To appear in *Language*.

A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese

Abstract[*]:

Within the framework of Optimality Theory (Prince and 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, only 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. Overall, this paper has a general theoretical implication that phonetic perceptibility can directly influence patterns in a phonological grammar.

[*] Acknowledgements: 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 paper. I am especially grateful to José Benkí, Kathryn Flack, Brian Joseph, John Kingston, John McCarthy, Joe Pater, Donca Steriade, and an anonymous reviewer for their extensive comments on earlier versions of this paper. All remaining errors are mine.

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 and Halle (1968: chapter 9), Stampe (1973; Natural Phonology), Hooper (1976; Natural Generative Phonology), and Archangeli and Pulleyblank (1994; Grounded Phonology) among many others. With the advent of Optimality Theory (OT; Prince and 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 place distinctions are not salient in codas (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 salient perceptible change an alternation involves, the higher-ranked faithfulness constraint it violates. For example, a voicing contrast is more saliently perceived prevocalically than preconsonantally—i.e. the contrast between [pa]~[ba] is more salient and perceptible than the contrast between [apta]~[abta] (Steriade 1997). Based on this knowledge of similarity, speakers project the faithfulness ranking Faith(voi)$_{/}$ _ V » Faith(voi)$_{/}$ _ C. As a consequence, a voicing contrast is more prone to neutralization in preconsonantal position than in prevocalic position, because 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 paper 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). On the other hand, as the words in 2 show, voiced singletons do not devoice even when they co-occur with another voiced obstruent.

(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 (Faith(voi)$_{Sing}$) and one for geminates (Faith(voi)$_{Gem}$), and that Faith(voi)$_{Sing}$ be ranked higher than Faith(voi)$_{Gem}$. 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 two experiments which show that this prediction of the P-map hypothesis is indeed borne out. This paper thus overall provides empirical endorsement for the P-map's central claim. A larger implication of this paper 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 (Prince and Smolensky 2004). I show that Faith(voice)$_{Gem}$ is ranked lower than Faith(voice)$_{Sing}$. 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 and Mester 1995, 1999, Kuroda 1965), they are allowed in recent loanwords.[1] In recent Japanese loanwords from foreign languages (mainly English), coda

---

[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 words, as reflected by the fact that they use different orthography systems for these two lexical classes. See also Amano and Moreton (1999) and Gelbart and

consonants which 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 English (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 paper, 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'

Kawahara (2005) for evidence from perceptual experiments which supports the psychological reality of lexical stratification in Japanese.

[2] The Japanese data reported in this paper 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 (Katamaya 1998).

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 devoice when they appear with another voiced obstruent.

gebberusu	~	gepperusu	‘Göbbels’
guddo	~	gutto	‘good’
beddo	~	betto	‘bed’
doreddo	~	doretto	‘dredlocks’
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'

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 86 hours of spoken Japanese, including both formal and spontaneous speech. In the spontaneous speech transcribed in the corpus, out of 54 DVDDV words, which canonically have voiced geminates, 34 of them appear in this corpus with voiceless geminates (57.4%). On the other hand, only 1 out of 27 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.

Unlike voiced geminates in DVDDV words, when there are two singleton voiced

---

[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.

obstruents in a word, devoicing is impossible. This generalization 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]

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

		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 I: IS DEVOICING CATEGORICAL?[5] Before developing an analysis of the patterns summarized in Table 1, I address the question of whether the devoicing which 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

[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.

remainder of this subsection reports an acoustic experiment which demonstrates this claim.

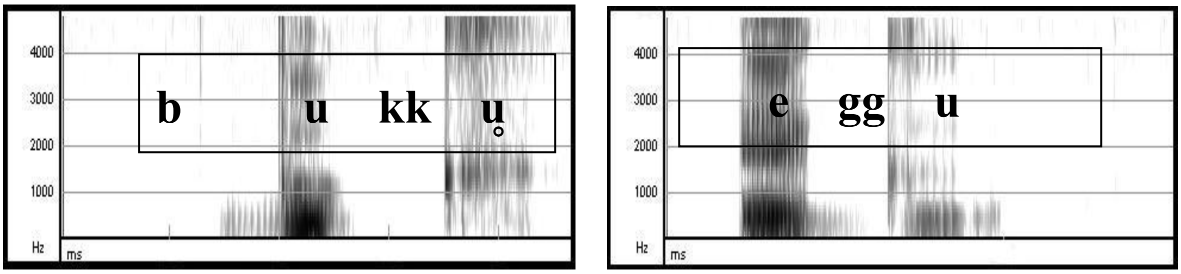
A male native speaker of Tokyo Japanese was recorded. He was in his early thirties, and was paid for his time. The speaker was naïve to the purpose of this study. His speech was recorded through a microphone (MicroMic II C420 by AKG) by a CD-recorder (TASCAM CD RW-700) at a 44.1KHz sampling rate, in a sound-attenuated booth at University of Massachusetts, Amherst. The recorded tokens were then downsampled to 22.050KHz 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, the stimuli were all 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'

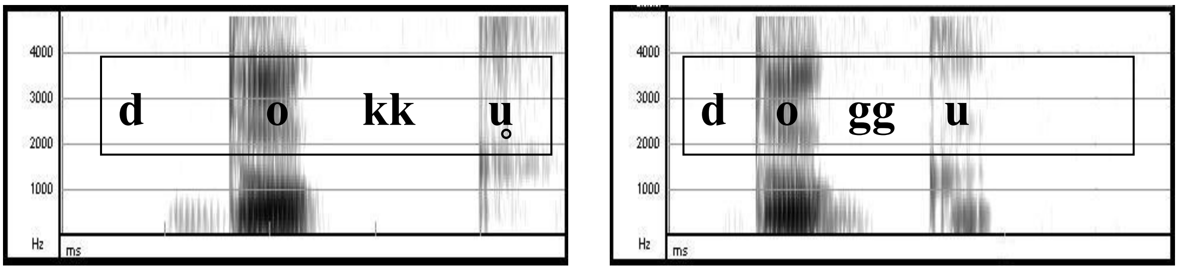
In addition to these target words, other 15 real words were added as fillers. Each stimulus was embedded in a different frame sentence. The words that followed the target words began with a voiceless consonant in order to facilitate devoicing—e.g. [bukku totte] 'please hand me a book'. The speaker was 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 had been 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 30 minutes. In this experiment, the speaker pronounced devoiced variants of 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 Figure 1a, the word-final [u] in [bukku] has aperiodic energy, and shows no acoustic reflexes of regular glottal pulses, indicating that the [u] is devoiced. In contrast, in Figure 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. The time scale is the same (750ms). A high vowel devoices after [kk] (left), but not after [gg] (right).



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 (left), but not after a voiced geminate (right).

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 Figure 2 we observe that the word-final [u] is devoiced only in Figure 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.

We can observe in Figure 1b and Figure 2b that much of the oral closure interval in [gg] has no voicing—i.e. 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, Zsiga 1995, 1997 among others for the distinction between gradient phonetic processes and categorical phonological processes).

In addition to the evidence from the devoicing patterns of high vowels, there are several other pieces of acoustic evidence for neutralization of underlying /gg/ to [kk]; namely, the closure duration and closure voicing duration. Comparing the spectrograms in Figure 1, we can see 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]. Now comparing the spectrograms in Figure 2 the same patterns emerge: 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 point where F3 disappears. The results are summarized in Figure 3 and Figure 4.

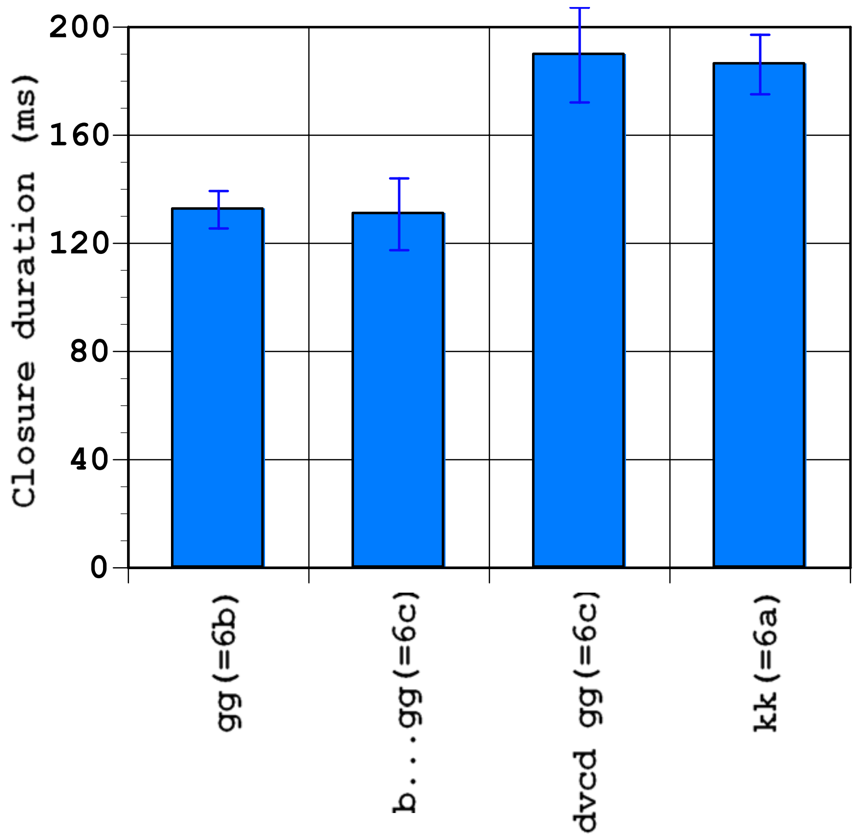


Figure 3: Mean closure duration of the words in 6. Error bars represent 95% confidence intervals. The first bar represents [gg] in TVDDV words (=6b). The second bar represents [gg] in DVDDV words (=6c). The third bar represents [kk] in DVDDV words (6c). The fourth bar represents [kk] derived from /kk/ (=6a).

As seen in Figure 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 ($t(22)=.35$, $p=.73$). Furthermore, phonologically devoiced /gg/ (the third bar) is longer than voiced /gg/ (the first and second bars, respectively), and the difference is statistically significant ($t(28)=7.77$, $p<.001$). In short, [kk] derived from /gg/ patterns with [kk] derived from /kk/, not with [gg] derived from /gg/.

The same pattern emerges in terms of closure voicing duration, summarized in Figure 4.

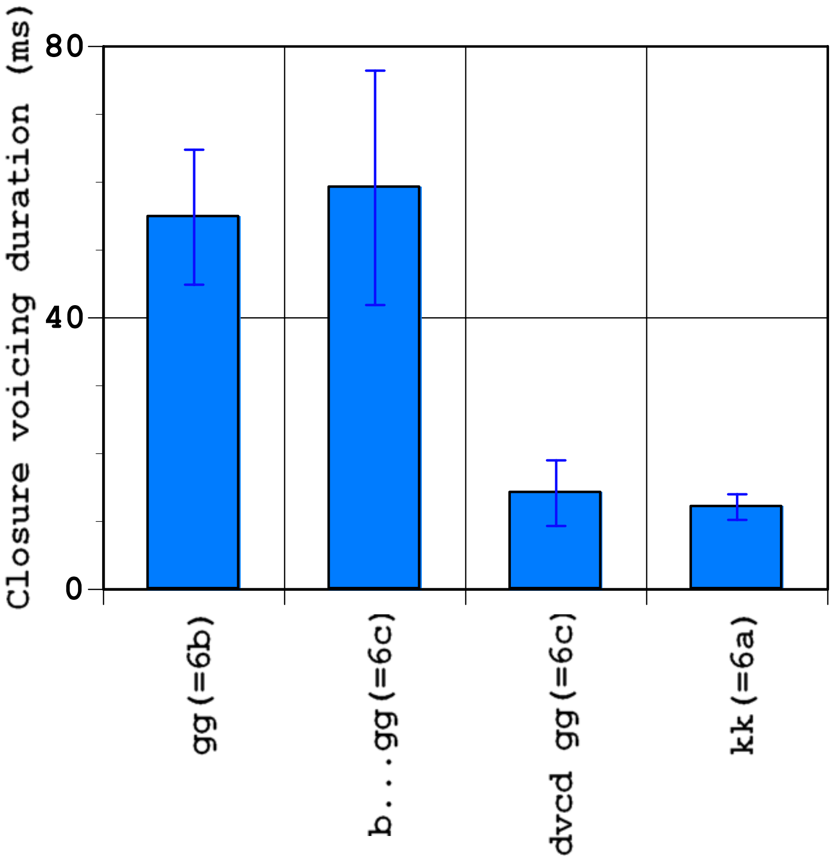


Figure 4: Mean closure voicing duration of the words in 6.

In terms of closure voicing duration, there is no difference between devoiced /gg/ (the third bar) and [kk] derived from /kk/ (the fourth bar) ($t(22)=1.09$, $p=.29$). Furthermore, closure voicing is longer for voiced /gg/ (the first and second bar) than for devoiced /gg/ (the third bar) ($t(28)=4.73$, $p<.001$). All of these acoustic observations again suggest that devoicing of the geminates in 6c, as shown in Figure 2b, is complete 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, this section presents a phonological analysis of the behavior of voiced consonants in Japanese loanwords. This subsection first discusses some assumptions crucial to the proposed analysis.

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

In addition, OT is suitable for the current purpose because of the central role of faithfulness in the theory. In addition to markedness constraints that 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 theory of faithfulness, I argue below that in Japanese, the faithfulness constraints which 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 paper is not the process of gemination in loanword adaptation, but the behavior of a [+voice] feature in the synchronic loanword phonology of Japanese. 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 e.g. Katayama (1998), Lovins (1973), Takagi and 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. On the other hand, [+voice] is well perceived in singletons, and it is therefore protected by a higher-ranked constraint (see Boersma 1998, Côté2004, Fleischhacker 2001, Jun 2004, Padgett 2002, Steriade 1997, 2001a, b, Zhang 2000, Zuraw 2005 for related proposals).

To formally express the proposal, I employ the Ident family of constraints which regulate featural changes (McCarthy and 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.

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].

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, in Japanese loanwords the ranking Ident(+voi)$_{\text{Sing}}$ » Ident(+voi)$_{\text{Gem}}$ holds, and it is grounded in perceptibility of [+voice] in Japanese singletons and geminates. See §6.2 for discussion of (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 only devoicing but not voicing, unlike a more general Ident(voi), which prohibits both devoicing and voicing (see Itô and Mester 2003: Chapter 7, Pater 1999, Walker 2001 for related discussion). In theory, either formulation works; however, Ident(+voi) is chosen to reflect the fact that this paper primarily deals 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 [+voice] 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 and 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, the restriction against two voiced obstruents also 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'. However, if the second stem already contains a voiced obstruent, this voicing process is blocked, as in /nise-taba/ → [nise-taba], *[nise-daba] 'fake bill'.

Itô and Mester (1986) formalize the restriction against two voiced obstruents as an effect of the Obligatory Contour Principle (OCP: Goldsmith 1973, 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 restriction as OCP(+voi).

---

[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.

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

Using OCP(+voi) and the two separate faithfulness constraints $\text{Ident}(+\text{voi})_{\text{Sing}}$ and $\text{Ident}(+\text{voi})_{\text{Gem}}$ proposed above, the patterns presented in §2.1 can be easily accounted for. First, since singleton consonants do not devoice under the duress of OCP(+voi) in DVDV words, $\text{Ident}(+\text{voi})_{\text{Sing}}$ dominates OCP(+voi), as shown by the tableau in (9).

(9) $\text{Ident}(+\text{voi})_{\text{Sing}}$ » OCP(+voi).

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

In contrast, in DVDDV words, the winning candidate satisfies OCP(+voi) while violating $\text{Ident}(+\text{voi})_{\text{Gem}}$ i.e. devoicing takes place. Thus, OCP(+voi) must be ranked higher than $\text{Ident}(+\text{voi})_{\text{Gem}}$, as shown in the tableau in 10. To account for optionality of devoicing, OCP(+voi) and $\text{Ident}(+\text{voi})_{\text{Gem}}$ can be left unranked; see e.g. Antilla (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 the tableau 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. In §2.1, we saw that voiced singletons and voiced geminates react differently to OCP(+voi). I argued above that this difference 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 (2003) analysis uses *VoiObsGem, which directly prohibits voicing in geminates. Since voiced geminates themselves are not devoiced, Ident(+voi) must be ranked above *VoiObsGem. Also, 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 the tableau in 12 where the wrong winner is indicated by '(☞)'.

(12) The desired candidate [bakku] fails to win.

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

Nishimura (2003) 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 *VoiObsGem are conjoined in the domain of stems, yielding $\{OCP(+voi)\&*VoiObsGem\}_{Stem}$, and this constraint is ranked above Ident(+voi), as shown in the tableau in 13.

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

/baggu/	{OCP(+voi)& *VoiObsGem}$_{Stem}$	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 several problems.

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

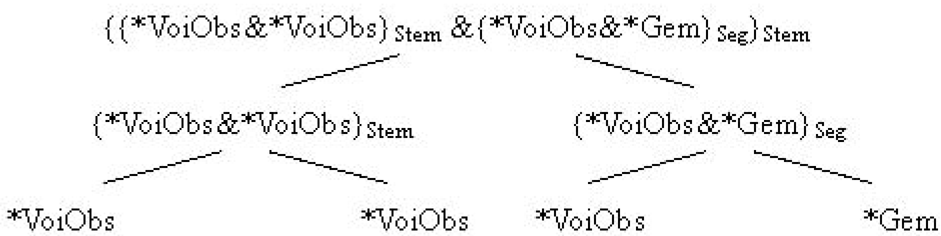


Figure 5: Recursive local conjunction.

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 and Prince 1986). However, by recursively self-conjoining *X, it is possible to derive a language that counts the instances of X within a domain D via a constraint like $\{\{*X\&*X\}_{D}\&*X\}_{D}$ (Itô and Mester 1998: footnote 17). This prediction is undesirable.

The second problem is that this approach allows non-identical 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 of such constraints are given in 14.

(14) Predicted conjoined constraints that are unattested.

$\{$*Lab & NoCoda$\}_{Stem}$:	A labial and a coda consonant cannot co-occur within a stem.
$\{$Max & NoCoda$\}_{Stem}$:	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$\}_{Stem}$:	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. This section discusses the theoretical consequences of the proposal presented here, focusing on its typological predictions. Optimality Theory (Prince and Smolensky 2004) is inherently typological, since the set of constraints is assumed to be universal and all variation between languages comes from the language-particular rankings of these constraints. Therefore, it is necessary to consider cross-linguistic 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 and 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 Ident(+voi)$_{Sing}$ and Ident(+voi)$_{Gem}$ is grounded in the perceptibility of [+voice] in singletons and geminates, this ranking restriction should also be grounded in the perceptibility of [+voice] in singletons and geminates as well. See §6.2 for further discussion of this point.

Second, with the elaboration of faithfulness constraints proposed in this paper, it is desirable in terms of theoretical parsimony to eliminate *VoiObsGem, which prohibits voiced geminates but not voiced singletons (Hayes and Steriade 2004, Itô and 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 only by devoicing, but also by other processes as well. For example, as shown in 15a, 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. However, as in 15b, when the second consonant is a voiced obstruent, 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 *VoiObs 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 the tableau 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 paper to argue for the absence of such a process with certainty, or to reanalyze all 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 can 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

---

[7] Thanks to the anonymous reviewer 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 and Steriade 2004, Jaeger 1978, Ohala 1983, Westbury 1979).

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] indeed differs 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 is 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 that this paper provides is necessarily limited—this paper shows that [+voice] is less perceptible in geminates than in singletons, and that phonologically, [+voice] is more easily neutralized in geminates than in singletons. No attempt is made, 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. A challenge to this position is the fact that there is a language like Japanese that nasalizes coda consonants to repair underlying voiced geminates (see 15), as well as a language like Endegeň that 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 II: ACOUSTICS OF [±VOICE] IN JAPANESE.[8] To examine whether a voicing contrast is indeed harder to perceive in geminates than in singletons, this experiment investigated (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. Next, 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 mean 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

---

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

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 44.1KHz sampling rate, in a sound-attenuated booth. The recorded tokens were then downsampled to 22.050KHz 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 45 minutes.

The stimuli consisted of 36 words, which were mostly nonce words.[10] In addition, 36 other nonce words were added 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 for both the first and second syllable. Some examples are [kepe], [kabba], [kete], [koddo], [kaga], [kekke] etc. 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 in katakana orthography, which is conventionally used for loanwords. Katakana orthography was used because voiced geminates are found

---

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

[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 real words.

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 and 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 whose names were the target words.

All measurements were done using Praat (Boersma and Weenink 2005). Following the past literature on acoustic and perceptual correlates of a [±voice] contrast (Kingston and Diehl 1994, Lisker 1986, Raphael 1981, Stevens and 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 and Abramson 1964 *et seq*.), but it was not measured because [±voice] in Japanese is not signaled by aspiration. The way in which each acoustic property was measured is illustrated in Figure 6, using a representative spectrogram of [kobbo].

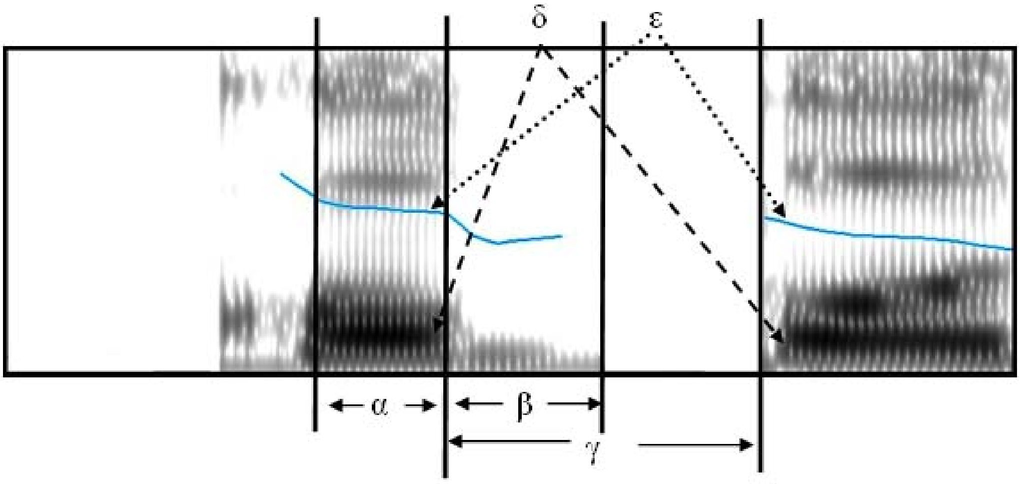


Figure 6: An illustrative spectrogram of [kobbo], which shows how each acoustic value was measured. α =V1 duration, β=closure voicing duration, γ=closure duration, δ=F1 at vowel edges, ε=F0 at vowel edges.

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 a voicing contrast (2-level) and consonantal length (2-level) as independent variables. These variables were chosen because we are interested in how a [±voice] difference 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 inter-speaker 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. This subsection first reports the acoustic differences between voiced and voiceless consonants, and then shows that there are at least three ways in which the [±voice] contrast in geminates is attenuated.

One of the most important correlates of [±voice] is closure voicing duration, the extent to which voicing continues into the closure (Lisker 1978, 1986, Raphael 1981, Stevens and Blumenstein 1981). Closure voicing is acoustically realized as a voice bar—i.e. 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 while the second pair shows geminate values. Within each pair, the first bar represents voiced consonants, and the second bar represents 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 Type1 error adjustment was applied. This is because error bars are provided to give an idea of the accuracy of the mean estimations, rather than for the sake of post-hoc multiple comparisons.

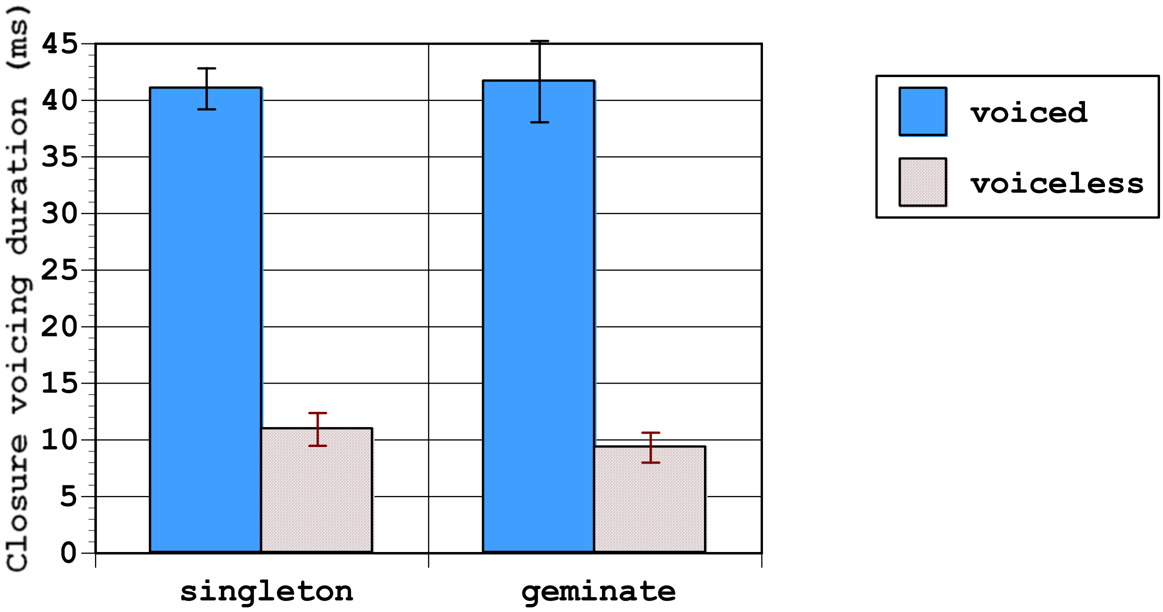


Figure 7: Mean closure voicing duration (in milliseconds). Error bars represent 95% confidence intervals, calculated as $t(\text{df})_{.05} \times$ s.e.

As seen in Figure 7, voiced consonants, regardless of whether they are singletons or geminates, have on average about 40ms of closure voicing. Voiceless consonants, on the other hand, have about only 10ms of closure voicing. The difference between voiced and voiceless consonants in terms of closure voicing is statistically significant ($F(1, 611)=720.41$, $p<.001$). On the other hand, there does not seem to be any effect of geminacy on closure voicing ($F(1, 611)<1$). The interaction between the two variables is not significant either ($F(1, 611)<1$). The overall results thus show that the closure voicing duration is longer in voiced consonants than in voiceless consonants, and the size of the differences is about the same between singleton pairs and geminate pairs.

However, as implied by the size of the error bars in Figure 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 ($t(290)=5.49$, $p<.001$).[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.

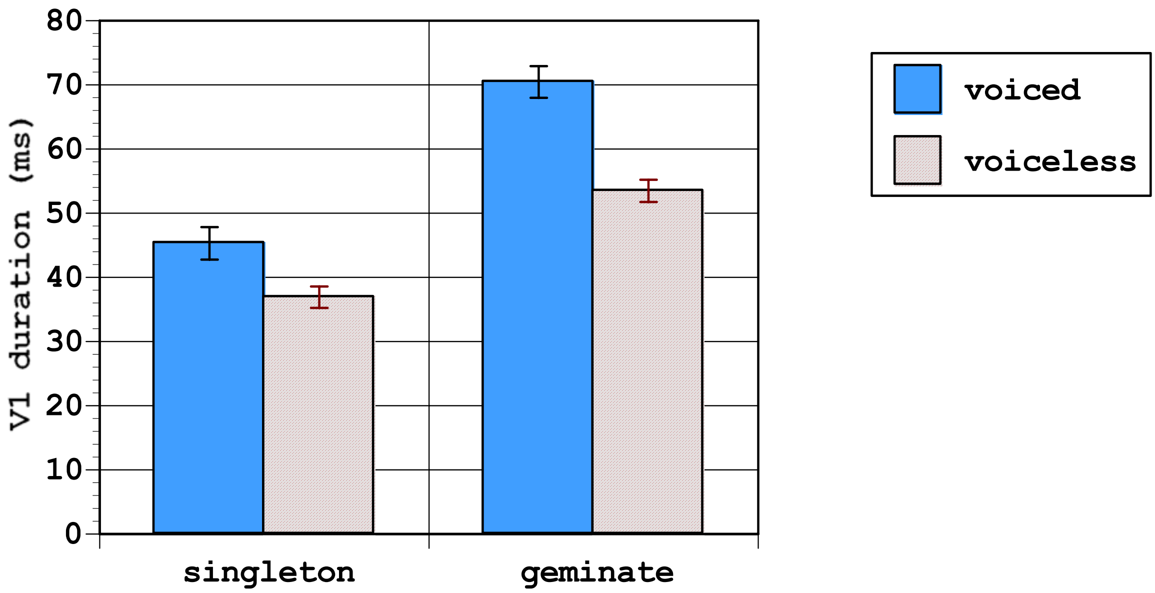


Figure 8: Mean V1 duration (ms).

As expected from reports for other languages (Chen 1970, Kingston and Diehl 1994, Raphael 1981), vowels are longer before voiced consonants than before voiceless consonants (the difference is on average 12.69ms; $F(1, 603)=166.34$, $p<.001$). Next, as previously reported by Han (1994), vowels are also longer before geminates than before singletons (the difference is 20.85ms; $F(1, 603)=453.18$, $p<.001$). The interaction of these two variables is significant ($F(1, 603)=19.49$, $p<.001$). This significant interaction effect arises because the V1 duration

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

difference is larger before geminates than before singletons (by about 8.5ms). This larger difference in V1 duration 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.

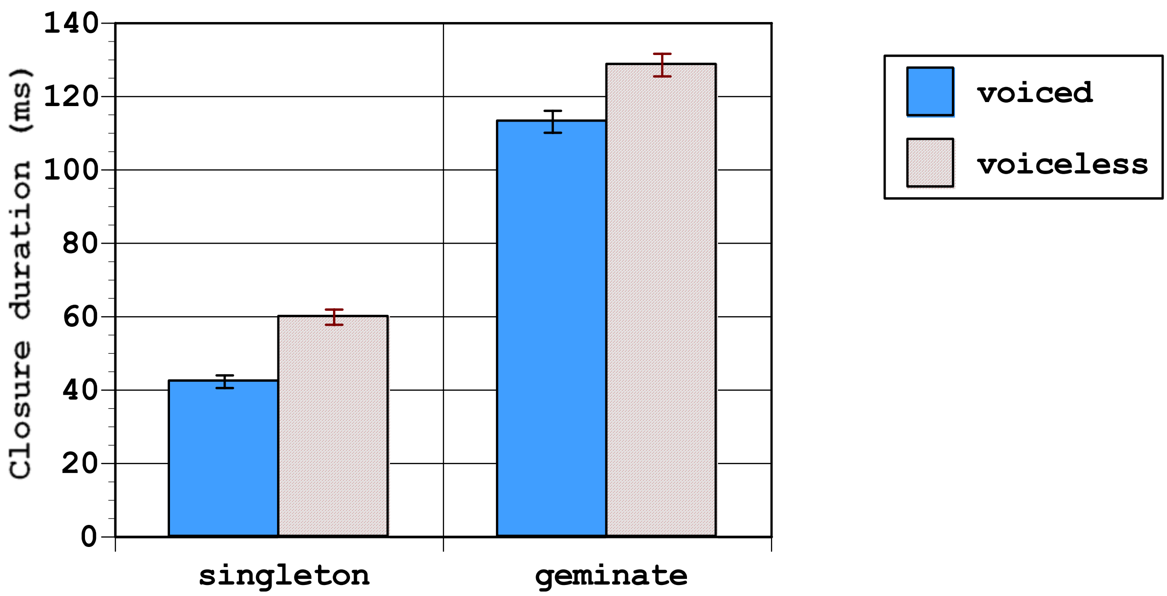


Figure 9: Mean closure duration (ms).

Again, as expected from reports for other languages (Kingston and Diehl 1994, Lisker 1957, Ohala 1983, Westbury 1979), consonants are longer when they are voiceless than when they are voiced (the difference is 16.54ms; $F(1, 603)=182.94$, $p<.001$). Geminates are on average longer than singletons by 69.77ms ($F(1, 603)=3220.48$, $p<.001$). 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 pairs and geminate pairs.

We have seen that voiced consonants exhibit shorter closure duration and longer V1 duration. These opposite effects of [+voice] on closure duration and V1 duration 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 and Diehl 1994, Port and Dalby 1982). The C/V duration ratios for the Japanese case at hand are summarized in Figure 10.

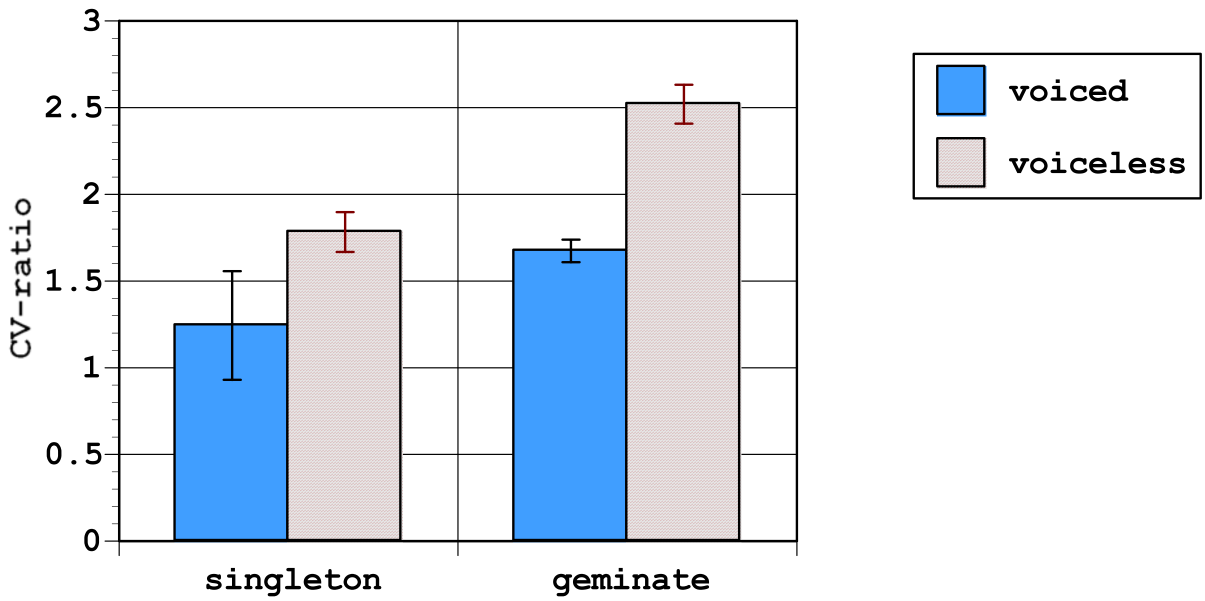


Figure 10: Mean CV-ratio, calculated as duration ratio of consonant/vowel.

An ANOVA shows that geminates have a higher CV-ratio than singletons (on average, singletons=1.52; geminates=2.10; $F(1, 587)=59.89$, $p<.001$). Voiceless consonants, as expected, also have a higher CV-ratio than voiced consonants (voiceless=2.15; voiced=1.46; $F(1, 587)=84.13$, $p<.001$). The interaction is marginally significant ($F(1, 587)=3.78$, $p=.052$), and this interaction 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 ($t(586)=2.33$, $p<.05$). 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 and Diehl 1994 and references cited therein). First, starting with V2, as expected, F0 is higher after voiceless consonants, as illustrated in Figure 11.

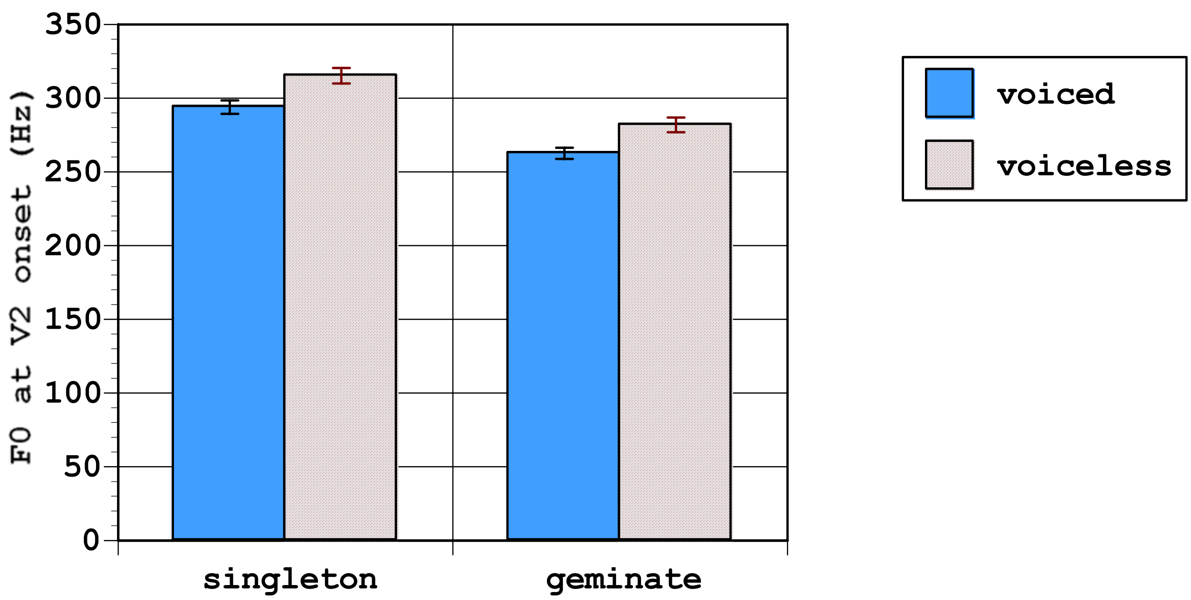


Figure 11: Mean F0 at V2 onset (Hz).

F0 is on average 20.32Hz higher after voiceless consonants than after voiced consonants (*F*(1, 604)=175.95 p<.001). On the other hand, F0 is on average 32.36Hz lower after geminates than after singletons ($F(1, 604)=365.28$, p<.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=.18), which indicates that the F0 frequency difference due to a [±voice] contrast is more or less constant between post-singleton and post-geminate positions.

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

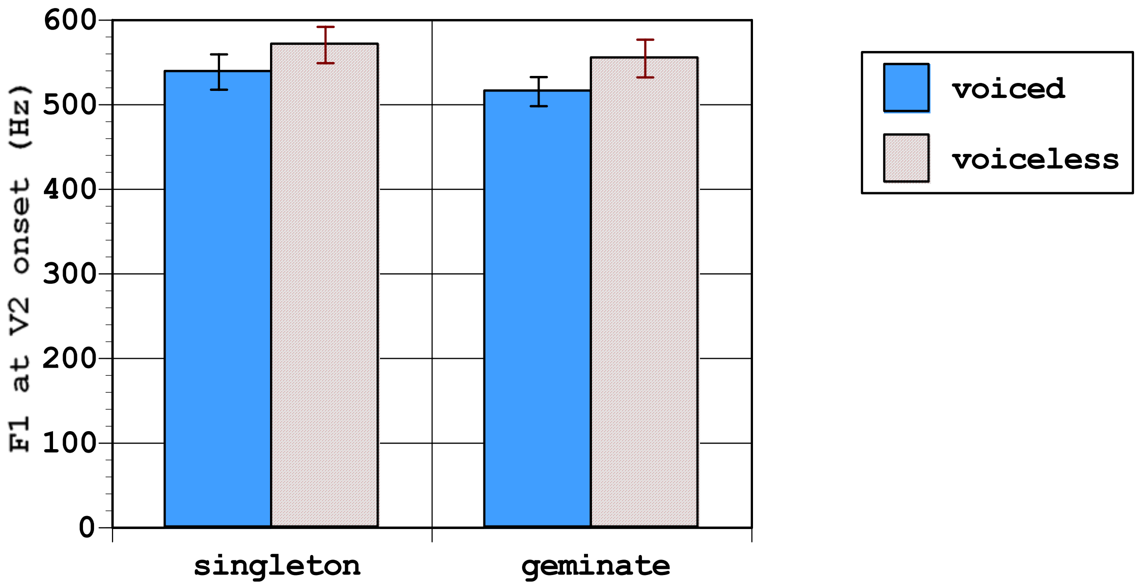


Figure 12: Mean F1 at V2 onset (Hz).

The difference due to a [±voice] difference is about 35.68Hz, which is statistically significant ($F(1, 600)=14.56$, p<.001). The effect of geminacy is only marginally significant ($F(1, 600)=3.18$, p=.075); on average, F1 is 19.56Hz lower after geminates. The interaction is not significant ($F(1, 600)<1$). Again, the lack of a significant interaction effect shows that F1 differences due to a [±voice] contrast are about the same size in post-singleton and post-geminate environments.

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

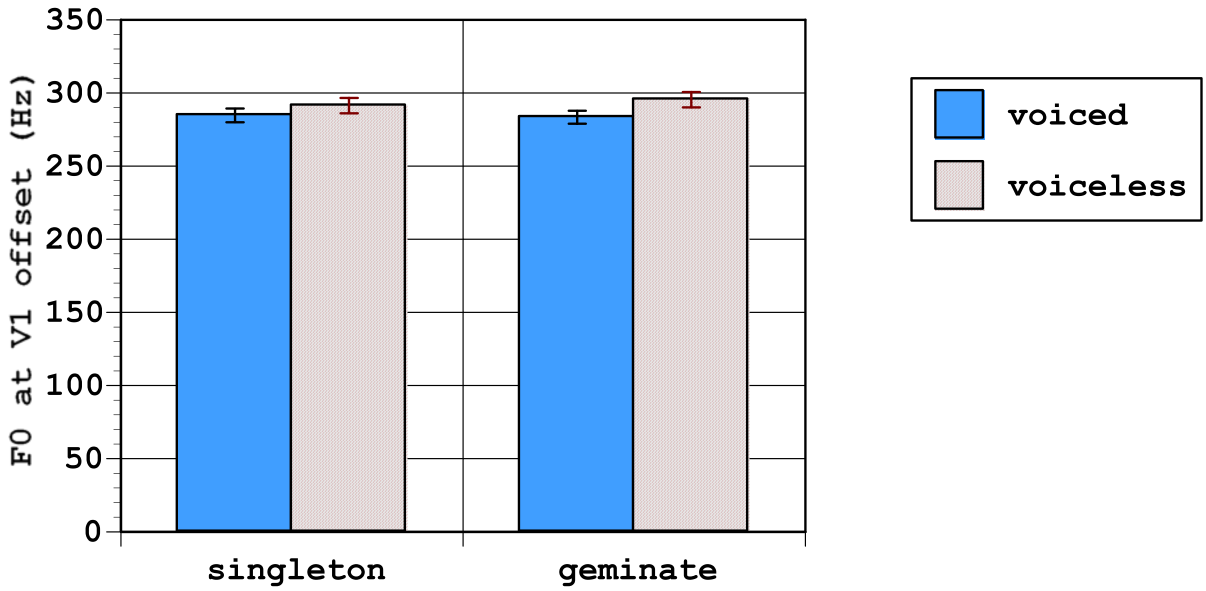


Figure 13: Mean F0 at V1 offset (Hz).

F0 is 9.30Hz higher before voiceless consonants than before voiced consonants (F(1, 601=71.29, p<.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<.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 pre-singleton and pre-geminate 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.

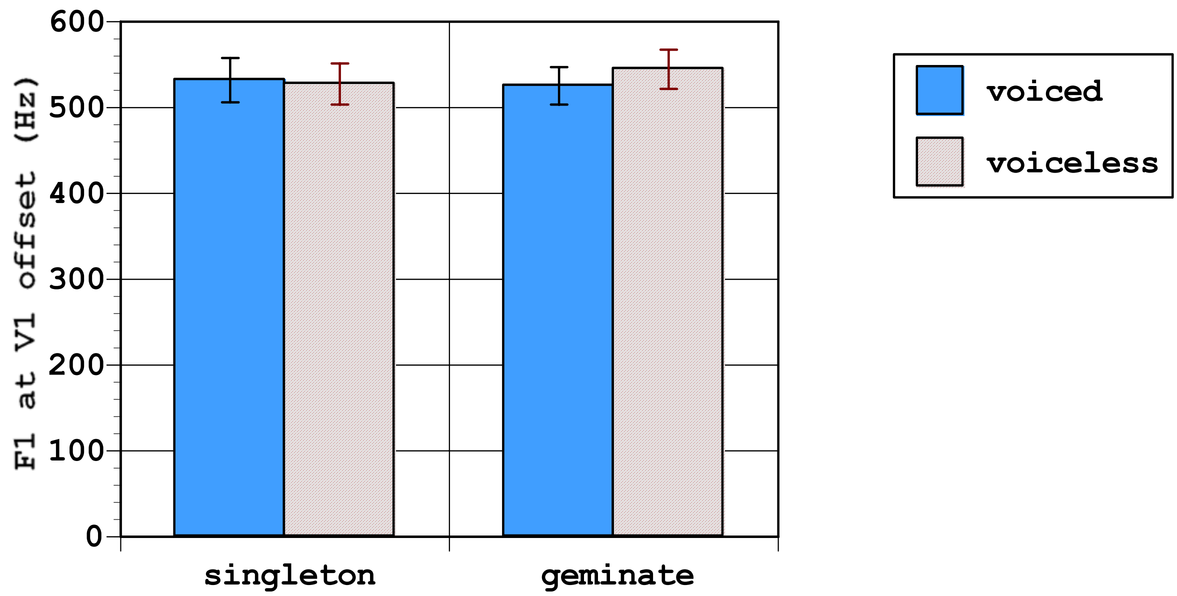


Figure 14: Mean F1 at V1 offset (Hz).

One might suspect from Figure 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 ($F(1, 322)=1.47$, $p=.23$).

We have seen that a voicing contrast in Japanese is associated with many of the cues that are known to signal a [±voice] distinction cross-linguistically. 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.

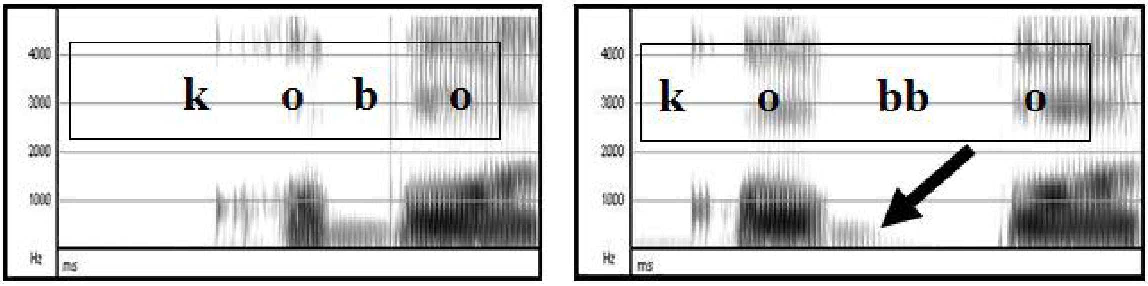


Figure 15: Spectrograms of a singleton [b] and geminate [bb]. The time scales are the same (350ms).

As seen in Figure 15, while voicing is fully maintained in the singleton [b] (left), partial devoicing is observed after the arrow in the geminate [bb] (right). This asymmetry between singletons and geminates is consistent across the three speakers. All of the speakers maintain full voicing for almost all singleton tokens. On the other hand, they rarely produced fully-voiced geminates: one speaker produced two tokens of fully-voiced [bb] out of 54 tokens of voiced geminates; another speaker produced one fully-voiced [gg]; and the third speaker produced no fully-voiced geminates.

To quantify the degree of partial devoicing, the proportion of closure voicing duration with respect to closure duration was calculated. The results appear in Figure 16.

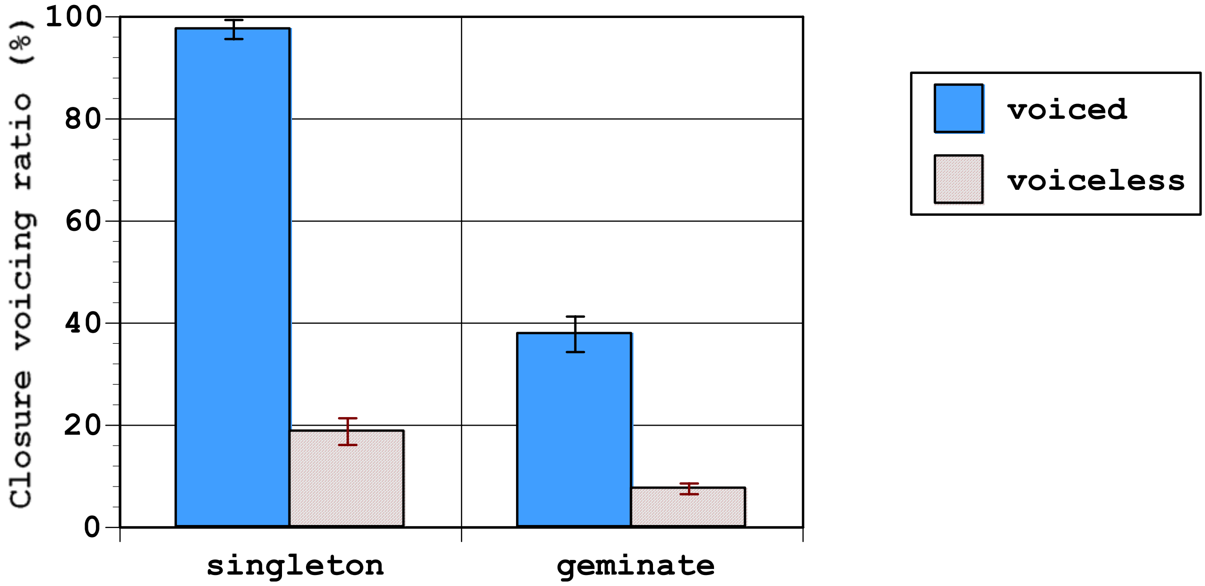


Figure 16: Closure voicing ratio, calculated as closure voicing duration divided by corresponding closure duration.

We observe in Figure 16 that 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 in maintaining voicing during obstruent 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 and Steriade 2004, Jaeger 1978, Ohala 1983, Westbury 1979).

Since the last 60% of a voiced geminate is phonetically voiceless, the percept of [+voice] in geminates may be undermined, since closure voicing is one of the most important cues signaling [+voice] (Lisker 1978, Parker et al. 1986, Raphael 1981). In particular, Lisker (1978) has shown that consonants with 120ms closure duration and 40ms 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 pairs and geminate pairs (Figure 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 pairs and geminate pairs in terms of closure duration, the proportion of voiced consonants' closure duration with respect 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<.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), who 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, whereas voiced geminates resist such spirantization. This contrast is illustrated in Figure 17.

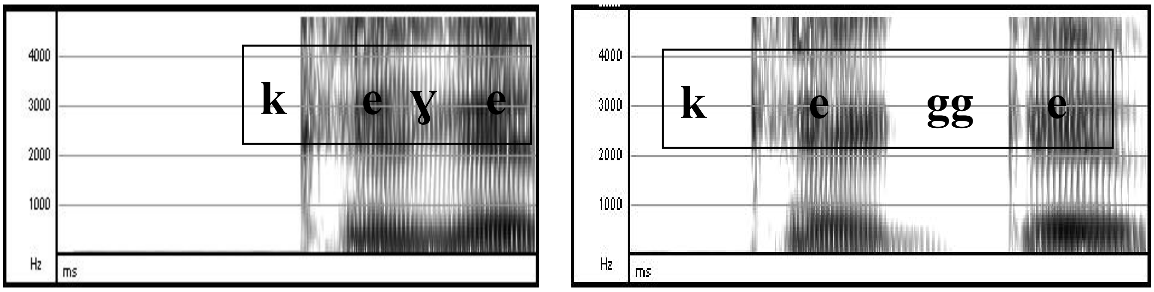


Figure 17: Spectrograms of a singleton /g/ realized as [ɣ] and a geminate /gg/ realized as [gg]. The time scales are the same (350ms).

As seen in Figure 17, a singleton /g/ is lenited almost to an approximant (left), as the visible formant energy during the constriction indicates, whereas lenition does not take place at all for geminates (right). Since voiceless consonants never spirantize, in singleton pairs, a [±voice] contrast 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, on the other hand, do not have this advantage because neither voiced nor voiceless consonants spirantize, and therefore [+voice] geminates are not distinguished from [-voice] geminates in terms of frication noise.

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 Figure 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 8Hz). It is therefore unlikely that attenuated cues in geminates are sufficiently compensated for, and the next perceptual experiment 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, this section has first shown that the Japanese [±voice] contrast is signaled by a number of the acoustic parameters that are known to cue a [±voice] distinction cross-linguistically. In addition, it has been shown that, 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 III: 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. So as 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. However, had natural tokens and nothing else been used, 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 $\text{Ident(+voi)}_{\text{Sing}}$ » $\text{Ident(+voi)}_{\text{Gem}}$ 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 36 types of stimuli (3 vowels × 3 places of articulation × 2 consonantal lengths × 2 [±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 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 to mask voicing, it was necessary to use speech-like noise with energy in low spectra ranges: 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 3-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, respectively. As a result, the average amplitude of the stimuli and that of the noise became 71.90dB and 72.35dB, 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 17 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. They are two possible orthographic representations of the stimuli, minimally different in [±voice] of the second consonant—e.g. 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 for the visual stimuli so that the subjects would be encouraged to perceive the stimuli as non-native words (recall that voiced geminates are allowed only in loanwords). In order to make sure that the subjects would respond to all of the stimuli, no time limits were enforced. The listeners were not given feedback about the correctness of their response.

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 36 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 24 times (3 speakers × 8 sessions). The subjects were encouraged to take short breaks once or twice during the experiment. Including the instructions at the beginning and the post-experiment 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 and Creelman 2005) which directly represents the perceptual distance between two stimuli—i.e. 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 and 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*(hit)–*z*(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 of d' for singletons across all of the 17 listeners is 3.79, which is significantly different from zero ($t(16)=34.15$, $p<.001$), and the average d' for geminates is 0.71, which again is significantly different from zero ($t(16)=11.47$, $p<.001$). 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

---

[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 and 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.

singletons than for geminates; a paired *t*-test comparing d' for singletons and geminates reveals a significant difference ($t(16)=27.27$, $p<.001$). 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 is, the more likely the geminate at that place is to devoice. The correlation 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 obtained data 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 and Creelman 2005). The bias function c is the sum of the *z*-scores of the hit and false alarm rates multiplied by -.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 .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 .08, which does not significantly deviate from zero ($t(16)=1.01$, $p=.33$). On the other hand, the mean c for geminates is 1.08, which is significantly different from zero ($t(16)=5.57$, $p<.001$). 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 in addition, 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 are voiced singletons in the overall 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 and 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: The frequency of voiceless singletons, voiced singletons, voiceless geminates, and voiced geminates in Amano and Kondo (2000). The numbers represent the token frequencies. The type frequencies are listed 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 are perceptual bias toward hearing acoustically ambiguous signals as more frequent possibilities than less frequent possibilities (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 result of the experiment alone, it is not clear which factor(s) is responsible for the bias against the [+voice] percept given geminates stimuli.

6. General discussion and conclusion.

6.1. Theoretical implications. In this paper, I have first shown 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(+voi)$_{Sing}$ and Ident(+voi)$_{Gem}$, and that these are ranked as Ident(+voi)$_{Sing}$ »

Ident(+voi)$_{Gem}$. 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 in a spectrogram shown in Figure 18.[13]

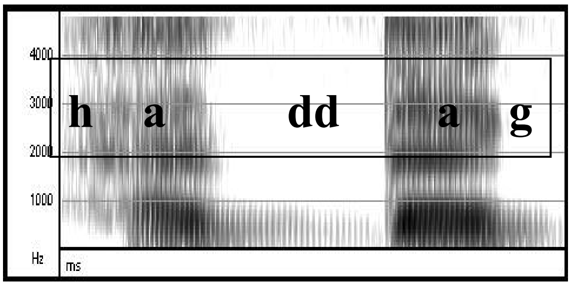


Figure 18: The spectrogram of a nonce Arabic word [haddag] pronounced by a female native speaker of Egyptian Arabic.

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 (i.e. partial devoicing) can affect a phonological pattern (i.e. 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 or never contrastive in phonology (Boersma 1998, Browman and Goldstein

[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.

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, which is related to the discussion above, is whether the ranking Ident(+voi)$_{Sing}$ » Ident(+voi)$_{Gem}$ is universal or not. The P-map hypothesis suggests that the ranking is not necessarily universal. As discussed above, the lesser perceptibility of [+voice] in Japanese voiced geminates might be due to the language-specific way in which 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 Figure 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 1028 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%). On the other hand, there are 811 roots whose second consonant is a voiced obstruent, and 325 of these occur in the verb pattern where the second consonant is geminated (=40.0%). The difference between these two ratios is not statistically significant (by approximating to a normal distribution, z=.89, p=.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)$_{Sing}$ and Ident(+voi)$_{Gem}$ are ranked in the same position: we would not observe a phonological split between [+voice] in singletons and [+voice] in geminates, like the one we observe in Japanese loanwords. This prediction remains to be tested.

A further prediction of the theory advanced in this paper 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)$_{Gem}$ » Ident(+voi)$_{Sing}$. Given this ranking, *VoiObs, a general constraint against voiced obstruents, could be sandwiched between these two faithfulness constraints. The result is a language that permits only voiced geminates, and not voiced singletons, but no such language exists (Hayes and Steriade 2004). Therefore, Ident(+voi)$_{Sing}$ must be universally ranked either as high as or higher than Ident(+voi)$_{Gem}$. 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)$_{Gem}$ is ranked as high as Faith(+voi)$_{Sing}$, and (ii) [+voice] is never more perceptible in geminates than in singletons. Whether these predictions of the P-map hypothesis are borne out should be tested cross-linguistically 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, several more 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 and 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 in this paper 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)$_{Gem}$. This paper thus overall provides empirical support for the P-map hypothesis, according to which a faithfulness ranking can be projected from a perceptibility scale.

Appendix: Acoustic values of the tokens used in Experiment II.

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							
	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	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
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. 2006. Faithfulness and perception in loanword adaptation: A case study from Hawaiian. Lingua, to appear.

Alderete, John. 1998. 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, MA: University of Massachusetts, Amherst dissertation.

Benkí, José. 2003. Analysis of English nonsense syllable recognition in noise. Phonetica 60.129-157.

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 6.201-252.

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 22.129-159.

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, NY: 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. Proceedings of ICPhS99.

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. Proceedings of the 1995 Southwestern Workshop on Optimality Theory.

Fougeron, Cecile, and Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. Journal of the Acoustical Society of America 106.3728-3740.

Fujimura, Osamu; Marian Macchi; and L. A. Streeter. 1978. Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. Language and Speech 21.337-346.

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, MA: 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. Voicing in Japanese, ed. by Jeroen van der Weijer, Kensuke Nanjo, and Tetsuo Nishihara, to appear. Berlin: Mouton de Gruyter.

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

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

Hayes, Bruce, and Donca Steriade. 2004. Introduction: The phonetic bases of phonological markedness. In Hayes, Kirchner, and Steriade 2004, 1-33.

Hooper, Joan Bybee. 1976. An introduction to Natural Generative Phonology. New York, NY: 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, 817-838. Oxford: Blackwell.

Itô, Junko, and Armin Mester. 1997. German truncation and sympathy theory. University of Maryland Working Papers in Linguistics 5.117-139.

Itô, Junko, and Armin Mester. 1998. Markedness and word structure: OCP effects in Japanese. Santa Cruz, CA: 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, 261-295. Cambridge, MA: Cambridge University Press.

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

Jaeger, Jeri. 1978. Speech aerodynamics and phonological universals. Chicago Linguistics Society 14.311-329.

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

Jun, Jongho. 2004. Place assimilation. In Hayes, Kirchner, and Steriade 2004, 58-86.

Katayama, Motoko. 1998. Loanword phonology in Japanese and Optimality Theory. Santa Cruz, CA: 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. Amherst, MA: University of Massachusetts, Amherst, ms.

Keating, Patricia. 1996. The phonology-phonetics interface. UCLA Working Papers in Phonetics 92.45-60.

Kingston, John, and Randy Diehl. 1994. Phonetic knowledge. Language 70.419-455.

Kirchner, Robert. 1997. Constrastiveness and faithfulness. Phonology 14.83-111.

Kohler, Klaus. 1979. Dimensions in the perception of fortis and lenis plosives. Phonetica 36.332-343.

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 Publishers.

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-55/56.251-259.

Lisker, Leigh. 1981. On generalizing the rapid-rabid distinction based on silent gap duration. Haskins Laboratories Status Report on Speech Research SR-65.251-259.

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, IL: University of Chicago dissertation.

Macmillan, Neil, and Douglas Creelman. 2005. Detection theory: A user's guide. 2nd Edition. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

McCarthy, John. 1986. OCP Effects: Gemination and antigemination. Linguistic Inquiry 17.207-263.

McCarthy, John. 1999. Sympathy and phonological opacity. Phonology 16.331-399.

McCarthy, John. 2003. Comparative markedness. Theoretical Linguistics 29.1-52.

McCarthy, John, and Alan Prince. 1986. Prosodic morphology. Amherst, MA, and New Brunswick, NJ: University of Massachusetts and 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 27.338-352.

Miura, Akira. 1993. English in Japanese. New York, NY: 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. Proceedings of the 6th European Conference on Speech Communication and Technology.

Myers, Scott. 1997. Expressing phonetic naturalness in phonology. Derivations and constraints in phonology, ed. by Iggy Roca, 125-152. Oxford: Oxford University Press.

Myers, Jerome, and Arnold Well. 2003. Research design and statistical analysis. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

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 62.353-383.

Ohala, John. 1983. The origin of sound patterns in vocal tract constraints. The production of speech, ed. by Peter MacNeilage, 189-216. New York, NY: Springer Verlag.

Padgett, Jaye. 2002. Constraint conjunction versus grounded constraint hierarchy in Optimality Theory. Santa Cruz, CA: 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 39.129-142.

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, 310-343. Cambridge, MA: Cambridge University Press.

Pierrehumbert, Janet. 1990. Phonological and phonetic representation. Journal of Phonetics 18.375-394.

Port, Robert, and John Dalby. 1982. Consonant/vowel ratio as a cue for voicing in English. Perception and Psycholinguistics 32.141-152.

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 38.126-147.

Sanders, Nathan. 2003. Opacity and sound change in the Polish lexicon. Santa Cruz, CA: 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 14.311-326.

Smith, Jennifer. 2006. Loan phonology is not all perception: Evidence from Japanese loan doublets. Japanese/Korean Linguistics 14, to appear.

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, CA: University of California, Santa Cruz dissertation.

Stampe, David. 1973. A dissertation on Natural Phonology. Chicago, IL: University of Chicago dissertation.

Steriade, Donca. 1995. Positional neutralization. Los Angeles, CA: University of California, Los Angeles, ms.

Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. Los

Angeles, CA: University of California, Los Angeles, ms.

Steriade, Donca. 2000. Paradigm uniformity and the phonetics/phonology boundary. Papers in laboratory phonology V, ed. by Michael Broe, and Janet Pierrehumbert, 313-334. Cambridge, MA: 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, 219-250. San Diego, CA: Academic Press.

Steriade, Donca. 2001b. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Los Angeles, CA: 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 22.343-356.

Tsuchida, Ayako. 1995. English loans in Japanese: Constraints in loanword phonology. Working Papers of the Cornell Phonetics Laboratory 10.145-164.

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 19.827-878.

Wehr, Hans. 1971. A dictionary of modern written Arabic: The 3rd edition. New York, NY: Spoken Language Services.

Westbury, John. 1979. Aspects of the temporal control of voicing in consonant clusters in English. Austin, TX: University of Texas, Austin dissertation.

Zhang, Jie. 2000. Non-contrastive features and categorical patterning in Chinese diminutive suffixation: Max[F] or Ident[F]? Phonology 17.427-478.

Zhang, Jie. 2004. The role of contrast-specific and language specific phonetics in contour tone distribution. In Hayes, Kirchner, and Steriade 2004, 157-190.

Zsiga, Elizabeth. 1995. An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. Papers in laboratory phonology IV, ed. by Bruce Connell and Amalia Arvaniti, 282 - 302. Cambridge, MA: Cambridge University Press.

Zsiga, Elizabeth. 1997. Features, gestures, and Igbo vowels: An approach to the phonology-phonetic interface. Language 73.227-274.

Zuraw, Kie. 2005. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog reduplication. Los Angeles, CA: University of California, Los Angeles, ms.