World Academy of Science, Engineering and Technology
International Journal of Cognitive and Language Sciences
Vol:12, No:9, 2018

# The Impact of Dialectal Differences on the Perception of Japanese Gemination: A Case Study of Cantonese Learners

Honghao Ren, Mariko Kondo



***Abstract*—**This study investigates the perceptual features of Japanese obstruent geminates among Chinese learners of Japanese, focusing on the dialectal effect of the checked-tone, a syllable that ends in a stop consonant or a glottal stop, which is similar to Japanese obstruent geminates phonetically. In this study, 41 native speakers of Cantonese are divided into two groups based on their proficiency as well as learning period of Japanese. All stimuli employed in this study are made into C[p,k,s]+V[a,e,i] structure such as /apa/, /eke/, /isi/. Both original sounds and synthesized sounds are used in three different parts of this study. The results of the present study show that the checked-tone does have the positive effect on the perception of Japanese gemination. Furthermore, the proportion of closure duration in the entire word would be a more reliable and appropriate criterion in testing this kind of task.

***Keywords*—**Dialectal differences, Cantonese learners of Japanese, acoustic experiment, closure duration.

## I. Introduction

JAPANESE is known as a mora-timed language, while Chinese is a typical syllable-timed language. This makes Chinese learners of Japanese extremely hard to identify the difference between singleton and geminate stops. This kind of difficulty appears in both production and perception parts which could somehow cause unexpected misunderstanding. For example, without geminate, "*Kitte ku da sa i*" (Please give me some stamps) will become "*Kite ku da sa i*" (Please come), which shows completely different meanings. Because the speech signal varies widely in the perception of a phoneme [1], it is more complex to discriminate singleton and geminate stops accurately.

Many scholars have explained the reasons which will bring about this difficulty for non-native speakers on their perceptual performance of Japanese geminates. Most of them agree that the non-existence of the lengthened consonant stops or "holds" should be considered as a main contributing factor [2], [3]. In most Mandarin dialects, there is no similar cluster to Japanese gemination, but in some southern dialects, syllables ending in voiceless stops, which are called "checked tones" still remain, such as Yue dialect, Min dialect, etc.

Wang [4] has argued that southern dialect speakers have more phonetically advantages than Mandarin speakers when they come with the special mora in learning Japanese. She has analyzed the phonetic features of some southern dialects from the aspect of inter-language study. Also, previous studies [5], [6] found that Cantonese learners of Japanese are adopting checked-tone as a possible cue in both production and perception of gemination. In [7], Zhang et al. have conducted a production experiment. They took Cantonese speakers as experimental group and Japanese native speakers as a control group. This study demonstrated that there is no significant difference between these two groups, meaning speakers those who have check tones in their dialects can acquire Japanese gemination more easily.

Besides dialectal differences, the majority of previous studies on Japanese gemination placed emphasis on timing factors, such as the length of neighboring segments or closure duration, etc. Nonetheless, absolute length of a syllable or a word is not a proper criterion while the speech rate changes. Idemaru & Anderson have attempted to apply relational timing in the production and perception of Japanese gemination. They found that relational timing is more stable across speaking rates to accurately classify singleton and geminate productions [8]. Also, the variation of sonority scale can be used as a phonological distinction between segments. Previous studies have already proved that sonority scale, as well as voice quality, i.e. fundamental frequency, intensity (dB) have fairly strong categorization power [9]-[11].

There are five vowels with different sonority scales in Japanese language (/a/ > /o/ > /e/ > /u/ > /i/). However, very few studies have employed stimuli with different vowels to look into the differences between them. In addition, previous studies about Japanese gemination acquisition of L2 learners pay more attention to high-skilled learners. In an earlier study [12], Ren investigated that intermediate learners have the most similar perception combining with Japanese native speakers, instead of advanced learners.

TABLE I
Overall Description of Participants in the Final Experiment 1

| Group | Gender | Age (Average) | JLPT Level | Learning Period (Average) |
|---|---|---|---|---|
| Beginner | Male: 1<br>Female: 20 | 18-20<br>(19.4) | N3: 1<br>Not have: 20 | 11-48 months<br>(14.1) |
| Advanced | Male: 6<br>Female: 14 | 19-23<br>(21.3) | N1: 20 | 24-42 months<br>(35.8) |

Based on these, the current study aims to examine the impact of dialectal differences on the perception of Japanese gemination among native Cantonese speakers. Therefore, this

Honghao Ren is a PhD student in linguistics at Waseda University, 1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, Japan; e-mail: haoer21262@fuji.waseda.jp).

Mariko Kondo is a professor in phonetics and speech communication at Waseda University, Japan (e-mail: mkondo@waseda.jp).

study will also explore whether there is a developmental change between learners with different Japanese proficiency.

## II. Methods

### A. Experiment 1

#### 1. Participants

Forty-one native speakers of Cantonese who were learning Japanese as their main major at Guangdong University of Technology, China, participated. They were compensated for their participation and self-reported that they have no symptoms of hearing disorder. All participants were parted into 2 groups basing on their JLPT level, learning period and the evaluation of their teachers (as shown in Table I).

#### 2. Materials

All materials involved 18 words (as shown in Table II). All of the test stimuli began and ended with /a/, /e/, /i/ were relatively divided into 3 groups, meaning each group contains 6 stimuli.

TABLE II
Test Word List Used in Experiments

| | Bilabial | | Alveolar | |
|---|---|---|---|---|
| | singleton | geminate | singleton | geminate |
| Plosive | *apa* | *aQpa* | *aka* | *aQka* |
| | *epe* | *eQpe* | *eke* | *eQke* |
| | *ipi* | *iQpi* | *iki* | *iQki* |
| Fricative | - | - | *asa* | *aQsa* |
| | - | - | *ese* | *eQse* |
| | - | - | *isi* | *iQsi* |

The medial consonants were stop /k/, /p/ which exist in Cantonese check tones and fricative /s/ which does not. All test stimuli were minimal pairs (e.g. *aka-akka, isi-issi*). Over half of these were meaningful words in spite of some were low pedagogical frequency words; for example, the word "*epe*" represents an event called epee in fencing.

The recordings were made in a sound-proof studio at Waseda University, Japan. Two native speakers of Japanese (1 male, 1 female) with Tokyo dialect recorded all these words isolatedly 3 times at natural tempo. The accent pattern was all limited as HL (singleton) or HLL (geminate). Based on these, a total of 108 tokens (18 words ×3 times × 2 speakers) were employed in experiment 1. Because we concerned the correction of original sounds, the duration of each stimulus was used as recorded. A 2msec-noiseless sound was added before and after each stimulus.

#### 3. Procedure

The collection of the data was carried out in a multi-media room at Guangdong University of Technology. Stimuli were presented randomly to all subjects via *Praat*. The subjects were asked to hear the test stimuli using noise-cancelling headphones and finish a two-alternative forced-choice; for example, they had to choose "いし" when they heard "*isi*", or choose "いっし" when they heard "*issi*".

Participants were required to take a break after finishing every 36 responses and finish all the experiment at their own speed and comfort. Subjects can replay every stimulus using "*Replay*" button and return to the former stimulus with "*Modoru* (back)" button. Each stimulus could be played 3 times at most.

### B. Experiment 2

#### 1. Participants

Same participants as experiment 1 were asked to take part in experiment 2 on another day and do these two parts separately.

#### 2. Materials

The test words we used in experiment 2 are completely as same as the former part (as shown in Table III). A female native speaker of Japanese (Tokyo dialect) did the database sound recordings. All the test words were produced in isolation with natural speech rate for six times at H(L)L accent. After all recordings were finished, spectrograms of all 108 data items (18 words ×6 times) were checked via *Praat*; simultaneously, preceding and following vowel of all test items were measured in millisecond. (as shown in Table III). From the table, the average length of preceding/following vowel is 102/153 msec. Then, waveforms of all test items were checked to seek out the most proper sounds for each vowel type and medial consonant type to synthesize.

We expanded or contracted length of preceding/following vowel of each selected sound to100/150 msec. Afterwards, the ratio of closure duration's length (*CDL*) to entire word's length (*EWL*) was processed from 15% to 60% (10 graduations) for each word (e.g. /*eke*/, as shown in Fig. 1). Also, we added the same 2msec-noiseless sound before and after each stimulus and each synthesized sound will appear randomly twice. Consequently, a total of 180 tokens (9-word types ×2 times ×10 graduations) were employed in experiment 2.

TABLE III
The Length of Preceding & Following Vowels (msec)

| Vowel | Preceding (max/min) | Following (max/min) | Average Length (Preceding/Following) |
|---|---|---|---|
| /a/ | 122 /102 | 167/ 137 | 104/ 157 |
| /e/ | 109/ 98 | 159/ 131 | 103/ 154 |
| /i/ | 110/ 93 | 156/ 138 | 98/ 147 |

#### 3. Procedure

Data collection was actualized at the same room as experiment 1. Participants were asked to hear each of the test stimuli three times maximum and choose the best candidate for each stimulus on the screen. They had to take a break after finishing every 30 responses. All the data collection was carried out by using *Praat*.

## III. Results

### 1. Experiment 1

Fig. 2 shows the perceptual identification accuracy of singleton/geminate for subjects of two Japanese proficiencies (beginner & advanced).

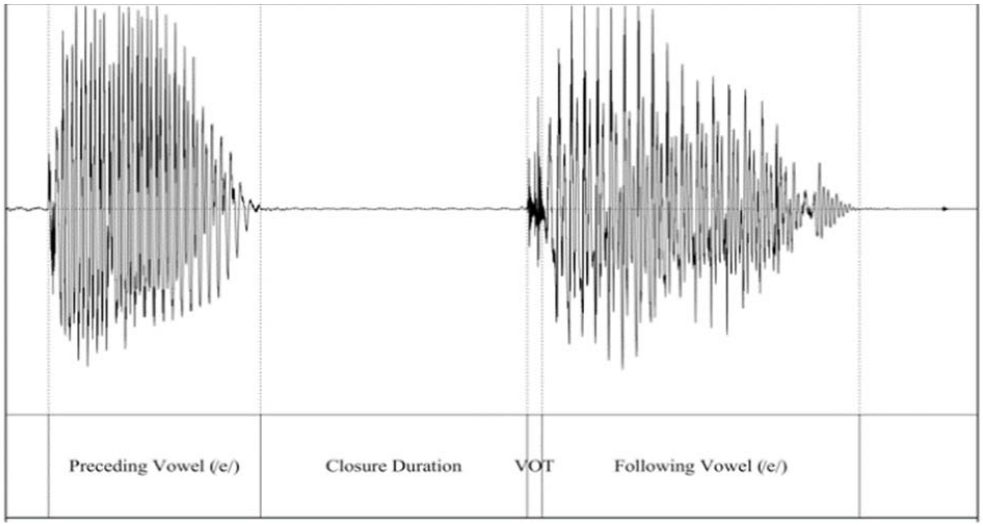


Fig. 1 Spectrogram of synthesized stimulus /eke/ without noiseless sound (CDL/EWL=35%)

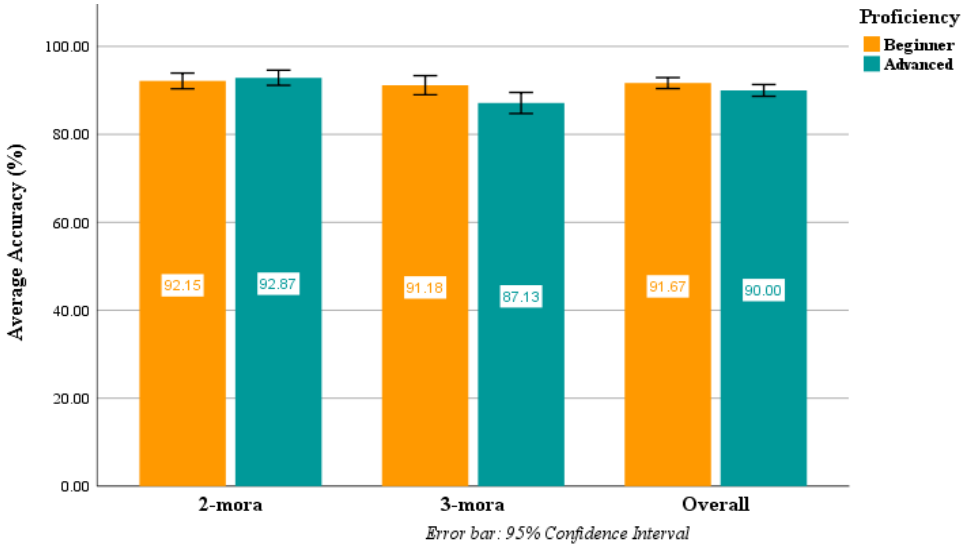


Fig. 2 Perceptual identification accuracy of singleton/geminate

As Fig. 2 shows, the average perceptual accuracy of each group was above 85%. Welch's Two Sample t-test on both singleton and geminate perception were preformed, showing significant difference between two learner groups only in 3-mora word (geminate) case. (Singleton: $t$ (366.918) =.424, $p$=.515 >.05, *n.s.*; Geminate: $t$ (316.91) = 2.6311, $p$=.014 < .05; Overall: $t$ (729.838) =2.097, $p$=.148 >.05, *n.s.* [13]). While the differences in Japanese proficiency exert a significant effect on the perception of geminate consonants, different consonant types (/k/, /p/, /s/) also had statistical differences when conducted repeated measures ANOVA [13] (as shown in Table IV). This result indicates that consonantal difference is a relevant factor for native Cantonese learners in perceiving Japanese gemination/non-gemination. The results are summarized in Fig. 2.

Repeated measures ANOVA did not show any significant interaction between vowel differences and perceptual identification accuracy in the two learner groups regardless of singleton or geminate.

TABLE IV
RESULTS OF REPEATED MEASURES ANOVA (EXPERIMENT 1)

| Group | Singleton | | | Geminate | | |
|---|---|---|---|---|---|---|
| | Df | *F* value | *p* | Df | *F* value | *p* |
| Beginner | 2,186 | 6.793 | .001*** | 2,186 | 18.439 | 4.95e-8*** |
| Advanced | 2,177 | 14.884 | .000*** | 2,177 | 15.57 | 5.92e-7*** |

## 2. Experiment 2

The responses by each participant were analyzed in terms of the rate of identification with singleton/geminate. For example, if one heard a stimulus as "aka (singleton)", the identification rate with singleton will be 100%, while if one judged it as "*akka* (geminate)", the rate with singleton will be coded as 0%. The average rates for each stimulus were visualize as Fig. 4. The number on the vertical-axis indicates the average percentage

among all participants who judged it as geminate (2 mora AVE) or non-geminate (3 mora AVE).

Meanwhile, boundary point (BP) was computed for each stimulus set responded by each participant. Boundary point shows the point of division, or the 50% perception between identification with a singleton or a geminate. The 50 % boundary point was calculated using logistic regression, which is an appropriate way to find boundary point of this type within the generalized linear model.

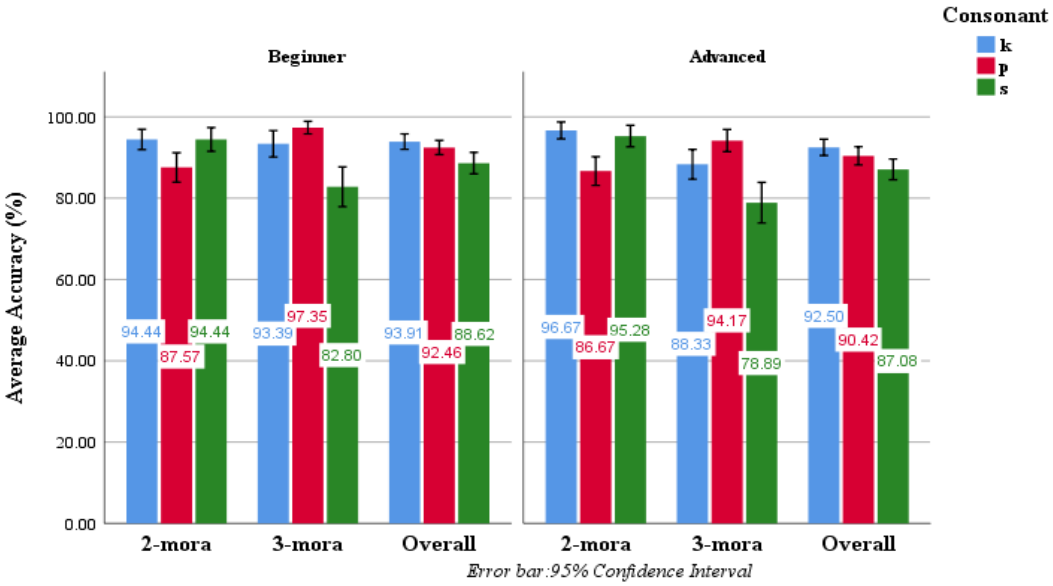


Fig. 3 Perceptual identification accuracy with different medial consonants /k(k)/, /p(p)/, /s(s)/

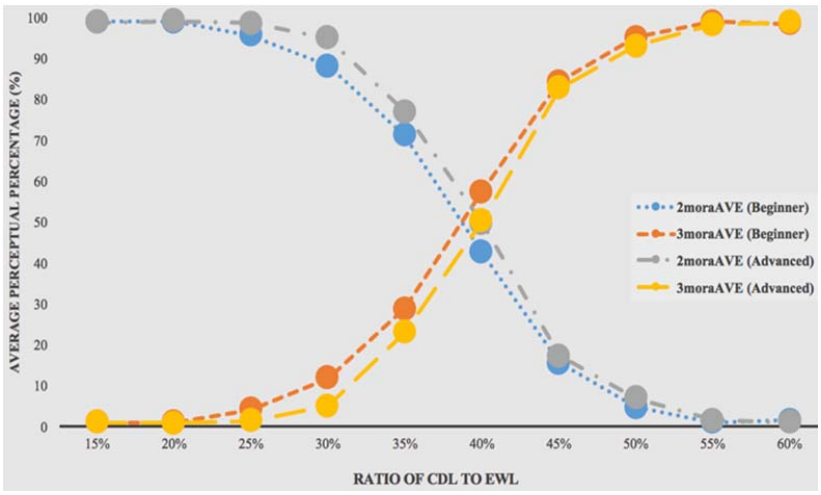


Fig. 4 The average perceptual percentage of identification test

TABLE V
AVERAGE BOUNDARY POINTS OF SINGLETON/GEMINATE CONTRAST (RATIO OF CDL TO EWL & MSEC)

| Group | Vowel | | | Consonant | | |
|---|---|---|---|---|---|---|
| | a | e | i | k | p | s |
| Beginner | 37.70% (151.28) | 37.35% (149.04) | 38.75% (158.16) | 35.95% (140.32) | 37.20% (148.09) | 41.45% (176.99) |
| Advanced | 38.24% (154.86) | 39.10% (160.51) | 40.10% (167.36) | 42.25% (182.90) | 39.00% (159.84) | 41.30% (175.89) |

Welch's Two Sample t-test on 50% boundary point were conducted again, showing significant difference between two learner groups ($t$ (366.27) = -2.2781, $p$ < .05 [13]). In order to know the interaction between vowel difference and perceptual BP, repeated measures ANOVA was also performed, showing statistical difference in the advanced-group case ($F$ (2,177) =3.77, $p$ < .05 [13]). However, no significant difference was shown in beginner group ($F$ (2,186) =0.658, $p$ > .1 [13]).

Finally, statistical analysis (ANOVA) was conducted on the BP of these two learner groups categorized by different medial consonants (*/k/, /p/, /s/*). Both beginner group ($F$ (2,177) =3.77, p <.05 [13]) and advanced group ($F$ (2,177) =19.51, p <.001 [13]) showed significant difference. Boundary points of singleton/geminate perception of each group are shown in Table V.

## IV. Conclusion

By comparing the perceptual identification accuracy, hypothesis that the existence of checked tones (/k/, /p/) in Yue dialect (Cantonese) do show positive transfer when they acquire Japanese gemination. The experimental data from this study also suggest that native Cantonese clearly differentiated singleton/geminate contrasts in ratio of length of closure duration to entire word length. However, vowel difference shows significant effect on perceptual boundary point only in advanced group of Cantonese learners of Japanese. Because we only used front vowels /a/, /e/, /i/ in this study, the question that whether voice quality will differentiate the perception of Japanese gemination still remains.

## Acknowledgment

This research was supported by the program of China Scholarships Council (No.201608210193).

## References

[1] S. Amano, Y. Hirata, "Perception and Production of Singleton and Geminate Stops in Japanese: Implications for the Theory of Acoustic Invariance," *Phonetica*, vol. 72, pp. 43–60, 2015.

[2] M. S. Han, "The timing control of geminate and single stop consonants in Japanese: a challenge for nonnative speakers," *Phonetica*, vol. 49, pp. 102- 127, 1992.

[3] T. Toda, *Second Language Speech Perception and Production: Acquisition of Phonological Contrasts in Japanese*. Lanham, MD: University Press of America.

[4] N. Wang, "Chinese Dialects Which Encourage Acquisition of Japanese Phonetic Recognition for Chinese-speaking Learners," *Journal of the Phonetic Society of Japan*, vol. 3 No.3, pp. 36- 42, 1999.

[5] W. Zhang, "The perception and production of Japanese gemination among Cantonese Learners: Focusing on the impact of checked-tone", *Waseda University*, 2011.

[6] X. Li, "Articulation and Perception of Geminate Consonants by Cantonese Learners of Japanese," *Studies in Japanese Language and Culture. Osaka University,* vol. 33, pp. 69-80, 2014.

[7] L. Zhang, Y. Liu, Y. Shi, "For Japanese Learners Whose Native Language Has Entering Tone, Is It Easier to Acquire Double Consonants: A Case Study of Native Cantonese Speakers," *The Society of Humanities. Osaka Prefecture University*, vol. 33, pp. 191-201, 2015.

[8] K. Idemaru, S.G. Anderson, "Relational Timing in the Production and Perception of Japanese Singleton and Geminate Stops," *Phonetica*, vol. 67, pp. 25-46, 2010.

[9] K. Idemaru, S. G. Anderson, "Acoustic covariants of length contrast in Japanese stops," *Journal of the International Phonetic Association*, vol. 38 No.2, pp. 167–186, 2008.

[10] M. Sadakata, M. Shingai, S. Sulpizio, A. Brandmeyer, K. Sekiyama, "Language specific listening of Japanese geminate consonants: a cross-linguistic study," *Front Psychol,* vol.5, pp.1422, 2014.

[11] D. M. Hardison, M. M. Saigo, "Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy," *Applied Psycholinguistics*, vol.32, pp. 81-99, 2010.

[12] H. Ren, "An experimental study of the influence of constriction duration on the perception of Japanese gemination among Chinese learners," *Journal of Japanese Language Study and Research*, vol.1, pp. 55-63, 2017.

[13] R version 3.4.2: https://cran.r-project.org.