PITCH CUES TO THE PERCEPTION OF CONSONANT LENGTH IN JAPANESE

Haruo Kubozono[a], *Hajime Takeyasu*[b], *Mikio Giriko*[a] *& Manami Hirayama*[c]

[a]National Institute for Japanese Language and Linguistics, Japan;
[b]Mie University, Japan; [c]Ritsumeikan University, Japan
kubozono@ninjal.ac.jp; giriko@ninjal.ac.jp;
thonsei@yahoo.co.jp; manahira@fc.ritsumei.ac.jp

ABSTRACT

It is generally assumed in the literature that contrasts in phonological length are signaled primarily by phonetic durations. Consonant length in Japanese is no exception and is believed to be phonetically correlated with the closure duration (p, t, k) or frication duration (s, h) of the relevant consonant. This paper presents new experimental evidence that the phonological contrast between geminate and single consonants in Japanese is signaled not only by duration but by pitch as well. This implies that native speakers of Japanese employ pitch cues as a secondary cue to distinguish geminate consonants from singletons.

Keywords: consonant length, geminate consonant, Japanese, pitch cues, perception

1. INTRODUCTION

Many languages exhibit a contrast in vowel length and/or consonant length. These contrasts in phonological length are generally distinguished on the basis of phonetic durations, so that long vowels and consonants are substantially longer than their short counterparts. Japanese, which displays a contrast both in vowel and consonant length as shown in (1) and (2), is no exception to this. Long vowels are reported to be substantially longer than short ones, and long (geminate) consonants are considerably longer than short consonants, i.e. singletons (Fujisaki & Sugito [3]).

(1) vowel length
/biru/ ‘building’—/biiru/ ‘beer’
/obasan/ ‘aunt’—/obaasan/ ‘grandmother’
(2) consonant length:
/kako/ ‘past’—/kakko/ ‘bracket’
/saki/ ‘edge’—/sakki/ ‘a little while ago’
/bagu/ ‘bug’—/baggu/ ‘bag’

On the other hand, it is also well known that a single phonological contrast is often manifested with more than one phonetic parameter. A good example is Japanese consonant length to be discussed in this paper, which correlates with the durations of surrounding vowels as well as the duration of the consonant itself: vowels before geminate consonants are significantly longer than those before single consonants, whereas vowels tend to be shorter when they follow geminate consonants than when they follow single consonants (Han [4]).

In many languages, length contrasts can be signaled by cues other than duration. In Estonian, for example, the three-way contrast in vowel length is putatively dependent on pitch as well as duration; notably, so-called overlong and long vowels are produced with different pitch patterns (Durand [2]).

Pitch is known to play a certain role in vowel length distinction in Japanese, too. According to Kinoshita, et al. [7], vowels with a falling pitch are perceived as long vowels at a higher rate than those with a level pitch if the durational factor is properly controlled. This perceptual fact squares well with the independent acoustic fact that vowels and words with a falling pitch are physically shorter than those with a level pitch if other things are equal (see Mori [9] for Tokyo Japanese and Ishihara [5] for Kagoshima Japanese).

While pitch seems to play a certain role in the production and perception of vowel length in Japanese, it remains unclear whether or not pitch interacts with consonant length in the same manner. This paper addresses this unsolved question by looking at how Japanese listeners react to stimuli with an identical consonant duration but different pitch patterns. If Japanese listeners employ the same strategy for vowels and consonants alike, it is expected that they perceive words with a falling pitch as involving a geminate consonant at a higher rate than those with a level pitch.

2. EXPERIMENT

2.1. Method

2.1.1. Stimuli

We first had a male native speaker of Standard American English produce three monosyllabic words in isolation at a natural speed. They are all nonsense words with the same vowel.

(3) tep, ket, pek

These three monosyllabic words were produced with two pitch patterns, falling and level (Fig 1 and 2).

Figure 1: /tep/ (falling)

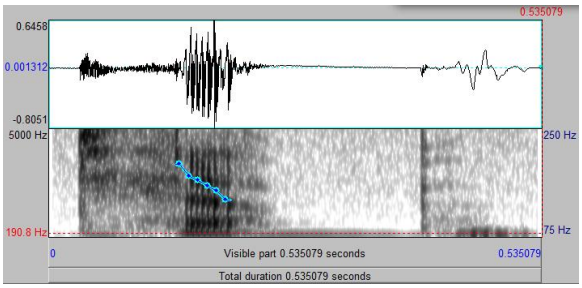
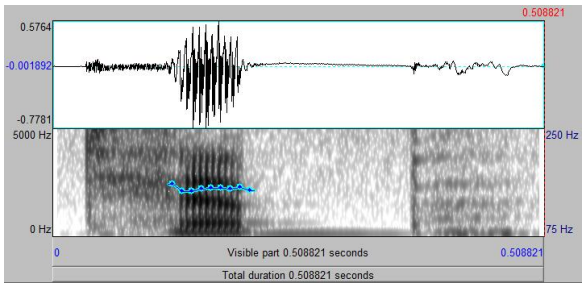


Figure 2: /tep/ (level)



The recordings ware made in a sound-proof studio, with a Marantz PMD660 digital recorder (24bit, 48kHz) and a SHURE SM58 microphone.

The recorded materials were then edited with Praat (Boersma & Weenink [1]). The monosyllabic words were edited so that the two members of each pair, e.g. /tep/ with a falling pitch and /tep/ with a level pitch, have the same durations with respect to the VOT of the initial consonant, the duration of the nuclear vowel, and the closure duration and release of the final consonant. This created three pairs of monosyllabic stimuli that are identical in temporal structure but different in pitch pattern:

(4) {tep, ket, pek} x two pitch patterns

These stimuli were copied four times, yielding a total of 24 stimuli (6 stimuli x 4 copies), which were subsequently mixed with some dummy stimuli and randomized.

2.1.2. Subjects

A total of twenty three native speakers of Japanese participated in the perception test. They are relatively young speakers ranging in age from 19 and 43. Most of them are native speakers of Kinki (Osaka/Kyoto) Japanese. None of them had a reported hearing problem.

2.1.3. Task

The subjects were asked to hear each of the test stimuli twice and to mark the best candidate for each stimulus on an answer sheet, which gives three candidates: (a) a form with a geminate consonant, (b) a form without a geminate, and (c) 'other'. The subjects were asked to give an answer when they chose the candidate (c). All the candidates were given in Japanese *katakana* letters on the answer sheet, as shown in (5).

(5) three candidates for the stimulus /tep/
  a. teppu (テップ)
  b. tepu (テプ)
  c. other ( )

2.2. Results

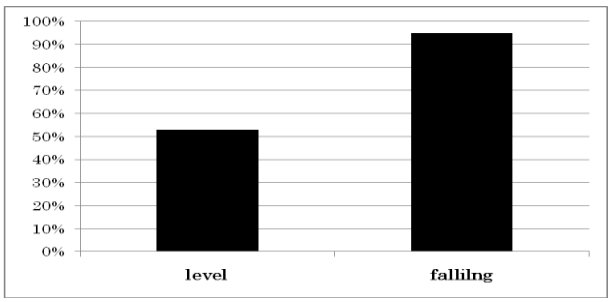
Table 1 shows the extent to which the stimuli were heard with or without a geminate consonant for the two pitch patterns (level and falling).

Table 1: Effects of pitch on geminate perception

	[-geminate]	[+geminate]	Total
LEVEL	130 (47.1%)	146 (52.9%)	276 (100%)
FALLING	14 (5.1%)	262 (94.9%)	276 (100%)
Total	144 (26.1%)	408 (73.9%)	552 (100%)

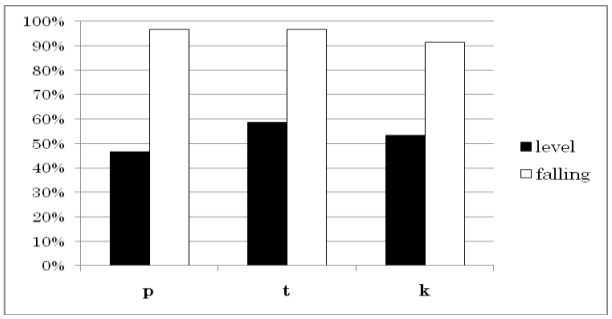
As this table shows, CVC forms with a falling pitch pattern were perceived as having a geminate consonant for the most part (94.9%). On the other hand, CVC forms with a level pitch pattern showed geminate and non-geminate responses equally well (52.9% vs. 47.1%). A statistical test (Repeated measures ANOVA) shows a significant difference between the two pitch patterns: $F(1, 110) = 105.364$, $p < 0.001$. This means that pitch patterns affect the perception of geminate consonants significantly, with the falling pitch pattern prompting the subjects to hear a geminate consonant at a higher rate than the level pitch pattern. This can be seen from Figure 3, too.

Figure 3: Pitch patterns (level vs. falling) and the percent of geminate/non-geminate perception (y axis)



While the differences in pitch pattern exert a significant effect on the perception of a geminate consonant, the three consonant types in the coda position, i.e. /p, t, k/, did not show any significant effect on the geminate/singleton perception ($F(2, 110) = 0.869$, $p = 0.422$ (*n.s.*)), indicating that place of articulation of the coda consonant is not a relevant factor. This is schematically shown in Figure 4.

Figure 4: Percent of geminate perception for /p, t, k/ with different pitch patterns (level vs. falling)



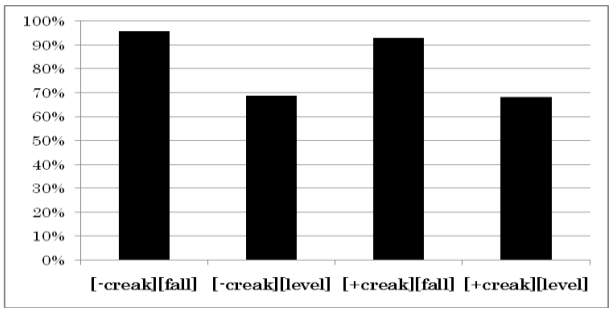
2.3. Effect of voice quality

The experimental results given and discussed above indicate that native Japanese listeners reacted differently to monosyllabic stimuli with different pitch patterns: if other things are equal, they perceived a geminate consonant more consistently in stimuli involving a falling pitch pattern than in those involving a level pitch pattern. This suggests that pitch as well as duration exerts a noticeable effect on the perception of geminate consonants as against singletons in Japanese.

On the other hand, it is also possible that the observed pitch effect on geminate perception might be attributable to some factor other than pitch. Notably, monosyllabic words in English are often accompanied with creakiness in final position, and this might affect the perception of geminate consonants in the coda position. In order to check this alternative possibility, we conducted a supplementary experiment where we used two sets of monosyllabic stimuli: one is the original monosyllabic stimuli with a creaky voice in final position and the other is its edited version without the voice quality feature. We had sixteen native speakers of Japanese listen to these stimuli in the same way as in the perception experiment outlined above. They were asked to mark the one they hear on an answer sheet.

The results of this supplementary perceptual experiment are summarized in Figure 5. As this figure shows, the presence or absence of creaky voice in the stimuli did not affect the presence or absence of a geminate consonant in perception. The factor of voice quality did not exert a significant effect on geminate responses (($F(1, 45) = 0.282$, $p = 0.598$ (*n.s.*)). On the other hand, the pitch patterns did exert a significant effect regardless of the presence or absence of voice quality: the stimuli with a falling pitch pattern showed a significantly higher ratio of geminate responses than those with a flat pitch pattern (($F(2, 110) = 63.456$, $p < 0.001$). An interaction between voice quality and pitch patterns was not significant ($F(1, 45) = 0.102$, $p = 0.751$ (*n.s.*)). This is compatible with the results of the first experiment described above, thus confirming the effect of pitch on the perception of geminate (vs. single) consonants.

Figure 5: The percent of geminate/non-geminate perception for two pitch patterns (fall/level) and two voice qualities (with/without creaky voice)



3. DISCUSSION AND CONCLUSION

From the discussion in the preceding section, it can be said that native Japanese listeners respond sensitively to differences in pitch pattern when perceiving a geminate consonant. This suggests that pitch is an additional cue to the perception of geminate consonants as opposed to single ones in Japanese.

This said, it may be argued that it is not totally clear if native Japanese listeners do actually rely on this second cue when they perceive spontaneous

speech. For one thing, Japanese has pairs of 'unaccented' words that contrast in consonant length: e.g. /kakoo/ 'descent' vs. /kakkoo/ 'appearance', /syutyoo/ 'assertion' vs. /syuttyoo/ 'business trip'. These pairs of words are pronounced with relatively flat pitch but are nevertheless readily distinguished in both production and perception on the basis of the geminate/single distinction. The existence of these pairs of words suggests that pitch may only be a potential cue to consonant length that is not actually used in speech perception.

On the other hand, there is evidence which suggests that pitch is actually used as a second perceptual cue. This evidence comes from loanword phonology of Japanese. Loanwords borrowed from English exhibit a positional asymmetry with respect to geminate consonants, by which coda consonants are geminated in Japanese if they are in word-final syllables in the source words, but not in non-final syllables (Kawagoe & Arai [6], Kubozono, et al. [8]). (6) and (7) give some typical examples: coda consonants underlined in (6) are geminated in Japanese, whereas those in (7) are not.

(6) picnic → pikunikku
technique → tekunikku
tax [tæks] → takkusu
box [bɑks] → bokkusu
fax [fæks] → fakkusu
sax [sæks] → sakkusu

(7) picnic → pikunikku
picture → pikutyaa
technique → tekunikku
taxi [tæksi] → takusii
boxer [bɑksə] → bokusaa
facsimile → fakusimiri
saxophone [sæksəfoun] → sakisofon

To take one example, the coda consonant of the second syllable in *picnic* undergoes gemination in Japanese, whereas the corresponding consonant in the first syllable does not. This peculiar positional asymmetry can be accounted for in a rather straightforward manner if the pitch patterns of the source words are taken into account as they are pronounced in isolation. Specifically, word-final syllables in English are produced with a falling pitch as in Figure 1 above, whereas non-final syllables are produced with a relatively flat pitch. This pitch difference nicely correlates with the phonological contrast between (6) and (7) regarding the perception or non-perception of a geminate consonant in Japanese loanwords.

If this pitch-based account is plausible, as it seems it is, it is realistic to assume that native speakers of Japanese make use of pitch cues when distinguishing geminate consonants from single ones in perceiving speech.

4. ACKNOWLEDGMENTS

The work reported in this paper was supported by the NINJAL collaborative research project 'Phonological characteristics of the Japanese lexicon' and the JSPS grant-in-Aid for Scientific Research (A) (Grant no. 22242011).

5. REFERENCES

[1] Boersma, P., Weenink, D. 2009. *Praat: doing phonetics by computer* (Version 5.0.46).

[2] Durand, M. 1939. Durée phonétique et durée phonologique. *Proceedings of the 3rd ICPhS Ghent*, 261-265.

[3] Fujisaki, H., Sugito, M. 1977. Sokuon no buturiteki seisitu (Physical properties of geminate consonants). In Ohno, S., Sibata, T. (eds.), *On'in*. Tokyo: Iwanami, 63-98.

[4] Han, M.S. 1994. Mora timing in Japanese. *JASA* 96(1), 73-82.

[5] Ishihara, S. 2004. *An Acoustic-phonetic Descriptive Analysis of Kagoshima Japanese Tonal Phenomena*. Doctoral dissertation, The Australian National University.

[6] Kawagoe, I., Arai, M. 2002. Syakuyoogo ni okeru sokuon (Geminate consonants in loanwords). *The Journal of the Phonetic Society of Japan* 6(1), 53-66.

[7] Kinoshita, K., Behne, D.M., Arai, T. 2002. Duration and F0 as perceptual cues to Japanese vowel quantity. *Proc. of the International Conf. on Spoken Language Processing (ICSLP)* Denver, 757-760.

[8] Kubozono, H., Itô, J., Mester, A. 2008. Consonant gemination in Japanese loanword phonology. *Proceedings of the 18th International Congress of Linguists*, Seoul.

[9] Mori, Y. 2001. Akusento no final lengthening-eno eikyoo (effect of word accent on final lengthening). *The Journal of the Phonetic Society of Japan* 5(1), 92-106.