in the learning process are: (i) The learner receives a learning input (a surface form produced by an adult); (ii) the learner determines the underlying form of the learning input, and submits this underlying form to his/her current grammar; and (iii) the learner compares the output generated by his/her current grammar to the learning input. If these two forms differ (i.e., if the learner's grammar generates an error), the learner adjusts his/her grammar to increase the likelihood that the grammar will generate an output identical to the learning input. In step (ii) of the learning cycle, the grammar is used to generate an output. In our modeling of learning above, this step did not include weight scaling. An alternative model of acquisition could incorporate weight scaling during this stage of grammar learning. The final state of the grammar that will be learned if weight scaling is incorporated during learning will be comparable to the final weight scaled grammars that we developed above. The most important difference between these two approaches is expected to be in the path of acquisition—i.e., how the grammar changes slowly during the learning period.

Although we did not incorporate frequency scaling during learning in our model, we can speculate about what would be expected from a model in which this is done. We use *t/d*-deletion as an example, but we expect the same basic pattern to be observed also in the acquisition of other variable processes. During the earlier stages of learning, when the learner has not yet built up a large corpus of learning inputs, chances are that the learner would have encountered mostly more frequent words. A child learning English, for instance, is more likely to hear a frequent word like 'want' (CELEX log frequency = 4.1) than an infrequent word like 'gourd' (CELEX log frequency = 0.9). Since more frequent words have higher deletion rates overall, and since the child is expected to hear mostly more frequent words, the corpus of learning inputs to which the child is exposed will have a higher *t/d*-deletion rate than the actual, complete adult production corpus. If the child aims to replicate the deletion rate in the learning corpus that he/she is exposed to, we would expect the child to show a higher overall deletion rate than what adults actually produce overall. This prediction agrees with the fact that child speech is often characterized by more reduction and simplification than adult speech.

Additionally, since during early acquisition the child will mostly be exposed to words from the higher end of the frequency spectrum, the range of the frequency distribution in the child's learning corpus is expected to be smaller than that in the actual adult speech corpus. (The range between the highest and lowest frequency words in the child's corpus is expected to be smaller than that in an adult's speech corpus.) In the model that we developed above, weight scaling is done based on how much the usage frequency of a specific word differs from the reference frequency in the corpus. In the child's early learning corpus, the usage frequencies are expected to differ less than in the adult corpus. The expectation is hence that usage frequency will have less of an influence during the early stages of acquisition than in an adult grammar. During early acquisition, all words are expected to be treated more or less the same. Only during the later stage of acquisition will the difference between how frequent and infrequent words are treated emerge more clearly. To the best of our knowledge, there is no study that specifically investigates how usage frequency interacts with first language acquisition of variation. There is, however, suggestive evidence from second language acquisition that learners make less fine distinctions in terms of usage frequency than native speakers. Lacoste (2008) studies the acquisition of standard Jamaican English by Jamaican primary school children, with a focus on words that end in *-Ct/-Cd* clusters (i.e., exactly the words to which *t/d*-deletion could apply). She shows that the teachers make at least a three-level distinction in terms of usage frequency (2008:198), while children in the early stages of acquisition make only a two-level frequency distinction (2008:190).

Ultimately, more research is necessary to probe in detail how children acquire variation and to track specifically how the production of individual words changes during the course of acquisition. Similarly, the learning algorithm needs to be augmented to include weight scaling. Only once both of these things have been done will it be possible to go beyond speculation with regard to how variation is acquired, and with regard to how well the predictions of the model developed above matches the actual acquisition trajectory.

5.4 Final remark: integrating generative and usage-based grammars

In this paper, we developed a model of phonological variation that incorporates influences from both grammatical and non-grammatical factors. Our model retains some of the core characteristics of a classic generative grammar, while also embracing insights from usage-based and exemplar models of grammar. In the phonological literature, the generative approach and the usage-based/exemplar approaches have often been presented as opposites and as incompatible with each other. We believe this to be a false dichotomy. Not only is it possible to integrate these approaches, but such an integration also enables phonological theory to account better for many phenomena than what either of the two approaches could do in isolation. If such an integration is indeed the correct route to go, then future research will have to focus on two issues. First, the proper way to integrate the contributions from the two types of models needs to be determined. This paper contains one proposal, and the success of this proposal leads us to believe that it has merit. But other ways of integration are possible, and more research is necessary to determine all of the viable options, and to evaluate their success. Secondly, more targeted data collection would need to be performed. The data on phonological variation that are currently available are usually not suited to address the questions raised by an integrated model such as that proposed in this paper. We hope that the line of research reported in this paper will stimulate research into these issues.

Acknowledgements The ideas expressed in this paper were presented in various forms at NAPhC 5, NELS 38, NELS 41, the University of Michigan, the University of Massachusetts, Michigan State University, Stanford University, and SUNY Stony Brook. The feedback and reaction of the audiences at these events contributed significantly to the development of our thoughts. This work has also been discussed in detail with many individuals, and we acknowledge our gratitude for their contribution. This list includes Joe Pater, John McCarthy, John Kingston, Anne-Michelle Tessier, Pam Beddor, San Duanmu, Ricardo Bermúdez-Otero, William Labov, Paul Smolensky, Matt Goldrick, Colin Wilson, Kevin McGowan, and Susan Lin. We also acknowledge the help of Amelia Compton in running many of the *Praat* simulations in this paper. The three reviewers and the associate editor similarly helped us to improve the paper and to express our ideas more clearly. As always, any remaining errors and shortcomings are our own.

References

Akaike, Hirotugu. 1973. Information theory as an extension of the maximum likelihood principle. In *Second international symposium on information theory*, eds. Boris N. Petrov and Frigyes Csaki, 267–281. Budapest: Akademiai Kiado.

Akaike, Hirotugu. 1983. Information measures and model selection. *International Statistical Institute* 44: 277–291.

Albright, Adam. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology* 26: 9–41.

Amano, Shigeaki, and Tadahisa Kondo. 2000. *NTT database series: Lexical properties of Japanese*, 2nd release. Tokyo: Sanseido.

Anttila, Arto. 1997. Deriving variation from grammar. In *Variation, change and phonological theory*, eds. Frans Hinskens, Roeland van Hout, and Leo Wetzels, 35–68. Amsterdam: John Benjamins.

Anttila, Arto. 2002a. Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory* 20: 1–42.

Anttila, Arto. 2002b. Variation and phonological theory. In *Handbook of language variation and change*, eds. Jack K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 206–243. Oxford: Blackwell.

Anttila, Arto. 2006. Variation and opacity. *Natural Language & Linguistic Theory* 24: 893–944.

Anttila, Arto. 2007. Variation and optionality. In *The Cambridge handbook of phonology*, ed. Paul de Lacy, 519–536. Cambridge: Cambridge University Press.

Anttila, Arto, Vivienne Fong, Stefan Benus, and Jennifer Nycz. 2008. Variation and opacity in Singapore English consonant clusters. *Phonology* 25: 181–216.

Auger, Julie. 2001. Phonological variation and Optimality Theory: Evidence from word-initial vowel epenthesis in Vimeu Picard. *Language Variation and Change* 13: 253–303.

Baayen, R. Harald, Richard Piepenbrock, and Leon Gulikers. 1995. *The CELEX Lexical Database (CD-ROM)*. Philadelphia: Linguistic Data Consortium.

Baese-Berk, Melissa, and Matt Goldrick. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24: 147–185.

Bane, Max. to appear. A combinatoric model of variation in the English dative alternation. In *Proceedings of the 36th annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Bane, Max. 2011. Deriving the structure of variation from the structure of non-variation in the English dative. In *Proceedings of the 28th annual meeting of the West Coast conference on formal linguistics*, eds. Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, 42–50. Somerville: Cascadilla Press.

Bayley, Robert. 1995. Consonant cluster reduction in Tejano English. *Language Variation and Change* 6: 303–326.

Bayley, Robert. 2002. The quantitative paradigm. In *The handbook of language variation and change*, eds. Jack K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 117–141. Oxford: Blackwell.

Bell, Alan, Jason Brenier, Michelle Gregory, Cynthia Girand, and Daniel Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60: 92–111.

Benua, Laura. 2000. *Phonological relations between words*. New York: Garland.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. *University of Amsterdam Institute of Phonetic Sciences Proceedings* 21: 43–58.

Boersma, Paul. 2008. Emergent ranking of faithfulness explains markedness and licensing by cue. Ms., University of Amsterdam. www.fon.hum.uva.nl/paul/papers/EmergeFaith.pdf. Accessed 11 July 2012.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45–86.

Boersma, Paul, and Joe Pater. 2008. Convergence properties of a Gradual Learning Algorithm for Harmonic Grammar. Ms., University of Amsterdam and University of Massachusetts, Amherst. www.fon.hum.uva.nl/paul/papers/boersmaPaterHGGLA.pdf. Accessed 11 July 2012.

Boersma, Paul, and David Weenink. 2009. *Praat: Doing phonetics by computer version 5.1.20.* [Computer Program.] http://www.praat.org. 31 Accessed October 2009.

Browman, Catherine P., and Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in laboratory phonology I: Between the grammar and physics of speech*, eds. John Kingston and Mary E. Beckman, 341–376. Cambridge: Cambridge University Press.

Burnham, Kenneth P., and David R. Anderson. 2004. Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research* 33: 261–304.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: Benjamins.

Bybee, Joan L. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language*, eds. Michael Barlow and Suzanne Kemmer, 65–85. Stanford: CSLI.

Bybee, Joan L. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan L. 2002. Word frequency and context of use in lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14: 261–290.

Bybee, Joan L. 2006. From usage to grammar: the mind's response to repetition. *Language* 82: 711–733.

Bybee, Joan L. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Byrd, Dani. 1992. A note on English sentence-final stops. In *UCLA Working Papers in Phonetics*, Vol. 81, eds. Pat Keating and Dani Byrd, 37–38. Los Angeles: Department of Linguistics, UCLA.

Carpenter, Angela. 2006. Acquisition of a natural versus an unnatural stress system. Ph.D. diss., University of Massachusetts.

Carpenter, Angela. 2010. A naturalness bias in learning stress. *Phonology* 27: 345–392.

Coetzee, Andries W. 2004. What it means to be a loser: Non-optimal candidates in Optimality Theory. Ph.D. diss., University of Massachusetts.

Coetzee, Andries W. 2005. The Obligatory Contour Principle in the perception of English. In *Prosodies*, eds. Sónia Frota, Marina Vigário, and Maria João Frietas, 223–245. Berlin: Mouton de Gruyter.

Coetzee, Andries W. 2006. Variation as accessing "non-optimal" candidates. *Phonology* 23: 337–385.

Coetzee, Andries W. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84: 218–257.

Coetzee, Andries W. 2009a. An integrated grammatical/non-grammatical model of phonological variation. In *Current issues in linguistic interfaces: Volume 2*, eds. Young-Se Kang, Jong-Yurl Yoon, Hyunkyung Yo, Sze-Wing Tang, Yong-Soon Kang, Youngjun Jang, Chul Kim, Kyoung-Ae Kim, and Hye-Kyung Kang, 267–294. Seoul: Hankookmunhwasa.

Coetzee, Andries W. 2009b. Learning lexical indexation. *Phonology* 26: 109–145.

Coetzee, Andries W. 2009c. Phonological variation and lexical frequency. In *NELS 38*, Vol. 1, eds. Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman, 189–202. Amherst: GLSA.

Coetzee, Andries W. 2012. Variation: Where laboratory and theoretical phonology meet. In *Oxford handbook of laboratory phonology*, eds. Abigail C. Cohn, Cécile Fougeron, and Marie K. Huffman, 62–75. Oxford: Oxford University Press.

Coetzee, Andries W., and Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language & Linguistic Theory* 26: 289–337.

Coetzee, Andries W., and Joe Pater. 2011. The place of variation in phonological theory. In *Handbook of phonological theory: 2nd Edition*, eds. John Goldsmith, Jason Riggle, and Alan Yu, 401–434. Cambridge: Blackwell.

Coetzee, Andries W., and Rigardt Pretorius. 2010. Phonetically grounded phonology and sound change: The case of Tswana labial plosives. *Journal of Phonetics* 38: 404–421.

Côté, Marie-Hélène. 2004. Syntagmatic distinctness in consonant deletion. *Phonology* 21: 1–41.

Crawford, Clifford James. 2009. Adaptation and transmission in Japanese loanword phonology. Ph.D. diss., Cornell University.

Diehl, Randy L., and Margaret A. Walsh. 1989. An auditory basis for the stimulus-length effect in the perception of stops and glides. *The Journal of the Acoustical Society of America* 85: 2154–2164.

Eek, Arvo, and Einar Meister. 1995. The perception of stop consonants: locus equations and spectral integration. In *ICPhS XIII: Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, Sweden, Vol. 1, 18–21.

Fasold, R. 1972. *Tense marking in Black English*. Arlington: Center for Applied Linguistics.

File-Muriel, Richard J. 2010. Lexical frequency as a scalar variable in explaining variation. *The Canadian Journal of Linguistics* 55: 1–25.

Fowler, Carol A. 1994. Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception and Psychophysics* 55: 597–610.

Frank, Austin F., and T. Florian Jaeger. 2008. Speaking rationally: Uniform Information Density as an optimal strategy for language production. In *Proceedings of the 30th annual meeting of the cognitive science society*, eds. Brad C. Love, Ken McRae, and Vladimir M. Sloutsky, 939–944. Austin: Cognitive Science Society.

Fruchter, David, and Harvey M. Sussman. 1997. The perceptual relevance of locus equations. *The Journal of the Acoustical Society of America* 102: 2997–3008.

Gahl, Susanne. 2008. *Time* and *thyme* are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84: 474–496.

Gahl, Susanne, and Alan Yu, eds. 2006. *Special issue on exemplar-based models in linguistics*. Vol. 23(3) of *The Linguistic Review*. Berlin: De Gruyter Mouton.

Goeman, Ton. 1999. T-deletie in Nederlandse dialecten. Kwantitatieve analyse van structurele, ruimtelijke en temporele variate. Ph.D. diss., Vrije Universiteit, Amsterdam. The Hague: Holland Academic Graphics.

Goeman, Ton, and Pieter van Reenen. 1985. *Word-final t-deletion in Dutch dialects. The roles of conceptual prominence, articulatory complexity, paradigmatic properties, token frequency and geographical distribution*. Amsterdam: Vakgroep Algemene Taalwetenschap Vrije Universiteit.

Goldinger, Stephen D., Paul A. Luce, David B. Pisoni, and Joanne K. Marcario. 1992. Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 18: 1211–1238.

Goldsmith, John. 1993. Harmonic phonology. In *The last phonological rule: Reflections on constraints and derivations*, ed. John Goldsmith, 21–60. Chicago: Chicago University Press.

Goldsmith, John, ed. 1995. *The handbook of phonological theory*. Oxford: Blackwell.

Gupta, Arjun K., and Saralees Nadarajah, eds. 2004. *Handbook of the beta distribution and its applications*. New York: Marcel Dekker.

Guy, Gregory R. 1991. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3: 1–22.

Guy, Gregory R. 2011. Variability. In *The Blackwell companion to phonology: Volume 4*, eds. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 2190–2213. Oxford: Blackwell.

Guy, Gregory R., and Charles Boberg. 1997. Inherent variability and the Obligatory Contour Principle. *Language Variation and Change* 9: 149–164.

Hammond, Michael. 1994. An OT account of variability in Walmatjari stress. Ms., University of Arizona.

Hayes, Bruce, and Zsuzsa Cziráky Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23: 59–104.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphological change. In *Current progress in historical linguistics*, ed. William M. Christie, 95–105. Amsterdam: North-Holland.

Hyman, Larry. 2001. On the limits of phonetic determinism in phonology: *NC revisited. In *The role of speech perception phenomena in phonology*, eds. Elizabeth Hume and Keith Johnson, 141–185. New York: Academic Press.

Itô, Junko. 1988. *Syllable theory in prosodic phonology*. New York: Garland.

Itô, Junko, and Armin Mester. 2001. Covert generalizations in Optimality Theory: The role of stratal faithfulness constraints. *Studies in Phonetics, Phonology, and Morphology* 7: 273–299.

Jaeger, T. Florian. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61: 23–62.

Jesney, Karen. 2007. The locus of variation in weighted constraint grammars. Poster presented at the Workshop on Variation, Gradience and Frequency in Phonology. Stanford University, July 2007 [Downloaded on December 27, 2007 from http://people.umass.edu/kjesney/papers.html].

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the emergence of linguistic structure*, eds. Joan L. Bybee and Paul Hopper, 229–254. Amsterdam: Benjamins.

Kaisse, Ellen M., and Patricia A. Shaw. 1985. On the theory of lexical phonology. *Phonology Yearbook* 2: 1–30.

Kaneko, Emiko, and Gregory K. Iverson. 2009. Phonetic and other factors in Japanese on-line adaptation of English final consonants. In *Papers from the eighth annual conference of the Japanese Society for*

Language Science, eds. Shunji Inagaki and Makiko Hirakawa, Vol. 8 of *Studies in language sciences*, 179–185. Tokyo: Kuroshio Publications.

Kang, Hye-Kyung. 1994. Variation in past-marking and the question of the system in Trinidadian English. In *CLS 30: Papers from the 30th regional meeting of the Chicago Linguistic Society Volume 2: The parasession on variation in linguistic theory*, ed. Katherine Beals, 15–164. Chicago: Chicago Linguistic Society.

Katayama, Motoko. 1998. Optimality Theory and Japanese loanword phonology. Ph.D. diss., University of California, Santa Cruz.

Kawahara, Shigeto. 2005. Voicing and geminacy in Japanese: An acoustic and perceptual study. In *Univeristy of Massachusetts occasional papers in linguistics*, Vol. 31, eds. Katherine Flack and Shigeto Kawahara, 87–120. Amherst: GLSA.

Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82: 536–574.

Kawahara, Shigeto. 2008. Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* 18: 317–330.

Kawahara, Shigeto. 2011a. Aspects of Japanese loanword devoicing. *Journal of East Asian Linguistics* 20: 169–194.

Kawahara, Shigeto. 2011b. Japanese loanword devoicing revisited: A wellformedness rating study. *Natural Language & Linguistic Theory* 29: 705–723.

Kempen, Gerard, and Karin Harbusch. 2008. Comparing linguistic judgments and corpus frequencies as windows on grammatical competence: A study of argument linearization in German clauses. In *The discourse potential of underspecified structures*, ed. Anita Steube, 179–192. Berlin: Walter de Gruyter.

Kewley-Port, Diane. 1983. Time-varying features as correlates of place of articulation in stop consonants. *The Journal of the Acoustical Society of America* 73: 322–335.

Kewley-Port, Diane, David Pisoni, and Michael Studdert-Kennedy. 1983. Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants. *The Journal of the Acoustical Society of America* 73: 1779–1793.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2: 85–138.

Kiparsky, Paul. 1993. An OT perspective on phonological variation. Ms. Stanford University. Paper presented at the Rutgers Optimality Workshop. October, 1993. http://www.stanford.edu/~kiparsky/Papers/nwave94.pdf. Accessed 11 July 2012.

Labov, William. 1966. *The Social stratification of English in New York City*. Washington: Center for Applied Linguistics.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715–762.

Labov, William. 1989. The child as linguistic historian. *Language Variation and Change* 1: 85–97.

Labov, William. 2004. Quantitative analysis of linguistic variation. In *Sociolinguistics: An international handbook of the science of language and society, volume 1: 2nd edition*, eds. Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier, and Peter Trudgill, 6–21. Berlin: Mouton de Gruyter.

Lacoste, Véronique. 2008. Learning the sounds of Standard Jamaican English: Variationist, phonological and pedagogical perspectives on 7-year-old children's classroom speech. Ph.D. diss., University of Essex.

Lahiri, Aditi, Letitia Gewirth, and Sheila E. Blumstein. 1984. A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study. *The Journal of the Acoustical Society of America* 76: 391–404.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Can connectionism contribute to syntax? Harmonic Grammar, with an application. In *Proceedings of the 26th regional meeting of the Chicago Linguistic Society*, eds. Michael Ziolkowski, Manuela Noske, and Karen Deaton, 237–252. Chicago: Chicago Linguistic Society.

Lin, Susan, Patrice S. Beddor, and Andries W. Coetzee. 2011. Gestural reduction and sound change: An ultrasound study. In *ICPhS XVII: Proceedings of the 17th International Congress of Phonetic Sciences*, eds. Wai-Sum Lee and Eric Zee, 1250–1253.

Lindblom, Bjorn. 1990. Explaining phonetic variation: A sketch of the H and H theory. In *Speech production and speech modeling*, eds. William J. Hardcastle and Alain Marchal, 403–439. Dordrecht: Kluwer Academic.

Luce, Paul A., and David B. Pisoni. 1986. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* 19: 1–35.

Martínez-Celdrán, Eugenio, and Xavier Villalba. 1995. Locus equations as a metric for place of articulation in automatic speech recognition. In *ICPhS XIII: Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, Sweden, Vol. 1, eds. Kjell Elenius and Peter Branderud, 30–33. Stockholm: Stockholm University.
Malécot, André. 1958. The role of releases in the identification of released final stops: A series of tape-cutting experiments. *Language* 34: 274–284.
McCarthy, John J. 2003. Sympathy, cumulativity, and the Duke-of-York gambit. In *The syllable in Optimality Theory*, eds. Caroline Féry and Ruben van de Vijver, 23–76. Cambridge: Cambridge University Press.
McCarthy, John J., and Alan Prince. 1995. Faithfulness and reduplicative identity. In *Papers in Optimality Theory*, eds. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey. Vol. 18 of *University of Massachusetts occasional papers in linguistics*, 249–384. Amherst: GLSA.
McNamara, Timothy P. 2005. *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press, Taylor and Francis.
Merchant, Nazarré, and Bruce Tesar. 2005. Learning underlying forms by searching restricted lexical subspaces. In *CLS 41: Proceedings from the 41st annual meeting of the Chicago Linguistic Society*, Vol. 2, eds. Rodney L. Edwards, Patrick J. Midtlyng, Colin L. Sprague, and Kjersti G. Stensrud, 33–48. Chicago: CLS.
Mitterer, Holger, and Mirjam Ernestus. 2006. Listeners recover /t/'s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics* 34: 73–103.
Moreton, Elliott. 2008. Learning bias as a factor in phonological typology. *Phonology* 25: 83–127.
Moreton, Elliott. 2010. Underphonologization and modularity bias. In *Phonological argumentation: Essays on evidence and motivation*, ed. Stephen Parker, 79–101. London: Equinox.
Nearey, Terrance M., and Sherrie E. Shammas. 1987. Formant transitions as partly distinctive invariant properties in the identification of voiced stops. *Canadian Acoustics* 15: 17–24.
Nevins, Andrew. 2007. Review of Scheer 2004. *Lingua* 118: 425–434.
Nishimura, Kohei. 2003. Lyman's law in loanwords. M.A. thesis, Nagoya University.
Nishimura, Kohei. 2006. Lyman's Law in loanwords. *Phonological Studies [Onin Kenkyuu]* 9: 83–90.
Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33: 999–1035.
Pater, Joe, and Anne-Michelle Tessier. 2006. L1 phonotactic knowledge and the L2 acquisition of alternations. In *Inquiries in linguistic development: Studies in honor of Lydia White*, eds. Roumyana Slabakova, Silvina A. Montrul, and Philippe Prévost, 115–131. Amsterdam: Benjamins.
Patrick, Peter L. 1992. Creoles at the intersection of variable processes: t, d deletion and past-marking in the Jamaican mesolect. *Language Variation and Change* 3: 171–189.
Patterson, David, and Cynthia M. Connine. 2001. Variant frequency in flap production: A corpus analysis of variant frequency in American English flap production. *Phonetica* 58: 254–275.
Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60: 320–342.
Phillips, Betty S. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In *Frequency and emergence of linguistic structure*, eds. Joan L. Bybee and Paul Hopper, 123–136. Amsterdam: John Benjamins.
Phillips, Betty S. 2006. *Word frequency and lexical diffusion*. New York: Palgrave Macmillan.
Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency effects and the emergence of lexical structure*, eds. Joan L. Bybee and Paul Hopper, 137–157. Amsterdam: Benjamins.
Pitt, Mark A., Laura Dilley, Keith Johnson, Scott Kiesling, William D. Raymond, Elizabeth Hume, and E. Fosler-Lussier. 2007. *Buckeye corpus of conversational speech*, 2nd Release. Columbus: Department of Psychology, Ohio State University. www.buckeyecorpus.osu.edu.
Postal, Paul. 1966. Review of "Elements of general linguistics" by André Martinet. *Foundations of Language* 2: 151–186.
Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper and Row.
Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in Generative Grammar. Ms., New Brunswick, Rutgers University.
Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in Generative Grammar*. Oxford: Blackwell.
Pullum, Geoffrey K. 2003. Learnability. In *The international encyclopedia of linguistics*, ed. William J. Frawley, 431–434. Oxford: Oxford University Press.
Raymond, William D., Robin Dautricourt, and Elizabeth Hume. 2006. Word-medial */t,d/* deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change* 18: 55–97.

Reynolds, Bill. 1994. Variation and phonological theory. Ph.D. diss., University of Pennsylvania.
Santa Ana, Otto. 1991. Phonetic simplification processes in the English of the barrio: A cross-generational sociolinguistic study of the Chicanos of Los Angeles. Ph.D. diss., University of Pennsylvania.
Scarborough, Rebecca. 2004. Coarticulation and the structure of the lexicon. Ph.D. diss., UCLA.
Scarborough, Rebecca. 2010. Lexical and contextual predictability: Confluent effects on the production of vowels. In *Papers in laboratory phonology X: Variation, phonetic detail and phonological modeling*, eds. Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio, and Nathalie Vallée, 557–586. Berlin: Mouton de Gruyter.
Schouten, Marten E.H. 1982. T-deletie in de stad Utrecht: Schoolkinderen en grootouders. *Forum der Letteren* 23: 282–291.
Schouten, Marten E.H. 1984. T-deletie in Het Zuiden van die provincie Utrecht. *Taal en Tongval* 36: 162–173.
Smolensky, Paul, and Géraldine Legendre, eds. 2006. *The harmonic mind: From neural computation to Optimality-Theoretic grammar, Volume 1: Cognitive architecture, Volume 2: Linguistic and philosophical implications*. Cambridge: MIT Press.
Steriade, Donca. 1999. Phonetics in phonology: The case of laryngeal neutralization. In *UCLA working papers in linguistics 2 (Papers in phonology 3)*, ed. Matthew K. Gordon, 25–146. Los Angeles: Department of Linguistics, UCLA.
Steriade, Donca. 2001. Directional asymmetries in place assimilation. In *The role of speech perception in phonology*, eds. Elizabeth Hume and Keith Johnson, 219–250. San Diego: Academic Press.
Stevens, Kenneth N., and Sheila E. Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America* 64: 1358–1368.
Stevens, Kenneth N., and Samuel J. Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65: 81–106.
Sussman, Harvey M., Helen A. McCaffrey, and Sandar A. Matthews. 1991. An investigation of locus equations as a source of relational invariance for stop place categorization. *The Journal of the Acoustical Society of America* 90: 936–946.
Tanaka, Shin-Ichi. 2009. The eurhythmics of segmental melody. *Journal of the Phonetic Society of Japan* 13: 44–52.
Tesar, Bruce. 2006. Learning from paradigmatic information. In *NELS 36: Proceedings of the thirty-sixth annual meeting of the North East Linguistic Society*, Vol. 2, eds. Christopher Davis, Amy-Rose Deal, and Youri Zabbal, 619–638. Amherst: GLSA.
Tesar, Bruce, and Paul Smolensky. 1996. *Learnability in Optimality Theory (long version).* Technical report JHU-CogSci-96-4, Department of Cognitive Science, The Johns Hopkins University, Baltimore, Md.
Tesar, Bruce, and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29: 229–268.
van Oostendorp, Marc. 1997. Style levels in conflict resolution. In *Variation, change and phonological theory*, eds. Frans Hinskens, Roeland van Hout, and Leo Wetzels, 207–229. Amsterdam: John Benjamins.
Versace, Rémy, and Brigitte Nevers. 2003. Word frequency effect on repetition priming as a function of prime duration and delay between the prime and the target. *British Journal of Psychology* 94: 389–408.
Vitevitch, Michael S., and Paul A. Luce. 1998. When words compete: Levels of processing in spoken word recognition. *Psychological Science* 9: 325–329.
Vitevitch, Michael S., and Paul A. Luce. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40: 374–408.
Walsh, Margaret A., and Randy L. Diehl. 1991. Formant transition duration and amplitude rise time as cues to the stop/glide distinction. *The Quarterly Journal of Experimental Psychology A* 43: 603–620.
Zsiga, Elisabeth. 2000. Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28: 69–102.

Fukazawa Haruka, Shigeto Kawahara, Mafuyu Kitahara and Shin-ichiro Sano (2015) Two is too much: Geminate devoicing in Japanese. In Phonological Society of Japan (ed.) Phonological Studies 18. Tokyo: Kaitakusha. pp.3-10.

Two Is Too Much: Geminate Devoicing in Japanese*

Haruka Fukazawa[1], Shigeto Kawahara[1], Mafuyu Kitahara[2], and Shinichiro Sano[3]
[1]Keio University, [2]Waseda University, [3]Okayama Prefectural University

ABSTRACT. This paper analyzes devoicing of geminates induced by singleton [p] in Japanese loanwords, as well as that induced by voiced obstruents (Nishimura 2003 *et seq.*). Corpus analyses and judgment experiments reveal that [p] can optionally devoice a voiced obstruent geminate in Japanese loanwords, as in /**p**irami**dd**o/ → [**p**irami**tt**o] "pyramid" (Kawahara and Sano 2014). An analysis of type and token frequency of the entire Japanese lexicon shows that [p] and voiced geminates are the two most rare segments in the Japanese lexicon. We argue that their low lexical frequencies may be responsible for the [p]-induced devoicing. To formalize this idea, we propose a new constraint, OCP(unfamiliar), within the framework of Harmonic Grammar (Pater 2009). Not only does the proposal offer a plausible analysis of geminate devoicing, it brings up a new theoretical perspective: different lexical strata in a single language should be treated as part of a single grammatical system (Fukazawa et al. 1998; Ito and Mester 2003).

Keywords: geminate devoicing, lexical strata, lexical frequency, corpus study, [p], Harmonic Grammar

1. Introduction

It is often observed in a language with two or more lexical strata that some phonological structures are banned in one stratum, but allowed in others. In Japanese, three phonological structures instantiate such stratum-dependent distributions: (i) two voiced obstruents within a morpheme, (ii) a voiced (obstruent) geminate, and (iii) a singleton [p], as shown and exemplified in (1). These three phonological structures are prohibited in the phonotactics of the Yamato (=native) stratum (Ito and Mester 1999, 2003). Additional evidence for (1i) comes from rendaku, a phenomenon in which the initial segment of the second member of a compound becomes voiced. In case it results in two voiced obstruents within a morpheme, rendaku is blocked; i.e., [kita-**k**a**z**e] "north wind," not *[kita-**g**a**z**e]—this blockage of rendaku has long been known as Lyman's Law.

(1) The prohibition against three types of structures in the Yamato stratum

Constraints:	(i) *[d…d]	(ii) *[bb, dd, gg, zz]	(iii) *[p]
Examples:	*[kita**g**a**z**e]	*[to**gg**e]	*[**p**ikari]

On the other hand, these three structures are allowed in the Foreign stratum, which consists of recent loanwords from other languages, mainly from English. The loanwords with those structures are all acceptable, as exemplified in (2).

(2) Example words containing these structures in the Foreign stratum
(i) [**b**a**g**u] "bug" (ii) [e**gg**u] "egg" (iii) [**p**agu] "pag dog"

In previous work on lexical stratification in Optimality Theory (OT: Prince and Smolensky 1993/2004), such differences between the Yamato and Foreign strata have been explained by positing the constraint ranking in (3), in which the FAITH constraint for the Foreign stratum outranks three MARKEDNESS constraints—OCP[voice] (=Lyman's Law), *DD, and *[p]—which in turn outrank the FAITH constraint for the Yamato stratum (e.g., Fukazawa et al. 1998; Ito and Mester 1999, 2003). OCP[voice] prohibits two voiced obstruents within the same morpheme; *DD prohibits voiced geminates; and *[p] forbids a singleton [p].

(3) The OT constraint ranking in Japanese
FAITH(Foreign) » MARKEDNESS (OCP[voice], *DD, *[p]) » FAITH(Yamato)

Although (3) is a single invariant ranking, it can be viewed as consisting of two subparts. The first subpart—FAITH(Foreign) » MARKEDNESS—explains why the three phonological structures in (1) are possible in the Foreign stratum, while the second sub-ranking—MARKEDNESS » FAITH (Yamato)—accounts for the prohibition against them in Yamato words.

However, this is not the end of the story. Although each of the structures in (1) is possible in foreign words, as in (2), when two of them co-occur, then the co-occurrence may become impossible even in the Foreign stratum. A well-known case is when a voiced geminate occurs with another voiced obstruent within a morpheme; this geminate can be devoiced, as in [**d**o**gg**u] "dog" becoming [**d**o**kk**u] (Nishimura 2003; Kawahara 2014, for a review). The cause of this devoicing may be the simultaneous violation of OCP[voice] and *DD.

A similar, but less well-known, case is the observation that a geminate can devoice when a singleton [p] co-exists within the same morpheme, such as [**p**a**dd**o] "pad" optionally becoming [**p**a**tt**o] (Kawahara and Sano 2014). This [p]-driven devoicing of geminates is the focus of this paper. We suggest that [p]-driven devoicing occurs because Yamato-specific restrictions still have an effect on the acquisition of the phonology of the Foreign stratum. In other words, the data of geminate devoicing suggests that Yamato and Foreign strata belong to the same language, namely, Japanese. This conclusion may sound trivial at first, but it is not, given that in OT, the difference between the Yamato and Foreign strata is explained in the same way as the difference between Japanese and, say, English. In other words, although it is true that the Japanese phonological lexicon may need to be stratified (Ito and Mester 1999), it is also important to treat these separate strata as belonging to a single grammatical system (see also Fukazawa et al. 1998; Ito and Mester 2003:135–136).

In what follows, we first describe and analyze the data of geminate devoicing, particularly in words that contain both a singleton [p] and a voiced geminate. To this end, section 2 summarizes the results of Kawahara and Sano (2014), who present both corpus and experimental evidence bearing on the [p]-driven geminate devoicing pattern. Section 3 analyzes the devoicing pattern within the framework of Harmonic Grammar (Legendre et al. 1990; Pater 2009), by proposing a new constraint OCP(unfamiliar). Section 4 examines an alternative analysis with Local Conjunction (Smolensky 1993), and shows why the current analysis is superior. Section 5 discusses general theoretical implications of our analysis.

2. Geminates devoice with a singleton [p]: Kawahara and Sano (2014)

2.1 An impressionistic observation

Let us first examine the basic patterns of geminate devoicing caused by [p]. When a singleton [p] co-occurs with a voiced geminate, the geminate can devoice, as exemplified in the words in (4). (One may wonder what would happen to a word with a geminate [pp] and a voiced geminate: unfortunately, words with two geminates barely exist: Ito and Mester 2003:49–50.) The data in (4) are based on the authors' intuition, as per the traditional generative approach.

(4) Geminate devoicing in words of the form [p…dd]

/**p**a**dd**o/	→	[pa**tt**o]	"pad"
/kyu**p**i**dd**o/	→	[kyuupi**tt**o]	"cupid"
/ra**p**i**dd**o/	→	[rapi**tt**o]	"rapid"
/**p**irami**dd**o/	→	[pirami**tt**o]	"pyramid"
/ai**p**o**dd**o/	→	[aipo**tt**o]	"iPod"
/tetora**p**o**dd**o/	→	[tetorapo**tt**o]	"tetra pod"

2.2 An analysis of the CSJ

This intuition-based data is further corroborated by the analysis of the Corpus of Spontaneous Japanese (the CSJ: Kokuritsu Kokugo Kenkyujo 2008), which shows that geminates devoice more often when they co-occur with a singleton [p] or with a voiced obstruent than elsewhere (Kawahara and Sano 2014). As shown in (5), in the CSJ, geminates appear as devoiced 66% of the time with a voiced obstruent, and 27% of the time with a singleton [p], while only 4% of the time elsewhere. The differences between the [p]-conditions and the other two conditions are statistically significant (Fisher's exact test: [p] vs. [+voi], $p < 0.001$; [p] vs. elsewhere, $p < 0.01$).

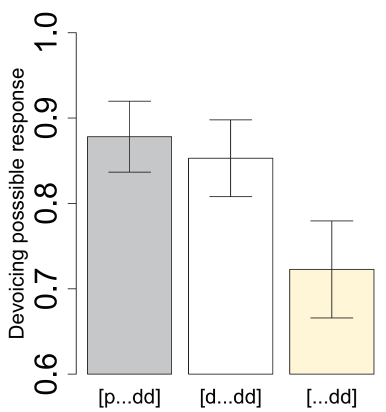
(5) The CSJ analysis: the number of occurrences of the faithful (= voiced) renditions and devoiced renditions of geminates, according to the three phonological environments. See Kawahara and Sano (2014) for details.

Trigger	**Faithful**	**Devoiced**	**% Devoiced**
[p]	11 (e.g., [**p**a**dd**o])	4 (e.g., [**p**a**tt**o])	27%
voiced obstruent	163 (e.g., [ri**j**i**dd**o])	313 (e.g., [ri**j**i**tt**o])	66%
elsewhere	689 (e.g., [sumo**gg**u])	28 (e.g., [sumo**kk**u])	4%

2.3 Experimental data

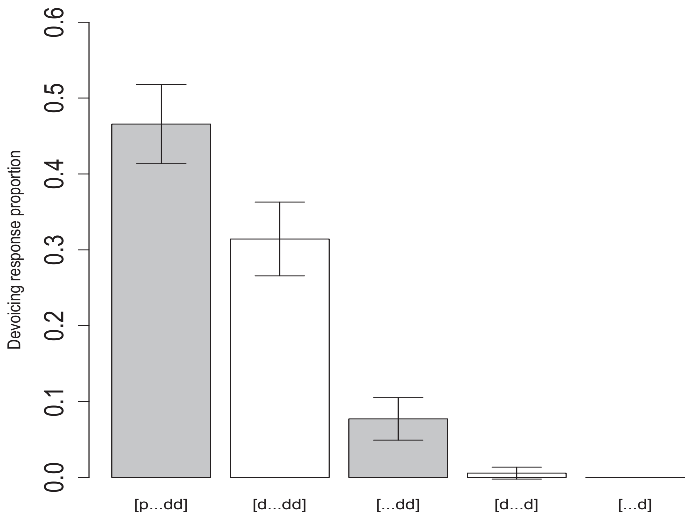
This section briefly reviews two experiments reported in Kawahara and Sano (2014). In their Experiment I, Japanese speakers were asked whether devoicing is possible or not for each lexical item. There were three conditions: (i) words with a singleton [p] and a voiced geminate (e.g., [**p**a**dd**o]); (ii) words with a voiced obstruent and a voiced geminate (e.g., [**b**a**dd**o]); (iii) words with a voiced geminate but without [p] or a voiced obstruent (e.g., [ma**dd**o]). The results in (6) show that participants judged devoicing of geminates to be possible 87% of the time with a singleton [p], 85% with a voiced obstruent, and 72% elsewhere. The differences between the first two conditions and the elsewhere condition were statistically significant.

(6) Experiment I (Kawahara and Sano 2014)



In Experiment II, Japanese speakers were asked which form, voiced or devoiced, they use for a given pair of forms. For instance, [paddo] and [patto] were presented, and then they judged which form they actually use. In this experiment, in addition to the three conditions used in Experiment I, two conditions were added; (iv) words with two singleton voiced obstruents, and (v) words with one singleton obstruent. As shown in (7), more than 30% of the responses were the devoiced responses in the first two conditions, while less than 10% were in the other three conditions. Almost half of the responses were the devoiced responses, when a geminate appears with a singleton [p] (the leftmost bar). The difference between the [p…dd] condition and the […dd] condition is thus more pronounced in Experiment II than in Experiment I.

(7) Experiment II (Kawahara and Sano 2014)



The overall results of Kawahara and Sano (2014), briefly reviewed here, thus show that a geminate tends to be more devoiced with a singleton [p] or with a voiced obstruent than elsewhere. The next section analyzes these patterns of geminate devoicing within the framework of Harmonic Grammar (HG: Legendre et al. 1990).

3. An HG Account for geminate devoicing

3.1 Geminate devoicing in [D + DD] words

First, as a background, we review the analysis of geminate devoicing pattern developed by Pater (2009), in which he analyzed only devoicing of geminates co-occurring with another voiced obstruent (see also Kawahara 2014). Pater (2009) argues that HG allows us to derive the effect of geminate devoicing caused by another voiced obstruent without a stipulation (cf. Nishimura 2003 who uses local conjunction to derive this effect). We show in section 3.2, however, that even in this framework, [p]-driven devoicing does not arise automatically.

Unlike OT in which the constraints are ranked, each constraint in HG is assigned some *weight*. A harmonic score for each candidate is a weighted sum of each constraint violation (Harmony=$\sum w_i * v_i$ where *w*=weight; *v*=violation profiles). The candidate with the highest harmonic score wins. As shown in the tableaux in (8), let the weight of FAITH for voicing be 2, and that of the other constraints (*DD, *[p], and OCP[voiced]) be 1.5. In this analysis, the precise values of these weights are not relevant, as long as two arithmetic conditions are met: (i) the weight of FAITH is higher than those for the markedness constraints, and (ii) the sum of the weights of *DD and OCP[voiced] is higher than that of FAITH, as we will see.

For the case of /bado/, the first faithful candidate [bado] violates OCP[voice], since it has two voiced obstruents, [b] and [d]. Therefore, its harmonic score is -1.5 (=1.5 times -1). On the other hand, the second candidate [bato] with devoicing violates FAITH, resulting in the harmonic score of -2. Thus, the first candidate [bado] becomes a winner, because its harmonic score is higher—closer to zero—than that of [bato] (-1.5 > -2).

The second input is /heddo/, which has a voiced geminate. The faithful candidate [heddo] violates *DD, resulting in the harmonic score of -1.5. The second candidate [hetto] violates FAITH, and its harmonic score is -2. Therefore, the faithful one [heddo] becomes the actual output, because of its higher harmonic score (again, -1.5 > -2).

In contrast to these two cases, when a voiced obstruent and a voiced geminate co-occur, the first candidate [beddo] which is faithful to the input violates both *DD and OCP[voice]; therefore, its harmonic score is -3.0 (= -1.5 + -1.5). On the other hand, a devoiced candidate [betto] violates only FAITH, with the resulting harmonic score of -2.0. Hence, the devoiced candidate [betto] wins in this evaluation. What we observe here is a so-called "gang-effect" in which lower weighted constraints gang up to triumph over a higher-weighted constraint. A gang effect occurs because one violation of FAITH simultaneously satisfies the violations of two lower-weighted markedess constraints.

(8) An HG analysis of geminate devoicing by Pater (2009)

		FAITH	*DD	*[p]	OCP[voi]	Harmony
		2	1.5	1.5	1.5	
/bado/	☞bado				-1	-1.5
	bato	-1				-2.0
/heddo/	☞heddo		-1			-1.5
	hetto	-1				-2.0
/beddo/	beddo		-1		-1	-3.0
	☞betto	-1				-2.0

3.2 Geminate devoicing in [p] + DD words

3.2.1 OCP(unfamiliar)

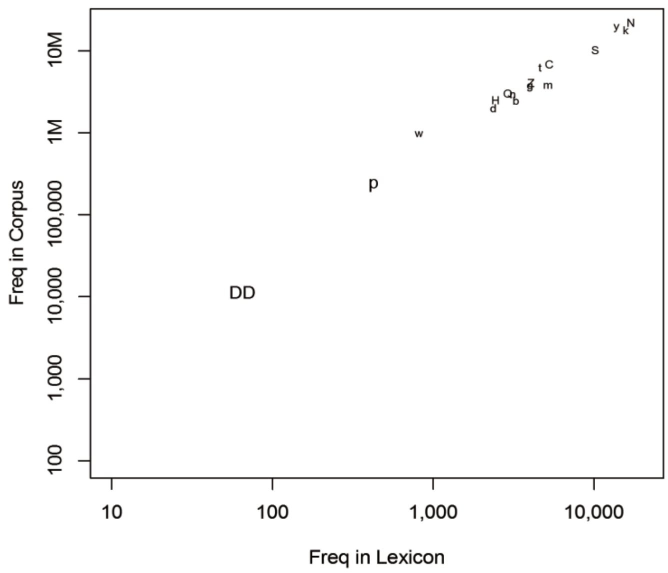
Now if we were to apply the same analysis for [p]-driven geminate devoicing, OCP[voice] does not work, since we do not have two voiced obstruents. Therefore, we propose a new constraint OCP(unfamiliar) which bans two "unfamiliar" elements within a word. This constraint is based on two intuitive ideas: (i) both singleton [p]s and voiced geminates are, despite being allowed in loanwords, still foreign, infrequent and unfamiliar, and (ii) it may be that Japanese speakers do not like to tolerate the co-occurrence of two such infrequent, unfamiliar structures. We conjecture that the constraint OCP(unfamiliar) itself is universal, but the choice of the unfamiliar elements is language specific. An underlying assumption for this mechanism is the hypothesis that children can identify infrequent, unfamiliar structures in the process of their language acquisition.

We now show that a singleton [p] and a voiced geminate are both indeed infrequent structures in Japanese.

3.2.2 Supporting data for the unfamiliarity of [p] and DD

As an objective support of this proposal, the figure in (9) shows the lexical frequency of each phoneme in Japanese. The y-axis shows the (token) frequency of the corpus data consisting of 10 years from the Asahi Newspaper, and the x-axis shows the (type) frequency data taken from the 80,000-word Shinmeikai Kokugo Dictionary (Amano and Kondo, 1999). Both the type and token frequencies of voiced geminate [DD] and singleton [p] are the lowest two among all phonemes in Japanese. It therefore seems safe to assume that children acquiring Japanese would think that both a singleton [p] and a voiced geminate are unfamiliar elements in Japanese. With this, OCP(unfamiliar) in Japanese would penalize any morpheme that contains both [p] and a voiced geminate.

(9) Type and token frequency of phonemes in Japanese



3.3.3 Geminate devoicing in [p + DD] words

With the new constraint OCP(unfamiliar), we now develop an HG analysis of geminate devoicing caused by a singleton [p]. In addition to those constraints used in (8), the tableaux in (10) have OCP(unfamiliar). We assign the same weight to OCP(unfamiliar) as that of OCP[voice] (1.5; however, as shown below, any weight higher than 0.5 would actually work). Recall that OCP(unfamiliar) penalizes the structure [p…dd].

Given an input like /paddo/, a devoiced candidate [patto] has a better harmonic score than the faithful candidate [paddo], because the latter candidate violates three markedness constraints *DD, *[p], and OCP(unfamiliar), resulting in the harmonic score -4.5, while the devoiced candidate [patto] violates only FAITH and *[p], resulting in -3.5. Importantly, without OCP(unfamiliar), [paddo] would receive a harmonic score of -3, and hence would not lose against [patto]. Just deploying HG—without OCP(unfamiliar)—does not account for [p]-driven devoicing by itself, unlike Pater's (2009) analysis of OCP(voice)-driven geminate devoicing. A gang effect does *not* occur here, because devoicing a geminate in words like [paddo] does not simultaneously resolve a violation of OCP[voice] and that of *DD. (For this reason, our analysis is directly translatable to OT, unlike that of Pater (2009).)

In words without a voiced geminate, a faithful candidate [pagu] has a better harmonic score than the devoiced candidate [paku], because the faithful one violates only *[p]. Devoicing here is fortuitous.

(10) Devoicing in [p + DD]

		FAITH	*DD	*[p]	OCP[voi]	OCP(unfam)	Harmony
		2	1.5	1.5	1.5	1.5	
/paddo/	paddo		-1	-1		-1	-4.5
	☞patto	-1		-1			-3.5
/pagu/	☞pagu			-1			-1.5
	paku	-1		-1			-3.5

The tableaux in (11) show that the patterns of geminate devoicing examined in (8) still hold with OCP(unfamiliar). Since none of the structures—/bado/, /heddo/, and /beddo/—

contains two unfamiliar elements (a voiced geminate or a singleton [p]), the constraint OCP(unfamiliar) is satisfied, and hence its addition does not affect their harmonic scores.

(11) Devoicing in [D + DD] with OCP(unfamiliar)

		FAITH	*DD	*[p]	OCP[voi]	OCP(unfam)	Harmony
		2	1.5	1.5	1.5	1.5	
/bado/	☞bado				-1		-1.5
	bato	-1					-2
/heddo/	☞heddo		-1				-1.5
	hetto	-1					-2
/beddo/	beddo		-1		-1		-3
	☞betto	-1					-2

4. On the analysis with Local Conjunction

Let us now address one obvious alternative: local conjunction (Smolensky 1993). Nishimura (2003) analyzed the geminate devoicing caused by OCP[voice] with a locally conjoined constraint; i.e., {OCP[voice] & *DD}$_{stem}$. Similarly, it would be possible to analyze [p]-driven geminate devoicing with a constraint like {*[p] & *DD}$_{stem}$. The analyses with such a locally-conjoined constraint would formally be equivalent to our HG analyses: given our case, OCP(unfamiliar) and {*[p] & *DD}$_{stem}$ would assign the same violation profiles to the candidates considered here.

However, we contend that OCP(unfamiliar) is better than the locally-conjoined constraint {*[p] & *DD}$_{stem}$, because unrestricted use of local conjunction will lead to an unrestrictive theory (e.g., Fukazawa and Lombardi 2003). To the extent that constraints like [p] and *DD can be conjoined in a domain as large as stem, the theory would have to allow any two constraints to be conjoined in a domain of stem. Such a grammatical model would be too powerful and too unrestrictive. For example, there is nothing that would prevent us from positing a constraint like {*[+nasal] & *[pp]}$_{stem}$, which prohibits geminate [pp] only when there is a nasal consonant within the same stem. Unlike the local conjunction analysis, the current analysis ties [p] and voiced geminates together in terms of their low lexical frequencies, or their perceived unfamiliarity. There would be no such ties in the local conjunction analysis.

5. Conclusion

This paper has analyzed the patterns of geminate devoicing, both when triggered by singleton [p] and when triggered by a voiced obstruent. We have shown that geminate devoicing data illustrate how the three phonological structures—two voiced obstruents in a word, a voiced geminate, and a singleton [p]—behave in the Foreign stratum: one of the three elements is possible, but *two can be too much* (notwithstanding that devoicing of singleton consonants is still impossible). The analysis suggests that structures like singleton [p] and voiced geminates, although they are allowed in the Foreign stratum, are still marked: their effects are still tangible in the patterns of geminate devoicing. We have demonstrated that HG with OCP(unfamiliar) provides a plausible analysis of the devoicing patterns.

One prominent remaining issue, which was set aside in the current paper, is the optionality —and probability—of devoicing. Future studies should address the optionality of geminate devoicing. Some promising lines of research have already been hinted at by Coetzee and Kawahara (2013) who use Noisy Harmonic Grammar, and Coetzee and Pater (2011) who use a MaxEnt Grammar. These models would allow us to calculate exact constraint weight values based on the probabilities of devoicing in the actual production pattern. However, we would like to leave this topic for future research, as we think it merits its own paper.

Before closing the paper, we would like to make one final concluding remark. In OT or HG, cross-linguistic differences are explained by differences in constraint rankings or weightings. Ultimately, whether a structure is allowed or not boils down to a difference between "FAITHFULNESS » MARKEDNESS", or "MARKEDNESS » FAITHFULNESS": the difference between the two languages (e.g., Japanese vs. English) and the difference between two strata within one language are explained by the same mechanism.

However, the analysis of geminate devoicing in Japanese shown above connects the Foreign stratum to the Yamato stratum. The markedness constraints that are active in the native words—*DD and *[p]—would make some structures infrequent or unfamiliar in the entire lexicon of Japanese; through OCP(unfamiliar), these markedness constraints would impact on a pattern in loanword phonology. Therefore, despite being stratified, Foreign and Yamato must belong to one single grammatical system, i.e., Japanese.

Notes

* We are grateful to two anonymous reviewers, the members of the Keio phonetics/phonology study group (Hinako Masuda, Jeff Moore, Miho Sasaki, Yukiko Sugiyama, Yoko Sugioka) as well as the participants of FAJL 7 and the Phonology Forum 2014. Thanks to Donna Erickson for her proofreading of the pre-final draft. This project is supported by a grant #26284059 to the 2nd, 3rd and 4th authors.

References

Amano, Shigeaki and Tadahisa Kondo. 1999. *Nihongo no goitokusei* [Psylex: Lexical Properties of Japanese]. Tokyo: Sanseido.

Coetzee, Andries and Shigeto Kawahara. 2013. Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 30(1).47–89.

Coetzee, Andries and Joe Pater. 2011. The place of variation in phonological theory. *The handbook of phonological theory, 2nd Edition*, ed. by John A. Goldsmith, Jason Riggle and Alan Yu, 401–431. Oxford: Blackwell-Wiley.

Fukazawa, Haruka, Mafuyu Kitahara and Mitsuhiko Ota. 1998. Lexical stratification and ranking invariance in constraint-based grammars. *CLS 34: The Panels*, 47–62.

Fukazawa, Haruka and Linda Lombardi. 2003. Complex constraints and linguistic typology in Optimality Theory. *The Linguistic Review* 20.195–215.

Ito, Junko and Armin Mester. 1999. The phonological lexicon. *The handbook of Japanese linguistics*, ed. by Natsuko Tsujimura, 62–100. Oxford: Blackwell.

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

Kawahara, Shigeto. 2014. Geminate devoicing in Japanese loanwords: Theoretical and experimental investigations. Tokyo: Keio University, ms. Revision submitted to *Language and Linguistic Compass*. lingbuzz/002161.

Kawahara, Shigeto and Shin-ichiro Sano. 2014. /p/-driven geminate devoicing in Japanese: Corpus and experimental evidence. Tokyo: Keio University, ms. lingbuzz/002188.

Kokuritsu-Kokugo-Kenkyuujo. 2008. *Nihongo hanashikotoba koopasu* [The Corpus of Spontaneous Japanese]. Ver. 2. Database.

Legendre, Geraldine, Yoshiro Miyata and Paul Smolensky. 1990. Harmonic Grammar: A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. *Proceedings of the twelfth annual conference of the Cognitive Science Society*, 388–395. Cambridge, MA: Lawrence Erlbaum.

Nishimura, Kohei. 2003. *Lyman's law in loanwords*. Nagoya: Nagoya University MA thesis.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33.999–1035.

Prince, Alan and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden and Oxford: Blackwell.

Smolensky, Paul. 1993. *Optimality, markedness, and underspecification*. Paper presented at the Rutgers University Optimality Workshop, New Brunswick, NJ.

Kawahara, Shigeto and Shin-ichiro Sano (2014) /p/-driven geminate devoicing in Japanese: Corpus and experimental studies. Ms., Keio University.

/p/-driven geminate devoicing in Japanese: Corpus and experimental evidence

Abstract

In Japanese loanwords, voiced geminates can be devoiced in the presence of another voiced obstruent (e.g. /**d**o**gg**u/ → /**d**o**kk**u/). This devoicing pattern has been studied extensively in the recent phonological literature, in terms of theoretical modeling as well as from the perspective of experimentation and corpus studies. Less well-known is the observation that /p/ may cause devoicing of geminates as well (e.g. /**p**irami**dd**o/ → /**p**irami**tt**o/), although to date no objective evidence has been offered to confirm this observation. The current study thus attempts to test this observation objectively, by way of a corpus study and two phonological judgment experiments. All the results generally support the idea that /p/ causes devoicing of geminates in loanwords. In addition to this descriptive discovery, throughout the paper we discuss intriguing task effects in phonological experimentation, by comparing the corpus data and the results of the two experiments. Although our aim is primarily descriptive, we offer some analytical possibilities for the /p/-driven devoicing of geminates toward the end of the paper.

1 Introduction

It is well known since Nishimura's (2003) discovery that in Japanese loanword phonology, voiced geminates optionally devoice when they co-occur with another voiced obstruent; e.g. /**d**o**gg**u/ 'dog' can be pronounced as /**d**o**kk**u/, but /e**gg**u/ 'egg' cannot be pronounced as */e**kk**u/ (see Kawahara 2015c for a review).[1] In other words, a restriction against two voiced obstruents—a constraint known as Lyman's Law or the Obligatory Contour Principle on [+voice] (OCP(voice)) in the native phonology (Ito & Mester, 1986, 2003; Vance, 2007)—causes devoicing of geminates. What makes this pattern even more interesting is the fact that OCP(voice) does not seem to devoice singletons (e.g. /**b**a**g**u/ → */**b**a**k**u/ 'bug').

[1] This paper uses phonemic transcription (Vance, 1987, 2008) rather than IPA transcription for the sake of typographical ease, as the phonetic details do not matter much in this paper.

This devoicing of geminates has been studied extensively within various theoretical frameworks (Coetzee & Kawahara, 2013; Coetzee & Pater, 2011; Crawford, 2009; Farris-Trimble, 2008; Ito & Mester, 2008; Hayes, 2009; Kawahara, 2006, 2008; McCarthy, 2008; Nishimura, 2003, 2006; Pater, 2009, to appear; Rice, 2006; Tesar, 2007; Tsujimura, 2014), and has been used to argue for several theoretical apparatuses, such as local conjunction (Nishimura, 2003, 2006), Harmonic Grammar (Farris-Trimble, 2008; Pater, 2009, to appear), and MaxEnt Grammar (Coetzee & Pater, 2011), among others. This OCP-driven devoicing pattern has also been studied experimentally (Kaneko & Iverson, 2009; Kawahara, 2011a,b, 2013), as well as from the perspective of corpus-based studies (Kawahara & Sano, 2013; Sano, 2013; Sano & Kawahara, 2013). It thus seems safe to say that this OCP-driven devoicing pattern has received substantial attention in the field in the last ten years or so, not exclusively among those who are interested in Japanese phonology *per se*, but also among those who work in the field of phonological theory in general.

Against this background, this paper points out a less well-known—but yet important—complication to this devoicing pattern. That is, it seems to be the case, at least according to our intuition, that geminates can devoice when they co-occur with /p/ as well; e.g., /kyuu**p**i**tt**o/ 'cupid' and /**p**irami**tt**o/ 'pyramid'. If our intuition is correct, then it is not only the OCP(voice), but also the co-occurrence with /p/, that can cause devoicing of geminates. This /p/-driven devoicing would probably come as a surprise to many phonologists, because having /p/ and having a voiced geminate seem phonologically mutually irrelevant. At the same time, there is a sense in which /p/ is special in Japanese, in that singleton /p/s are allowed only in loanwords and onomatopoetic words but are banned in the native phonology (Ito & Mester, 1995, 1999, 2008). It may be the case that /p/ causes geminate devoicing because of this special property. At any rate, if /p/-driven geminate devoicing does indeed exist in the contemporary phonology of Japanese, it would also require significant revisions to the theoretical analyses of the OCP-driven geminate devoicing pattern cited above, because the co-occurrence of /p/ and voiced geminates should not violate the OCP(voice).

The pattern of /p/-driven devoicing, however, would probably be taken to be surprising by many practicing phonologists, and can be viewed with suspicion, because as far as we are aware, no other languages show a pattern of devoicing caused by /p/ at a distance. In order to check our intuition on this matter, we conducted a search using Jeff Mielkes P-Base (http://pbase.brohan.ca/query) to examine whether such an alternation exists in other languages. The search results did not find any example of /p/ causing devoicing. The P-Base contains 7318 patterns from 629 languages, indicating that /p/-driven devoicing is rare at best—and possibly hitherto unattested—in natural languages. Therefore, this apparently surprising pattern should probably not be used for phonological argumentation when it is purely based on the authors' own intuition (Kawahara, 2015a; Labov, 1996; Ohala, 1986; Schütze, 1996). Therefore, this /p/-driven devoicing pattern requires careful empirical scrutiny.

To summarize, this paper aims to verify this /p/-driven devoicing of geminate using objective methods, because (i) this /p/-driven devoicing of geminates is a surprising non-local interaction between two phonological structures, and because (ii) if this /p/-driven devoicing is a real process, then the previous analyses of geminate devoicing need to be revised. To that end we report one corpus-based study and two judgment experiments to explore the reality of the /p/-driven devoicing. The judgment experiments build on the previous judgment experiments on OCP-driven devoicing of geminates (Kawahara, 2011a,b, 2013).

The rest of this paper is structured as follows. A corpus-based analysis is reported in Section 2. Experiment I used a yes/no format, which shows that /p/ does induce more "devoicing possible" responses than other consonants. Experiment II asked the participants which form—voiced or devoiced—they would actually use in a forced-choice format, which again supports the reality of /p/-driven devoicing. In addition, the comparison of these two experiments reveals an interesting task effect in phonological experimentation. Section 5 discusses some possible theoretical analyses. The final section is a conclusion.

2 A corpus-study based on the CSJ

We first analyzed the Corpus of Spontaneous Japanese (the CSJ) (NINJAL, 2008). This corpus is one of the largest corpora of spoken Japanese, containing about 7.5 million words, which is equipped with rich annotation systems. It also encodes underlying forms as well as actual pronounced forms, and this feature allows us to access the devoicing status of geminates (see Kawahara & Sano 2013; Maekawa 2004; Maekawa et al. 2000 for further details of the CSJ).

We first extracted all the words containing a voiced geminate, and classified them into three categories: (i) those that occur with /p/, (ii) those that occur with a voiced obstruent, and (iii) those that would fit neither of the structural descriptions (henceforth the "elsewhere" condition).[2] We then examined whether these geminates appear as voiced or devoiced in the corpus for each condition, based on the transcription provided by the CSJ. The results are shown in Table 1.

The bottom row shows that geminates rarely appear as devoiced when they do not

[2]We put double quotes around "elsewhere", in order to distinguish it from the Elsewhere Condition, one of the principles deployed in Lexical Phonology (Kiparsky, 1982)

Table 1: The results of the search of the CSJ: the second column shows the number of geminates appearing as voiced; the third column the number of geminates appearing as devoiced. The rightmost column shows the percentages of devoiced tokens.

	voiced	devoiced	total	percent devoiced
/p/ (e.g. /paddo/)	11	4	15	26.7%
[+voice, -son] (e.g. /beddo/)	163	313	476	65.8%
elsewhere (e.g. /heddo/)	689	28	717	3.9%

appear either with /p/ or a voiced obstruent. This result shows that context-free devoicing of geminates rarely occurs. The second row shows that geminates appear as voiceless 65.8% of the time when they appear with a voiced obstruent. The top row shows that although /p/ does not cause devoicing as much as a voiced obstruent, the devoicing percentage is higher than the third row, the "elsewhere" condition.

We fully admit that the N for the /p/-condition is small (the top row; N=15), and that we should not be conclusive about the productivity of /p/-driven devoicing based on these data alone. Nevertheless, Fisher's Exact tests show that /p/'s devoicing proportion is higher than the "elsewhere" condition ($p < .01$), although it is lower than the devoicing proportion by a voiced obstruent ($p < .01$). Again, we should not conclude based on this data that /p/ causes devoicing of geminates, because the relevant number of items is small. We thus followed up this corpus-based study with phonological judgment experiments.

3 Experiment I: A yes/no judgement task

Experiment I used a task in which the participants judged, for each given item, whether devoicing is possible or not in a yes/no format. This experiment followed the methodology of Kawahara (2013).

3.1 Method

3.1.1 Task

Within each trial, the participants were given one word containing a geminate. They were then asked if devoicing that geminate was possible or not. For example, they were asked: "given the word *kyuupiddo*, is it possible to pronounce it as /kyuupitto/?"

3.1.2 Stimuli

The stimuli consisted of a set of real words and another set of nonce words. Within each set, there were three conditions: (i) those that contain /p/, (ii) those that contain a voiced obstruent, and (iii) those that contain neither.[3] Seven items were included in all conditions; we could find only seven real words that fit the structural description of (i) (at the time of the experiment). The place of articulation was controlled across the three conditions; six items contained geminate /dd/ and one item contained geminate /gg/ for the real word stimuli; this unbalance was necessitated by the fact that there are not many words that contain /p...gg/. For the nonce word stimuli, for each condition, there were four items containing /dd/ and three items containing /gg/.

The experimental items for the real words and nonce words are provided in Table 2 and Table 3, respectively. The real words for the /p/-condition were largely based on those that were found in the CSJ.

Table 2: The real word stimuli.

/p...dd/		/b...dd/		/...dd/	
/kyuupiddo/	‘cupid’	/baddo/	‘bad’	/heddo/	‘head’
/paddo/	‘pad’	/beddo/	‘bed’	/reddo/	‘red’
/aipaddo/	‘i-pad’	/deddo/	‘dead’	/uddo/	‘wood’
/aipoddo/	‘i-pod’	/guddo/	‘good’	/kiddo/	‘kid’
/supureddo/	‘spread’	/daddo/	‘dad’	/maddo/	‘mad’
/piramiddo/	‘pyramid’	/goddo/	‘god’	/roddo/	‘rod’
/piggu/	‘pig’	/biggu/	‘big’	/eggu/	‘egg’

Table 3: The nonce word stimuli.

/p...dd/	/b...dd/	/...dd/
/piddo/	/biddo/	/meddo/
/poddo/	/buddo/	/ruddo/
/puddo/	/boddo/	/yoddo/
/peddo/	/doddo/	/taddo/
/paggu/	/boggu/	/uggu/
/puggu/	/goggu/	/oggu/
/peggu/	/gaggu/	/noggu/

[3]One may wonder what would happen to words containing a geminate /pp/—not a singleton /p/—and a voiced geminate. Unfortunately, Japanese prohibits words with two geminates (Ito & Mester, 2003), which does not allow us to test such structures.

3.1.3 Procedure

The experiment was run online using Surveymonkey.[4] The participants were told that within each trial, they would be given one form, and another pronunciation, and would be asked if the latter form is a possible pronunciation of the first form. For the real word stimuli, the participants were told that all the stimuli would be existing loanwords. For the nonce word stimuli, they were told the stimuli were non-existing words in Japanese. All the stimuli were written in the *katakana* orthography, which is usually used for loanwords and nonce words. The real words and nonce words were separated into two different blocks, with a self-timed break in-between, and the block for the real word stimuli was presented first. Within each block, the order of the stimuli were randomized per participant by Surveymonkey.

In the instructions, the participants were asked to read the stimuli in their head, and use the auditory impression to respond to the questions.

3.1.4 Participants

The participants were recruited by word of mouth and through advertisement on a social networking service. Thirty-four native speakers of Japanese completed the online experiment.

3.1.5 Statistics

Since the response was binary (devoicing possible or devoicing impossible), ANOVA was avoided, and instead logistic linear mixed effects model analyses were run (Baayen et al., 2008; Jaeger, 2008). Subjects and items were encoded as random factors. Both slopes and intercepts of random effects were included in the models to have the maximal random structure, following the recent suggestions by Barr (2013) and Barr et al. (2013). R was used to implement the statistical analysis (R Development Core Team, 1993–2015).

3.2 Results

Figure 1 shows the proportions of devoicing possible responses for each condition for real word stimuli, with error bars representing 95% binomial confidence intervals. It shows that the first two bars show higher devoicing proportions than the third bar (87.8%, 85.3% vs. 72.3%).

The first linear mixed model comparing all the three conditions shows that there is a statistically significant difference among the three conditions ($z = 2.52, p < .05$). Subsequent contrast analyses show that the difference between the first condition (the /p/-condition)

[4]Several experiments have shown that online experimentation is as reliable as experiments conducted in a laboratory setting (Reips, 2002; Sprouse, 2011; Yu & Lee, 2014).

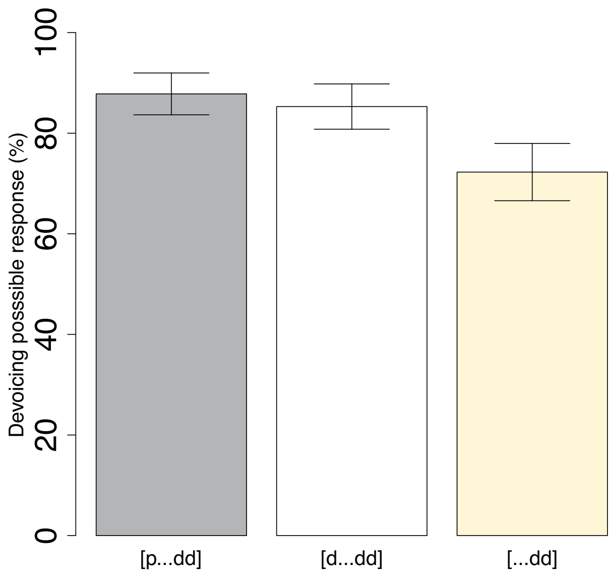


Figure 1: The proportions of devoicing possible responses for each condition: Real words.

and the third condition (the "elsewhere" condition) is significant ($z = 2.50, p < .05$). The difference between the first two conditions was not significant, however ($z = 0.05, n.s.$). This result supports that /p/ indeed causes devoicing of geminates in Japanese, which is also compatible with the patterns found in the corpus, shown in Table 1.

Despite the fact that the overall patterns are compatible with what is expected from the patterns in the corpus, as well as our own intuition, the devoicing possible responses are overall unexpectedly high. This issue is taken up on in the discussion section as well as in Experiment II.

Figure 2 shows the results of nonce words. Although the third condition shows slightly low devoicing possible responses compared to the first two (79.8%, 80.7% vs. 77.3%), the differences were very small; about 2.5%. In fact, a linear mixed model analysis comparing the three different conditions shows no statistically significant differences ($z = 0.03, n.s.$).

3.3 Discussion

The results of the real words show that naive native speakers of Japanese have an intuition that /p/ can cause devoicing at least more than the "elsewhere" condition, just like our own intuition (which could have been biased).

One immediate point to notice, however, is the overall differences between the results

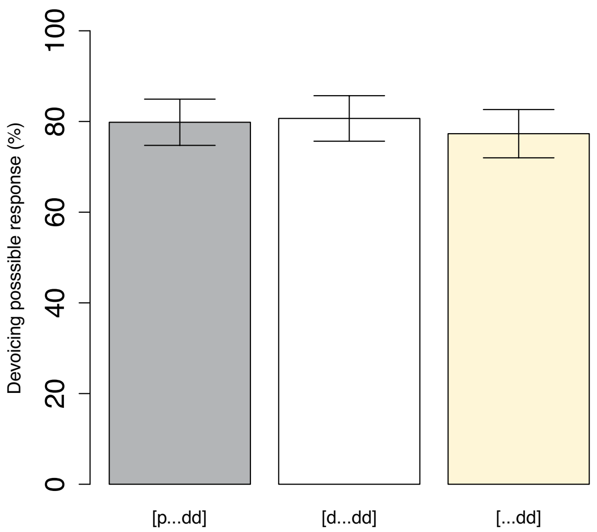


Figure 2: The proportions of devoicing possible responses for each condition: Nonce words.

of the experiment and the patterns in the corpus. For example, /p/-driven devoicing was judged to be possible 87.8% of the time in experiment, but in the corpus, actually 26.7% of geminates appear as devoiced. This discrepancy is more notable in the "elsewhere" condition; in the experiment 72.3% of them were considered to be devoicable, whereas in the actual corpus, only 3.9% of them appear as devoiced. Why is it that devoicing possible responses were so high in the experiment, across all the conditions?

This discrepancy may point to an important lesson about general methodology in phonological judgment experimentation. If the participants were asked whether some phonological pattern is *possible or not*, they may be inclined to be "more forgiving"—or more willing to accept a phonological change—than what they actually do in their actual speech behavior (cf. Labov 1996). This task effect may arise partly because speakers know that other speakers may do what they do not do themselves, possibly due to dialectal or speech style differences. The current participants may have thought that *some* speakers may devoice these geminates in the stimuli, even if they themselves would pronounce them as voiced.

In fact, this task effect of the possible vs. impossible judgment paradigm may not be new. Kawahara (2013) used the same methodology as Experiment I and asked about the devoicability of singletons and geminates in Japanese. The results were that geminates in the "elsewhere" condition were judged to be devoicable 62% of the time, again higher than the corpus data presented above. Singletons were judged to be devoicable 34% of the time when there is another voiced obstruent, and 22% of the time when there is not. We do not have corpus data regarding how often singletons are devoiced in actual utterances, but these percentages seem unrealistically high (see the results of Experiment II). This issue of the task effect observed in Experiment I will be addressed in Experiment II by directly asking the participants what they would actually do.

The results of nonce words, albeit being null results, may be informative as well. Although we should not make a conclusion based on null results, one possible explanation is that the /p/-driven devoicing is "phonologically too outlandish" that it may not be internalized as a productive process in the minds of Japanese speakers, and hence is not extended to nonce words (see Becker et al. 2011 and Hayes et al. 2009 for related discussion). This possibility is discussed in further depth in the general discussion section.

An alternative explanation for the results of the nonce word stimuli is possible. It is independently observed that, for some reason, differences between different grammatical conditions in phonological judgment experiments become smaller when the participants make a judgment about nonce words than when they make judgements about real words (Kawahara, 2010, 2013). Furthermore, responses become closer to a chance level for nonce words than for real words (Kawahara, 2010, 2013). These general characteristics of nonce words in phonological experimentation may have diminished the potential differences between the three conditions in the current experiment.

4 Experiment II: A forced-choice judgment task

In Experiment I, the participant judged devoicing of geminates to be possible much more frequently than what is observed in the corpus as well as our intuition-based expectations. As discussed above, this result may be due to a task effect by which the participants were more accommodating about a possible phonological process when asked if the process is possible or not. In order to address this possibility, Experiment II asked the participants to choose a form that they would actually use.

4.1 Method

4.1.1 Task

As with Experiment I, within each trial, the participants were given one word containing a geminate, but this time presented with two forms, one "faithful" rendition and the other "devoiced rendition". For example, they were asked: "given the word *kyuupiddo*, which pronunciation would you actually use, /kyuupiddo/ or /kyuupitto/?" This task can be considered as a "head-to-head" or "forced-choice" task in which the participants are presented with two forms, and are asked to choose one form that sounds more grammatical (see Daland et al. 2011, Kawahara & Sano 2014 and Kawahara 2015b for the use of this paradigm in phonological experimentation; see also Sprouse & Almeida 2012 and others for the use of this test in experimental syntax).

In this paradigm, the participants may have been biased toward choosing the faithful form, because it is also given as a base form, and also because they may think that devoicing a geminate is "prescriptively not correct". However, the results showed that there were enough non-faithful, devoiced responses, as we will see below.

4.1.2 Stimuli

The current methodology cannot be used for nonce words, because the participants would not know how they actually pronounce words that do not exist. This restriction, however, allowed us to include more conditions than in Experiment I. Therefore, Experiment II included five conditions: (i) geminates appearing with /p/ (e.g. /paddo/), (ii) geminates appearing with another voiced obstruent (e.g. /baddo/), (iii) geminates without either /p/ or a voiced obstruent (e.g. /heddo/—the "elsewhere" condition above), (iv) voiced singletons appearing with another voiced singleton (e.g. /baado/), and (v) voiced singletons without another voiced singleton (e.g. /haado/). This experiment included singleton conditions, because devoicability of singletons has never been tested in this task format.

The experimental items for the real words and nonce words are provided in Table 4.[5] The stimuli for the first three conditions were almost identical to that of Experiment I, except that those words with /gg/ were replaced with those with /dd/. By mistake, /gaado/ was included twice in the fourth condition, and the responses for the first occurrence of /gaado/ were excluded from the analysis.

4.1.3 Procedure and statistics

The procedure and the statistical analyses were identical to those of Experiment I.

4.1.4 Participants

Fifty native speakers of Japanese completed the online experiment. Since the participants were invited partly through a social network service, there may be some overlap of the participants between the two experiments. However, since the two experiments were conducted more than a year apart, we expected little or no influences.

[5]Singleton consonants tend to be preceded by a long vowel or a diphthong because gemination is very common after a short vowel in the phase of loanword adaptation (Katayama, 1998; Kubozono et al., 2008).

Table 4: The stimuli for Experiment II.

/p...dd/		/b...dd/		/...dd/	
/kyuupiddo/	'cupid'	/baddo/	'bad'	/heddo/	'head'
/paddo/	'pad'	/beddo/	'bed'	/reddo/	'red'
/aipaddo/	'i-pad'	/deddo/	'dead'	/uddo/	'wood'
/aipoddo/	'i-pod'	/guddo/	'good'	/kiddo/	'kid'
/supureddo/	'spread'	/daddo/	'dad'	/maddo/	'mad'
/piramiddo/	'pyramid'	/goddo/	'god'	/roddo/	'rod'
/tetorapoddo/	'tetrapod'	/budda/	'Buddha'	/middo/	'mid'

/d...d/		/....d/	
/baraado/	'ballad'	/haado/	'hard'
/baado/	'bird'	/raado/	'lard'
/bideo/	'video'	/kaado/	'card'
/gaado/	'guard'	/koodo/	'cord'
/boodo/	'board'	/roodo/	'road'
/gaaden/	'garden'	/saido/	'side'
		/huudo/	'hood'

4.2 Results

Figure 3 shows the proportions of devoiced forms being selected, with error bars representing 95% binomial confidence intervals. We observe declining order of devoiced forms being selected from left to right (46.6%, 31.4%, 7.7%, 0.6%, 0%).

Setting aside the singleton consonants, which showed few or no devoiced responses, a linear mixed model comparing the first three conditions, all containing geminates, showed a significant effect ($z = 3.75, p < .001$). The difference between the first bar (the /p/-condition) and the third condition (the "elsewhere" condition) is significant ($z = 4.31, p < .001$); the difference between the first two conditions did not reach significance ($z = 1.83, n.s.$), although it is near significance ($p = .067$).

4.3 Discussion

First of all, the experimental results show that /p/ does cause devoicing of geminates, as much as—possibly more than—a voiced obstruent does. This result supports our initial intuition that words of the form /p...dd/ may be pronounced as /p...tt/ with fifty naive native speakers of Japanese. Although the same results were obtained in Experiment I for real words, the differences between the /p/ condition and the "elsewhere" condition were more pronounced in the current study.

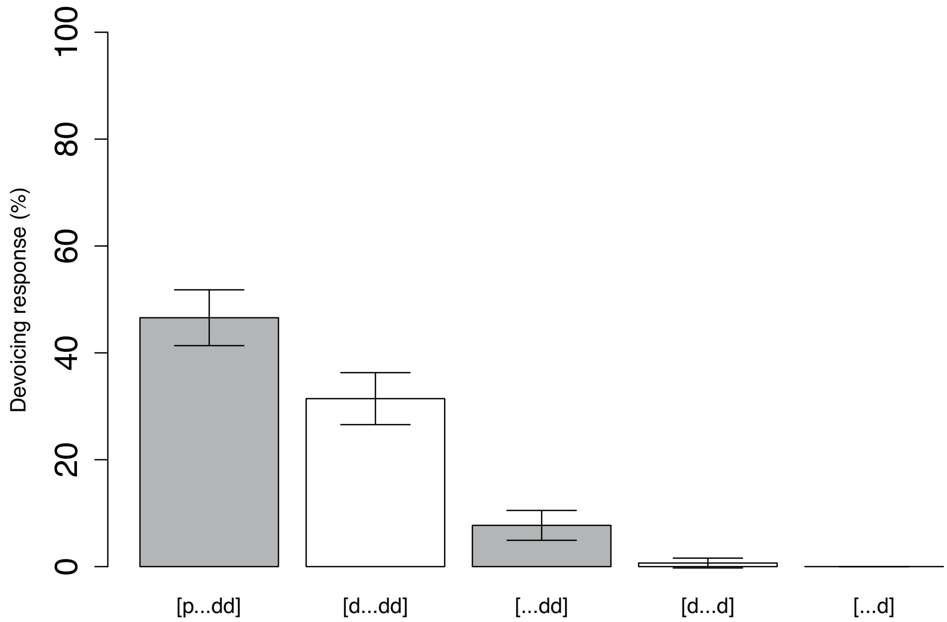


Figure 3: The proportions of devoicing responses for each condition.

Compared to Experiment I, the current experimental paradigm yielded values that are closer to those that are found in the corpus studies, at least for the /p/ condition and the "elsewhere" condition. Recall also that the participants did not choose devoiced singletons in Experiment II, while Kawahara (2013) found that in a possible vs. impossible format, singletons were judged to be devoicable about 20-30% of the time. Which experiment better reflects the true production of voiced singletons by Japanese speakers needs to be checked against actual production patterns, although we are inclined to say, admittedly based on our own intuition, that the current paradigm yielded results that are closer to what Japanese speakers actually do.

This task effect is perhaps not too surprising because the current experiment asked the participants about what they would actually do. This result again highlights the importance of probing task effects in phonological experimentation. It may be the case that this type of methodology is better than the possible vs. impossible format deployed in Experiment I and Kawahara (2013), especially when we are interested in what the speakers actually do. However, this paradigm has a drawback of not being able to use nonce word stimuli.

Another point to be noted about this experimental paradigm is that this experiment was unexpectedly successful in that it elicited enough "non-faithful" responses. Given that the devoicing process is optional, there was a concern that the participants may not choose the devoiced options at all.[6] Recall that the participants were presented with a word A, and they were asked if they would pronounce it as A (faithful) or A' (non-faithful).[7] It would not be surprising if all the participants would have chosen A throughout, especially if they think they should be "prescriptively correct". The experimental results show that the head-to-head methodology is possible when probing an optional process, because participants do choose unfaithful options. This methodology can and perhaps should be applied to other optional phonological processes, such as t/d-deletion in English (Coetzee & Pater, 2011; Coetzee & Kawahara, 2013; Guy, 1991).[8]

5 Possible analyses

The corpus-based study as well as the judgement experiments have shown that /p/ does cause devoicing of geminates in Japanese loanwords. Although the focus of this paper is descriptive (i.e. to attempt to find objective evidence for /p/-driven devoicing), we briefly entertain possible ways of modeling this pattern theoretically. Overall, modeling the interaction between /p/ and voiced geminates is challenging, because /p/ and voiced geminates do not have anything in common, at least superficially. At least none of the theoretical analyses that are developed for devoicing of geminates due to another voiced obstruent predicts this /p/-driven devoicing (Coetzee & Kawahara, 2013; Coetzee & Pater, 2011; Farris-Trimble, 2008; Kawahara, 2006; Nishimura, 2003; Pater, 2009; Rice, 2006). This is because essentially all the analyses assume or posit that the cause of devoicing is OCP(voice), but /p/ should not induce a violation of OCP(voice).[9]

5.1 Prohibition against two rare structures?

One possible analysis is to build on the observation that both singleton /p/'s and voiced geminates are allowed only in loanwords (Ito & Mester, 1995, 1999, 2008). In this sense, /p/

[6]If the target phenomenon is a morphophonological process involving morpheme concatenation, then this paradigm is equivalent to a well-known wug-test in a forced choice format (Berko, 1958).

[7]A comparable question in English would be: "given a word like *west*, would you usually pronounce it as /west/ or /wes/?"

[8]Anonymous (p.c.) pointed out that "[f]rom a sociolinguistic perspective, the finding that participants are willing to acknowledge devoicing indicates that this is not a stigmatized variant associated with non-standard speech, in contrast to the other processes suggested, such as t/d deletion in English".

[9]Or should it? We could potentially postulate that /p/ and /b/ are both [+voice], and distinguish them in terms of [spread glottis], for example. As far as we know, there is no independent evidence that /p/ is [+voice] in Japanese, however. The fundamental problem here is that /p, b, d, g/ do not form a natural class for other phonological patterns in Japanese.

and voiced geminates share the property of being unfamiliar or non-frequent in the entire lexicon of Japanese. Fukazawa et al. (2015) show based on the lexical search of Amano & Kondo (1999) that /p/ and voiced geminates are the two most infrequent sounds in the whole Japanese lexicon. They argue that there may be a sort of OCP constraint, independent of OCP(voice), which prohibits the occurrence of two unfamiliar segments within a word; i.e. OCP(unfamiliar). This theory predicts that there should be other languages which prohibit a word that contains two sounds that are infrequent or two sounds that are found only in loanwords. This prediction is yet to be explored in other languages.

5.2 Local conjunction?

A related possibility is to posit a locally conjoined constraint (Smolensky 1993 *et seq.*), in the spirit of Nishimura (2003), like $\{^*/\text{p}/\&^*\textsc{VoiceObsGem}\}_{stem}$. However, this constraint seems nothing more than a restatement of the observation, and predicts that, cross-linguistically, two irrelevant markedness constraints can be conjoined with a domain as large as a stem. Allowing this sort of local conjunction would probably result in too much theoretical power (Kawahara, 2006; McCarthy, 2003; Pater, 2009, to appear), although Blust (2012) recently argues that this sort of powerful local conjunction is necessary after all. The prediction of this approach is that any segment can be a trigger of geminate devoicing, as long as there exists a markedness constraint against that segment. This prediction needs to be tested in further detail, although we are suspicious of this prediction at this point.

Anonymous (p.c.) pointed out an interesting follow-up of this hypothesis. Although /p/ is voiceless, /p/ is the 'worst' of the voiceless stops for aerodynamic reasons—cession of vocal fold vibration is hardest for /p/ because it has the largest intraoral space (Hayes, 1999; Ohala, 1983). In fact, Maddieson (2013) documents a number of languages that lack /p/, just like the native phonology of Japanese (Ito & Mester, 1995, 1999, 2008). Thus, both /p/ and voiced geminates are in some sense 'laryngeally marked', and /p/-driven geminate devoicing could be understood as a pattern that avoids the presence of too many laryngeally marked things in the same word. To put it another way, even though /p/ and voiced geminates do not have the same value of the feature [voice], they are both marked with respect to that feature. To formalize the idea, we can formulate a constraint like OCP(LaryngealyMarked).

5.3 An orthography-based explanation?

A more radical alternative analysis is possible based on the Japanese orthography: voiced obstruents and /p/ are shown with diacritic marks on the upper right corner, the former with *dakuten* and the latter with *han-dakuten*, as illustrated by some examples in Figure 4.

ハ = /ha/

パ = /pa/

バ = /ba/

Figure 4: The Japanese orthography symbols for /ha/, /pa/, and /ba/. Those for /pa/ and /ba/ are shown with a diacritic mark on the upper right corner.

Therefore both /p/'s and voiced obstruents are written with an orthographic diacritic; in this sense, /p/ and voiced obstruents form a natural class. It may be the case then that OCP(voice) is actually OCP(diacritic), which accounts for both /p/-driven devoicing and OCP-driven devoicing at the same time. This analysis is radical in the sense that it shifts the burden of explanation from sounds to letters—a move that should be cautiously taken from the viewpoint of phonological theory (though see Ito et al. 1996 for a similar example of an orthography-based explanation of a phonological observation). This theory makes a specific prediction, namely in terms of the behavior of pre-literature children. To the extent that Lyman's Law or OCP(voice) is a matter of orthography, so as to target the configuration /p...dd/, then preliterate Japanese-speaking children should not show the evidence for Lyman's Law, or at least /p/-driven devoicing. In fact, Fukuda & Fukuda (1994) do find evidence to this effect—under-learning of Lyman's Law. A more targeted longitudinal experiment is called for to fully defend this prediction, however, especially to examine whether pre-literate children fail to show /p/-driven devoicing.[10]

5.4 It is not phonological after all?

A final possible explanation is to say that /p/-driven devoicing occurs in existing loanwords, but this pattern is too "outlandish" as a phonological pattern so that it is not phonologized in the minds of the contemporary Japanese speakers (cf. Becker et al. 2011). This idea would capitalize on the null results of the nonce words in Experiment I. Although this analysis is not impossible, it does not explain why the existing words show evidence for /p/-driven devoicing. This devoicing pattern of geminates emerged spontaneously in the loanword phonology, because Japanese did not used to have voiced geminates in the native phonology. This theory thus fails to explain why /p/-driven devoicing emerged in the first place. This

[10]This theory makes another specific prediction about rendaku and its blockage by Lyman's Law. To the extent that Lyman's Law is a prohibition against two diacritics, then /p/ should block rendaku as well. Rendaku generally applies only to native words, but singleton /p/s do not appear in native words. There are words, however, that undergo rendaku despite the presence of geminate /pp/; e.g. /ama+**zupp**ai/ 'sweat and sour'.

theory also has a danger of relying on null results.

5.5 Summary

Since there are no knock-out arguments for one over the other, we remain neutral about the best explanation of the /p/-devoicing pattern in this paper. We reiterate that the value of this paper mostly lies in its new descriptive discovery, in addition to the new findings about task effects in phonological experimentation.

We will close this section by raising one final question, which may bear on the theoretical analyses: are /p/-driven devoicing and OCP-driven devoicing the same phonological pattern or are they different? The corpus study shows that OCP-driven devoicing is more common; neither conditions in Experiment I revealed differences between the two; Experiment II shows that /p/-driven devoicing tends to be more common, although the difference did not reach statistical significance. At this point, the evidence seems so mixed that this issue needs to be resolved in a future study.

6 Conclusion

The main goal of this paper was to objectively examine, beyond our own intuition, whether /p/ causes devoicing of geminates in Japanese loanword phonology. A corpus study as well as two judgment experiments show that /p/ does cause devoicing, at least in existing words. This new descriptive discovery poses a challenge to the theoretical analysis of geminate devoicing patterns in Japanese phonology in general.

In addition to this new descriptive discovery, we found non-trivial task effects in phonological experimentation. Experiment I shows that when the participants were asked if a phonological process is possible or not, they may be more inclined to say yes, perhaps more than they actually do. Experiment II shows that it is possible to ask what they do in their phonological behavior of an optional phonological process.

All in all, it seems safe to conclude that /p/, just like voiced obstruents, causes devoicing of geminates in Japanese. We look forward to seeing future theoretical analyses developed to account for these devoicing patterns.

References

Amano, Shigeaki & Tadahisa Kondo (1999) Nihongo-no goitokusei [Psylex: Lexical Properties of Japanese]. Tokyo: Sanseido.

Baayen, Harald R., Doug.J. Davidson, & Douglas. M. Bates (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**: 390–412.

Barr, Dale J. (2013) Random effects structure for testing interactions in linear mixed-effects models. *Frontiners in Psychology* **4**: 328.

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.

Becker, Michael, Nihan Ketrez, & Andrew Nevins (2011) The surfeit of the stimulus: Grammatical biases filter lexical statistics in Turkish voicing deneutralization. *Language* **88**(2): 231–268.

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

Blust, Robert (2012) One mark per word? Some patterns of dissimilation in Austronesian and Australian languages. *Phonology* **29**(3): 355–381.

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.

Crawford, J. Clifford (2009) *Adaptation and transmission in Japanese loanword phonology.* Doctoral dissertation, Cornell University.

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

Farris-Trimble, Ashley (2008) *Cumulative faithfulness effects in phonology.* Doctoral dissertation, Indiana University.

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**: 3–10.

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.

Guy, Gregory (1991) Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* **3**: 1–22.

Hayes, Bruce (1999) Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In *Functionalism and Formalism in Linguistics, vol. 1: General Papers*, Michael Darnell, Edith Moravscik, Michael Noonan, Frederick Newmeyer, & Kathleen Wheatly, eds., Amsterdam: John Benjamins, 243–285.

Hayes, Bruce (2009) Embedding grammar in a quantitative framework: Case studies from phonology and metrics. A handout for a minicourse at Brown University (available at http://www.linguistics.ucla.edu/people/hayes/brownminicourse/).

Hayes, Bruce, Kie Zuraw, Péter Siptár, & Zsuzsa Londe (2009) Natural and unnatural constraints in Hungarian vowel harmony. *Language* **85**(4): 822–863.

Ito, Junko, Yoshihisa Kitagawa, & Armin Mester (1996) Prosodic faithfulness and correspondence: Evidence from a Japanese Argot. *Journal of East Asian Linguistics* **5**: 217–294.

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 (1995) Japanese phonology. In *The Handbook of Phonological Theory*, John Goldsmith, ed., Oxford: Blackwell, 817–838.

Ito, Junko & Armin Mester (1999) The phonological lexicon. In *The Handbook of Japanese Linguistics*, Natsuko Tsujimura, ed., Oxford: Blackwell, 62–100.

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.

Jaeger, Florian T. (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**: 434–446.

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.

Katayama, Motoko (1998) *Optimality Theory and Japanese Loanword Phonology.* Doctoral dissertation, University of California, Santa Cruz.

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 (2010) Modes of phonological judgments. Ms. Rutgers University.
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 (2015a) Can we use rendaku for phonological argumentation? *Linguistic Vanguard* : 1–12.
Kawahara, Shigeto (2015b) Comparing a wug-test and a naturalness rating test: An exploration using rendaku. *Language Sciences* **48**: 42–47.
Kawahara, Shigeto (2015c) Geminate devoicing in Japanese loanwords: Theoretical and experimental investigations. *Language and Linguistic Compass* **9**(4): 168–182.
Kawahara, Shigeto & Shin-ichiro Sano (2014) Identity Avoidance and Lyman's Law. *Lingua* **150**: 71–77.
Kawahara, Shigeto & Shinichiro Sano (2013) A corpus-based study of geminate devoicing in Japanese: Linguistic factors. *Language Sciences* **40**: 300–307.
Kiparsky, Paul (1982) From cyclic phonology to lexical phonology. In *The Structure of Phonological Representations*, vol. 1, Harry van der Hulst & Norval Smith, eds., Dordrecht: Foris, 131–175.
Kubozono, Haruo, Junko Ito, & Armin Mester (2008) Consonant gemination in Japanese loanword phonology. In *Current Issues in Unity and Diversity of Languages. Collection of Papers Selected from the 18th International Congress of Linguists*, The Linguistic Society of Korea, ed., Republic of Korea: Dongam Publishing Co, 953–973.
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.
Maddieson, Ian (2013) Voicing and gaps in plosive systems. In *The World Atlas of Language Structurse Online, Chapter 5*, Matthew Dryer & Martin Haspelmath, eds., Max Planck Institute for Evolutionary Anthropology.
Maekawa, Kikuo (2004) Nihongo hanashikotoba koopasu-no gaiyoo [An overview of the Corpus of Spontaneous Japanese]. *Nihongo Kagaku* **15**: 111–133.
Maekawa, Kikuo, Hanae Koiso, Sadaoki Furui, & Hitoshi Isahara (2000) Spontaneous speech corpus of Japanese. *Proceedings of the Second International Conference of Language Resources and Evaluation* : 947–952.
McCarthy, John J. (2003) Comparative markedness. *Theoretical Linguistics* **29**: 1–51.
McCarthy, John J. (2008) *Doing Optimality Theory*. Oxford: Blackwell-Wiley.
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.
Nishimura, Kohei (2006) Lyman's Law in loanwords. *On'in Kenkyu [Phonological Studies]* **9**: 83–90.
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. (1986) Consumer's guide to evidence in phonology. *Phonology* **3**: 3–26.
Pater, Joe (2009) Weighted constraints in generative linguistics. *Cognitive Science* **33**: 999–1035.
Pater, Joe (to appear) Universal grammar with weighted constraints. In *Harmonic Grammar and Harmonic Serialism*, John J. McCarthy & Joe Pater, eds., London: Equinox.
R Development Core Team (1993–2015) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
Reips, Ulf-Dietrich (2002) Standards for internet-based experimenting. *Experimental Psychology* **49**(4): 243–256.
Rice, Keren (2006) On the patterning of voiced stops in loanwords in Japanese. *Toronto Working Papers in Linguistics* **26**: 11–22.
Sano, Shin-ichiro (2013) Patterns in avoidance of marked segmental configurations in Japanese loanword

phonology. *Proceedings of GLOW in Asia* : 245–260.
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 (1996) *The empirical base of linguistics: Grammaticality judgments and linguistic methodology.* Chicago: University of Chicago Press.
Smolensky, Paul (1993) *Optimality, markedness, and underspecification.* Paper presented at the Rutgers University Optimality Workshop, New Brunswick, NJ.
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.
Sprouse, Jon & Diogo Almeida (2012) Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics* **48**: 609–652.
Tesar, Bruce (2007) A comparison of lexicographic and linear numeric optimization using violation difference ratios. Ms. Rutgers University.
Tsujimura, Natsuko (2014) *An Introduction to Japanese Linguistics, 3rd Edition.* Oxford: Blackwell-Wiley.
Vance, Timothy (1987) *An Introduction to Japanese Phonology.* New York: SUNY Press.
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 (2008) *The Sounds of Japanese.* Cambridge: Cambridge University Press.
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.