



# Perception of consonant length in familiar and unfamiliar languages by native speakers of Mandarin, Italian and Japanese

Kimiko Tsukada<sup>1</sup>, John Hajek<sup>2</sup>

<sup>1</sup>Macquarie University/The University of Melbourne

<sup>2</sup>The University of Melbourne

kimiko.tsukada@gmail.com, j.hajek@unimelb.edu.au

## Abstract

This study builds on our previous research and provides additional analyses to determine if there is a relationship between the ability to process consonant length in familiar and unfamiliar languages for learners of Japanese whose native language is Italian or Mandarin. The emphasis is on L2-L3 (second-third language) phonetic influence. Japanese and Italian use consonant length contrastively, but not Mandarin. We thus asked if Mandarin learners with higher proficiency in Japanese are more or less accurate in length identification than Italian learners with first language (L1) experience of consonant length. Specifically, we focused on finding out if learners who accurately identify Japanese consonant length might also be accurate in their identification of the length category in Italian. Four groups of listeners differing in their L1 (Italian x 2 groups, Japanese, Mandarin) and experience with consonant length participated in forced-choice identification experiments. L1 Japanese and L1 Italian listeners identified the length category more accurately in their L1 than in the foreign language (FL). The ability to identify consonant length in Japanese and Italian by 18 advanced Mandarin-speaking learners of Japanese seemed unrelated. This suggests that speech processing skills acquired in one FL may not automatically transfer to another FL.

**Index Terms:** consonant length, Japanese, Italian, Mandarin, L2-L3 phonetic influence

## 1. Introduction

With increasing globalization, opportunities to communicate in non-native languages also increase. Efficiency in spoken language processing is crucial in face-to-face communication. However, for adults who have already gained mastery in their first language (L1), becoming fluent in subsequent languages is not easy. Multilingualism may be a reality in many societies in a rapidly globalizing world, but our current understanding of how individuals process typologically unrelated multiple languages remains insufficient [1], although research interest is increasing over time. Our aim is to contribute to the advancement of not only L1 but also foreign language (FL) transfer effects in cross-language speech processing and FL learning by comparing four groups of participants differing in their L1 and experience with consonant length contrasts.

Languages such as Japanese and Italian use consonant length (i.e. singleton vs geminate) contrastively. In Japanese, *ika* and *ikka* mean ‘below’ and ‘lesson one’ while, in Italian, *eco* and *ecco* mean ‘echo’ and ‘here (it is)’, respectively. This is an area of pronunciation that is known to pose difficulties to non-native learners [2-4]. Research on FL speech learning

focusing on non-segmental features such as length is increasing [5-7], but still relatively limited, in particular, for target languages other than English [see, however, 8, 9 for length in Estonian and Finnish].

In our previous research involving non-native learners of Japanese from Italian and Mandarin backgrounds [10, 11], we reported that L1 experience with length contrasts was more helpful than Japanese learning experience for consonant length identification, but that there is an additional benefit of Japanese learning for L1 Italian listeners, for whom positive transfer may be assumed. The current study builds on the earlier work and explores the question of general phonetic talent, i.e. whether there is a relationship between Mandarin-speaking learners’ consonant length perception in a familiar FL, Japanese, and an unfamiliar FL, Italian. If a positive correlation (or L2-L3 (second-third language) phonetic influence) is found in learners’ abilities to process speech sounds in two unrelated FLs, it has implications for FL pronunciation pedagogy and multilingualism.

## 2. Methods

### 2.1. Stimuli preparation

#### 2.1.1. Speakers

Seven (4 males, 3 females) native speakers of Japanese and three (2 males, 1 female) native speakers of Italian in their 20-60s participated in the recording sessions lasting between 45 and 60 minutes. All Japanese speakers spoke standard Japanese, having been born or having spent most of their life in the Kanto region surrounding the Greater Tokyo Area. The three Italian speakers originally came from different parts of Italy (including Rome and Sicily) and were long-term residents of Sydney, Australia. All of them were highly experienced university teachers of Italian with a clear understanding of pronunciation norms. The authors with expertise in phonetics/phonology of these target languages auditorily confirmed that all the speakers clearly differentiated the singleton and geminate consonants by duration. The speakers were recorded in a recording studio at a university in Sydney, Australia or at a research institute in Tokyo, Japan. They received \$20 (or equivalent in Japanese yen) for their participation. None of these speakers participated in the perception experiments. According to self-report, they had normal hearing.

#### 2.1.2. Speech materials

Tables 1 and 2 show examples of Japanese and Italian (non)words analyzed in this study. The Japanese items included (C)VC(C)V (non)word tokens where the underlined medial C(C) was /t k ʃ/. The Italian items included (C)VC(C)V

(non)word tokens where the underlined medial C(C) was /p t k b d g dʒ f s v/. Italian uses length contrasts for a wide variety of phonemes differing in voicing, place and manner of articulation [5, 12-15]. In Japanese, on the other hand, voiced geminates are disfavoured and their occurrence is limited mostly to loanwords [16, 17].

Table 1: Examples of Japanese test words used.

Manner	Singleton	Geminate
stop	<i>kato</i> ‘transition’	<i>katto</i> ‘cut’
stop	<i>haka</i> ‘a grave’	<i>hakka</i> ‘mint’
affricate	<i>ichi</i> ‘one’	<i>icchi</i> ‘agreement’

Table 2: Examples of Italian test words used.

Manner	Singleton	Geminate
stop	<i>sete</i> ‘thirst’	<i>sette</i> ‘seven’
fricative	<i>casa</i> ‘a house’	<i>cassa</i> ‘a case’
affricate	<i>agio</i> ‘ease’	<i>aggio</i> ‘premium’

To record the stimuli to be used in the perception study, each word was presented on a computer screen in random order and was produced in two separate conditions: one in isolation and the other in a carrier sentence (/diko X di nwovo/ ‘I say X again’ for Italian and /sokowa X to jomimasu/ ‘You read it as X there’ for Japanese). The pace of presentation was controlled by the experimenter (the first author). The speech materials were digitally recorded at a sampling rate of 44.1 kHz and the target words were segmented and stored in separate files. To avoid inter-speaker variation in fluency (specifically, the duration of a pause before and after the target (non)word), only tokens produced in isolation were used as experimental stimuli in this study.

## 2.2. Stimuli characteristics

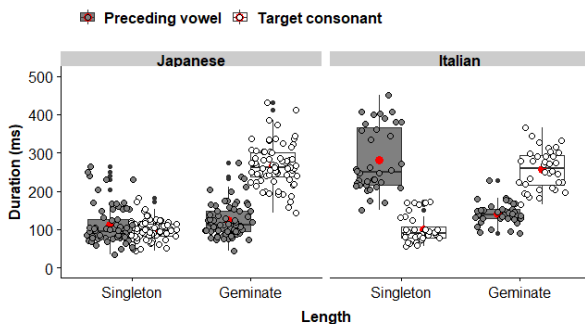


Figure 1: The durational distributions of Japanese and Italian stimuli as a function of consonant length category. The bold horizontal line and the red circle in each box indicate the median and mean, respectively. The bottom and top lines of the box indicate the first and third quartiles. The smaller black points outside the box are outliers.

Figure 1 shows the durational distributions of the stimuli. On average, geminates were more than 2.5 times as long as singletons in both languages (Japanese: 99 ms ( $sd = 26.2$  ms) vs 267 ms ( $sd = 55.2$  ms), Italian: 101 ms ( $sd = 34.4$  ms) vs 257 ms ( $sd = 51.4$  ms)). The mean duration of the Japanese vowels preceding geminates was 123 ms ( $sd = 43.4$  ms) and the mean duration of the Japanese vowels preceding singletons was 111 ms ( $sd = 50.5$  ms). The difference in vowel duration in the two contexts was not statistically significant. The mean duration of the Italian vowels preceding geminates was 142 ms ( $sd = 28.0$  ms) and the mean duration of the Italian vowels preceding

singletons was 281 ms ( $sd = 84.1$  ms). Unlike Japanese, there is a substantial difference of 139 ms in vowel duration in Italian according to the length category of the following consonant. This finding confirms durational compensation reported in previous work [13, 15].

## 2.3. Participants

Four groups of listeners participated in forced-choice consonant length identification experiments. The first two groups consisted of L1 Italian (NNJ-Italian, 7 males, 7 females, *mean age* = 23.4, *sd* = 1.6) and L1 Mandarin (NNJ-Mandarin, 3 males, 15 females, *mean age* = 24.7, *sd* = 4.5) learners of Japanese. Five (1 male, 4 females) of the NNJ-Mandarin learners were originally from Taiwan. The NNJ-Mandarin group took part in the study in Japan and had a mean length of residence of 1.3 (*range* = 0.1-5.0) years in Japan.

Two of the NNJ-Italian learners had passed the JLPT (Japanese Language Proficiency Test, according to which, the easiest level is N5 and the most difficult level is N1) at N2 level, six had passed N3 and one had passed N4, respectively. On the other hand, all the NNJ-Mandarin learners had passed N1 and were considered more advanced than the NNJ-Italian learners. While JLPT has a listening element, we acknowledge that the NNJ-Mandarin learners’ knowledge of Chinese characters (*Kanji*) may have substantially contributed to their passing N1. Crucially, however, only Italian, unlike Mandarin, uses consonant length contrastively. The other two groups consisted of L1 Italian (NI, 7 males, 7 females, *mean age* = 28.9, *sd* = 7.6) and L1 Japanese (NJ, 5 males, 5 females, *mean age* = 25.5, *sd* = 6.9) listeners who served as controls. All the NI or NJ listeners were naïve to the other target language. Except for three NI listeners who took part in the study in Sydney, Australia, the two (NNJ-Italian, NI) Italian groups consisted of students at a university in Turin, Italy. The NJ listeners were recruited from the student/staff populations at universities or from the local communities in Australia or Japan.

## 2.4. Procedures

The listeners participated in a forced-choice identification task and listened to a total of 252 ( $84 \times 3$  blocks) Japanese tokens followed by one block of 84 Italian tokens. For the present study, a subset of these tokens (144 tokens for Japanese and 72 tokens for Italian) with intervocalic obstruents only was included in the analyses.

The listeners’ task was to decide whether the medial consonant was short/singleton or long/geminate and indicate their choice on the computer. The listeners were allowed to replay the stimulus tokens multiple times in order to reduce their anxiety and were asked to guess if uncertain. Once a choice was made, the listeners were unable to change their decision and the next token was presented automatically. The listeners were tested individually in a sound-attenuated booth or quiet classroom on the university campus in their country of residence. The experimental session was self-paced, but typically lasted between 30 and 40 minutes. They heard the stimuli at a self-selected, comfortable amplitude level over the headphones on a computer and received \$20 (or equivalent) for their participation.

## 3. Results and discussion

We used R version 3.6.0 for statistical analyses and data visualization reported below [18].

### 3.1. Overall results

Figure 2 shows the distributions of percentages of correct identification for the Japanese and Italian stimuli by the four groups of listeners. The L1 Japanese and L1 Italian groups were at ceiling and more accurate in their L1 than FLs. For the Japanese stimuli, the NJ group ( $mean = 99\%$ ) was more accurate than all three non-native groups. However, the NNJ-Italian group ( $mean = 96\%$ ) scored higher than the other two groups ( $mean = 90\%$  for NI vs  $mean = 86\%$  for NNJ-Mandarin). For the Italian stimuli, NI ( $mean = 99\%$ ) and NNJ-Italian ( $mean = 99\%$ ) groups were equally accurate and scored higher than the two non-native groups ( $mean = 96\%$  for NJ vs  $mean = 88\%$  for NNJ-Mandarin).

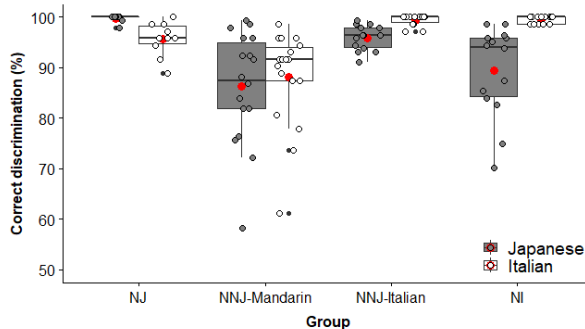


Figure 2: The distributions of correct identification scores (%) for the Japanese and Italian stimuli by four groups of listeners (NJ, NNJ-Mandarin, NNJ-Italian, NI).

A two-way repeated-measures analysis of variance (ANOVA) with group (G: NJ, NNJ-Mandarin, NNJ-Italian, NI) as a between-subjects factor and stimulus language (L: Japanese, Italian) as a within-subjects factor yielded significant main effects of G and L and a significant two-way interaction effect. Table 3 shows the results of a two-way ANOVA. A significant two-way interaction is expected, because L1 Japanese and L1 Italian listeners were more accurate in their own L1 than in the unfamiliar FL as clearly seen in Figure 2.

Table 3: Results of Group x Language ANOVA.

Factor	df	F-value	p-value	$\eta_G^2$
G	3, 52	10.9	< .001	.32
L	1, 52	13.4	< .001	.06
G x L	3, 52	9.0	< .001	.11

Table 4: Results of one-way ANOVA assessing the effects of Group and multiple comparison tests (significance level at .05).

Language	df	F	p	Between-group comparisons
Japanese	3, 25.2	22.7	< .001	NJ > NNJ-Italian, NI, NNJ-Mandarin; NNJ-Italian > NNJ-Mandarin
Italian	3, 23.7	11.9	< .001	NI, NNJ-Italian > NJ > NNJ-Mandarin

The simple effect of Language was significant for both Japanese and Italian. Table 4 shows the results of one-way ANOVA which assessed the effect of Group (not assuming equal variances) and of Dunnett's Modified Tukey-Kramer pairwise multiple comparison *post hoc* tests.

For the Japanese stimuli, the NJ group was significantly more accurate than all three non-native groups as expected. However, it is notable that the NNJ-Italian learners identified the consonant length category more accurately than did the NNJ-Mandarin learners who were more advanced in Japanese proficiency. The NI listeners who were naïve to Japanese and the NNJ-Mandarin learners did not differ from each other. Thus, it seems a combination of L1 Italian and Japanese learning experience is necessary for attaining more accurate consonant length identification. It needs to be pointed out, however, that the NNJ-Mandarin learners in this study were highly advanced in Japanese (at JLPT N1). If learners with less Japanese experience had participated in the study, the NI listeners may have outperformed the NNJ-Mandarin group.

For the Italian stimuli, both groups of L1 listeners, who were equally accurate, identified consonant length more accurately than did the two groups of non-native listeners. Despite the absence of experience with Italian, the NJ group outperformed the NNJ-Mandarin group. This suggests that although the NJ listeners can use their L1 knowledge of Japanese length distinctions to perceive novel Italian length distinctions, the NNJ-Mandarin learners cannot use their FL knowledge of Japanese length distinctions in the same way, at least not as effectively.

### 3.2. Consonant length identification in Japanese

Figure 3 shows the distributions of percentages of correct identification for the Japanese stimuli by the four groups of listeners as a function of the length category of the target obstruent. While the NNJ-Mandarin group had higher mean identification accuracy for the geminate (90%) than for the singleton (83%) obstruents, the other three groups were relatively balanced in their mean identification accuracy for the two length categories (NJ: 99% vs 100%, NNJ-Italian: 95% vs 97%, NI: 89% vs 91% for geminate and singleton, respectively).

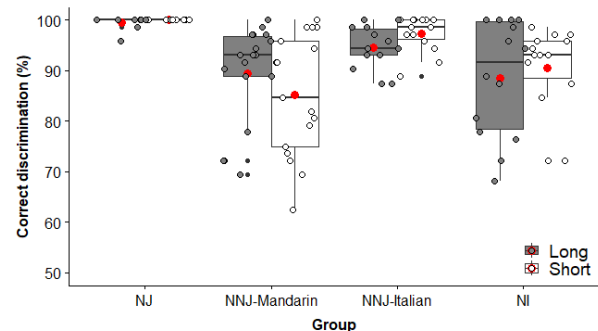


Figure 3: The distributions of correct identification scores (%) for the Japanese stimuli differing in consonant length by four groups of listeners (NJ, NNJ-Mandarin, NNJ-Italian, NI).

Table 5: Results of Group x Length ANOVA.

Factor	df	F-value	p-value	$\eta_G^2$
G	3, 52	8.2	< .001	.26
L	1, 52	0.4	ns	.00
G x L	3, 52	3.6	< .05	.05

A two-way repeated-measures ANOVA with group (G: NJ, NNJ-Mandarin, NNJ-Italian, NI) as a between-subjects factor and length (L: long, short) as a within-subjects factor yielded a significant main effect of G and two-way interaction effect (Table 5). The simple effect of Group was significant for both

length categories. Table 6 shows the results of one-way ANOVA which assessed the effect of Group (not assuming equal variances) and of Dunnett's Modified Tukey-Kramer pairwise multiple comparison *post hoc* tests.

For geminates, the NJ group outperformed the three non-native groups, who did not significantly differ from one another. For singletons, the NJ listeners never misperceived any tokens and could not be included in the analysis. Of the three non-native groups, the NNJ-Italian group identified Japanese singletons significantly more accurately than did the NI and NNJ-Mandarin groups, who did not differ from each other. Thus, while the NNJ-Italian learners diverged from the NJ listeners in identifying geminates, they were more native-like than the other groups in identifying singletons. Even with advanced knowledge of Japanese, the NNJ-Mandarin learners identified both singletons and geminates less accurately than did the NNJ-Italian learners (significantly so for singletons) and did not outperform the NI listeners who were naïve to Japanese.

Table 6: Results of one-way ANOVA assessing the effects of Group and multiple comparison tests (significance level at .05).

Length	df	F	p	Between-group comparisons
Geminate	3, 26.0	13.7	< .001	NJ > NNJ-Italian, NNJ-Mandarin, NI
Singleton	2, 23.7	10.4	< .001	NNJ-Italian > NI, NNJ-Mandarin

### 3.3. Consonant length identification in Italian

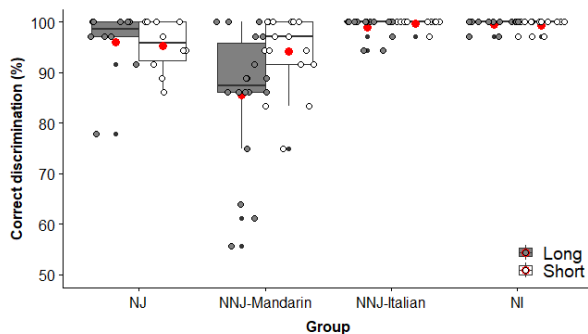


Figure 4: The distributions of correct identification scores (%) for the Italian stimuli differing in consonant length by four groups of listeners (NJ, NNJ-Mandarin, NNJ-Italian, NI).

Figure 4 shows the distributions of percentages of correct identification for the Italian stimuli by the four groups of listeners as a function of the length category of the target obstruent. The two groups of Italian listeners accurately identified L1 tokens for both geminates (*mean* = 99% for both NNJ-Italian and NI) and singletons (*mean* = 99% for both NNJ-Italian and NI). Interestingly, the NNJ-Mandarin group had higher mean identification accuracy for the singleton (91%) than for the geminate (86%) obstruents, which is the reverse pattern of results compared to their length identification in Japanese. The NJ listeners were relatively balanced in their mean length identification accuracy for both categories (96% for geminate vs 95% for singleton) as they were in Japanese.

A two-way repeated-measures ANOVA with group (G: NJ, NNJ-Mandarin, NNJ-Italian, NI) as a between-subjects factor and length (L: long, short) as a within-subjects factor reached

significance for the Group factor only [ $F(3, 52) = 14.8, p < .001, \eta_p^2 = 0.25$ ] (see 3.1. above).

### 3.4. Consonant length identification in Japanese and Italian by NNJ-Mandarin learners

Figure 5 shows the percentages of correct consonant length identification for the Japanese (x-axis) and Italian (y-axis) stimuli by 18 NNJ-Mandarin learners.

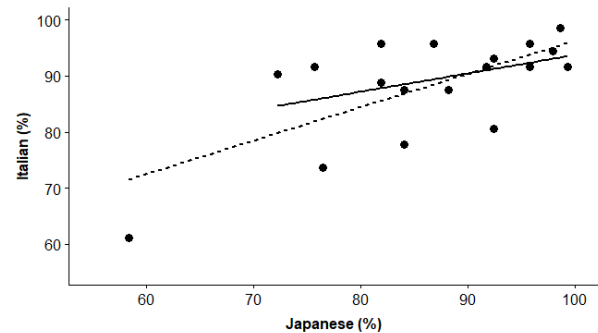


Figure 5: Correlation between correct identification in Japanese and Italian by NNJ-Mandarin learners with (solid line) and without (dotted line) an outlier participant.

While there appears to be a positive correlation between NNJ-Mandarin learners' perceptual accuracy for Japanese and Italian consonant length [ $\rho = .53, p < .05$ ], this was driven by one outlying learner whose accuracy was lowest for both Japanese (58%) and Italian (61%). When this learner's data was removed, the correlation became non-significant [ $\rho = .44, p < .08$ ]. Thus, as a group, there may be no relationship between the NNJ-Mandarin learners' consonant length perception in Japanese and Italian. Nevertheless, the top 5 learners who scored above 95% in Japanese also scored high (> 90%) in Italian. This suggests there is a subgroup of phonetically talented learners who are generally skilled in spoken language processing beyond their advanced proficiency in Japanese.

## 4. Conclusions

We observed that non-native listeners who have learned to efficiently process the consonant length contrasts in FL Japanese may not necessarily be skilled in the consonant length identification in an unfamiliar FL Italian. This, together with the finding that the NNJ-Mandarin learners showed an opposite pattern of results for Japanese and Italian, is consistent with our previous study [7] and suggests that some learners, according to their L1, may categorize length in the two quantity languages differently rather than uniformly applying the concept of  $[\pm\text{long}]$ . We tentatively conclude that while speech processing skills acquired in one's L1 may positively transfer to an unfamiliar FL as was the case with L1 Italian listeners, the phonetic skills acquired in an FL in adulthood may not automatically aid pronunciation learning of subsequent FLs as was the case L1 Mandarin listeners.

## 5. Acknowledgements

This research was supported by the 2018 Endeavour Research Fellowship to the first author. We thank Valentina De Iacovo for research assistance and participants for their co-operation. Finally, we thank three anonymous reviewers for their time and input.

## 6. References

- [1] I. L. Chan and C. B. Chang, "Perception of nonnative tonal contrasts by Mandarin-English and English-Mandarin sequential bilinguals," *The Journal of the Acoustical Society of America*, vol. 146, no. 2, pp. 956–972, 2019.
- [2] B. De Clercq, E. Simon and C. Crocco, "Rosa versus rossa: The acquisition of Italian geminates by native speakers of Dutch," *Phrasis: Studies in Language and Literature*, vol. 2, pp. 3–29, 2014.
- [3] M. S. Han, "The timing control of geminate and single stop consonants in Japanese: A challenge for non-native speakers," *Phonetica*, vol. 49, pp. 102–127, 1992.
- [4] H. Kubozono, "Introduction to the special issue on Japanese geminate obstruents," *Journal of East Asian Linguistics*, vol. 22, pp. 303–306, 2013.
- [5] H. Altmann, I. Berger and B. Braun, "Asymmetries in the perception of non-native consonantal and vocalic length contrasts," *Second Language Research*, vol. 28, pp. 387–413, 2012.
- [6] R. McAllister, J. E. Flege and T. Piske, "The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian," *Journal of Phonetics*, vol. 30, pp. 229–258, 2002.
- [7] K. Tsukada, F. Cox, J. Hajek and Y. Hirata, "Non-native Japanese learners' perception of consonant length in Japanese and Italian," *Second Language Research*, 34, vol 2, pp. 179–200, 2018.
- [8] L. Meister and E. Meister, "Perception of the short vs. long phonological category in Estonian by native and non-native listeners," *Journal of Phonetics*, vol. 39, pp. 212–224, 2011.
- [9] S. Ylinen, A. Shestakova, P. Alku and M. Huottilainen, "The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish," *Language and Speech*, vol. 48, pp. 313–338, 2005.
- [10] K. Tsukada and J. Hajek, "Cross-language perception of Italian and Japanese consonant length contrasts: A comparison of native Italian listeners with and without Japanese language learning experience," In S. Calhoun, P. Escudero, M. Tabain and P. Warren (eds.) *Proceedings of the 19<sup>th</sup> International Congress of Phonetic Sciences, Melbourne, Australia, 2019*, Canberra, Australia: Australasian Speech Science and Technology Association Inc., pp. 82–86.
- [11] K. Tsukada and J. Hajek, "Cross-language perception of Italian and Japanese length contrasts: A comparison of Italian and Mandarin-speaking learners of Japanese," *New Sounds 2019*, Tokyo, Japan.
- [12] A. Esposito and M. G. Di Benedetto, "Acoustical and perceptual study of gemination in Italian stops," *The Journal of the Acoustical Society of America*, vol. 106, pp. 2051–62, 1999.
- [13] E. M. Payne, "Phonetic variation in Italian consonant gemination," *Journal of the International Phonetic Association*, vol. 35, pp. 153–181, 2005.
- [14] D. Rogers and L. d'Arcangeli, "Italian," *Journal of the International Phonetic Association*, vol. 34, pp. 117–21, 2004.
- [15] J. Hajek, M. Stevens and G. Webster, "Vowel duration, compression and lengthening in stressed syllables in Italian," In J. Trouvain and W. J. Barry (eds) *Proceedings of the 16<sup>th</sup> International Congress of Phonetics Sciences (ICPhS XVI), Saarbrücken, Germany, 2007*, pp. 1057-1060.
- [16] Q. Hussein and S. Shinohara, "Partial devoicing of voiced geminate stops in Tokyo Japanese," *The Journal of the Acoustical Society of America*, vol. 145, pp. 149–163, 2019.
- [17] S. Kawahara, "The phonetics of *sokuon*, or geminate obstruents," In: H. Kubozono (ed), *Handbook of Japanese Phonetics and Phonology*. Berlin: Walter de Gruyter, pp. 43–78, 2015.
- [18] R Core Team, "R: A language and environment for statistical computing", R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> 2019.