一般社団法人 電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report

[招待講演] Amplitude drops facilitate categorization and discrimination of length contrasts

Shigeto KAWAHARA†

† Department of Linguistics and RuCCs, Rutgers University

18 Seminary Pl. Department of Linguistics,
Rutgers University, New Brunswick, NJ 08901, USA

E-mail: † kawahara@rci.rutgers.edu

Abstract Kato et al. showed that given streams of sounds, listeners use amplitude changes to demarcate segmental boundaries [5]. This study builds on this result and argues that amplitude changes facilitate categorization and discrimination of consonantal short/long contrasts in intervocalic position. Given non-speech stimuli mimicking VCV-sequences, larger amplitude changes in VC- and CV-transitions improved the perceptibility of the length of C. The result implies that in real speech, short/long contrasts are harder to perceive for more sonorous consonants. This conclusion in turn accords well with the phonological, cross-linguistic observation that geminates of more sonorous consonants are disprefered more strongly in natural languages [7; 13; 14], instantiating a case of contrast dispersion [3; 9].

Keywords Perception of duration, geminates, short/long contrast, contrast dispersion, amplitude changes

1. Introduction

Previous cross-linguistic phonological studies of geminates—or long consonants—have identified two typological tendencies: (i) many languages lack sonorant geminates, and (ii) within sonorant geminates, more sonorous geminates are more strongly disfavored [7; 13; 14]. (This paper assumes the following, standard sonority scale: obstruents < nasals < liquids < glides, [12][1]). Previous studies argue that the first observation—the cross-linguistic rarity of sonorant geminates—has its roots in the difficulty of perceiving singleton/geminate contrasts. Since segmental boundaries are not (psycho)acoustically clear-cut for sonorants because of their spectral continuity with surrounding vowels, it is difficult to perceive sonorants' constriction duration accurately. This difficulty leads to lower perceptibility of singleton/geminate contrasts for sonorants than for obstruents. Since languages disfavor contrasts that are not highly perceptible [3; 9], languages tend to avoid sonorant geminates.

Building on previous work [7; 13; 14], the current study addresses the second observation about sonorant geminates: within sonorant geminates, languages tend to disfavor more sonorous geminates. This paper builds on Kato et al.'s [5] result to explore the phonetic grounding of this cross-linguistic tendency. Kato et al. showed that given streams of sounds, listeners use amplitude changes to demarcate segmental boundaries. The current study shows that amplitude changes facilitate categorization and discrimination of short/long contrasts. Given non-speech stimuli mimicking VCV-sequences, larger amplitude changes in VC- and CV-transitions improve the perceptibility of the length of C.

The current results may thus offer a phonetic explanation for why geminates of more sonorous consonants are disfavored more strongly in (at least some) natural languages. The results of the current experiments show that consonantal short/long contrasts in VCV sequences are less perceptible when there are smaller

[1] The relative sonority distinctions among different segments within liquids may be complicated [12]. We will observe later in section 2 that rhotic geminates tend to be more dispreferred than lateral geminates. See footnote 3 for more discussion.

amplitude changes in VC- and CV-transitions. Since more sonorous consonants (e.g. glides) involve smaller amplitude changes with respect to surrounding vowels, the current results imply that the singleton/geminate distinction should be harder to perceive for more sonorous consonants. If we embrace the principle of contrast dispersion in which languages disfavor contrasts that are hard to perceive [3; 9], the current results offer an explanation as to why some languages disprefer geminates of more sonorous consonants more strongly.

2. Phonological observation

First we review evidence from three languages which shows that, within sonorant geminates, languages disfavor geminates of higher sonority consonants more strongly. The first example comes from Japanese, which has lexical singleton/geminate contrasts in nasals, but not in liquids or glides [6]. For example, Japanese has words like *samma* 'fish name' and *konna* 'this' with a long nasal, as opposed to *sama* 'Mr./Mrs./Ms' and *kona* 'powder' with a short nasal. On the other hand, Japanese lacks words with long liquids or glides.

Japanese also provides evidence from a phonological alternation for a restriction against liquid geminates and glide geminates. The suffix [-ri] causes gemination of root-final consonants of mimetic roots,[2] as in (1) [8]. However, when the root-final consonant is [r], it fails to show gemination. When it is a glide, it shows insertion of a coda nasal, instead of gemination. The examples in (2) show that Japanese actively avoids creating liquid geminates and glide geminates.

(1) Root-final gemination by [-ri]
- a. /ba**t**a-(bata)/ → [ba**tt**ari]
- b. /si**n**a-(sina)/ → [si**nn**ari]

(2) Liquids and glides do not geminate
- a. /ko**r**o-(koro)/ → [ko**r**ori], *[ko**rr**ori]
- b. /hu**w**a-(huwa)/ → [hu**Nw**ari], *[hu**ww**ari]
- c. /bo**y**a-(boya)/ → [bo**Ny**ari], *[bo**yy**ari]

Ilokano offers another example of a case in which more sonorous geminates are avoided more strongly ([4]: 270-271). Ilokano resolves hiatus by changing the first vowel into a glide, and this glide formation causes compensatory gemination of the preceding consonant. This gemination process usually applies to obstruents, as in (3). In the same environment, gemination is marginally possible for nasals and [l], as in (4). According to [4], gemination of these consonants is optional, possibly with lexical variation. Gemination never applies to [r, w, y], as in (5). The difference between (4) and (5) shows that Ilokano speakers avoid more sonorous geminates more strongly. (See footnote 3 for more on rhotic geminates.)

(3) Obstruents usually become geminates after gliding of vowels
- a. /lú**t**o-én/ → [lu**tt**wén]
- b. /pag-ʔá**s**o-án/ → [pagʔa**ss**wán]
- c. /kina-ʔa**p**ó-án/ → [kinaʔa**pp**wán]
- d. /bá**g**i-én/ → [ba**gg**yén]
- e. /pag-ʔatá**k**e-án/ → [pagʔatá**kk**yán]

(4) Nasals and [l] only sporadically become geminates
- a. /dá**m**o-én/ → [da**m**wén], ?[da**mm**wén]
- b. /na-ʔali**n**o-án/ → [naʔali**n**wán], ?[naʔali**nn**wán]
- c. /pag-ʔali**ŋ**ó-án/ → [pagʔali**ŋ**wán], ?[pagʔali**ŋŋ**wán]
- d. /bá**l**e-án/ → [ba**ly**án], ?[ba**ll**yán]

(5) [r, w, y] never become geminates
- a. /pag-ʔá**r**i-an/ → [pagʔa**r**yán]
- b. /ká**r**o-án/ → [ka**r**wán]
- c. /ʔá**y**o-én/ → [ʔa**y**wén]
- d. /babá**w**i-én/ → [baba**w**yén]

The final example comes from Berber. In Berber, to derive incomplete forms, medial consonants become geminates, as shown in (6). However, when medial consonants are [R] or [w], they become stop geminates, as in (7) ([2]: 194-195).

(6) Gemination in incomplete form
- a. /n**k**r-μ/ → [n**kk**r]
- b. /l**d**i-μ/ → [l**dd**i]
- c. /n**g**i-μ/ → [n**gg**i]
- d. /n**s**a-μ/ → [n**ss**a]
- e. /n**z**a-μ/ → [n**zz**a]

(7) Stopping of sonorant geminates
- a. /n**R**a-μ/ → [n**qq**a]
- b. /ṛ**R**a-μ/ → [ṛ**qq**a]
- c. /r**w**l-μ/ → [r**gg**ʷl]
- d. /n**w**a-μ/ → [n**gg**ʷa]

(8) [l] becomes a geminate without hardening
- a. /j**l**a-μ/ → [j**ll**u]

[l] becomes a geminate without hardening in this position, as in (8). Berber thus instantiates yet another

[2] These roots are usually reduplicated when used in isolation.

case in which geminates with higher sonority are more strongly avoided, as [w] is more sonorous than [l] (see footnote 3 for a caveat about rhotic geminates.)

In summary, the examples reviewed in this section show that even in these languages that allow sonorant geminates, geminates with higher sonority are disfavored more strongly. This paper pursues the hypothesis that this effect of sonority derives from the fact that amplitude changes facilitate the perception of segmental boundaries [5].[3] 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. Since the singleton/geminate distinction primarily relies on the perception of segmental durations (see [6] for a review), listeners have more difficulty in accurately perceiving the singleton/geminate distinction of more sonorous consonants.[4] To test this hypothesis, the current study uses identification and discrimination experiments to investigate whether larger amplitude drops facilitate the categorization and discrimination of short/long contrasts.

3. Experiment I: Discrimination Experiment

3.1. Method

3.1.1. Stimuli

The current experiment used non-speech sine waves to control for factors other than amplitude drops. All the stimuli mimicked VCV structures in natural languages, as illustrated in Figures 1 and 2. All components were made out of pure sine waves, but with different amplitudes. The vocalic intervals were 100ms and 70dB. The consonants were either short (80ms) or long (145ms) with 10 ms transitions on each side. In one condition, the consonant was 64dB (henceforth the 6dB drop condition, Figure 1) and in the other condition the consonant was 52dB (the 18dB drop condition, Figure 2).

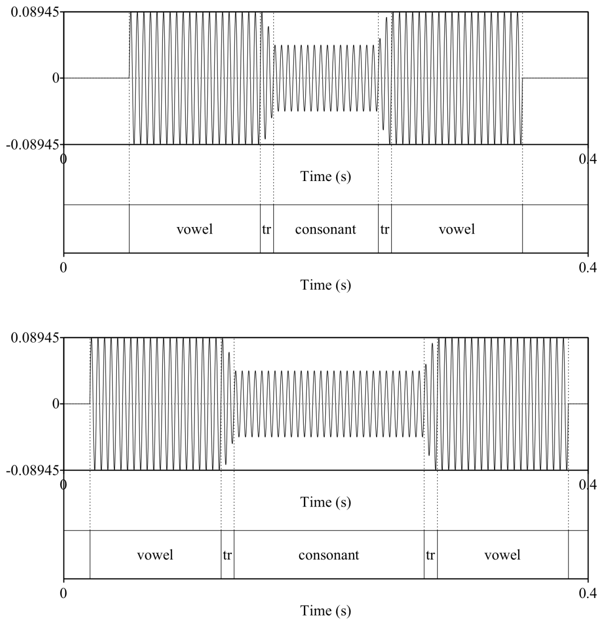

Figure 1.The 6dB drop condition. The top panel=short; the bottom panel=long. Time scale=400ms.

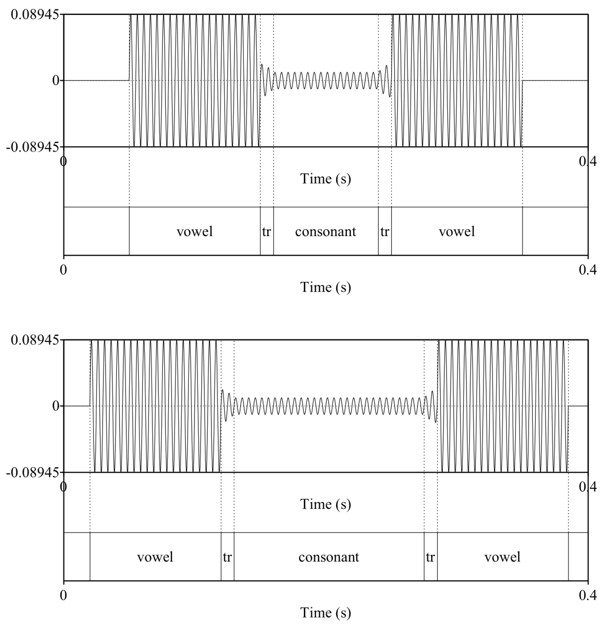

Figure 2.The 18dB drop condition. The top panel=short; the bottom panel=long. Time scale=400ms.

[3] It is acknowledged that gemination of rhotics is likely to also involve, in addition to the perceptual problem hypothesized here, articulatory difficulties: it is articulatorily difficult/impossible to lengthen a tap because a tap involves a short closure in the first place; a trill faces its own articulatory/aerodynamic difficulty [16]. A clear comparison can still be made between nasals and glides, for example.

It is also noted that this confusability problem is not all that governs phonological properties of geminates in all natural languages: other phonetic and phonological factors are likely to be relevant. This paper therefore focuses on explaining why the three languages reviewed in section 2 show stronger dispreference against sonorant geminates with high sonority.

[4] The hypothesis pursued here focuses on the perception of singleton/geminate distinctions in intervocalic position (or at least inter-sonorantal position, as in the case of Berber). The case of non-intervocalic geminates is set aside here. See [1; 11] for discussion of word-initial geminates.

3.1.2. Procedure

The task was a same-different discrimination (AX-discrimination) experiment. The experiment had four pairs of combinations of S(hort) and L(ong) stimuli—SS (same), LL (same), SL (different), LS (different)—for each of the two conditions. The ISI (inter-stimulus interval) was 400ms. Given each pair of stimuli, the participants were asked whether the two sounds were the same or different. The ITI (inter-trial interval) was 500ms. Superlab (ver 4.0, Cedrus) was used to present the stimuli and feedback. All the participants wore high quality headphones (Sennheiser HD 280 Pro), and registered their responses using an RB-730 response box (Cedrus).

Participants first went through all the stimuli once in the practice block and received feedback in the form of the correct answer (i.e. same or different). The order of the stimuli was randomized. 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 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 received feedback during the main session as well. The experiment took place in a sound-attenuated phonetics laboratory at Rutgers University.

3.1.3. Participants

Twenty-three native speakers of English participated in this experiment. They received extra course credit for their classes. English does not have singleton/geminate contrasts, so their native language knowledge should not have made one particular short/long contrast easier to discriminate than the other contrast.

3.1.4. Analysis

To analyze the results, *d'*-values were calculated as a measure of discrimination to tease apart sensitivity from bias. Given the roving mode of the experiment in which different types of pairs were presented in one session, the *d'*-values were calculated based on a differentiating mode of discrimination [10: 221-225], using `psyphy` package of R [15]. The *d'*-values between the two conditions were compared using a within-subject t-test.

3.2. Results and Discussion

Figure 3 illustrates the results of the discrimination experiment. The scatterplot compares *d'*-values in the two different conditions. Each point 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 18dB drop condition. We observe that almost all listeners showed higher degree of sensitivity to a short/long contrast in the 18dB drop condition.

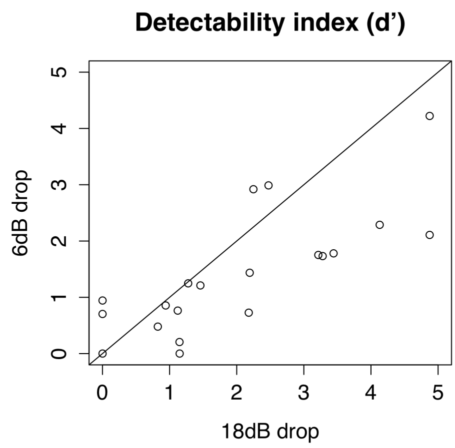

Figure 3: The distributions of *d'*-values: the discrimination experiment

The average *d'*-values were statistically higher for the 18dB drop condition than the 6dB drop condition (1.94 vs. 1.35, $t(22) = 5.54$, $p < .001$). The experiment thus shows that larger amplitude changes do facilitate the discrimination of singleton/geminate contrasts.

If we interpret the results in terms of actual speech, short/long contrasts should be less perceptible for sonorants with higher sonority (e.g. glides) than for sonorants with lower sonority (e.g. nasals). Consonants with higher sonority involve less of an amplitude change from surrounding vowels, which would make the perception of their duration—and consequently their short/long contrasts—more difficult to perceive. This result matches well with the actual phonological observation that geminates with higher sonority are disfavored more strongly in Japanese, Ilokano and Berber (section 2).

4. Experiment II: Identification Experiment

The previous discrimination experiment 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. The second experiment followed up on this result with an identification experiment, which addressed whether smaller amplitude changes make it harder to learn the short and long categories. Although a discrimination experiment has an advantage for experimental participants in that they do not need to learn two categories, an identification experiment may emulate the language acquisition situation more closely. During the course of language acquisition, learners need to learn the short and long categories based on tokens presented in isolation.

4.1. Method

4.1.1. Procedure

The identification experiment used the same set of stimuli as the discrimination experiment. However, in this experiment, VCV tokens (each panel of Figures 1 & 2) were presented in isolation. Listeners learned two categories differing in the duration 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 (the 6dB drop condition and the 18dB drop condition). 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.

Since a pilot experiment showed that it is difficult to learn the two categories (A and B) for several types of non-speech sounds at the same time, each type of stimulus (the 6dB drop condition and the 18dB drop condition) was blocked into smaller, separate sessions, each with its own practice phase and testing phase. The order of learning the 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 category with feedback. The main session contained 90 tokens each of the short and long stimuli (a total of 180 stimuli). The order of the 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.

Similar to the discrimination experiment, Superlab was used to present the stimuli and feedback. All the participants wore high quality headphones, and registered their responses using an RB-730 response box. The experiment took place in a sound-attenuated phonetics laboratory at Rutgers University.

4.1.2. Participants

Twenty native speakers of English participated in this experiment. They all received extra credit for their class. There is no overlap between the participants of Experiment I and those of Experiment II.

4.1.3. Analysis

D′-values were calculated as a measure of sensitivity using the formula, *z(Hit) − z(False Alarm)*.[5] The *d′*-values in the two conditions were compared using a within-subject t-tests.

4.2. Results and discussion

Figure 4 illustrates the distribution of *d′*-values for each listener in the identification experiment.

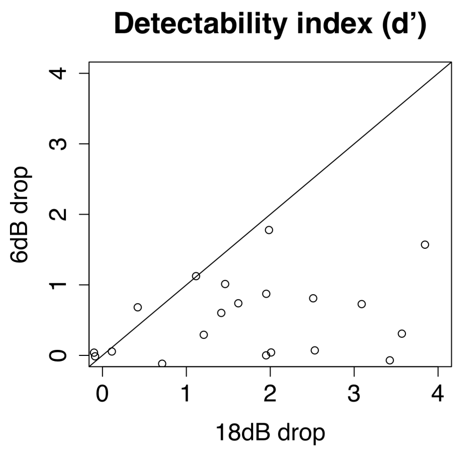

Figure 4: The distributions of *d′*-values: the identification experiment.

As expected from the results of the discrimination experiment, listeners learned the contrast between short and long better in the 18dB drop condition than in the 6dB drop condition. Almost all listeners showed higher *d′*-values in the 18dB 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 is a larger amplitude drop.

[5] Hit is the probability of saying A given A, and False Alarm is the probability of saying A given B.

5. General discussion

The results of the two experiments show that larger amplitude changes facilitate both discrimination and categorization of short/long contrasts. The results are as expected given the previous results that amplitude changes facilitate perceptual demarcation of segmental boundaries [5]. 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 in more sonorous consonants. The results thus accord well with the cross-linguistic tendency that within sonorants, the more sonorous a consonant is, the more strongly it is avoided. With the principle of contrast dispersion in which languages disfavor contrasts that are hard to perceive [3; 9], the results may explain why geminates of more sonorous consonants are more strongly disfavored.

Acknowledgements

I am grateful to the lab assistants at the Rutgers phonetics laboratory for their assistance with this research. For comments on this draft, I thank Aaron Braver, Chris Kish, and Melanie Pangilinan. This research is partially supported by a Research Council Grant from Rutgers University. Remaining errors are mine.

References

[1] Abramson, A. S. (1999). Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. *Proceedings of ICPhS,* XIV:591–594.

[2] Elmedlaoui, M. (1995). *Aspects des Representations Phonologiques dans Certaines Langues Chamito-Semitiques*. Imprimerie Najah El Jadida, Casablanca.

[3] Flemming, E. (1995). *Auditory Representations in Phonology*. Doctoral dissertation, University of California Los Angeles.

[4] Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry,* 20:253–306.

[5] Kato, H., Tsuzaki, M., and Sagisaka, Y. (1997). Acceptability for temporal modification of consecutive segments in isolated words. *Journal of the Acoustical Society of America,* 101:2311–2322.

[6] Kawahara, S. (to appear). The phonetics of obstruent geminates, *sokuon*. In Kubozono, H., editor, *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*. Mouton.

[7] Kawahara, S., Pangilinan, M., and Garvey, K. (2011). Spectral continuity and the perception of duration. Ms. Rutgers University (A talk presented at International Conference on Geminates 2011, Kobe University, Jan 2011).

[8] Kuroda, S.-Y. (1965). *Generative Grammatical Studies in the Japanese Language*. Doctoral dissertation, MIT.

[9] Liljencrants, J. and Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language,* 48:839–862.

[10] Macmillan, N. and Creelman, D. (2005). *Detection Theory: A User's Guide. 2nd Edition*. Lawrence Erlbaum Associates Publishers, Mahwah.

[11] Muller, J. (2001). *The Phonology and Phonetics of Word-initial Geminates*. Doctoral dissertation, Ohio State University.

[12] Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics,* 36:55–90.

[13] Podesva, R. (2000). Constraints on geminates in Burmese and Selayarese. In Bilerey-Mosier, R. and Lillehaugen, B. D., editors, *Proceedings of West Coast Conference on Formal Linguistics 19,* pages 343–356. Cascadilla Press, Somerville.

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

[15] R Development Core Team (1993–2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

[16] Solé, M.-J. (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics,* 30:655–688.