Spectral continuity, amplitude changes, and perception of length contrasts

Abstract

Japanese deploys a singleton-geminate contrast in obstruents and nasals, but not in glides. Even though Japanese allows lexical nasal geminates, patterns of emphatic gemination show that Japanese avoids creating nasal geminates. Japanese therefore disfavors sonorant geminates in general, and glide geminates in particular. These phonological patterns of geminates are actually found in other languages as well, such as Ilokano (Hayes, 1989). This paper tests hypotheses about why speakers of these languages show these preferences. Concerning the distinction between obstruent geminates and sonorant geminates, Podesva (2002) hypothesizes that the phonological dispreference against sonorant geminates exists because these geminates are easily confused with corresponding singletons. This confusability problem arises because sonorants are spectrally continuous with flanking vowels, and consequently their constriction durations are difficult to perceive. Two non-speech perception experiments, Experiments I and II, confirm this hypothesis by showing that length distinctions of consonant intervals that are spectrally continuous with surrounding segments are difficult to perceive. Concerning the difference between nasal geminates and glide geminates, this paper builds on the finding by Kato et al. (1997) that given streams of sounds, listeners use amplitude changes to demarcate segmental boundaries. Experiments III and IV show that amplitude changes facilitate categorization and discrimination of short/long contrasts of consonantal intervals. These results are compatible with the fact that several languages disfavor glide geminates more than nasal geminates. Overall, the results of the four perception experiments reported here accord well with the cross-linguistic phonological patterning of geminates. We close this paper by discussing what the current results imply about how the phonetics-phonology works.

1 Introduction

This paper begins with a phonological observation concerning the cross-linguistic patterning of geminates. We start by discussing Japanese in some detail in section 1.1, followed by discussion of other languages in section 1.2.

1.1 Japanese

Japanese uses a lexical singleton-geminate (short vs. long) contrast, and this phonological contrast is primarily cued by a difference in consonantal duration (Kawahara, 2015b). This contrast is limited to (voiceless) obstruents and nasals, as shown by the examples in (1) (Kawagoe, 2015). On the other hand, Japanese does not allow lexical singleton-geminate contrasts in glides.[1]

(1) Lexical singleton/geminate contrasts in Japanese

ka**t**a	'frame'	ka**tt**a	'bought'
i**s**o	'shore'	i**ss**o	'rather'
ko**n**a	'powder'	ko**nn**a	'such'

Even though Japanese allows nasal geminates, one phonological process shows that Japanese avoids creating nasal geminates as well. Nasu (1999, 2005) points out that given reduplicated sound-symbolic, mimetic C_1VC_2V-C_3VC_4V forms, in order to create their emphatic forms, Japanese speakers predominantly geminate C_2 when it is a stop, as in (2):

(2) Emphatic forms via gemination of C_2 when C_2 is a stop

a. /pa**t**a-pata-μ/ $\rightarrow$ [pa**tt**apata] 'running'

b. /pi**k**a-pika-μ/ $\rightarrow$ [pi**kk**apika] 'shining'

However, when C_2 is a nasal and C_3 is a stop, speakers prefer to target C_3 for emphatic gemination, as in (3). In the experiment reported in Kawahara (2013), C_3 was fixed as stops, and C_2 was varied among stops, fricatives and nasals. When asked to create emphatic forms of nonce mimetic words, when C_2 and C_3 are both stops, Japanese speakers chose C_2-gemination about 80% of the time, supporting the preference in (2). However, when C_2 is a nasal and C_3 is a stop, they chose C_2-gemination only about 35% of the time and instead resort to C_3-gemination, as exemplified in (3). The flopping of gemination locus in (3) shows that Japanese avoids nasal geminates when possible.

(3) When C_2 is a nasal and C_3 is a stop, speakers prefer C_3-gemination

a. /ka**n**o-kano-μ/ $\rightarrow$ [kano**kk**ano] (nonce word)

[1] Japanese lacks geminate [rr] as well (Kawahara, 2015a; Labrune, 2014), and we will return to the discussion of liquid geminates in section 2.2.

b. /ki**n**a-kina-μ/ → [kina**kk**ina] (nonce word)

Furthermore, in loanword gemination pattern in which word-final consonants in the source languages are borrowed as geminates, oral stops undergo gemination, but word-final nasal stops do not (Katayama, 1998). This asymmetry is shown in (4) and (5):

(4) Gemination of oral stops in loanword adaptation

a. "stop" → [suto**pp**u]

b. "top" → [to**pp**u]

c. "rap" → [ra**pp**u]

(5) Nasals do not geminate in the same environment[2]

a. "Tom" → [to**m**u]

b. "ham" → [ha**m**u]

c. "lamb" → [ra**m**u]

We acknowledge that loanword adaptation patterns should be used in phonological argumentation with caution (de Lacy, 2006, 2009), because loanword adaptation is non-trivially affected by non-phonological—e.g. perceptual, orthographic, and sociolinguistic—factors (e.g. Irwin 2011; Kang 2011; Peperkamp 2005; Peperkamp & Dupoux 2003; Peperkamp et al. 2008; Silverman 1992; Takagi & Mann 1994; Vendelin & Peperkamp 2006). At any rate, the asymmetry between (4) and (5) is at least compatible with the view that Japanese avoids nasal geminates, although this argument is admittedly not a very strong one.

In summary, Japanese avoids glide geminates the most in that it does not allow them at all to make lexical contrasts. Japanese allows nasal geminates to signal lexical contrasts, but nevertheless avoids creating them in gemination process(es). The preferential hierarchy in the phonology of Japanese is therefore: obstruent geminates > nasal geminates > glide geminates.

1.2 Other languages

While this preferential hierarchy of geminates is clearly observed in Japanese, we observe the same hierarchy in other languages as well. Some languages avoid sonorant geminates entirely, whereas others avoid glide geminates in particular, just like Japanese.

One example that instantiates the avoidance of sonorant geminates comes from gemination blocking in Selayarese (Podesva, 2000, 2002). When the prefix /taʔ-/ is attached to a root that begins with a voiceless obstruent, the prefix-final glottal stop assimilates to the following consonant, resulting in a geminate, as shown in (6) (Mithun & Basri 1986: 243):

[2] Word-final [n] in English is borrowed as a moraic coda nasal without epenthesis (and without gemination); e.g. "run" is borrowed as [raɴ]. Whether gemination fails because of the lack of epenthesis or the markedness of [nn] is not clear (though see Peperkamp et al. 2008).

(6) Gemination when root-initial consonants are voiceless obstruents

a. /ta**ʔ-p**elaʔ/ → [ta**pp**elaʔ] 'get lost'

b. /ta**ʔ-t**uda/ → [ta**tt**uda] 'bump against'

c. /ta**ʔ-k**apula/ → [ta**kk**alupa] 'faint'

d. /ta**ʔ-s**ambaŋ/ → [ta**ss**ambaŋ] 'stumble, trip'

The gemination fails when root-initial consonants are sonorants, as in (7) (Mithun & Basri 1986: 244). Since there are no glides in Selayarese in the first place, we cannot tell whether glides undergo gemination or not.[3]

(7) Gemination is blocked when root-initial consonants are sonorants

a. /ta**ʔ-m**uri/ → [ta**ʔm**uri] 'smile'

b. /ta**ʔ-n**oʔnoso/ → [ta**ʔn**oʔnoso] 'to be shaken'

c. /ta**ʔ-ŋ**oaʔ/ → [ta**ʔŋ**oaʔ] 'to yawn'

d. /ta**ʔ-l**esaŋ/ → [ta**ʔl**esaŋ] 'to be removed'

e. /ta**ʔ-r**iŋriŋ/ → [ta**ʔr**iŋriŋ] 'to be walled'

Another example comes from Ilokano (Hayes 1989: 270-271), which is just like Japanese. Ilokano resolves hiatus by gliding a first vowel, and this formation of a glide causes compensatory gemination of the preceding consonant. This gemination process usually applies to obstruents, as in (8). In the same environment, gemination is marginally possible for nasals as in (9)—according to Hayes (1989: 270), gemination of these consonants is optional, possibly with lexical variation. Gemination never applies to [w, y], as in (10).

(8) Obstruents usually geminate after gliding of vowels

a. /lú**t**o-én/ → [lu**tt**wén] 'cook GOAL-FOCUS'

b. /pag-ʔá**s**o-án/ → [pagʔa**ss**wán] 'place where dogs are raised'

c. /kina-ʔa**p**ó-án/ → [kinaʔa**pp**wán] 'leadership qualities'

d. /bá**g**i-én/ → [ba**gg**yén] 'to have as one's own'

e. /pag-ʔatá**k**e-án/ → [pagʔatá**kk**yán] 'place where an attack takes place'

(9) Nasals only sporadically geminate

a. /dá**m**o-én/ → [da**m**wén], ?[da**mm**wén] 'to be new to something'

[3]Gemination also fails when root-initial consonants are voiced stops. The dispreference against voiced stop geminates is well motivated phonetically: stop closure raises intraoral airpressure and therefore it is difficult to maintain transglottal airpressure drop to sustain voicing during stop closure. This aerodynamic problem is particularly challenging for geminates because of their long constriction (Hayes & Steriade, 2004; Ohala, 1983; Westbury, 1979). However, this aerodynamic challenge does not explain the dispreference against sonorant geminates, because the airway is not significantly occluded in sonorants—the interoral airpressure should not rise so much as to hinder the airflow across the glottis. See also section 2.2 for related discussion.

b. /na-ʔali**n**o-án/ → [naʔali**n**wán], ?[naʔali**nn**wán] 'to become sensitive'

c. /pag-ʔali**ŋ**ó-án/ → [pagʔali**ŋ**ẃán], ?[pagʔali**ŋŋ**ẃán] 'place where boars are found'

(10) Glides never geminate

a. /ʔá**y**o-én/ → [ʔa**y**wén] 'cheer-up GOAL-FOCUS'

b. /babá**w**i-én/ → [baba**w**yén] 'regret GOAL-FOCUS'

Finally, Icelandic (Games, 1976) and Classical Nahuatl (Andrews, 1975) (both cited in Hansen & Myers 2014) are examples of languages which, like Japanese, lack length contrasts in glides. See Hansen & Myers (2014) and Kawahara et al. (2011) for other potentially relevant examples from other languages.

To summarize the observation in this section, we seem to find the following preferential hierarchy in the phonology of several languages: obstruent geminates > nasal geminates > glide geminates. The question is why this hierarchy holds across different languages. The experiments reported below attempt to address this question.

2 The phonetic grounding of the dispreference against sonorant geminates

2.1 The hypothesis

This paper first addresses the distinction between obstruent geminates and sonorant geminates. The hypothesis being tested in the following sections is not ours: Podesva (2002) proposes that sonorant geminates are dispreferred because they are perceptually confusable with corresponding singletons (see also Podesva 2000). The logic goes as follows: sonorants have blurry transitions into and out of flanking vowels, because sonorants are spectrally continuous with surrounding vowels. It is thus hard to pin down where sonorants begin and where they end (Myers & Hansen, 2005; Turk et al., 2006). As a result, their constriction durations are hard to perceive. Since the difference in constriction duration serves as the primary cue for singleton-geminate contrasts (e.g. Kawahara 2015b for a recent review), singleton-geminate distinctions are hard to distinguish for sonorants.

For the sake of illustration, Figures 1 and 2 show waveforms of singleton-geminate contrasts in stops and glides in Arabic (Kawahara, 2007). While stops have clear boundaries with the surrounding vowels, glides have very blurry boundaries. It is therefore difficult to know where the glides begin and where they end. For this reason, it is expected that the constriction durations are harder to accurately perceive for sonorants than for obstruents.

Besides the acoustic blurriness of segmental boundaries of sonorants, another factor that may work against the accurate perception of duration of sonorants is the fact that changes in amplitudes—or changes in perceived loudness—facilitate the detection of segmental boundaries (Kato & Tsuzaki, 1994; Kato et al., 1997). Because sonorantal boundaries with spectral continuity involve less amplitude/loudness changes than obstruent boundaries, sonorants have yet another disadvantage in signaling their boundaries. This issue is more fully addressed in Experiments III and IV.

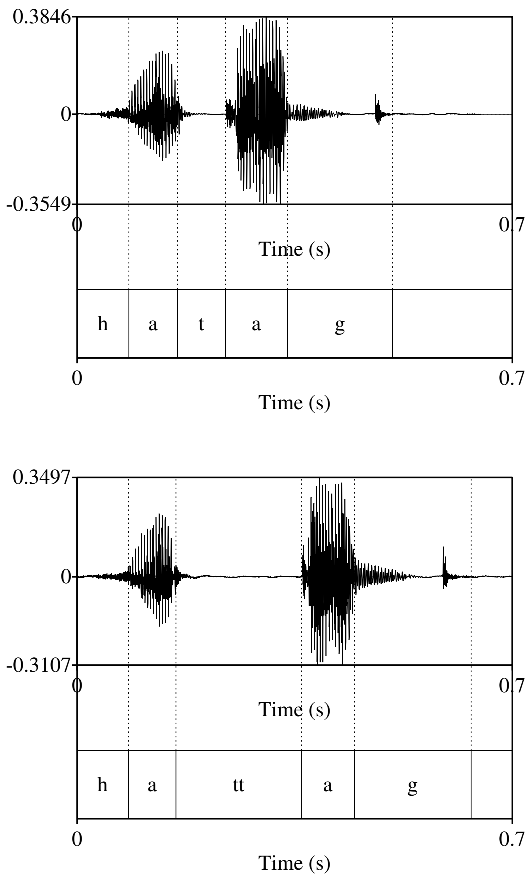

Figure 1: Arabic [t]-[tt] pair.

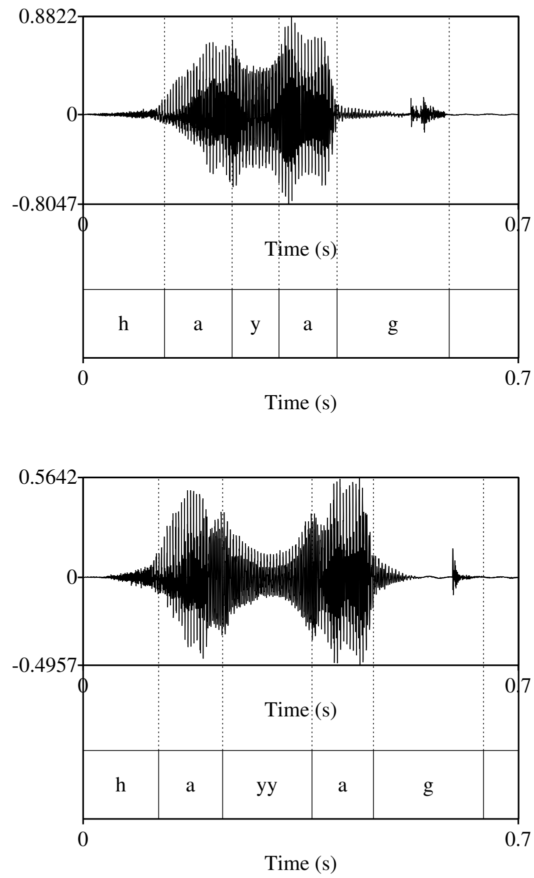

Figure 2: Arabic [y]-[yy] pair.

As summarized here, Podesva (2002) offers an interesting and plausible story about the perceptual grounding of the dispreference against sonorant geminates. However, no perception experiments have been reported to directly test this hypothesis. Partly to address this problem, Kawahara (2007) created continua from geminates to singletons for each type of geminates in Arabic, and presented them to Arabic speakers for an identification task. The results show that the identification functions were steeper for obstruents than for sonorants—more of the continuum was consistently categorized for obstruents than for sonorants. However, the relationship between the steepness of identification functions and the distinctiveness of singleton-geminate contrasts does not seem straightforward to interpret. Moreover, the experiment used speech sounds of Arabic as stimuli and Arabic listeners as participants. Therefore, the effect of factors other than sonority—such as lexical frequencies of each type of geminates or transitional probabilities from preceding consonant to each of singletons and geminates—remained uncontrolled, and possibly worked as confounds.[4]

Experiments I and II thus more directly test the relative non-distinctiveness of singleton-geminate contrasts in sonorants. To control for phonetic factors other than spectral continuity, the experiments used non-speech sounds that mimicked singleton-geminate contrasts in obstruents and sonorants.

2.2 Some caveats

A few remarks are in order before proceeding to the description of the experiments, first on the theoretical context of the current experiments. Podesva's (2000, 2002) general idea is couched within the framework of Adaptive Dispersion Theory (Liljencrants & Lindblom, 1972; Lindblom, 1986) (see also Engstrand & Krull 1994; Schwartz et al. 1997a,b; Zygis & Padgett 2010). This theory claims that languages generally prefer to use contrastive pairs that are perceptually dissimilar to each other; using perceptually distinct set of sounds is important in order for speakers not to be misunderstood by listeners (Lindblom et al., 1995). The Adaptive Dispersion Theory is further developed in recent years, as it was incorporated into generative phonology (Flemming, 1995, 2004; Ito & Mester, 2006; Padgett, 2002, 2009; Ní Chiosáin & Padgett, 2009) via Optimality Theory (OT) (Prince & Smolensky, 2004). OT allows a formal grammatical theory to incorporate the insights of the Adaptive Dispersion Theory, since OT can directly encode phonetic naturalness into the formulation of constraints (Hayes & Steriade, 2004; Ito & Mester, 2003; Kawahara, 2006a; Kager, 1999; Myers, 1997).

Within the OT version of the Dispersion Theory, singleton-geminate pairs of sonorants are marked—or disfavored by languages—because they are not perceptually distinct. This disperference can be directly expressed as a constraint that prevents a length contrast in sonorants; sonorant geminates, not singletons, are subsequently prohibited, because geminates are in general more marked than singletons. In this theory, it is the singleton-geminate contrasts in sonorants, not the sonorant geminates *per se*, that are marked; see the references cited above for formal implementations of this idea (see also Boersma 1998). An alternative is to encode these sorts of perceptual effects on phonology through diachrony (Barnes, 2002; Blevins, 2004a,b; Yu, 2004).

This paper is not intended to solve this debate about whether perceptibility should be encoded synchronically or diachronically. Instead, the aim of Experiments I and II is to test the assumption behind Podesva's (2002) hypothesis—the non-distinctiveness of sonorant singleton-geminate pairs—but we do not commit ourselves to any particular theoretical implementation of this idea.[5]

[4] Various lexical factors can impact the categorization of speech sounds, which include the distinction between word vs. non-word (Ganong, 1980), lexical frequency differences (Connine et al., 1993), neighborhood densities (Vitevitch & Luce, 1999), transitional probabilities (McQueen & Pitt, 1996), and phonotactic restrictions (Massaro & Cohen, 1983; Moreton, 2002).

[5] Though see de Lacy & Kingston (2013), Hayes & Steriade (2004), Hura et al. (1992), Kawahara (2006b), Martin & Peperkamp (2011), Moreton (2008), Steriade (2008), Wilson (2006), and Zsiga (2011) for arguments for encoding

Second, the confusability problem between singletons and geminates may not be the only source of the avoidance of sonorant geminates. For example, given intervocalic geminate glides (e.g. [iyyi]), it is conceivable that the first part of the geminate can be confused as a part of a preceding (long) vowel (cf. Myers & Hansen 2005). Also concerning rhotic geminates, it would be impossible to prolong the duration of a tap or a flap, and they would instead have to turn into a trill, in order to become a geminate while keeping its rhoticity. However, a trill requires a very precise articulatory coordination (Ladefoged & Maddieson, 1996; Solé, 2002). In short, we do not intend to claim that low distinctiveness of singleton-geminate pairs is the only phonetic problem for sonorant geminates.

Nor is it the case that sonorant geminates are the only kinds of geminates that are avoided for a phonetic reason. For example, voiced obstruent geminates are known to be avoided in many languages because it is difficult to maintain voicing during obstruents for a long stretch of time for aerodynamic reasons (Hayes & Steriade 2004; Ohala 1983; Westbury & Keating 1986).[6] Likewise, Kirchner (1998) argues that fricative geminates are articulatorily challenging, and indeed there are languages that avoid fricative geminates; in Wolof, for example, fricatives occlusivize to stops when they undergo a gemination process (Ka, 1994).

3 Experiment I: Discrimination experiment, obstruents vs. sonorants

3.1 Introduction

The first experiment tested whether sonorantal spectral continuity makes a short-long pair difficult to distinguish. The stimuli were non-speech analogues mimicking singleton-geminate pairs of stops, fricatives, and sonorants. The experiment used non-speech stimuli so as to control for acoustic parameters other than spectral continuity, such as preceding vowel duration, intensity of surrounding vowels, and duration of consonant intervals themselves. In experiments using real speech, on the other hand, it is difficult to control for the duration of consonant intervals, because the duration of glides is difficult to measure for the spectral continuity problem discussed in section 2 (see also Turk et al. 2006). Using non-speech sounds also avoided perceptual bias effects, such as lexical bias (Ganong, 1980) or transitional probability bias (McQueen & Pitt, 1996) (see footnote 4 for the full list).

---

phonetic factors in synchronic phonological systems. We will briefly come back to this issue in 7.2.

[6] It may as well be the case that spectral continuity at low frequency range in voiced stops makes the perception of duration harder for voiced stops than for voiceless stops, because spectral continuity at low frequency range can "shrink" the percept of that interval (Parker et al., 1986). However, this paper sets this hypothesis aside, because the aerodynamic challenge of voiced stops geminates is well-established. There can be more than one phonetic problem for some phonological structure.

3.2 Method

3.2.1 Stimuli

The three types of consonantal stimuli were non-speech analogues of stops, fricatives, and sonorants. All the stimuli had VCV structure in which the duration of C was varied. Non-speech analogues of vowels were complexes of sine waves (Kingston et al., 2009). They consisted of 50 sine waves ranging from 100 Hz to 16 kHz and separated by equal natural log intervals.[7] The amplitude of each sine wave negatively correlated with its frequency. The peak amplitude of the vowel analogues was set to 0.8 Pascal by Praat (Boersma, 2001; Boersma & Weenink, 1999–2015). Both vocalic intervals were 100 ms.

Consonant intervals mimicked the acoustic properties of stops, fricatives and sonorants (particularly glides); i.e., silence for stops, white noise filtered between 2 kHz and 22 kHz for fricatives, and the same interval as the vocalic interval with half of its peak energy for sonorants. Figures 3 illustrate the stimuli of the current experiment. Figure 4 is shown next to Figures 3 to illustrate the parallel between the non-speech stimuli and the corresponding speech forms.[8] The duration of the short consonants was set to 100 ms and that of long consonants was set to 150 ms. These two values were chosen because the short-long contrasts based on these values were neither too easy nor too difficult to discriminate in pilot studies.[9] For the discrimination experiment, two VCV sequences were concatenated with 400 ms inter-stimulus intervals (ISI).

3.2.2 Procedure

The task was a same-different discrimination (AX-discrimination) experiment. Four pairs of combinations of S(hort) and L(ong) stimuli—SS (same), LL (same), SL (different), LS (different)—were created for each condition. Participants went through all the stimuli once in the practice block while receiving feedback. An experimenter stayed with the participants during the practice run so that if the participants had remaining questions, they could be answered.

The main session presented 25 repetitions of all the stimuli, thus a total of 300 pairs (25 repetitions * 4 same-different pairs * 3 conditions). The participants kept receiving feedback during

[7]Interested readers are welcome to contact the first author to get speech samples.

[8]These two figures are placed next to each other for the sake of comparison. It is not the case that particular acoustic parameters are extracted from Arabic speech to create the non-speech stimuli. Relatedly, an anonymous reviewer asked why the non-speech analogue of the "fricative" intervals have lower peak amplitude than the surrounding vowels, whereas in Arabic speech samples, the fricatives have higher peak amplitude than the surrounding vowels. The reason is as follows. Since the vocalic analogues had a very simple spectral structure, they hit only limited portions of our auditory drums (recall that it consists of only 50 sine waves); i.e., they sound much quieter than natural vowels. We therefore needed to lower the amplitude of the fricative analogues accordingly.

[9]The main participants of the pilot studies were research assistants working for the Rutgers phonetics laboratory during the time of the experiments, including the second author of the paper. They were all native speakers of New Jersey English. They tried out a variety of durational settings, and decided on the values deployed in the current experiment. The same holds for Experiments III and IV.

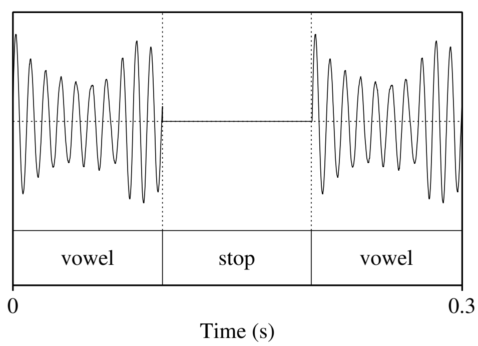

Figure 3: The stimuli. Top=stop; middle=fricative; bottom=nasal.

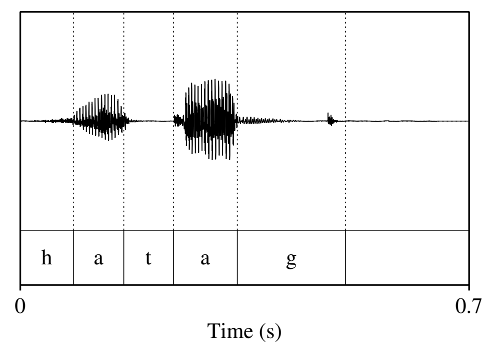

Figure 4: Corresponding speech forms (in Arabic).

the main session in the form of the correct answer (i.e. Same or Different). Superlab (ver 4.0) was used to present the stimuli and feedback (Cedrus Corporation, 2010). The order of the stimuli was randomized. All the participants wore high quality headphones (Sennheiser HD 280 Pro), and registered their responses using a Cedrus RB-730 response box. The experiment took place in a sound-attenuated laboratory.

3.2.3 Participants

Twenty-five native speakers of English participated in this experiment. These participants—and also those for Experiments II, III, and IV—were all undergraduate students at Rutgers, the State University of New Jersey, who were mostly native speakers of New Jersey English. They received course credit for their classes for participating in the experiment. (The participants of all four experiments received extra course credit, and hence this information is not repeated below.) No participant took part in more than one experiment reported in this paper. English does not have singleton-geminate contrasts, and hence their native language knowledge should not make one particular singleton-geminate contrast easier to discriminate than the other contrasts.

3.2.4 Analysis

d'-values were deployed as a measure of discriminatbility to tease apart sensitivity from bias. Using d'-values is particularly important in this experiment, because in AX-discrimination tasks, listeners are often biased to saying "Same", unless they hear a clear difference. Given the roving mode of the experiment in which different types of pairs were presented in one session, a differentiating mode of discrimination was assumed (Macmillan and Creelman 2005: 221-225). d'-values were calculated using `psyphy` package (Knoblauch, 2009) of R (R Development Core Team, 1993–2015). In a few cases, hit rates were lower than false alarm rates. In that case, negative d'-values were replaced with zero. Two listeners showed lower hit rates than false alarm rates in two out of three conditions, so their data were excluded. d'-values across three conditions were compared using a within-subject t-test. The alpha level was Bonferroni-adjusted according to the number of comparisons (.05/3=.017).

3.3 Results

Figure 5 illustrates the results of Experiment I. Each scatterplot compares d'-values in two different conditions. Each point within a scatterplot represents the pair of d'-values for each participant. Any point that is to the left of the diagonal axis shows that the listener had a higher d'-value for the condition represented in the y-axis; any point that is to the right of the diagonal axis shows that the listener showed a higher d'-value for the condition that is represented in the x-axis.

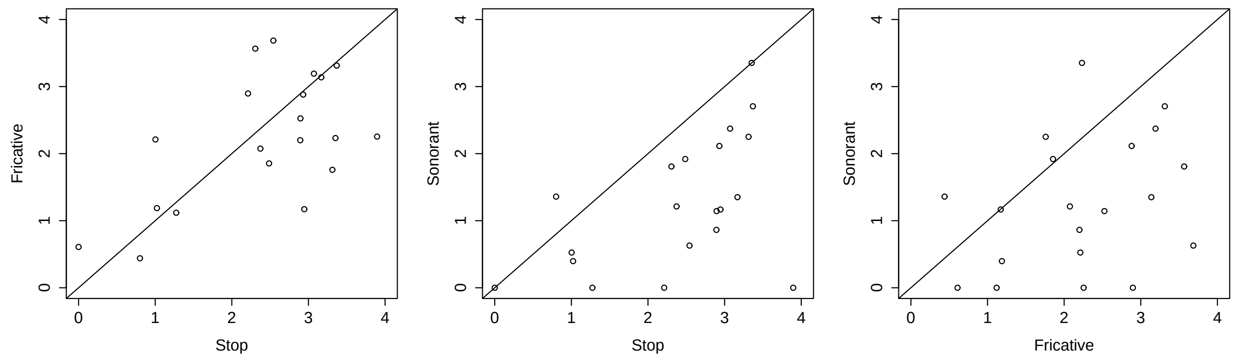

Figure 5: The distribution of d'-values in each condition in Experiment I.

In the stop-fricative comparisons, some listeners showed higher d'-values in the stop condition, while others showed the opposite pattern. The stop condition and the fricative condition thus did not differ significantly (the averages: stop 2.63 vs. fricative 2.25; $t(22) = 1.95, n.s.$). In the other two panels, most, if not all, listeners showed lower d'-values in the sonorant condition than in the stop or the fricative conditions (the average for the sonorant condition=1.38). Statistically, the sonorant condition was different from the stop condition ($t(22) = 5.71, p < .001$), and the fricative condition ($t(22) = 3.56, p < .01$).

3.4 Discussion

The result shows that sonorantal spectral continuity does make the short-long pair less discriminable. This result supports Podesva's (2002) hypothesis that sonorantal spectral continuity makes the duration of the consonantal intervals difficult to distinguish, and hence make the short-long pair harder to discriminate. This result is in turn compatible with the observation that several languages disfavor sonorant geminates, because short/long contrasts for sonorant consonants should be hard to discriminate. The conclusion further implies that languages generally disprefer contrasts that are hard to perceive, in the spirit of Dispersion Theory (see section 2.2 and also McCrary 2004).

Admittedly the current experiment, or any experiment for that matter, cannot prove the *causality* relationship between the low discriminability of a durational contrast of spectrally continuous intervals and the fact that some languages avoid sonorant geminates. However, the experiment does show the *correlation* between the two observations. It therefore seems reasonable to speculate that the avoidance of sonorant geminates may have its root in the discriminiability problem of the singleton-geminate contrasts, in some way or another.

4 Experiment II: Identification experiment, obstruents vs. sonorants

4.1 Introduction

Experiment I shows that it is hard to distinguish a short-long pair when the consonant interval is spectrally continuous with surrounding intervals. Experiment II followed up on this result with an identification experiment, which addressed whether spectral continuity makes it challenging to learn the short and long category. Although a discrimination experiment has an advantage in that the participants do not need to learn two categories, an identification experiment emulates the language acquisition situation more closely. During the course of acquisition, language learners need to learn the short and long categories based on tokens presented in isolation—parents do not usually present minimal pairs of short-long contrasts to children.

4.2 Method

4.2.1 Procedure

The identification experiment used the same set of stimuli as Experiment I. Listeners learned two categories in the practice phase, and were tested on how well they learned each category in three different conditions. Listeners were not told that the two categories were based on durational differences; instead the short category was labeled as A and the long category was labeled as B—this format again emulates the actual language acquisition situation: learners are not explicitly told that a singleton-geminate contrast is based on differences in duration.

Since a pilot experiment showed that it is difficult to learn the two categories for three types of non-speech sounds at the same time, each type of stimuli (stop, fricative, and sonorant) was blocked into small, separate sessions, each with its own practice phase and testing phase. Since the order of learning these three categories might influence their performance, the order of the presentation of the three blocks was controlled by a Latin Square design. Group 1 went through the experiments in the order of stop, fricative, and sonorant; Group 2 in the order of fricative, sonorant, and stop; Group 3 in the order of sonorant, stop, fricative.

The practice session consisted of three phases. The first phase presented five repetitions of A-B chains, followed by five repetitions of B-A chains. The second phase presented five repetitions of A in isolation and five repetitions of B in isolation. In the final practice phase, the participants were tested on 15 tokens of each with feedback. A main session contained 60 tokens of each of the short and long stimuli. The order of stimuli was randomized during the main sessions. Feedback was provided in the main session as well, because a pilot experiment without feedback resulted in performances near chance. All other aspects of the experimental methodology are the same as Experiment I.

4.2.2 Participants

Eight native English speakers participated in each Latin Square order (a total of twenty-four speakers). The general nature of the participants is identical to that of Experiment I; there is no overlap between the participants of Experiment I and those of Experiment II.

4.2.3 Analysis

As with the discrimination experiment, d'-values were used as a measure of sensitivity. Three listeners showed a negative d'-value in one of the three conditions; these values were replaced by 0. One listener showed negative d'-values for two out of the three conditions, and this person's data was therefore thrown out. Another listener was run to compensate for the gap. d'-values in the three conditions were compared using within-subject t-tests. Since the predictions were clear from the results of Experiment I, the alpha-level was not adjusted.

4.3 Results

Figure 6 illustrates the distribution of d'-value for each listener in Experiment II.

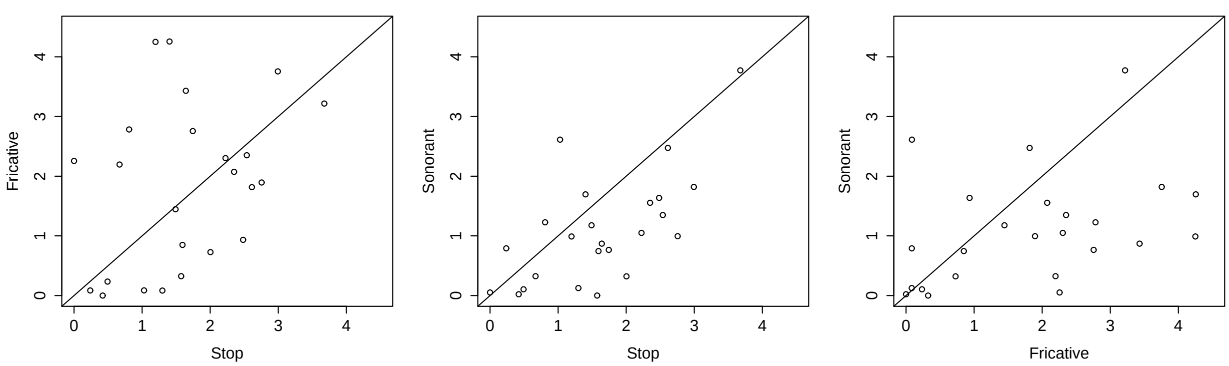

Figure 6: The distribution of d'-values in each condition in Experiment II.

Starting with the leftmost-figure, as was the case for Experiment I, some listeners were better at the stop condition, while others were better at the fricative condition; therefore, there was no significant difference between these two conditions (the averages: stop=1.63 vs. fricative: 1.84; $t(23) = -0.73, n.s.$). On the other hand, d'-values for the sonorant condition were generally lower than those for the stop condition (the middle panel: $t(23) = 3.29, p < .01$) or those for the fricative condition (the right panel: $t(23) = 2.68, p < .05$) (the average for the sonorant condition=1.10). In terms of the order effect, the average d'-values increase in successive blocks (1st block: 1.35; 2nd block: 1.47; 3rd block: 1.71), although this correlation did not reach significance ($\rho = .14, n.s.$).

4.4 Discussion

The results show that the short and long categories are generally harder to learn for the sonorant condition than the obstruent conditions. There was one exceptional listener who showed a very high d'-value in the sonorant condition (2.61) compared to the stop (1.03) or the fricative condition (0.09). This listener took the sonorant condition in the third block; therefore, it may be that this listener got used to identifying non-speech stimuli after the first two blocks.[10] All the other listeners showed a d'-value for the sonorant condition that is lower than or comparable to the d' values for the other two conditions. These results show that a duration contrast that is spectrally continuous with surrounding intervals is harder to learn than contrasts that are spectrally not continuous. The results of Experiments I and II may thus offer a phonetic explanation for why some natural languages avoid sonorant geminates.

5 Experiment III: Discrimination experiment, nasals vs. glides

The next two experiments tested the distinction between nasal geminates and glide geminates. Experiments III and IV pursued a hypothesis that this effect of sonority derives from the fact that amplitude changes facilitate the perception of segmental boundaries (Kato & Tsuzaki, 1994; Kato et al., 1997).[11] Given VCV sequences, larger amplitude changes in both VC- and CV-transitions should facilitate the demarcation of consonantal boundaries. If so, sonorants with high sonority (for example, glides) have a disadvantage in signaling their edges with respect to the surrounding vowels. In other words, glides have the problem of not having large amplitude changes, in addition to the blurriness problem identified in Experiments I and II. To test this hypothesis, Experiments III and IV tested whether larger amplitude drops facilitate the categorization and discrimination of singleton-geminate contrasts.

5.1 Method

5.1.1 Stimuli

As with Experiments I and II, Experiment III used non-speech sounds to control factors other than amplitude drops. All the stimuli mimicked VCV structures, as illustrated in Figures 7 and 8. All components were made out of pure sine waves, but with varying amplitudes. The vocalic intervals were 100 ms and were 70 dB. The consonants were either short (80 ms) or long (145 ms) with 10

[10] Recall that the order effect did not reach statistical significance. Therefore, this learning effect must have been strong specifically for this participant.

[11] Recall from section 2.2 that gemination of rhotics is likely to also involve, in addition to the perceptual problem hypothesized here, articulatory difficulties: it is articulatorily difficult to lengthen a tap because a tap involves a short closure in the first place, and a trill faces its own articulatory/aerodynamic difficulty (Ladefoged & Maddieson, 1996; Solé, 2002). A clear comparison can still be made between nasals and glides.

ms of transition on each side. In one condition, the consonant was 64 dB (the 6 dB drop condition, Figure 7) and in the other condition the consonant was 52 dB (the 18 dB drop condition, Figure 8).

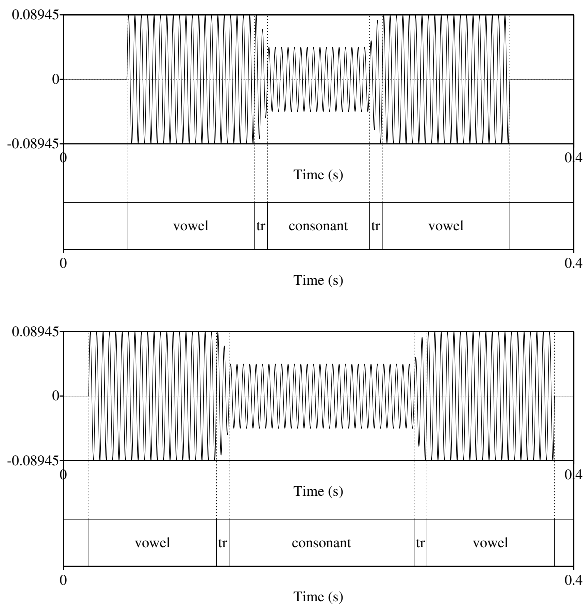

Figure 7: The 6 dB drop condition. The top panel=short; the bottom panel=long.

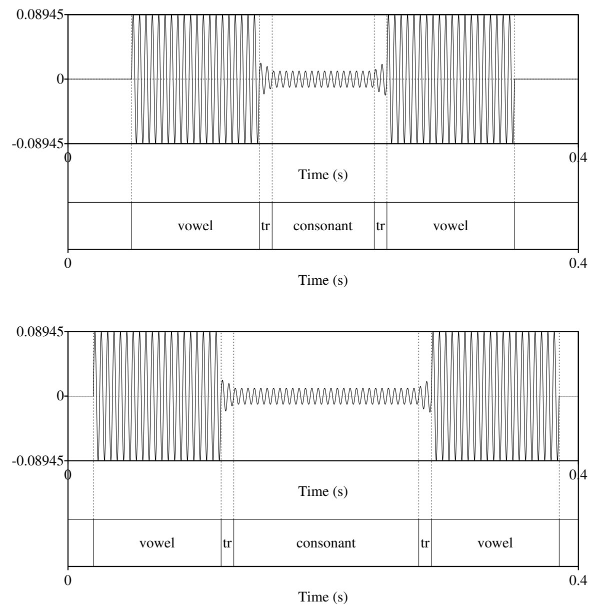

Figure 8: The 18 dB drop condition.The top panel=short; the bottom panel=long.

5.1.2 Other aspects of the experiment

The details of the procedure were similar those of Experiments I. The task was a same-different discrimination experiment. The experiment had four pairs of combinations of S(hort) and L(ong) stimuli—SS, LL, SL, LS—for each of the two condition. The ISI was 400 ms. Superlab was used to present the stimuli and feedback in the form of the correct answer.

Participants first went through all the stimuli once in the practice block while receiving feedback. The order of the stimuli was randomized. The main session presented 50 repetitions of all the stimuli, thus a total of 400 pairs (50 repetitions * 4 same-different pairs * 2 amplitude change conditions). The order of the stimuli was randomized in the main session, and the participants kept receiving feedback during the main session.

Twenty-three native speakers of English participated in this experiment. To analyze the results, d'-values were calculated as a measure of discriminatbility in the same way as Experiment I.

5.2 Results and discussion

Figure 9 illustrates the results of the discrimination experiment. The scatterplot compares d'-values in the two different conditions. Each point within a scatterplot shows a pair of d'-values for each participant. Any point that is to the right of the diagonal axis shows that the listener had a higher d'-value in the 18 dB drop condition. Almost all listeners showed higher degree of sensitivity to a short/long contrast in the 18 dB drop condition.

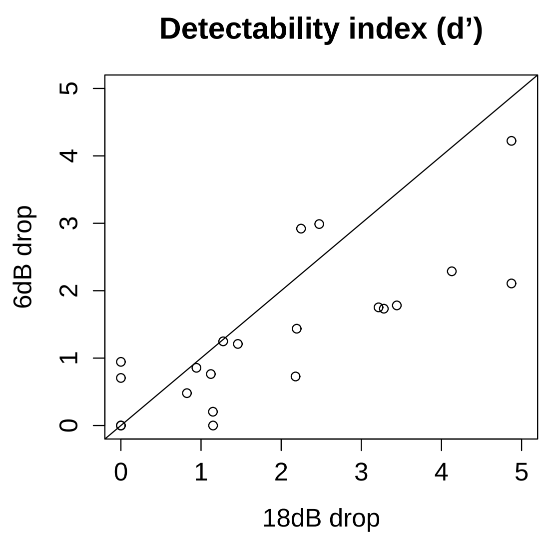

Figure 9: The distribution of d'-values in Experiment III.

The average d'-values were statistically higher for the 18 dB drop condition than the 6 dB drop condition (1.94 vs. 1.35, $t(22) = 5.54, p < .001$). Experiment III thus shows that larger amplitude changes facilitate the discrimination of singleton-geminate contrasts. As we interpret the results in terms of actual speech, short/long contrasts should be less perceptible for glides than for nasals: glides involve less of amplitude changes from surrounding vowels, which would make the perception of their duration—and consequently their short-long contrasts—difficult to perceive.

6 Experiment IV: Identification experiment, nasals vs. glides

Experiment III shows that it is harder to distinguish a short-long pair when the consonant intervals involve smaller amplitude changes with respect to the surrounding vocalic intervals. Experiment IV followed up on this result with an identification experiment, which addressed whether smaller amplitude changes make it more challenging to learn the short and long categories.

6.1 Method

Experiment IV used the same set of stimuli as Experiment III. However, in this experiment, VCV tokens were presented in isolation, as with Experiment II. Listeners learned two categories of the consonant interval (short or long) in the practice phase, and were tested on how well they learned each category in the two different conditions. Like Experiment II, the short category was labeled as A and the long category was labeled as B, and they were told nothing about durational differences between A and B.

Each type of stimuli (the 6 dB drop condition and the 18 dB drop condition) was blocked into small, separate sessions, each with its own practice phase and testing phase. The order of learning two categories was counterbalanced across the participants.

The practice session consisted of three phases. The first phase presented five repetitions of A-B chains, followed by five repetitions of B-A chains. The second phase presented five repetitions of A in isolation and five repetitions of B in isolation. In the final practice phase, the participants were tested on 15 tokens of each with feedback. The main session contained 90 tokens of each of the short and long stimuli. The order of stimuli was randomized during the main sessions. Feedback was provided in the main session, as was the case with Experiment II. Twenty native speakers of English participated in this experiment.

6.2 Results and discussion

Figure 10 illustrates the distribution of d'-value for each listener in Experiment IV.

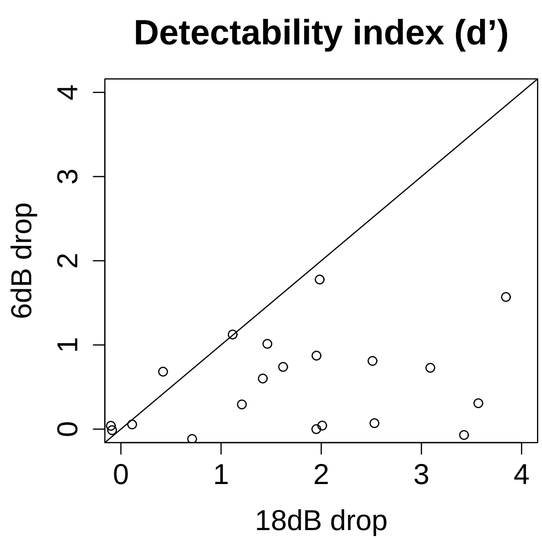

Figure 10: The distribution of d'-values in Experiment IV.

Listeners learned the contrast between short and long contrasts better in the 18 dB drop condition. Almost all listeners showed higher d'-values in the 18 dB drop condition, and the difference between the two conditions was significant (averages: 1.74 vs. 0.53, $t(19) = 4.68, p < 001$). The results thus show the short and long categories are easier to learn when there are larger amplitude drops.

The results of Experiment III and IV show that larger amplitude changes facilitate both discrimination and categorization of a short-long contrast. The results accord well with the claim that amplitude changes facilitate perceptual demarcation of segmental boundaries (Kato & Tsuzaki, 1994; Kato et al., 1997). The results also imply that, since more sonorous consonants (e.g. glides) involve smaller amplitude changes with respect to surrounding vowels, the singleton/geminate distinction would be harder to perceive for more sonorous consonants.

7 General discussion

7.1 Summary

This paper started with two phonological observations about geminates in Japanese and other languages. To take Japanese for example, (i) no glide geminates are allowed at all to make lexical contrasts, and (ii) nasal geminates are allowed, but nevertheless avoided by a phonological process. The preferential hierarchy is there obstruent geminates > nasal geminates > glide geminates. Moreover, this hierarchy seems to operate in languages that are genetically unrelated to Japanese. The question that arises is why this hierarchy is found in different languages.

To address this question, we started with Podesva's (2002) hypothesis that what may underlie the preferential hierarchy is the perceptibility of length contrasts. Specifically, sonorants may be at a disadvantage in signaling length contrasts because their durations are not easy to perceive precisely. We also expanded on Kato's observation (Kato & Tsuzaki, 1994; Kato et al., 1997) that amplitude changes facilitate the demarcation of segmental boundaries; given this observation, glides should have more disadvantage in signaling their length than nasals. To address these hypotheses, four experiments were conducted. The results show that spectral continuity and lack of amplitude changes make the perception of length contrasts difficult. The results are thus compatible with the perception-driven hypothesis about the geminate preferential hierarchy: glide geminates have both problems (the spectral continuity problem and small amplitude drop problem), while nasal geminates have the spectral continuity problem.

7.2 Implications for the phonetics-phonology interface

Taken together, the current experiments add to the growing body of the literature suggesting that phonological patterns are non-trivially affected by perceptibility of contrasts. This thesis is not new—in modern phonetics, the principle of Adaptive Dispersion was first formulated by Liljencrants & Lindblom (1972); after the advent of Optimality Theory, which allows to express phonetic grounding directly in the formulation of constraints, this thesis has received renewed interests. Now there are a plethora of cases in which languages seem to disfavor contrasts that are not highly perceptible (Flemming, 1995, 2004; Kingston, 2007; Liljencrants & Lindblom, 1972; Lindblom, 1986; Martin & Peperkamp, 2011; McCrary, 2004; Padgett, 2002, 2009; Schwartz et al., 1997a,b; Zygis & Padgett, 2010). The current work shows that the same principle may be operative at shaping cross-linguistic characteristics of geminates (see Engstrand & Krull 1994 in particular for evidence for the effect of contrast dispersion in length distinctions).

To be cautious, it is not possible to prove the *causality* relationship between the difficulty of perception of particular geminate types and their phonological behavior. However, the experimental results are *compatible* with what we would expect if languages disprefer contrasts that are hard to perceive. To the extent that the thesis of contrast dispersion is motivated elsewhere in the patterns of phonetics-phonology interface, which we believe it is (see above), we may be able to say that the perception problems of length contrasts do underlie the phonological patterns of geminates. Nevertheless, a question still remains as to how direct this relationship might be (to the extent, of course, that there is a relationship). It could be that this relationship is very direct and synchronic (Flemming, 1995; Padgett, 2002), or indirect and diachronic (Blevins, 2004a,b).

In the context of this debate, one point is worth mentioning (Kawahara et al., 2011)—recall from section 1 that sonorant geminates are avoided by several different ways. For example, Japanese speakers avoid creating nasal geminates when they make emphatic forms, and seek for another locus for gemination, as in (3). Therefore, phonetics may determine what is avoided in phonology, but the phonetic problem of a phonological structure may not determine how it is resolved in phonology. In other words, the structure [A] can be marked because [A] is confusable with [B], but it is not necessarily the case that [A] becomes [B]; in other words, the phonetic problem presented by [A] is independent of how it is resolved in phonology (Boersma, 2005; Dinnsen, 1980; Kawahara, 2006b; Keating, 1985).

7.3 Remaining questions

Finally, the current experiment opens up several opportunities for future experimentation. One question, raised by an anonymous reviewer, is how much we can generalize the current results to speakers of other languages. In particular, what would happen to speakers of language that use duration to make phonemic differences (like Japanese or Arabic)? Would Japanese speakers be particularly bad at distinguishing length contrasts for glide-analogues? Relatedly, both reviewers raised a question about the possible impact of their second language for the results of this experiment. Would exposure to L2 with actual length contrasts impact the behavior of English speakers? While addressing these questions is beyond the scope of this paper, the impact of their L1 or L2 knowledge on the perceptibility of length contrasts should be investigated in the future. Finally, the current stimuli did not involve a non-speech analogue of formant transitions, which are usually present in VCV sequences in natural languages, as long as they are not a homorganic sequence. How such formant transitions would help to demarcate perceptual boundaries of segments is to be examined in future studies.

Another remaining challenge is how to model the whole complexity of phonological patterns of geminates. We, in a sense, "distilled" patterns of geminates in such a way that we focused on languages that avoid sonorant geminates (and glide geminates in particular). However, as discussed in section 2.2, there are languages that avoid other types of geminates. Voiced obstruents geminates are a typical example, and there is a well-understood aerodynamic reason for the markedness of voiced obstruent geminates (Hayes & Steriade, 2004; Ohala, 1983; Westbury & Keating, 1986). Pharyngeal geminates are another kind of geminates that are avoided cross-linguistically, and Hansen & Myers (2014) argue that the perceptibility problem that is similar to what we discussed in this paper may lie behind the dispreference against pharyngeal geminates. Further research is necessary to fully understand what phonetic considerations underline phonological patterning of geminates.

Acknowledgments

Experiments I and II were reported in a manuscript circulated in 2011 as Kawahara et al. (2011). Experiments III and IV were reported in Kawahara (2012). Portions of this paper were presented at various occasions, including International Christian University, Rutgers University, the University of California, Santa Cruz as a distinguished alumnus lecture, GemCon at Kobe University, MAPLL at Yamagata University, and ICPP 2011 at Kyoto University. We are grateful for the comments that I received at these occasions, particularly those provided by Hiroaki Kato, John Kinston, Jaye Padgett, Julien Musolino and Yoshinori Sagisaka. We also thank two anonymous reviewers and Donna Erickson for providing useful comments on a previous version of this paper. At the final stage of writing, we received financial support from two JSPS Kakenhi grants #26770147 and #26284059, whose help we would like to acknowledge here. Any remaining errors are our responsibility.

References

Andrews, J. Richard (1975) *Introduction to Classical Nahuatl*. University of Texas Press.

Barnes, Jonathan (2002) *Positional Neutralization: A Phonologization Approach to Typological Patterns*. Doctoral dissertation, University of California, Berkeley.

Blevins, Juliette (2004a) *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette (2004b) Klamath sibilant degemination: Implications of a recent sound change. *International Journal of American Linguistics* **70**: 279–289.

Boersma, Paul (1998) *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Boersma, Paul (2001) Praat, a system for doing phonetics by computer. *Glot International* **5**(9/10): 341–345.

Boersma, Paul (2005) Phonology without markedness constraints. Slides for presentation at ICLaVE 3.

Boersma, Paul & David Weenink (1999–2015) Praat: Doing phonetics by computer. Software.

Cedrus Corporation (2010) Superlab v. 4.0. Software.

Connine, C. M., D. Titone, & J. Wang (1993) Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology* **19**: 81–94.

de Lacy, Paul (2006) *Markedness: Reduction and Preservation in Phonology*. Cambridge: Cambridge University Press.

de Lacy, Paul (2009) Phonological evidence. In *Phonological Argumentation: Essays on Evidence and Motivation*, Steve Parker, ed., London: Equinox, 43–77.

de Lacy, Paul & John Kingston (2013) Synchronic explanation. *Natural Language and Linguistic Theory* **31**(2): 287–355.

Dinnsen, Daniel (1980) Phonological rules and phonetic explanation. *Journal of Linguistics* **16**: 171–191.

Engstrand, Olle & Diana Krull (1994) Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of Adaptive Dispersion. *Phonetica* **51**: 80–91.

Flemming, Edward (1995) *Auditory Representations in Phonology*. Doctoral dissertation, University of California, Los Angeles.

Flemming, Edward (2004) Contrast and perceptual distinctiveness. In *Phonetically Based Phonology*, Bruce Hayes, Robert Kirchner, & Donca Steriade, eds., Cambridge: Cambridge University Press, 232–276.

Games, Sara (1976) *Quantity in Icelandic: Production and Perception*. Hamburg: Helmut Buske Verlag.

Ganong, W. F., III (1980) Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* **6**: 110–125.

Hansen, Benjamin B. & Scott Myers (2014) The consonant length contrast in Persian: Production and perception. Ms. Texas University.

Hayes, Bruce (1989) Compensatory lengthening in moraic phonology. *Linguistic Inquiry* **20**: 253–306.

Hayes, Bruce & Donca Steriade (2004) Introduction: The phonetic bases of phonological markedness. In *Phonetically Based Phonology.*, Bruce Hayes, Robert Kirchner, & Donca Steriade, eds., Cambridge: Cambridge University Press, 1–33.

Hura, Susan, Björn Lindblom, & Randy Diehl (1992) On the role of perception in shaping phonological assimilation rules. *Language and Speech* **35**: 59–72.

Irwin, Mark (2011) *Loanwords in Japanese*. Amsterdam and Philadelphia: John Benjamins.

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

Ito, Junko & Armin Mester (2006) Systemic markedness and faithfulness. In *Proceedings of Chicago Linguistic Society 39*, Jonathon E. Cihlar, David W. Kaiser, Irene Kimbara, & Amy Franklin, eds., Chicago Linguistics Society, 665–689.

Ka, Omar (1994) *Wolof Phonology and Morphology*. Lanham: University Press of America.

Kager, René (1999) *Optimality Theory*. Cambridge: Cambridge University Press.

Kang, Yoonjung (2011) Loanword phonology. In *The Blackwell companion to phonology*, Marc van Oostendorp, Colin J. Ewen, Hume Elizabeth, & Keren Rice, eds., Oxford: Blackwell-Wiley, 2258–2282.

Katayama, Motoko (1998) *Optimality Theory and Japanese Loanword Phonology*. Doctoral dissertation, University of California, Santa Cruz.

Kato, Hiroaki & Minoru Tsuzaki (1994) Intensity effect on discrimination of auditory duration flanked by preceding and succeeding tones. *Journal of the Acoustical Society of Japan* **15**(5): 349–351.

Kato, Hiroaki, Minoru Tsuzaki, & Yoshinori Sagisaka (1997) Acceptability for temporal modification of consecutive segments in isolated words. *Journal of the Acoustical Society of America* **101**: 2311–2322.

Kawagoe, Itsue (2015) Phonology of *sokuon*, or geminate obstruents. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton de Gruyter, 79–119.

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

Kawahara, Shigeto (2006b) Mimetic gemination in Japanese: A challenge for Evolutionary Phonology. *Theoretical Linguistics* **32**(1): 411–424.

Kawahara, Shigeto (2007) Sonorancy and geminacy. In *University of Massachusetts Occasional Papers in Linguistics 32: Papers in Optimality Theory 3.*, Leah Bateman, Michael O'Keefe, Ehren Reilly, & Adam Werle, eds., Amherst: GLSA, 145–186.

Kawahara, Shigeto (2012) Amplitude changes facilitate categorization and discrimination of length contrasts. *IEICE Technical Report. The Institute of Electronics, Information, and Communication Engineers.* **112**: 67–72.

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

Kawahara, Shigeto (2015a) Japanese /r/ is not featurelesss: A rejoinder to Labrune (2014). *Open Linguistics* **1**: 432–443.

Kawahara, Shigeto (2015b) The phonetics of *sokuon*, obstruent geminates. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton, 43–73.

Kawahara, Shigeto, Melanie Pangilinan, & Kelly Garvey (2011) Spectral continuity and the perception of duration. Ms. Rutgers University (Talk presented at International Conference on Geminates 2011, Kobe University, Jan 2011).

Keating, Patricia A. (1985) Universal phonetics and the organization of grammars. In *Phonetic Linguistic Essays in Honor of Peter Ladefoged*, Victoria A. Fromkin, ed., Orlando: Academic Press, 115–132.

Kingston, John (2007) The phonetics-phonology interface. In *The Cambridge Handbook of Phonology*, Paul de Lacy, ed., Cambridge: Cambridge University Press, 401–434.

Kingston, John, Shigeto Kawahara, Della Chambless, Daniel Mash, & Eve Brenner-Alsop (2009) Contextual effects on the perception of duration. *Journal of Phonetics* **37**(3): 297–320.

Kirchner, Robert (1998) *An Effort-Based Approach to Consonant Lenition*. Doctoral dissertation, University of California, Los Angeles.

Knoblauch, Kenneth (2009) `Psyphy`: Functions for analyzing psychophysical data in R. Software.

Labrune, Laurence (2014) The phonology of Japanese /r/: A panchornic account. *Journal of East Asian Linguistics* **23**(1): 1–25.

Ladefoged, Peter & Ian Maddieson (1996) *The Sounds of the World's Languages: 2nd Edition*. Oxford: Blackwell Publishers.

Liljencrants, Johan & Björn Lindblom (1972) Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* **48**: 839–862.

Lindblom, Björn (1986) Phonetic universals in vowel systems. In *Experimental Phonology*, John

Ohala & Jeri Jaeger, eds., Orlando: Academic Press, 13–44.

Lindblom, Björn, Susan Guion, Susan Hura, Seung-Jae Moon, & Raquel Willerman (1995) Is sound change adaptive? *Rivista di Linguistitca* **7**: 5–37.

Macmillan, Neil & Douglas Creelman (2005) *Detection Theory: A User's Guide. 2nd Edition*. Mahwah: Lawrence Erlbaum Associates Publishers.

Martin, Andrew & Sharon Peperkamp (2011) Speech perception and phonology. In *The Blackwell companion to phonology*, Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, & Keren Rice, eds., Oxford: Blackwell-Wiley, 2334–2356.

Massaro, Dominic & Michael Cohen (1983) Phonological context in speech perception. *Perception & Psychophysics* **34**: 338–348.

McCrary, Kristie (2004) *Reassessing the Role of the Syllable in Italian Phonology: An Experimental Study of Consonant Cluster Syllabification, Definite Article Allomorphy and Segment Duration*. Doctoral dissertation, University of California, Los Angeles.

McQueen, James & Mark Pitt (1996) Transitional probability and phoneme monitoring. *Proceedings of the 4th International Conference on Spoken Language Processing* **4**: 2502 – 2505.

Mithun, Marianne & Hasan Basri (1986) The phonology of Selayarese. *Oceanic Linguistics* **25**: 210–254.

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

Moreton, Elliott (2008) Analytic bias and phonological typology. *Phonology* **25**(1): 61–82.

Myers, Scott (1997) Expressing phonetic naturalness in phonology. In *Derivations and Constraints in Phonology*, Iggy Roca, ed., Oxford: Oxford University Press, 125–152.

Myers, Scott & Benjamin Hansen (2005) The origin of vowel-length neutralization in vocoid sequences. *Phonology* **22**: 317–344.

Nasu, Akio (1999) Chouhukukei onomatope no kyouchou keitai to yuuhyousei [Emphatic forms of reduplicative mimetics and markedness]. *Nihongo/Nihon Bunka Kenkyuu [Japan/Japanese Culture]* **9**: 13–25.

Nasu, Akio (2005) Asymmetries in mimetic phonology. *Onsei Kenkyu [Phonetic Studies]* **9**: 20–29.

Ní Chiosáin, Máire & Jaye Padgett (2009) Phonological evidence. In *Contrast, comparison sets, and the perceptual space*, Steve Parker, ed., London: Equinox, 103–121.

Ohala, John J. (1983) The origin of sound patterns in vocal tract constraints. In *The Production of Speech*, Peter MacNeilage, ed., New York: Springer-Verlag, 189–216.

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

Padgett, Jaye (2009) Systemic contrasts and Catalan rhotics. *The Linguistic Review* **26**(4): 431–463.

Parker, Ellen, Randy Diehl, & Keith Kluender (1986) Trading relations in speech and non-speech. *Perception & Psychophysics* **39**: 129–142.

Peperkamp, Sharon (2005) A psycholinguistic theory of loanword adaptation. In *Proceedings of the 30th Annual meeting of the Berkeley Linguistics Society*, M. Ettlinger, N. Fleischer, & M. Park-Doob, eds., Berkeley: Berkeley Linguistics Society, 341–352.

Peperkamp, Sharon & Emmanuel Dupoux (2003) Reinterpreting loanword adaptation: The role of perception. *Proceedings of ICPhS* **XV**: 367–370.

Peperkamp, Sharon, Inga Vendelin, & Kimihiro Nakamura (2008) On the perceptual origin of loanword adaptations: Experimental evidence from Japanese. *Phonology* **25**(1): 129–164.

Podesva, Robert (2000) Constraints on geminates in Burmese and Selayarese. In *Proceedings of West Coast Conference on Formal Linguistics 19*, Roger Bilerey-Mosier & Brook Danielle Lillehaugen, eds., Somerville: Cascadilla Press, 343–356.

Podesva, Robert (2002) Segmental constraints on geminates and their implications for typology. Talk given at the LSA Annual Meeting.

Prince, Alan & Paul Smolensky (2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden and Oxford: Blackwell.

R Development Core Team (1993–2015) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Valleé, & Christian Abry (1997a) The dispersion focalization theory of vowel systems. *Journal of Phonetics* **25**: 255–286.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Valleé, & Christian Abry (1997b) Major trends in vowel system inventories. *Journal of Phonetics* **25**: 233–253.

Silverman, Daniel (1992) Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* **9**(2): 289–328.

Solé, Maria-Josep (2002) Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics* **30**: 655–688.

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

Takagi, Naoyuki & 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.

Turk, Alice, S. Nakai, & Mariko Sugahara (2006) Acoustic segment durations in prosodic research: A practical guide. In *Methods in Empirical Prosody Research*, Stefan Sudhoff, Denisa Lenertova, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter, & Johannes Schliesser, eds., Berlin: Walter de Gruyter, 1–27.

Vendelin, Inga & Sharon Peperkamp (2006) The influence of orthography on loanword adaptations. *Lingua* **116**(7): 996–1007.

Vitevitch, Michael & Paul Luce (1999) Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* **40**: 374–408.

Westbury, John R. (1979) *Aspects of the Temporal Control of Voicing in Consonant Clusters in English*. Doctoral dissertation, University of Texas, Austin.

Westbury, John R. & Patricia Keating (1986) On the naturalness of stop consonant voicing. *Journal of Linguistics* **22**: 145–166.

Wilson, Colin (2006) Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* **30**(5): 945–982.

Yu, Alan (2004) Infixing with a vengeance: Pingding Mandarin infixation. *Journal of East Asian Linguistics* **13**: 39–58.

Zsiga, Elizabeth C. (2011) Local assimilation. In *The Blackwell companion to phonology*, Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, & Keren Rice, eds., Oxford: Blackwell-Wiley, 1919–1944.

Zygis, Marzena & Jaye Padgett (2010) A perceptual study of Polish fricatives, and its relation to historical sound change. *Journal of Phonetics* **38**: 207–226.