LEXICALLY CONDITIONED PHONETIC VARIATION: AN EXPERIMENTAL TEST WITH THE SINGLETON-GEMINATE CONTRAST IN JAPANESE

Keiichi Tajima[1], Mafuyu Kitahara[2], and Kiyoko Yoneyama[3]

[1]Hosei University, [2]Sophia University, [3]Daito Bunka University
tajima@hosei.ac.jp, mafuyu@sophia.ac.jp, yoneyama@ic.daito.ac.jp

ABSTRACT

The present study investigated the extent to which lexical factors, e.g. lexical status and competition, affect phonetic realization of phonological structure, by examining the singleton-geminate contrast in Japanese. Stimuli were two- and three-mora minimal pairs contrasting in singleton vs. geminate stops, half of which were real words, e.g. /kako/, while the other half were similar-sounding nonwords, e.g. */nako/. Furthermore, half of the items had a lexical competitor contrasting in quantity, e.g. /kako/-/kakko/, while the other half did not, e.g. /tako/-*/takko/. Thirty-two native Japanese speakers read the target items interspersed with filler items. Results revealed that the vowel preceding the singleton or geminate stop was shorter for real words than for nonwords. However, no significant effect of lexical competition was found. These results suggest that phonetic realization of words is affected by some lexical factors even in a non-stress-based language such as Japanese.

Keywords: Singleton-geminate contrast, lexical competition, word frequency, Japanese

1. INTRODUCTION

1.1. Lexical effects on phonetic realization

One major question in phonetics research concerns the degree to which different levels of linguistic structure affect how spoken language is produced and perceived. It has been shown that lexical factors, i.e. knowledge about the words in the mental lexicon, affect not just the perception of spoken language but also its production.

One lexical factor that has been extensively investigated is lexical competition, i.e. the presence of lexical neighbors that are phonologically similar, e.g. *pit* vs. *bit*. Previous studies have demonstrated that words with lexical competitors are hyper-articulated, i.e. produced with enhanced phonetic attributes. For example, words with many phonological neighbors are produced with a more expanded vowel space than words with fewer neighbors [13,15]. Similarly, the voicing contrast in word-initial position, as phonetically implemented by VOT, is enhanced when a voicing minimal-pair competitor exists [2,6,8].

Another lexical factor that has received wide attention is word familiarity or frequency, i.e. how often the word appears in ordinary language usage. Studies have shown that high-frequency words are more hypo-articulated, i.e. phonetically reduced, than low-frequency words. For example, word-final /t/ and /d/ in English words are reduced more often in high-frequency words than in low-frequency words [4]. Likewise, vowels are reduced more in high-frequency words, e.g. *astronomy*, *nursery*, than in low-frequency words, e.g. *gastronomy*, *cursory* [5,7].

Most of these previous studies on lexical effects have been conducted with stress-accent languages such as English. However, it is not immediately clear whether similar effects would be observed in non-stress-accent languages as well, such as Japanese, in which phonetic dimensions such as VOT and vowel space may not function in a similar manner as they do in stress-based languages.

A previous study [11] examined whether the lexical effects on VOT found in English [2,8] would be observed in Japanese, using the Corpus of Spontaneous Japanese (CSJ) [12]. Results revealed that VOT did not vary systematically by whether or not a voicing minimal pair existed. However, duration of the following vowel was significantly longer when a voicing minimal pair existed than when it did not, suggesting a slower, more careful articulation in the presence of a competitor.

1.2. Singleton-geminate contrast

The present study focuses on the singleton-geminate contrast in Japanese, and investigates whether lexical factors influence the phonetic realization of the quantity contrast.

Japanese words can be distinguished by consonant quantity, i.e. whether a consonant is singleton or geminate, e.g. /kako/ "past" vs. /kakko/ "parentheses." Geminates count as an extra mora, a basic rhythmic unit in Japanese, so /ka.ko/ and /ka.k.ko/ are 2- and 3-mora words, respectively. The primary acoustic correlate of the distinction for stops is the duration of the stop closure, with geminates having substantially

longer closure duration than singletons. Closure duration varies substantially as a function of speaking rate, but closure-to-word-duration ratio, which divides the closure duration by the total word duration, has been shown to reliably classify between singleton and geminate words even in the face of speaking rate variation [9]. In addition to closure duration, there are other secondary cues to the distinction, including duration of the preceding and following vowels, as well as fundamental frequency and intensity differences between the surrounding vowels [10].

The effect of lexical neighbors on the singleton-geminate contrast in Japanese has been investigated using the CSJ relational database [14]. Results showed that the distinction is enhanced when the singleton and geminate words form a minimal pair than when they do not. However, the effect of word frequency on the singleton-geminate contrast has not been closely examined.

The purpose of the present study is to investigate the degree to which phonetic realization of the singleton-geminate contrast is influenced by lexical factors, by way of a controlled speech production experiment. Two factors are examined: (1) lexical competition, i.e. the presence vs. absence of minimal-pair competitors, and (2) lexical status, i.e. whether a word is a real word or a nonword. Two hypotheses are tested. (1) The singleton-geminate contrast is phonetically enhanced in the presence of a minimal-pair competitor. (2) The singleton-geminate contrast is phonetically reduced for real words compared to nonwords.

2. METHODS

2.1. Participants

Participants were 32 native Japanese-speaking college students (16 females and males each) with an age range of 20~22.

2.2. Materials

The experimental design followed that of Celata et al. [6]. The target stimuli were organized into quadruplets of minimal pairs (see Table 1). Each minimal pair consisted of Japanese words of 2-3 moras in length that contrasted in singleton /k/ vs. geminate /kk/. All words were of the form /$C_1V_1C_2V_2$/ where C_2 was either /k/ or /kk/. The first pair consisted of two real words, e.g. /kaki/ "persimmon" vs. /kakki/ "vitality." The second pair consisted of two similar-sounding nonwords, e.g. */raki/ vs. */rakki/. The third and fourth minimal pairs consisted of one real word and one nonword, in which either the singleton or geminate member was a real word, e.g. /taki/ "waterfall" vs. */takki/ and

*/haki/ vs. /hakki/ "exertion, demonstration." This design enabled orthogonal manipulation of two factors, lexical status (real word vs. nonword), and competition (with vs. without a minimal-pair competitor). The real words had word familiarity ratings between 4.4 and 6.4 on a 7-point scale from 1 (not familiar at all) to 7 (highly familiar) [1], so they were reasonably familiar to native speakers.

Table 1: Quadruplets of minimal pairs used in the experiment. Each pair consisted of either real words or nonwords. Following each word, (y) indicates that the word has a real-word minimal-pair competitor contrasting in quantity, and (n) indicates that it does not have such a competitor.

singleton		geminate	
real word	nonword	real word	nonword
/kaki/ (y)		/kakki/ (y)	
	/raki/ (n)		/rakki/ (n)
/taki/ (n)			/takki/ (y)
	/haki/ (y)	/hakki/ (n)	

Two sets of quadruplets were prepared for the experiment, one set consisting of unaccented items as shown in Table 1, and another set consisting of initially accented items, for a total of 16 target items. Other items of 2-3 moras in length, which were not analyzed in the present study, were included as filler items. The minimal pairs were split into separate lists so that each participant read only one member of each minimal pair.

2.3. Procedure

The experiment was divided into two blocks. In each block, participants first practiced reading aloud words with one of the two accent types, using ten real Japanese words of 2-3 moras in length that were different from the test items. This was to ensure that speakers would read aloud all real words and nonwords in the block using the prescribed accent type. Immediately following the practice, participants read aloud 32 test items, consisting of four target items repeated twice along with 24 filler items, presented in a pseudo-randomized order. After completing the first block, participants moved on to the second block, in which they read aloud practice and test items with the other accent type. Presentation order of the test items and the accented/unaccented blocks were counter-balanced across participants.

The items were presented on a laptop computer screen using PowerPoint slides. Each item was presented sequentially on a separate slide, in the center of the slide in hiragana syllabary characters using a 132-point font. For real words, Chinese kanji characters were simultaneously presented above the

hiragana in a 40-point font. The slides were advanced by the experimenter. When a speaker mispronounced an item, e.g. due to hesitation or wrong accent placement, it was immediately repeated again.

Recordings were made using a portable recording device at a sampling rate of 44,100 Hz at 16-bit resolution. The experiment lasted approximately 10-20 minutes for each participant.

2.4. Analysis

The recorded tokens were segmented into acoustic segments by hand using Praat [3]. The word-initial consonant (C_1) was marked from the start of visible acoustic energy to the onset of the first vowel. Initial stops with prevoicing were marked from the start of prevoicing. Vowels (V_1 and V_2) were marked from the start to the end of visible formant structure and stable phonation. The medial consonant C_2 (/k/ and /kk/) was segmented into two intervals: (1) C_2 closure, marked from the end of the first vowel to the onset of release burst, and (2) C_2 release, marked from the onset of release burst to the onset of V_2. A total of 512 tokens were analyzed (32 speakers * 16 target items per speaker), with no missing data.

The data were analyzed using linear mixed effects models, implemented in the lmer and lmerTest packages in R ver. 3.4.0. Each model contained three fixed effects: (1) quantity (singleton, geminate), (2) word status (nonword, real word), and (3) competition (without competitor, with competitor). Random effects were added step by step to the model, and likelihood ratio tests were conducted after each step to arrive at a best-fitting model that still converged to a solution. The resulting model contained by-speaker random intercepts and random slopes for lexical status, competition, and quantity and by-word random intercepts. The dependent variables were the duration of each segment (C_1, V_1, C_2 closure, C_2 release, V_2) and of the entire word.

3. RESULTS

Table 2 shows means and standard deviations of the duration of each segment and of the entire word as a function of the three independent variables examined: lexical status, competition, and quantity.

Among the segments in Table 2, C_2 closure duration is most relevant to the singleton-geminate distinction. A mixed effects model fitted to the C_2 closure duration data indicated a significant main effect of quantity ($t = 19.84$; $p < .001$). Not surprisingly, closure duration was significantly longer for geminate stops (mean = 0.218s) than for singleton stops (mean = 0.104s). However, the main effects of lexical status and competition were not significant, nor were the interactions among the three variables.

In addition to C_2 closure duration, C_2 release duration was also examined. According to Table 2, C_2 release duration ranged from 0.033s to 0.038s. A mixed effects model indicated that the main effects of lexical status, competition, and quantity and the interactions among these variables were not statistically significant ($p > .05$).

Aside from duration of the stop closure, duration of the preceding and following vowels has also been claimed to be relevant to the singleton-geminate distinction. For the preceding vowel (V_1), Table 2 shows that means ranged from 0.055s to 0.085s depending on the condition. A mixed effects model indicated a significant main effect of lexical status (t

Table 2: Means and standard deviations of the duration of each acoustic segment and of the whole word across all speakers as a function of lexical status, lexical competition, and quantity. Durations are expressed in seconds.

interval		nonword				real word			
		without competitor		with competitor		without competitor		with competitor	
		sing.	gem.	sing.	gem.	sing.	gem.	sing.	gem.
C_1	mean	0.037	0.041	0.054	0.037	0.041	0.056	0.055	0.052
	s.d.	0.012	0.013	0.039	0.013	0.012	0.037	0.013	0.010
V_1	mean	0.083	0.085	0.065	0.067	0.057	0.071	0.055	0.064
	s.d.	0.020	0.014	0.022	0.016	0.013	0.018	0.016	0.014
C_2 closure	mean	0.106	0.217	0.105	0.219	0.101	0.215	0.104	0.221
	s.d.	0.019	0.035	0.020	0.025	0.020	0.030	0.019	0.030
C_2 release	mean	0.036	0.036	0.038	0.034	0.037	0.036	0.036	0.033
	s.d.	0.016	0.010	0.011	0.012	0.013	0.013	0.017	0.009
V_2	mean	0.101	0.094	0.096	0.099	0.099	0.099	0.098	0.095
	s.d.	0.023	0.025	0.021	0.026	0.023	0.023	0.025	0.024
word	mean	0.363	0.473	0.366	0.456	0.336	0.482	0.348	0.465
	s.d.	0.053	0.048	0.050	0.053	0.046	0.069	0.055	0.050

= -2.415, $p < .05$). On average, V_1 was significantly longer for nonwords (0.075s) than for real words (0.062s). The main effects of competition and quantity, as well as the interactions among the three variables, were not significant ($p > .05$). The distribution of V_1 durations across the eight stimulus conditions is shown as boxplots in Figure 1.

Figure 1: Boxplots of V_1 duration for singleton and geminate stops as a function of word status and competition.
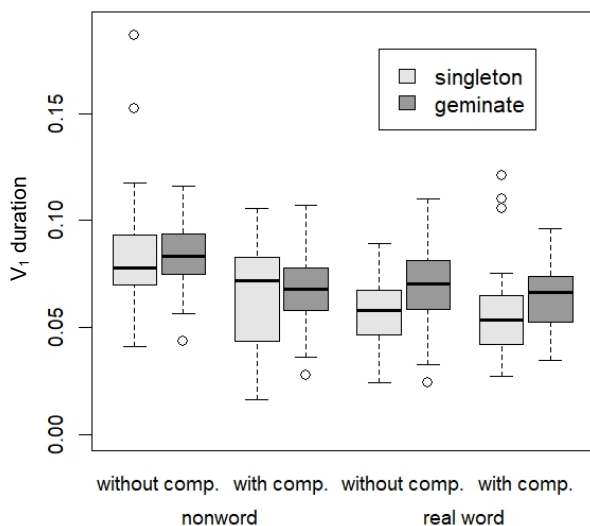
For the following vowel (V_2), Table 2 indicates that means ranged from 0.095s to 0.101s, suggesting that V_2 was generally longer but less variable than V_1. A mixed effects model conducted on the V_2 duration data indicated that the main effects of lexical status, competition, and quantity were not significant, nor were the interactions among the three variables.

For the sake of completeness, the other intervals in Table 2 (C_1 and word duration) were also analyzed using mixed effects models. The only significant effect was the main effect of quantity for word duration ($t = 3.40$, $p < .05$). Target words were significantly longer on average when they contained geminate stops (0.469s) than when they contained singleton stops (0.353s). No other main effects or interactions were significant.

Since speaking rate varied somewhat across speakers, further analyses were conducted with relative measures in which segment duration was divided or normalized by total word duration. A mixed effects model was fitted to relative measures for each of the five intervals in the target word (C_1, V_1, C_2 closure, C_2 release, V_2). The only significant effect was the main effect of quantity for C_2-closure-to-word-duration ratio ($t = 5.60$, $p < .001$). C_2-closure-to-word-duration ratio was significantly larger for words with geminate stops (0.467) than for words with singleton stops (0.298).

4. DISCUSSION AND CONCLUSION

The present study investigated the extent to which the phonetic realization of the singleton-geminate stop contrast in Japanese is influenced by lexical status and by the presence of minimal-pair competitors, by analyzing Japanese speakers' productions of words that varied orthogonally with respect to two lexical factors, lexical status (real word vs. nonword) and competition (with vs. without a minimal-pair competitor).

Results revealed that V_1 duration, i.e. duration of the vowel preceding the singleton or geminate stop, was shorter for real words than for nonwords. This result is consistent with past findings demonstrating that high-frequency or high-familiarity words in English tend to have phonetically reduced consonants and vowels [4,5,7]. However, it is not clear whether the difference in V_1 duration depicted in Figure 1 is a result of phonetic reduction of real words, phonetic enhancement of nonwords, or simply more careful pronunciation of nonwords. Since real words, but not nonwords, were presented with Chinese kanji characters in order to disambiguate the identity of the target words, this difference in stimulus presentation may have to some extent contributed to the observed effect of lexical status. This could be circumvented in future work by presenting both real words and nonwords in the same way, e.g. with hiragana syllabary only.

Meanwhile, results from the present study did not show systematic effects of lexical competition, i.e. presence of minimal-pair lexical neighbors. This is in contrast to studies that have found effects of minimal-pair competitors in English [2,8,13,15] and in other languages including Japanese [6,11,14]. This discrepancy may possibly be due to methodological differences among the studies. For example, the present study used controlled lab speech with target words produced in isolation, while some of the other studies used speech corpora with target words produced in continuous speech. Further research is needed to clarify how these factors affect lexically induced phonetic variation.

In short, results from the present study suggest the possibility that some lexical factors such as the lexical status of words potentially have phonetic consequences, in a hitherto understudied, typologically distinct, non-stress-accent language such as Japanese.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Amano, S., Kondo, T. 1999. *Lexical Properties of Japanese*, Tokyo: Sanseido.

[2] Baese-Berk, M. Goldrick, M. 2009. Mechanisms of interaction in speech production. *Lang. Cogn. Process.* 24, 527–554.

[3] [2.5] Boersma, P. Weenink, D. 2018. Praat: doing phonetics by computer. http://www.fon.hum.uva.nl/praat/.

]4] Bybee, J. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In Barlow, M., Kemmer, S. (eds.), *Usage-Based Models of Language*. Stanford: CSLI, 65–85.

[5] Bybee, J. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14, 261–290.

[6] Celata, C., Dmitrieva, O., Meluzzi, C. Valentina C. 2018. The effects of lexical status and lexical competitors on production of Italian stops. *Poster presented at LabPhon 16*, Lisbon, Portugal.

[7] Fidelholtz, J. 1975. Word frequency and vowel reduction in English. *Chicago Linguistic Society* 11, 200–213.

[8] Goldrick, M., Vaughn, C. Murphy, A. 2013. The effects of lexical neighbors on stop consonant articulation. *J. Acoust. Soc. Am.* 134, EL172-7.

[9] Hirata, Y., Whiton, J. 2005. Effects of speaking rate on the single/geminate stop distinction in Japanese. *J. Acoust. Soc. Am.* 118, 1647-1660.

[10] Idemaru, K., Guion, S. G. 2008. Acoustic covariants of length contrast in Japanese stops. *J. IPA*, 38, 167-186.

[11] Kitahara, M, Tajima, K., Yoneyama, K. 2019. The effect of lexical competition on realization of phonetic contrasts: A corpus study of the voicing contrast in Japanese. Submitted. In *Proc. ICPhS2019*.

[12] Maekawa, K. 2003. Corpus of Spontaneous Japanese: Its design and evaluation. In *Proc, ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003)*, Tokyo.

[13] Munson, B., & Solomon, N. P. 2004. The effect of phonological neighborhood density on vowel articulation. *J. Sp. Lang. Hear. Res*. 47, 1048-1058.

[14] Sano, S. 2017. Minimal pairs and hyperarticulation of singleton and geminate consonants as enhancement of lexical/pragmatic contrasts. *NELS 48*, Reykjavík, Iceland.

[15] Scarborough, R. A. 2010. Lexical and contextual predictability: Confluent effects on the production of vowels. *Lab. Phon.* 10, 557-586.