THE DISTRIBUTION OF SINGLETON/GEMINATE CONSONANTS IN SPOKEN JAPANESE AND ITS RELATION TO PRECEDING/FOLLOWING VOWELS

Shin-Ichiro Sano

Keio University, Japan
s-sano@keio.jp

ABSTRACT

By studying the *Corpus of Spontaneous Japanese* [1], this project offers new findings on the patterns in the distribution of singleton and geminate consonants in spoken Japanese. It also examines the relationship between singletons/geminates and preceding/following vowels.

The study yielded the following findings that would be difficult to observe with other approaches, such as intuition/dictionary-based studies and perception experiments. The patterns in the distribution and phonetic properties of singletons and geminates differ greatly according to place, manner, and voicing. Specifically, less sonorous sounds produce greater singleton-geminate contrasts and are thus more compatible with gemination. Furthermore, the distribution affects the duration of preceding/following vowels: 1) both preceding and following vowels are longer when next to geminates than to singletons; 2) the duration of the preceding and following vowels positively correlates with the duration of singletons/geminates. The study also uncovered hitherto unnoticed aspects, such as properties of affricate, fricative, and nasal geminates.

Keywords: duration, corpus, manner, continuancy, vowel

1. INTRODUCTION

1.1. Length contrasts and singleton/geminate consonants

In Japanese, segment duration plays an important role, both lexically and pragmatically. Modern Japanese has a variety of length contrasts. For vowels, the distinction between short and long carries lexical contrasts, as represented in examples (1) and (2):

(1) /tori/ 'bird' vs. /toori/ 'street'
(2) /biru/ 'building' vs. /biiru/ 'beer'

Similarly, the contrastive nature of consonant length is exemplified in such minimal pairs as shown in (3) and (4):

(3) /kako/ 'past' vs. /kakko/ 'parenthesis'
(4) /hato/ 'dove' vs. /hatto/ 'hat'

Regarding the length contrasts, short consonants are called singletons, while long consonants are called geminates [2-9]. The primary acoustic correlate of gemination in Japanese is closure duration: geminates are twice or three times as long as their singleton counterparts, although the duration differs according to place and voicing.

1.2. Previous studies of singleton/geminate consonants

Thus far, many studies have been conducted of singletons and geminates [10-21]. These studies have focused on the differences between singletons and geminates in an attempt to identify cues or factors that can affect the distinction between them. These phonetic studies have identified factors such as duration, (duration of) preceding/following consonants/vowels, and non-durational acoustic correlates like intensity, F0, and F1. Among many cues or factors, the difference in constriction duration has been argued to be the primary acoustic correlate of the singleton-geminate contrast.

Previous studies of singletons and geminates were also conducted from a variety of perspectives. These studies have made substantial contributions to phonetic and phonological studies, as well as to many theoretical debates.

1.3. Problems and goals

Previous studies were largely based on experimental data and focused mostly on perception. On the other hand, the acoustic correlates of geminates in spontaneous speech have not been sufficiently explored. Additionally, most of these studies have targeted stop geminates [6]. For example, the areas that need to be addressed are manner of articulation (such as fricative, affricate, and nasal geminates) and voicing (such as voiced geminates).

Given this background, the goals of this study are as follows. Focusing on production in actual language use, the current study (i) sheds light on the frequency distribution of singletons and geminates, in addition to their durational properties that have been reported in previous studies; and (ii) examines the segmental properties of singletons and geminates, such as manner (sonority) and its relation to preceding/following vowels.

2. METHOD

2.1. Corpus

This study employs the *Corpus of Spontaneous Japanese–Relational Database* (henceforth CSJ-RDB) [1]. The CSJ-RDB is a part of the CSJ core and has abundant annotations. This corpus consists of 201 speech samples, amounting to 45 hours of speech.

2.2. Data collection

I retrieved the target data from the CSJ-RDB by focusing on 12 speech samples by 11 speakers, amounting to about 34,000 words. Using MySQL (http://www.navicat.com), which implements the programming language SQL, this study employed the phonetic/phonological and morphological information annotated in the CSJ-RDB.

The CSJ-RDB adopts the traditional phonemic transcription of Japanese, where "*Q*" represents the closure (constriction) portion of oral geminates, while "*N*" represents the moraic nasal that can comprise the constriction portion of nasal geminates. This study targeted geminates (*Q*s followed by oral consonants and *N*s followed by nasal consonants) and singletons (consonants observed in other contexts).

It also focused only on singletons/geminates occurring word-medially (except when derived from *N* that can occur word-finally, as in *jibuN-no* 'self-GEN'): the word (utterance)-initial position may be preceded by a pause, making it difficult to accurately measure the closure duration of stops in particular. For consistency, this study excluded tokens of singletons/geminates in word-initial position.

The data retrieval instruction was given using a search formula specific to MySQL. From the retrieved data, I filtered out certain irrelevant information. Tokens were excluded from the dataset if the targeted segments were a part of filled pauses or word fragments.

The durations of the constriction and burst (only frication or constriction for continuants, such as fricatives and nasals) portions of singletons and geminates were calculated by referring to the onset time (StartTime) and offset time (EndTime) annotated for each segment, as shown in (6) and (7). Other segmental properties, such as manner, voicing, and place, were manually annotated for each target segment.

(6) duration of singleton:
  constr.: offset (preceding V) to onset (target C)
  burst: onset (target C) to offset (target C)

(7) duration of geminate:
  constr.: onset (Q or N) to onset (target C)
  burst: onset (target C) to offset (target C)

2.3. Dataset

An exhaustive search and filtering of the data from the CSJ-RDB resulted in 10,027 tokens, of which 8,407 (83.8%) were singletons and 1,620 (16.2%) were geminates. For the whole dataset, the mean duration of the singletons is 41.63 ms, the mean duration of the geminates is 115.47 ms, and the singleton-geminate ratio (SG ratio) is 2.77.

The retrieved tokens were then analyzed in terms of the segmental properties of singletons and geminates (place, manner, and voicing), and those of preceding vowels. The distributional skews in the results were tested by using the linear mixed-effects model (lmer) in R (version 3.3.2) [21-22] with speakers and items included as random effects.

3. ANALYSIS

3.1. The distribution of singleton/geminate consonants

This section presents details of the distribution of singletons and geminates observed in the CSJ-RDB. Most previous studies reported the durational properties of stops but not of other segments such as fricatives, affricates, and nasals [6, 10-12, 14-16, 23, 24].

Table 1 shows the mean duration and frequency distribution of singleton-geminate pairs by place, manner, and voicing. For each segment, the mean duration (ms) is followed by the parenthesized token frequency. Throughout this paper, the mean duration and the SG ratio were calculated for stops and affricates on the basis of the closure duration, and on the basis of the constriction duration for fricatives and nasals.

Table 1 suggests that (i) more geminates occur as stops; (ii) voiced segments ([b], [d], [g], [z], [dʒ]) are rarely geminated, which is consistent with the aerodynamic difficulty of producing voiced obstruent geminates [16, 25, 26], and [b] especially tends to resist gemination [27]; (iii) although it differs depending on each segment, the SG ratio is greater overall for stops and affricates (2.89) than for fricatives and nasals (2.13). Geminates of highly sonorous sounds such as glides, liquids, and approximants were not found. This offers an interesting contrast with some Romance languages; for example, Italian shows a high propensity for gemination of liquids and nasals [28].

Table 1: The mean duration and frequency distribution of singleton-geminate pairs by place, manner, and voicing in the CSJ-RDB.

	Singleton	Geminate	SG
stop			
[p]	48.94 (10)	118.00 (76)	2.41
[t]	49.17 (825)	116.98 (763)	2.38
[k]	43.95 (1,586)	124.78 (262)	2.84
[b]	38.51 (204)	NA	NA
[d]	25.52 (299)	96.01 (2)	3.76
[g]	20.82 (251)	126.11 (1)	6.06
fricative			
[f]	58.40 (19)	114.46 (2)	1.96
[s]	55.30 (1,015)	134.15 (33)	2.43
[ʃ]	54.10 (638)	118.46 (25)	2.19
[h]	55.08 (75)	NA	NA
[z]	48.91 (156)	NA	NA
affricate			
[tʃ]	35.69 (432)	71.17 (38)	1.99
[dʒ]	27.33 (36)	NA	NA
nasal			
[m]	51.54 (407)	112.88 (48)	2.19
[n]	43.93 (458)	108.35 (370)	2.47

The results overall support the previous observation that SG ratios follow the order of voiced stops > voiceless stops > fricatives. Additionally, singleton fricatives are longer than singleton stops [10, 24, 29], while the difference in duration between geminate fricatives and geminate stops is smaller. Consequently, the SG ratio is smaller for fricatives than for stops, which results in a less perceptible contrast in fricative pairs. Cross-linguistically, pairs that do not contrast straightforwardly are observed at lower frequency [16, 30, 31], which is attested also in the lower frequency of fricative geminates here.

Having confirmed the overall distribution in the corpus, the next step is to analyze the duration of geminates in detail. Previous studies argued that the primary acoustic cue for gemination is the closure duration; accordingly, the primary question to be addressed here is whether this holds in spontaneous speech.

Table 2 summarizes the durational ratios for closure and burst for geminates and singletons. Fricatives and nasals, which inherently do not have a closure, were excluded from the analysis. For each segment, the SG ratio is followed in parentheses by the mean duration (geminate/ singleton).

Table 2 shows the following: (i) for all segments, the closure portion is two to three times as long as for their singleton counterparts (t = 53.95, $p < 0.01$) (except for the voiced segments [d] and [g] that occur with low frequency), while the burst portion shows an SG ratio of around one, suggesting that the duration of singletons and geminates in the burst portion does not differ significantly (t = -0.864, n.s.); (ii) for stop geminates, the duration of the burst portion is shorter for geminates than for singletons (t = -6.04, $p < 0.01$), suggesting that the durations of closure and burst are inversely proportional to each other; (iii) for affricate geminates, in addition to the closure portion, the frication portion is longer as well, contra the claim that the primary acoustic correlate of affricate geminates seems to lie in the difference in the closure duration, and not in the frication duration [32].

Table 2: The ratio of singleton-geminate (closure and burst portions) pairs in the CSJ-RDB.

	SG ratio (closure)	SG ratio (burst)
stop		
[p]	2.41 (118.00/48.94)	1.17 (18.96/16.21)
[t]	2.38 (116.98/49.17)	0.94 (15.90/16.92)
[k]	2.84 (124.78/43.95)	1.02 (25.25/24.68)
[b]	NA	NA
[d]	3.76 (96.01/25.52)	1.13 (16.27/14.35)
[g]	6.06 (126.11/20.82)	1.29 (22.56/17.46)
affricate		(frication)
[tʃ]	1.99 (71.17/35.69)	1.40 (64.42/45.90)
[dʒ]	NA	NA

3.2. Manner

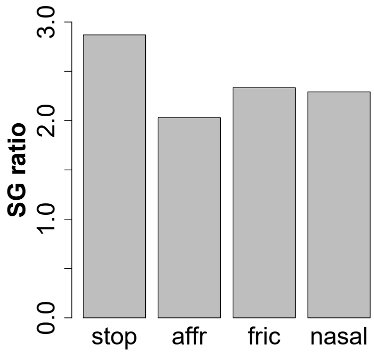
Previous studies have focused mainly on stop geminates; thus, fricative geminates, affricate geminates, and sonorant geminates have not yet been sufficiently explored. This section discusses the differences in the distribution of singletons and geminates according to manner. Figure 1 summarizes the distribution.

As Figure 1 shows, the SG ratio is higher in stops than in affricates, fricatives, and nasals (t = 7.2, $p < 0.01$). The order of the SG ratio is mostly consistent with the sonority hierarchy: stop > affricate > fricative > nasal > liquid, approximant, except for the lowest ratio in affricates. As we observed in Table 2, not only the closure portion but the frication portion of affricates is longer in geminates than in singletons (t = 2.89, $p < 0.01$, SG ratio in frication portion: 1.42), which makes the entire duration sufficiently long. This durational compensation may affect the lower SG ratio in closure portion shown in Figure 1.

The four manner categories can be divided by the continuancy: non-continuants (stops and affricates), which include a closure portion, and continuants (fricatives and nasals), with continuous airflow. The observation that non-continuants show higher SG ratios than continuants (Section 3.1) reflects the fact that the closure portion plays a crucial role in the perception of the singleton-geminate contrast.

This result confirms the claim that sounds with lower sonority can produce greater singleton-geminate contrasts [33]; geminates are preferred in contexts where they produce greater contrasts [16, 30, 31].

Figure 1: The distribution of SG ratio by manner in the CSJ-RDB.



3.3. Preceding/following vowels

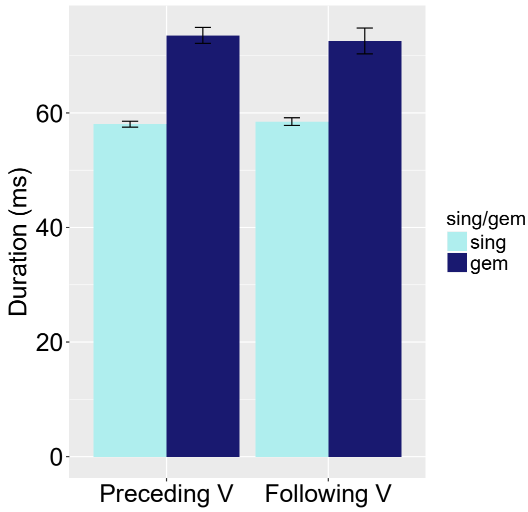
Studies of the durational properties of segments preceding or following singletons and geminates have been conducted experimentally [12, 15, 16, 23, 24, 29, 30, 34-36]. These studies focused on the effects of consonant duration on preceding or following consonants or vowels (i.e., $C_1V_1C_2V_2$, where C_2 is a singleton or geminate) on the perception of singletons and geminates. Additionally, various acoustic measures that may affect classification accuracy have been proposed and tested in perception experiments (e.g., raw C duration, C/V_1 ratio, C/Word [13, 37], C/V_2 ratio [35, 37]). The general understanding of Japanese gemination patterns is that pre-geminate vowels are longer than pre-singleton vowels. Post-geminate vowels, on the other hand, are shorter than post-singleton vowels, though the difference in duration of the preceding vowels is more substantial and more consistent than that of following vowels [35]. Hence, the current study examines the relationship between preceding and following vowels and the durations of singletons and geminates.

In Japanese, vowels are longer before geminates than before singletons. This pattern is also observed in Figure 2: preceding vowels are 15.41 ms longer before geminates (73.54 ms) than before singletons (58.13 ms) (t = 18.85, p < 0.01). In addition, the durations of preceding vowels are positively correlated with the durations of singletons and geminates (t = 33.21, p < 0.01, r = 0.32). This shows that longer consonants induce increased durations of preceding vowels, which is consistent with between-mora-timing compensation [11, 12, 38].

The durations of the following vowels show a similar pattern, contrary to the previous observation that post-geminate vowels are shorter than post-singleton vowels: following vowels are 14.09 ms longer after geminates (72.58 ms) than after singletons (58.48 ms) (t = 8.84, p < 0.01). Furthermore, a positive correlation between the durations of following vowels and singletons/geminates was observed (t = 27.73, p < 0.01, r = 0.27), which suggests the absence of within-mora-timing compensation.

These results support the status of the duration of preceding and following vowels as a secondary acoustic correlate of the singleton-geminate contrast.

Figure 2: The mean duration of vowels preceding/following singletons/geminates in the CSJ-RDB.



4. CONCLUSIONS

This study of singletons and geminates (i) described the frequency distribution and durational properties of singletons and geminates using actual speech data; (ii) examined the segmental properties of singletons and geminates, such as manner (sonority), and the relationship between singletons and geminates and preceding/following vowels.

The results of the current study confirm (i) the importance of closure duration in singleton-geminate contrasts and the lengthening of frication duration in the gemination of affricates; and that (ii) preceding and following vowels show the same pattern: vowels are longer before/after geminates than before/after singletons, contrary to previous observations, which suggests the presence of between-mora-timing compensation and the absence of within-mora-timing compensation, implying the existence of some temporal unit higher than mora in Japanese.

7. REFERENCES

[1] National Institute for Japanese Language and Linguistics. 2012. *CSJ-RDB*.

[2] Arisaka, H. *On'inron* [Phonology]. Tokyo: Sanseido, 1940.

[3] Hashimoto, S. 1950. *Kokugo On'in-no Kenkyū* [Studies on Japanese Phonology]. Tokyo: Iwanami Shoten.

[4] Hattori, S. 1960. *Gengogaku-no Hōhō* [Methodology in Linguistics]. Tokyo: Iwanami Shoten.

[5] Kawagoe, I. 2015. The phonology of sokuon, or geminate obstruents. In: Kubozono, H. (ed), *Handbook of Japanese Phonetics and Phonology*. Berlin: De Gruyter Mouton, 79–119.

[6] Kawahara, S. 2015. The phonetics of sokuon, or geminate obstruents. In: Kubozono, H. (ed), *Handbook of Japanese Phonetics and Phonology*. Berlin: De Gruyter Mouton, 43–73.

[7] Koizumi, T. 1978. *Nihongo-no Seishohō* [Orthography of Japanese]. Tokyo: Taishukan Shoten.

[8] Vance, T. 1987. *An Introduction to Japanese Phonology*. Albany: State University of New York.

[9] Vance, T. 2008. *The Sounds of Japanese*. Cambridge: Cambridge University Press.

[10] Beckman, 1982. M. Segment duration and the 'mora' in Japanese, *Phonetica* 39, 113–135.

[11] Han, M. 1962. The feature of duration in Japanese. *Onsei-no Kenkyū* [Study of Sounds] 10, 65–80.

[12] Han, M. 1994. Acoustic manifestations of mora timing in Japanese. *JASA* 96, 73–82.

[13] Hirata, Y., Whiton, J. 2005 Effects of speaking rate on the singleton/geminate stop distinction in Japanese. *JASA* 118, 1647–1660.

[14] Homma, Y. 1981. Durational relationship between Japanese stops and vowels. *J of Phonetics* 9, 273–281.

[15] Idemaru, K., Guion, S. 2008. Acoustic covariants of length contrast in Japanese stops. *J of IPA* 38(2), 167–186.

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

[17] Ridouane, R. 2010. Geminates at the junction of phonetics and phonology. In: Fougeron, C., Kuehnert, B., lmperio, M., Valleé, N. (eds.), *Laboratory Phonology* 10. Berlin: Mouton de Gruyter, 61–90.

[18] Kawahara, S., Sano, S. 2013. A corpus-based study of geminate devoicing in Japanese: Linguistic factors. *Language Sciences* 40, 300–307.

[19] Kawahara, S., Sano, S. 2017. /p/-driven geminate devoicing in Japanese: Corpus and experimental evidence. *Journal of Japanese Linguistics* 32, 57–77.

[20] Sano, S. 2017. A corpus-based study of phonological variation: The domain of OCP and morphological boundaries. *Proceedings of WCCFL 34*, Somerville: Cascadilla Press, 439–446.

[21] R Development Core Team. 1993–2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[22] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015 Fitting linear mixed-effects models using lme4. *J of Statistical Software* [Online], 67.1: 1–48.

[23] Hirose, A., Ashby, M. 2007. An acoustic study of devoicing of the geminate obstruents in Japanese. In: Trouvain, J., Barry, W. J. (eds.), *Proc. of ICPhS XVI*. Saarbrücken: Germany, 909–912.

[24] Port, R., Dalby, J., O'Dell, M. 1987. Evidence for mora timing in Japanese. *JASA* 81, 1574–1585.

[25] Hayes, B., Steriade, D. 2004. Introduction: The phonetic bases of phonological markedness. In: Hayes, B., Kirchner, R., Steriade, D. (eds.), *Phonetically Based Phonology*. Cambridge: Cambridge University Press, 1–33.

[26] Ohala, J. 1983. The origin of sound patterns in vocal tract constraints. In: MacNeilage, P. (ed.), *The Production of Speech*. New York: Springer, 189–216.

[27] Katayama, M. 1998. *Optimality Theory and Japanese loanword phonology*. Ph.D. Dissertation, University of California, Santa Cruz.

[28] Bortolini, U., Zampolli, A. 1979. Frequenza e distribuzione dei gruppi consonantici nella lingua italiana [Frequency and distribution of the consonantal groups in Italian:. *Acta Phoniatrica* 1, 195-208.

[29] Campbell, N. 1999. A study of Japanese speech timing from the syllable perspective. *Journal of the Phonetic Society of Japan* 3(2), 29–39.

[30] Kawahara, S. 2013. Emphatic gemination in Japanese mimetic words: a wug-test with auditory stimuli. *Language Sciences* 40(2), 24–35.

[31] Kawahara, S., Pangilinan, M. 2017. Spectral continuity, amplitude changes, and perception of length contrasts. In: Kubozono, H. (ed.), *The Phonetics and Phonology of Geminate Consonants*, Oxford: Oxford University Press, 13–33.

[32] Oba, R., Braun, A., Handke, J. 2009. The perception of Japanese geminates by native and non-native listeners. *ISPhS* 2009, 9–29.

[33] Kawahara, S. 2007. Sonorancy and geminacy. *UMass Occasional Papers in Linguistics 32: Papers in Optimality Theory III*, 145–186.

[34] Fukui, S. 1978. Nihongo heisaon-no enchō/tanshuku-niyoru sokuon/hisokuon toshite-no chōshu [Perception for the Japanese stop consonants with reduced and extended durations]. *Onsei Gakkai Kaihō [The Bulletin, The Phonetic Society of Japan]* 159, 9–12.

[35] Hirata, Y. 2007. Durational variability and invariance in Japanese stop quantity distinction: Roles of adjacent vowels. *J of the Phonetic Society of Japan* 11(1), 9–22.

[36] Takeyasu, H., Giriko, M. 2017. Effects of duration and phonological length of the preceding/following segments on perception of the length contrast in Japanese. In: Kubozono, H. (ed.), *The Phonetics and Phonology of Geminate Consonants*, Oxford: Oxford University Press, 85-117.

[37] Idemaru, K., Guion-Anderson, S. 2010. Relational timing in the production and perception of Japanese singleton and geminate stops. *Phonetica* 67, 25–46.

[38] Ota, M., Ladd, D. R., M., Tsuchiya, M. 2003. Effects of foot structure on mora duration in Japanese? *Proc. of the 15th ICPhS*, 459-462.