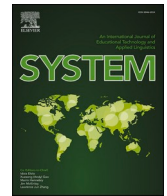




ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

System

journal homepage: www.elsevier.com/locate/system

Mandarin Chinese-speaking learners' acquisition of Italian consonant length contrast

Qiang Feng^{a,b,*}, M. Grazia Busà^b

^a Dalian University of Foreign Languages, Dalian, China

^b Università degli Studi di Padova, Padova, Italy

ARTICLE INFO

Keywords:

L2 speech acquisition
Consonant length contrast
Singletons
Geminates
Italian
Mandarin Chinese

ABSTRACT

The acquisition of Italian consonant length contrast by Mandarin Chinese-speaking learners (henceforth Chinese learners) has hardly been investigated. This study aims to fill this gap. Ten first-year, ten second-year and ten third-year Chinese undergraduate students majoring in Italian and ten native Italian-speaking controls took part in a perception and a production experiment. In the perception experiment, the participants had to identify ten Italian disyllabic minimal pairs contrasting in consonant length. In the production experiment, the participants were asked to read five out of the ten minimal pairs mentioned above; the duration values of the intervocalic consonants and the pre-consonantal vowels were measured and converted into duration ratios for statistical analyses. The results show that, in both perception and production, Chinese learners can distinguish between the two consonant length categories in Italian to a certain extent, but not in a native-like manner. Also, the duration interplay between Italian consonants and pre-consonantal vowels (i.e., longer vowels before singleton consonants and shorter vowels before geminate consonants) is entirely ignored by Chinese learners. Moreover, Chinese learners' increased learning experience does not appear to enhance their acquisition of Italian consonant length contrast. Some pedagogical implications are discussed.

1. Introduction

Acquiring L2 consonant length contrast can be quite challenging for Mandarin Chinese-speaking learners (henceforth Chinese learners), and in fact they have been defined as “quantity-insensitive” language speakers (Tsukada et al., 2014). However, investigations on Chinese learners' perception and/or production of L2 consonant quantity contrast have focused mainly on L2 Japanese (see §1.1 for a review), while the acquisition of consonant length contrast in L2 Italian is much less investigated. Since the actual implementation of consonant length contrast is rather language-specific and differs in Italian and Japanese to a considerable extent (Kawahara, 2015; Tanaka, 2017; Tsukada et al., 2018), by investigating the L1 Mandarin Chinese-L2 Italian language pair, this article aims to contribute to a better understanding of Chinese learners' acquisition mechanisms of L2 consonant quantity contrast.

1.1. Chinese learners' acquisition of Japanese consonant length contrast

In Japanese, length is used contrastively mainly for obstruent and nasal consonants (Kubozono, 2013). The time-related parameter

* Corresponding author. Via E. Vendramini 13, 35137, Padova, Italy.

E-mail addresses: qiang.feng@phd.unipd.it (Q. Feng), mariagrazia.busa@unipd.it (M.G. Busà).

(i.e., duration) is the primary acoustic cue that differentiates Japanese singleton consonants from their geminate counterparts (Kawahara, 2015). The geminate to singleton consonant duration ratio in Japanese is usually greater than 2:1 (2.8:1 in Han, 1992; 2.4:1 in Toda, 2003). The duration interplay between Japanese singleton/geminate consonants and preconsonantal vowels is the opposite to the typological tendency reported in Maddieson (1985), that is, singleton consonants are preceded by longer vowels and geminates by shorter vowels. In other words, Japanese geminate consonants are preceded by longer vowels, and singletons by shorter vowels (Han, 1994; Idemaru & Guion, 2008).

In the perception of L2 Japanese singleton/geminate consonant contrast, the identification accuracy of beginning Chinese learners seems to be affected by Japanese word pitch patterns. As shown by Minagawa and Kiritani (1996), when the pitch pattern is LH (low high), Chinese learners tend to make more errors in correctly identifying Japanese geminate consonants than singleton consonants (respectively about 75% accuracy vs. about 79% accuracy¹). On the other hand, when the pitch pattern is HL (high low), Chinese learners' misperception direction is the opposite: it is about 88% accurate for geminate consonants and about 65% accurate for singletons. As for advanced Chinese learners of Japanese, in Masuko and Kiritani (1990), the reported perceptual identification accuracy is 78% and 77% respectively for Japanese voiceless fricative and voiceless stop consonant length contrasts. In Tsukada and Hajek (2020), advanced Chinese learners have 90% accuracy for Japanese geminate obstruent consonants and 83% accuracy for singletons.

In producing L2 Japanese consonant length contrast, beginning Chinese learners tend to produce Japanese singleton stops that are longer than one mora unit. As a result, about 35% of the consonants that are intended to be produced as singletons are misperceived as geminates by native Japanese listeners (Minagawa-Kawai & Kiritani, 1998). Intermediate level Chinese learners of Japanese consistently produce geminate consonants that are longer than singleton consonants (Lee et al., 2018; Lu et al., 2016). However, they fail to approximate the native norms in producing preconsonantal vowel duration ratios (Lu et al., 2016). That is, their pre-singleton vowels do not differ in duration from their pre-geminate vowels (Lee et al., 2018). In Yamakawa et al. (2021), Chinese learners of Japanese at a higher level than intermediate mispronounce more singleton stops as geminates than geminates as singletons. Moreover, no duration differences are found between their pre-singleton and pre-geminate syllables.

The above studies show that, in acquiring L2 Japanese consonant length contrast, Chinese learners' perceptual accuracy is always above chance level. This means that Chinese learners are able to differentiate perceptually between Japanese singleton and geminate consonants to a certain extent. In production, Chinese learners generally produce geminates that are longer than singletons, which causes the majority of their consonant length production to be perceived as intended. However, in some cases, their singleton consonants sound so long that they are perceived as geminates by native Japanese listeners. As for preconsonantal vowels, Chinese learners make no difference in pre-singleton and pre-geminate vowel durations, unlike native Japanese speakers, who alternate between shorter vowels before singleton consonants and longer vowels before geminates.

1.2. Consonant length in Italian and Mandarin Chinese

Italian has a great number of disyllabic minimal pairs contrasting in word-internal intervocalic consonant length (e.g., *nono* 'ninth' vs. *nonno* 'grandfather', *caro* 'dear' vs. *carro* 'wagon', etc.). These consonants can be stops (/b, d, g, p, t, k/), affricates (/tʃ, dʒ/), fricatives (/f, v, s/), nasals (/m, n/) and liquids (/l, r/) (for stops, see Esposito & Di Benedetto, 1999; for affricates and fricatives, see Di Benedetto & De Nardis, 2021a; for nasals and liquids, see Di Benedetto & De Nardis, 2021b). In all cases, geminate consonants are longer than their singleton counterparts. At a regular speaking rate, geminate to singleton duration ratios vary from less than 2:1 (in affricates and fricatives) to greater than 2:1 (in stops, nasals and liquids).

Italian geminate consonants are usually considered ambisyllabic (Payne, 2005). That is, the two consonants are treated respectively as the coda of the preceding syllable and the beginning of the following one. Therefore, in Italian disyllabic minimal pairs contrasting in consonant length, the first syllables are closed in geminate words and open in singleton words. Since in Italian vowels are short in closed syllables and long in open syllables (Kramer, 2009), in disyllabic minimal pairs contrasting in singleton/geminate consonants, pre-geminate vowels are always shorter than the corresponding pre-singleton ones. The reduction ratios vary from -25% (in affricates) to -41% (in liquids).

On the other hand, in Mandarin Chinese there is no consonant length distinction at the phonemic level. Among all Mandarin Chinese consonants, only /n/ and /ŋ/ can occur in syllabic coda position; while /ŋ/ cannot occur in onset position (Duanmu, 2007). Therefore, when a syllable with nasal coda /n/ is followed by another syllable with nasal onset /n/ (e.g., <半年> *bannian*, 'half year', /pannjen/), together they form a "fake" or "derived" geminate consonant. However, it should be noted that "the potential for fake geminates in Mandarin is very restricted as it only occurs through concatenation of words with identical nasals across syllables in phrases or disyllabic compounds" (Meng et al., 2021, p. 3). In addition, in Mandarin Chinese, vowel length is never contrastive but only a by-product of syllable type. Specifically, vowels are long in open syllables and short in closed and weak syllables (Duanmu, 2007). Therefore, following Tsukada et al. (2014), Mandarin Chinese speakers can be considered "quantity-insensitive".

1.3. Theoretical framework

Among currently influential L2 speech acquisition theories, the two that have included duration properties in predicting L2 speech acquisition are the Second Language Linguistic Perception model (L2LP) proposed by Escudero (2005) and the Feature Model proposed

¹ In Minagawa and Kiritani (1996), the authors did not provide the participants' precise accuracy rates. The data reported here were drawn from Fig. 1 (p. 25).

by Brown (1998). According to the L2LP model, which focuses on L2 sound perception, at the initial stage of L2 learning, the learners tend to duplicate their L1 perception grammar in their L2 perception. Yet, through increased exposure to the L2 perceptual input, L2 learners modify their L2 perception grammar and ultimately become native-like. In our case, since in Mandarin Chinese phonology only singleton consonants exist, we speculate that Chinese learners with little Italian learning experience may tend to categorize both Italian singleton and geminate consonants as singletons (“a new scenario” according to the L2LP model) and therefore make more errors in identifying Italian geminate consonants than in identifying singletons. However, with an increased learning experience, Chinese learners may gradually approximate the L2 native norms in perceptually differentiating between Italian consonant length contrast.

According to the Feature Model, the features exploited/absent in L1 phonology can either facilitate or hinder learners’ acquisition of L2 segmental contrast, depending on the availability of redeployment in L2 pronunciation learning. Specifically, in the case of L2 consonant length contrast acquisition, if the feature $[\pm \text{long}]$ is exploited in the learners’ native language, it will be transferred and reused in L2 learning and thereby will favor learners’ perception and production accuracy. On the contrary, learners that are not familiar with the feature $[\pm \text{long}]$ through their L1 phonology will not have this inherent advantage. Since Chinese learners are “quantity-insensitive”, it is highly possible that they may have difficulty producing Italian singleton/geminate consonant contrast properly. However, the Feature Model does not take L2 learners’ learning experience into account. Since L2 production is usually believed to be closely related to perception (SLM; Flege, 1995), we expect that Chinese learners’ improved perception as a result of their increased learning experience will lead to more native-like production.

1.4. Research questions and hypotheses

Based on the above speculations, the research questions (RQ) and the corresponding hypotheses (H) are formulated as follows:

RQ1. How do Chinese learners perceive Italian consonant length contrast?

Our hypotheses are that Chinese learners with little Italian learning experience will tend to misidentify perceptually more geminates as singletons than singletons as geminates (H1); and the more experienced learners will outperform the less experienced ones in perceptual identification accuracy (H2).

RQ2. How do Chinese learners produce Italian singleton/geminate consonants and preconsonantal vowels?

Our hypotheses are that Chinese learners may not produce Italian consonant length contrast in a native-like manner (H3); but the more experienced learners will be closer to the L2 native norms than the less experienced ones (H4).

RQ3. What role does learning experience play in Chinese learners’ acquisition of Italian consonant length contrast?

We hypothesize that an increased learning experience will play a positive role in enhancing Chinese learners’ acquisition of Italian singleton/geminate consonant contrast (H5).

To verify the hypotheses, a perception experiment and a production experiment were carried out.

2. Perception experiment

2.1. Method

2.1.1. Participants

The perception experiment had 30 Chinese learners and 10 native Italian speakers as participants. The 30 Chinese learners were undergraduate students majoring in Italian at Dalian University of Foreign Languages in China; they were subdivided according to their Italian learning experience as follows: 10 first-year students (1 MC group: Female = 9, Male = 1, Mean age = 19.1), 10 second-year students (2 MC group: Female = 10, Male = 0, Mean age = 19.9), and 10 third-year students (3 MC group: Female = 8, Male = 2, Mean age = 20.9). The relatively restricted number of participants was due to the various restrictions imposed by the Covid-19 pandemic. The Chinese learners’ gender distribution reflects the imbalance of the students’ enrollment in the degree course. According to the Chinese learners’ self-reports, they had received little to no specific training on the pronunciation of L2 Italian consonant length contrast at the time of the experiment. The 10 native Italian speakers (NIT group: Female = 7, Male = 3, Mean age = 21.8) were undergraduate students from the Veneto region in the North-East of Italy. Though northern Italian speakers are conventionally believed to tend to degeminate consonants (e.g., Canepari & Giovannelli, 2008; Cavanaugh, 2005), Mairano and De Iacovo (2020) show that this is not the case. No participant reported any hearing impairment at the time of the experiment.

2.1.2. Materials

The target stimuli of the perception experiment were ten Italian disyllabic minimal pairs contrasting in consonant length (see Appendix A). The target consonants consisted of all types of consonants that can be geminated, namely stop ($/p, t, k, d/$), affricate ($/tʃ, dʒ/$), fricative ($/f, v/$), nasal ($/n/$) and liquid ($/l/$) consonants. Besides, an equal number of disyllabic minimal pairs contrasting in stop voicing (i.e., voiced vs. voiceless) served as distractors. The carrier sentence was *Leggo __ bene* (‘I read __ well’).

One female native Italian speaker (age = 24) from the Veneto region was recruited for the recording session. She was instructed to read the randomly arranged word stimuli first in isolation then in the carrier sentences at a normal speech rate. The recording took place in the Language and Communication Lab of the University of Padova in Italy, using a Roland R09 voice recorder with a sampling rate of 44.1 kHz and 16-bit resolution. A total of 80 trials ([10 target minimal pairs + 10 distracting minimal pairs] \times 2 words \times 2 contexts = 80) were produced. The auditory intelligibility of these trials was perceptually verified by the second author, who is a native Italian-speaking phonetician. Afterwards, the trials produced were segmented into isolated words and sentences, and saved as separate audio files. The intensity of the audio files was normalized to 70 dB.

2.1.3. Procedure

The audio files were uploaded to the Alchemer platform (www.alchemer.com/) to create an identification test, which consisted of two parts. In the first part, the 40 isolated words were ordered randomly. For each word, two options (i.e., a minimal pair) were given on the screen. The participants, seated in front of a computer with headphones on, had to click on the option corresponding to the word heard. Each word was repeated only once. The participants were asked to listen to each word for a maximum of three times. The second part of the experiment was structured similarly to the first one except that the stimuli were the 40 sentences. The Chinese learners took the identification test in a multimedia speech lab at Dalian University of Foreign Languages in China. The native Italian speakers took the test in the Language and Communication Lab of the University of Padova in Italy.

2.1.4. Analyses

In total, we obtained 1600 responses for the target trials (10 target minimal pairs \times 2 words \times 2 contexts \times 40 participants = 1600). A generalized linear mixed-effects model (GLMM) with a binomial link function was applied to the participants' responses (i.e., correct vs. incorrect) using the lme4 package 1.1.26 (Bates et al., 2015) in R 3.6.3 (R Core Team, 2020), with Group (four levels: 1 MC, 2 MC, 3 MC, NIT), Consonant Type (two levels: Singleton and Geminate), Context (two levels: In isolation and In carrier sentence) and their interactions as fixed factors, and Subject and Item as random intercepts. The coding scheme used was treatment coding. The assessment of the main effects of the fixed factors was performed with the Type II Wald chi-squared tests using the car package 3.0.10 (Fox & Weisberg, 2019). Post-hoc pairwise comparisons with FDR (false discovery rate) correction were carried out using the emmeans package 1.5.3 (Lenth, 2020).

2.2. Results

Fig. 1 illustrates the four groups' perceptual accuracy rates calculated by Consonant Type and Context. As can be seen, both in isolation and in the carrier sentences, the NIT group's accuracy rates reached the ceiling for both the singleton and geminate consonants. On the other hand, in both contexts, the 1 MC, 2 MC and 3 MC groups' accuracy rates for both types of consonants were all above chance level (hovered between 70% and 81%), but always lower than those of the NIT group. Moreover, the three learner groups' accuracy rates were rather close to each other.

The results of the GLMM applied to the four groups' responses are shown in Table 1. As can be seen, the GLMM yielded a significant main effect on Group, indicating that the participants' perceptual accuracy differed significantly by group. Post-hoc pairwise comparisons revealed that there were no significant differences across the three learner groups (1 MC vs. 2 MC: $\beta = 0.005$, $SE = 0.32$, $z = 0.02$, $p = 0.99$; 1 MC vs. 3 MC: $\beta = -0.06$, $SE = 0.32$, $z = -0.18$, $p = 0.99$; 2 MC vs. 3 MC: $\beta = -0.06$, $SE = 0.32$, $z = -0.19$, $p = 0.99$), which were all significantly different from the NIT group (1 MC vs. NIT: $\beta = 2.82$, $SE = 0.50$, $z = 5.67$, $p < 0.0001$; 2 MC vs. NIT: $\beta = 2.82$, $SE = 0.50$, $z = 5.66$, $p < 0.0001$; 3 MC vs. NIT: $\beta = 2.88$, $SE = 0.50$, $z = 5.78$, $p < 0.0001$). These results indicate that the Chinese learners with different learning experiences identified perceptually Italian consonant length contrast with an equal level of accuracy; and they were always significantly less accurate than the native Italian listeners. Also, it should be noted that there was no significant

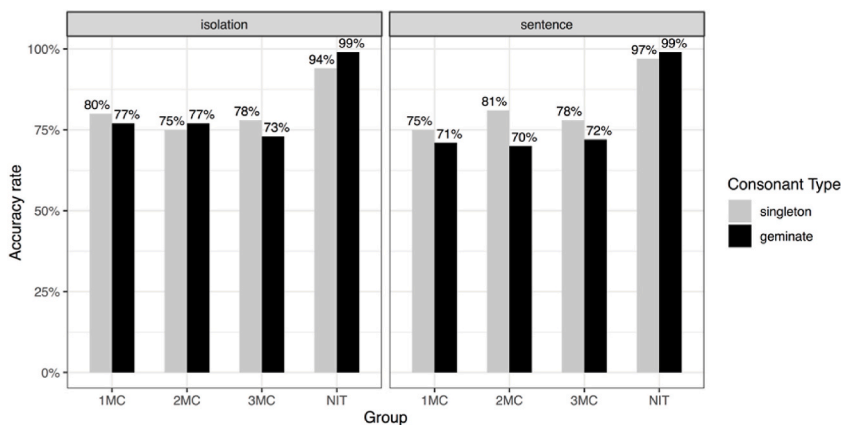


Fig. 1. The 1 MC, 2 MC, 3 MC and NIT groups' perceptual accuracy rates for the singleton and geminate consonants in isolation and in the carrier sentences.

Table 1Results of the GLMM of the participants' perceptual accuracy of Italian consonant length contrast (*p* values < 0.05 are in bold).

	Fixed effects			Random effects	
	χ^2	df	p	By Subject	By item
				SD	SD
Intercept	–	–	–	0.36	0.15
Group	31.95	3	< .001	–	–
Consonant Type	0.72	1	0.40	–	–
Context	0.48	1	0.49	–	–
Group × Consonant Type	5.32	3	0.15	–	–
Group × Context	1.92	3	0.59	–	–
Consonant Type × Context	1.19	1	0.27	–	–
Group × Consonant Type × Context	1.62	3	0.66	–	–

main effect on Consonant Type, indicating that the participants' identification accuracy did not vary by this factor. Moreover, the nonsignificant interaction of Group × Consonant Type × Context indicates that the ways in which the perceptual identification accuracy differed between groups did not vary by consonant type and reading context.

2.3. Discussion

As expected, compared to native Italian listeners, Chinese learners are less accurate in perceptually identifying Italian singleton and geminate consonants. However, some results are unexpected.

First, though Chinese learners are “quantity-insensitive” language speakers, in the majority of cases (more than 70% of instances), they succeed in correctly differentiating between Italian singleton and geminate consonants. We argue that though speakers' perceptual sensitivity to duration may be reduced if length distinctions are absent in their L1 (Gottfried & Beddor, 1988), it does not mean that in processing linguistic stimuli, the so-called “quantity-insensitive” language speakers completely lose the perceptual sensitivity to segment duration differences. In fact, “perceptual sensitivity to duration is needed in all languages as segment durations do not only signal lexical contrasts but may also convey information about prosodic structure” (Altmann et al., 2012, p. 389). For example, it has been noted that neutral-tone syllables in Mandarin Chinese are shorter than lexical-tone syllables (Lee & Zee, 2008; Lin & Yan, 1980), and duration is the primary cue by which Mandarin Chinese listeners differentiate the two types of syllables (Lin, 1985). Thus, Mandarin Chinese speakers may not be completely “quantity-insensitive” as is commonly believed. Moreover, the duration differences between Italian singleton consonants and their geminate counterparts are over the threshold of the Just Noticeable Difference (JND) for segment duration (Payne, 2005; see Klatt [1976] and Lehiste [1970] for JND), which means that the duration differences existing in Italian consonant length contrast are in most cases perceptually discernible. Therefore, though Chinese learners are not as sensitive as native Italian listeners to Italian consonant duration differences, they are able to distinguish perceptually between Italian singleton and geminate consonants to a certain extent.

Second, though geminate but not singleton consonants are absent in Mandarin Chinese, contrary to our H1, Chinese learners do not misperceive geminates more than singletons. We argue that this is because in the present study, even the Chinese learners with the least learning experience had studied Italian for 1 year at the time of the experiment. Therefore, it may not be appropriate to view them as beginning learners. In other words, they might have already passed the initial L2 learning stage as defined by the L2LP model (Escudero, 2005), that is, the stage at which L2 learners tend to duplicate their L1 perception grammar in their L2 perception, and have established two consonant length categories (i.e., short and long) to address the new duration distribution in L2 Italian. This fact may imply that Chinese learners' new duration category does not take a long time to be formed.

Third, contrary to our H2, Chinese learners' increased learning experience does not result in higher accuracy rates. This fact raises the question of what role Chinese learners' learning experience plays in their L2 quantity acquisition, which will be further discussed later.

To test how Chinese learners' perception of Italian consonant length contrast maps onto their production pattern, a production experiment was carried out.

3. Production experiment

3.1. Method

3.1.1. Participants

The participants of the production experiment were the same as those in the perception experiment. None of them reported any speech impairment at the time of the experiment.

3.1.2. Materials

To maintain consistency between the production and perception experiments as well as prevent unfamiliar sounds from affecting the Chinese learners' production, five minimal pairs that served as target stimuli in the perception experiment and had target

consonants available in both Italian and Mandarin Chinese phonology (i.e., /p, t, k, n, l/) were selected as target stimuli for the production experiment (see Appendix A). The distractors were the five minimal pairs contrasting in stop voicing that were used in the perception experiment. The carrier sentence was still *Leggo __ bene* ('I read __ well'). The target word stimuli were first repeated twice randomly in isolation. Then they were inserted in the carrier sentences and repeated twice in random order. The stimuli in isolation and in the carrier sentences were printed on a paper sheet for the subjects to read. After every 20 words/sentences, a cartoon picture was inserted to remind the participants to have a short break.

3.1.3. Procedure

In the production experiment, the participants were instructed to read the stimuli on the paper sheet at a normal speed. They were also asked to make a short break when the cartoon pictures told them to. The recordings of the native Italian speakers took place in the Language and Communication Lab of the University of Padova in Italy, using a Roland R09 voice recorder with a sampling rate of 44.1 kHz and 16-bit resolution. The recordings of the Chinese learners were administered in a quiet setting at Dalian University of Foreign Languages in China, using a Zoom H4n Pro voice recorder with a sampling rate of 44.1 kHz and 16-bit resolution.

3.1.4. Annotation and measurement

The target intervocalic consonants and the prenasal vowels were labeled in Praat (Boersma & Weenink, 2020). As shown by Esposito and Di Benedetto (1999), closure duration is the primary cue for gemination in Italian stops. Therefore, the intervocalic stop consonants in the present experiment were labeled from the offset of the periodic wave of the preceding vowel to the release burst of the stop consonant; the VOT was not included. For the target intervocalic lateral and nasal consonants, their boundaries with the adjacent vowels were identified where abrupt spectral changes were observed. For the prenasal vowels, boundaries were located at the onset and offset of periodicity in the acoustic waveform. Afterwards, the durations of the target intervocalic consonants and the prenasal vowels were extracted using a Praat script (Lennes, 2002).

3.1.5. Analyses

The duration values were converted into duration ratios for statistical analyses. A total of 1600 target tokens were elicited (10 target word stimuli \times 2 contexts \times 2 repetitions \times 40 participants = 1600). Of these tokens, half had intervocalic singleton consonants, and the other half had intervocalic geminate consonants. Therefore, we should have had 800 geminate to singleton consonant duration ratios and 800 pre-singleton to pre-geminate vowel duration ratios. However, in measuring duration values, 56 unmeasurable/misread tokens with intervocalic consonants (26 singletons, 30 geminates) and 61 unmeasurable/misread tokens with prenasal vowels (28 pre-singleton, 33 pre-geminate) were discarded. Because of these missing values, in converting duration values into duration ratios, we obtained 751 effective consonant duration ratios and 747 effective prenasal vowel duration ratios.

Two linear mixed models (LMM) were applied respectively to the consonant and prenasal vowel duration ratios using the lme4 package 1.1.26 (Bates et al., 2015) in R 3.6.3 (R Core Team, 2020), with Group (four levels: 1 MC, 2 MC, 3 MC, NIT), Context (two levels: In isolation and In carrier sentence) and their interaction as fixed factors, and Subject and Item as random intercepts. The coding scheme used was treatment coding. The assessments of the main effects of the fixed factors were performed with the Type II Wald chi-squared tests using the car package 3.0.10 (Fox & Weisberg, 2019). Post-hoc pairwise comparisons with FDR (false discovery rate) correction were carried out using the emmeans package 1.5.3 (Lenth, 2020).

3.2. Results

3.2.1. Consonant duration ratios

Table 2 lists the mean duration values of the consonants produced by the 1 MC, 2 MC, 3 MC and NIT groups and their mean geminate to singleton consonant duration ratios. Fig. 2 shows a visual representation of the distribution of the four groups' consonant duration ratios. As Table 2 and Fig. 2 show, both in isolation and in the carrier sentences, all four groups had longer durations for geminate than for singleton consonants. However, the NIT group had greater consonant duration ratios than the three learner groups. Besides, in terms of duration value, the singleton consonants produced by the 1 MC, 2 MC, 3 MC groups were consistently longer than those of the NIT group; while the four groups' geminate consonants had durations that did not differ to a great extent both in isolation and in the carrier sentences.

For the statistical analyses, the consonant duration ratios were normalized using the bestNormalize package 1.7.0 (Peterson & Cavanaugh, 2020). After fitting the LMM, the visual diagnostics of the histogram and the plot of residuals showed no drastic violations

Table 2

Mean consonant durations (in ms; SDs in parentheses) and mean geminate to singleton consonant duration ratios (G:S ratio; SDs in parentheses) of the 1 MC, 2 MC, 3 MC and NIT groups.

	In isolation			In carrier sentence		
	Singleton	Geminate	G:S ratio	Singleton	Geminate	G:S ratio
1 MC	131.2 (41.7)	171.6 (66.9)	1.37 (0.53)	100.3 (55.3)	122.7 (64.1)	1.32 (0.68)
2 MC	143.4 (45.0)	182.3 (58.7)	1.31 (0.35)	105.0 (39.6)	128.2 (53.2)	1.29 (0.52)
3 MC	140.0 (75.8)	169.5 (78.0)	1.37 (0.83)	106.7 (46.8)	134.3 (71.5)	1.38 (0.72)
NIT	80.5 (28.0)	183.5 (43.2)	2.61 (1.48)	69.6 (22.4)	133.3 (33.0)	2.11 (0.93)

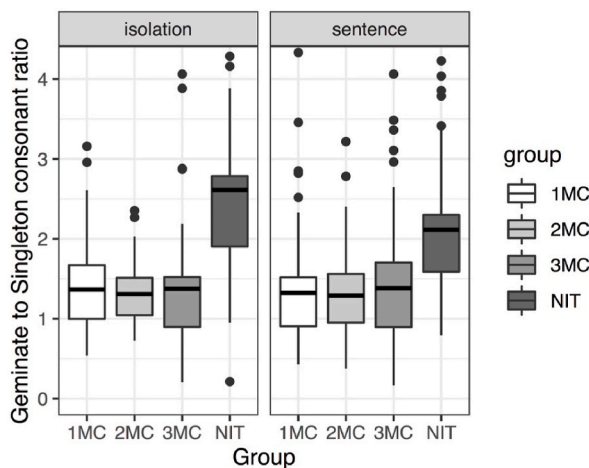


Fig. 2. Distribution of the geminate to singleton consonant duration ratios of the 1 MC, 2 MC, 3 MC and NIT groups.

of the assumptions of normality and homoscedasticity. Table 3 shows the results of the LMM. As can be seen, the LMM yielded significant main effects on Group and Context, indicating that the participants’ geminate to singleton consonant duration ratios differed significantly by these two factors. Post-hoc pairwise comparisons revealed that, first, all three learner groups were significantly different from the NIT group (1 MC vs. NIT: $\beta = -1.17, SE = 0.17, t(36.4) = -7.02, p < 0.0001$; 2 MC vs. NIT: $\beta = -1.14, SE = 0.17, t(35.3) = -6.94, p < 0.0001$; 3 MC vs. NIT: $\beta = -1.18, SE = 0.17, t(35.2) = -7.19, p < 0.0001$), indicating that though the Chinese learners were able to produce short-long differences for Italian consonant length contrast, their geminate to singleton consonant duration ratios were significantly smaller than those of the native Italian speakers. Second, there were no significant differences across the learner groups (1 MC vs. 2 MC: $\beta = -0.02, SE = 0.17, t(36.9) = -0.13, p = 0.92$; 1 MC vs. 3 MC: $\beta = 0.02, SE = 0.17, t(36.8) = 0.11, p = 0.92$; 2 MC vs. 3 MC: $\beta = 0.04, SE = 0.17, t(35.7) = 0.24, p = 0.92$), indicating that though the Chinese learners had different learning experiences, they produced Italian geminate to singleton consonant duration ratios very similarly. Third, the participants performed significantly differently in the two reading contexts (in isolation vs. in the carrier sentences: $\beta = 0.15, SE = 0.06, t(704) = 2.643, p = 0.0084$). That is, the participants had larger consonant duration ratios for the words produced in isolation than for those produced in the carrier sentences. Also, there was no significant interaction between Group and Context, indicating that the differences in consonant duration ratio between groups did not vary by reading context.

3.2.2. Preconsonantal vowel duration ratios

The mean duration values of the preconsonantal vowels produced by the 1 MC, 2 MC, 3 MC and NIT groups and their mean pre-singleton to pre-geminate vowel duration ratios are shown in Table 4. Fig. 3 shows a visual representation of the distribution of the four groups’ preconsonantal vowel duration ratios. As Table 4 and Fig. 3 show, only the NIT group had pre-singleton to pre-geminate vowel duration ratios greater than 1, both in isolation and in the carrier sentences. The preconsonantal vowel duration ratios of the three learner groups hovered around 1. This means that only the NIT group produced pre-singleton vowels that were longer than pre-geminate vowels. The three learner groups, on the other hand, produced pre-singleton and pre-geminate vowels with similar durations.

For the statistical analyses, the preconsonantal vowel duration ratios were normalized using the bestNormalize package 1.7.0 (Peterson & Cavanaugh, 2020). After fitting the LMM, the visual diagnostics of the histogram and the plot of residuals showed no drastic violations of the assumptions of normality and homoscedasticity. Table 5 shows the results of the LMM, which yielded significant main effects on all fix factors, namely Group, Context, and their interaction. Table 6 summarizes the results of the post-hoc pairwise comparisons. As can be seen, the three learner groups were always significantly different from the NIT group, indicating that while the native Italian speakers alternated between short vowels before geminate consonants and long vowels before singletons, the Chinese learners produced no difference in pre-singleton and pre-geminate vowel durations. Moreover, no significant between-group differences were found across the 1 MC, 2 MC and 3 MC groups, indicating that the Chinese learners’ different learning

Table 3

Results of the LMM of the participants’ production of geminate to singleton consonant duration ratios (*p* values < 0.05 are in bold).

	Fixed effects			Random effects	
	χ^2	df	p	By Subject	By item
				SD	SD
Intercept	–	–	–	0.10	0.01
Group	74.88	3	< .001	–	–
Context	7.32	1	0.0068	–	–
Group × Context	6.58	3	0.086	–	–

Table 4

Mean preconsontal vowel durations (in ms; SDs in parentheses) and mean pre-singleton to pre-geminate vowel duration ratios (PS:PG ratio; SDs in parentheses) of the 1 MC, 2 MC, 3 MC and NIT groups.

	In isolation			In carrier sentence		
	pre-singleton	pre-geminate	PS:PG ratio	pre- singleton	pre-geminate	PS:PG ratio
1 MC	189.8 (51.9)	198.9 (52.7)	0.97 (0.24)	154.4 (40.7)	161.1 (39.0)	0.97 (0.23)
2 MC	167.8 (52.9)	176.8 (62.7)	1.00 (0.23)	155.0 (48.3)	167.7 (60.0)	0.97 (0.21)
3 MC	147.0 (38.4)	152.5 (39.5)	0.99 (0.21)	145.4 (41.1)	151.5 (46.5)	1.00 (0.29)
NIT	149.5 (37.1)	103.7 (28.2)	1.48 (0.33)	119.3 (39.2)	90.0 (19.3)	1.32 (0.31)

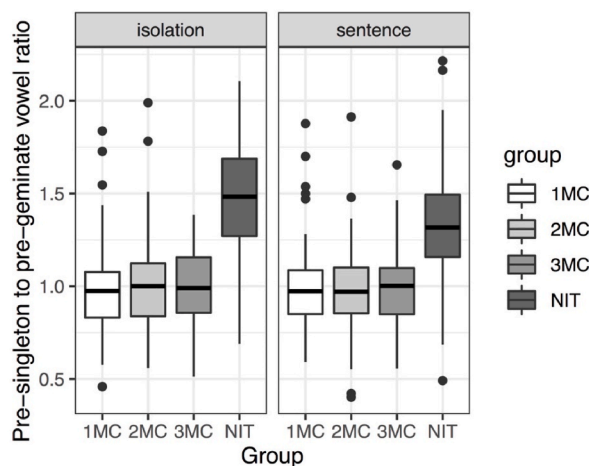


Fig. 3. Distribution of the pre-singleton to pre-geminate vowel duration ratios of the 1 MC, 2 MC, 3 MC and NIT groups.

Table 5

Results of the LMM of the participants' production of pre-singleton to pre-geminate vowel duration ratios (p values < 0.05 are in bold).

	Fixed effects				Random effects	
	χ^2	df	p	By Subject		By item
				SD	SD	SD
Intercept	–	–	–	0.06	–	0.01
Group	116.22	3	< .001	–	–	–
Context	6.14	1	0.013	–	–	–
Group × Context	9.88	3	0.020	–	–	–

experiences did not affect their productions of Italian preconsontal vowel duration ratios. Also, the within-group analysis showed that only the NIT group performed differently in the two reading contexts (i.e., in isolation vs. in the carrier sentences). That is, the duration differences between the pre-singleton and pre-geminate vowels produced by the native Italian speakers were more evident when the words were produced in isolation than in the carrier sentences.

3.3. Discussion

The productions of the native Italian speakers in the present experiment show that northern Italian speakers consistently and properly produce Italian geminate consonants - in contrast to the conventional claim that northern Italian speakers tend to degeminate consonants (Canepari & Giovannelli, 2008; Cavanaugh, 2005). This confirms what was found in Mairano and De Iacovo (2020) that attribute this fact to the progressive standardization of Italian in Italy.

Concerning Chinese learners' production of Italian consonant length contrast, the results show that Chinese learners succeed in making a distinction between shorter singleton and longer geminate consonants in Italian, even when they have only one year of Italian learning experience. This parallels what was found in our perception experiment, suggesting that Chinese learners manage to form new consonant duration categories in a short time and successfully apply the perceived duration differences to their productions. Besides, it is conceivable that the orthographic difference in Italian consonant quantity contrast (i.e., one letter for singleton consonants and two letters for geminates) may also contribute to a certain extent to guiding Chinese learners to produce Italian geminate consonants that are longer than their singleton counterparts. That the orthographic forms of singleton vs. geminate consonants may

Table 6

Results of the post-hoc pairwise comparisons of the pre-singleton to pre-geminate vowel duration ratios of the 1 MC, 2 MC, 3 MC and NIT groups (*p* values < 0.05 are in bold).

	Estimate	SE	df	t ratio	p value
In isolation					
1 MC vs. 2 MC	-0.10	0.17	68.8	-0.61	0.81
1 MC vs. 3 MC	-0.06	0.17	70.1	-0.37	0.81
2 MC vs. 3 MC	0.04	0.16	64.2	0.25	0.81
1 MC vs. NIT	-1.48	0.17	67.3	-9.00	<.0001
2 MC vs. NIT	-1.38	0.16	61.6	-8.59	<.0001
3 MC vs. NIT	-1.42	0.16	62.6	-8.80	<.0001
In carrier sentence					
1 MC vs. 2 MC	0.004	0.16	65.7	0.03	0.98
1 MC vs. 3 MC	-0.07	0.16	64.4	-0.46	0.78
2 MC vs. 3 MC	-0.08	0.16	62.5	-0.49	0.78
1 MC vs. NIT	-1.05	0.16	64.4	-6.45	<.0001
2 MC vs. NIT	-1.05	0.16	62.5	-6.53	<.0001
3 MC vs. NIT	-0.98	0.16	61.2	-6.08	<.0001
In isolation vs. In carrier sentence					
1 MC	0.006	0.12	702.9	0.05	0.96
2 MC	0.11	0.12	700.2	0.96	0.34
3 MC	-0.007	0.12	700.4	-0.06	0.95
NIT	0.44	0.11	699.6	3.88	0.0001

influence speakers' productions has been shown, for example, by Bassetti et al. (2018), who found that L1 Italian speakers, following English orthography, produce consonant length distinctions in their L2 English speech (though with smaller duration differences than in L1 Italian).

However, Chinese learners produce significantly smaller geminate to singleton consonant duration ratios than native Italian speakers. This shows that Chinese learners fail to approximate the L2 norms in producing Italian consonant length contrast, which is in line with our H3. Nevertheless, it should be noted that Chinese learners have smaller consonant duration ratios not because they produce shorter geminate consonants, but because they produce much longer singletons as compared to native Italian speakers. Since Italian consonant length is inversely related to speaking rates (Pickett et al., 1999), we argue that it is Chinese learners' slower speaking rates that lengthen both their singleton and geminate consonants. Thus, Chinese learners approach native Italian speakers' duration values in the production of geminate but not singleton consonants. This triggers the following questions: Will a certain number of these longer singletons be perceived as geminates by native Italian listeners as is reported for L2 Japanese (Minagawa-Kawai & Kiritani, 1998; Yamakawa et al., 2021)? Will these lengthened geminate consonants be perceived as "true" geminates by native Italian listeners? To answer these questions, a future perception experiment that involves native Italian listeners as raters is needed.

As for Chinese learners' production of preconsonantal vowels in Italian consonant length contrast, our data show that Chinese learners produce no duration differences between pre-singleton and pre-geminate vowels, unlike native Italian speakers that have shorter vowels in pre-geminate position. Given that in both Italian and Mandarin Chinese vowels are short in closed syllables, we hypothesize that this is due to the fact that Chinese learners do not treat the first syllables in Italian disyllabic geminate words as closed syllables. This is because, with the exception of /n/, none of the consonants that can be geminated in Italian can occur in syllable coda position in Mandarin Chinese. Thus, influenced by the L1 norms, Chinese learners fail to consider Italian geminate consonants ambisyllabic as native Italian speakers do, but treat them simply as a whole that acts as the initial part of the subsequent syllable. In other words, Chinese learners treat the first syllables of both Italian disyllabic geminate and singleton words as open syllables, and therefore make no difference in preconsonantal vowel durations.

Regarding the effects of learning experience, our data show that, contrary to our H4, an increase in learning experience brings no improvement in Chinese learners' production of Italian consonant length contrast, which parallels what was found in our perception experiment. We will further discuss this point in § 4.

4. General discussion and conclusions

This study set out to investigate the acquisition of Italian consonant length contrast by Mandarin Chinese-speaking learners. Regarding the first two research questions that respectively address how Chinese learners perceive Italian consonant length contrast and how they produce Italian singleton/geminate consonants and preconsonantal vowels, our results show that, in both perception and production, Chinese learners can distinguish between the two consonant length categories to a certain extent, but not in a native-like manner. Moreover, the duration interplay between Italian consonants and preconsonantal vowels (i.e., longer vowels before singleton consonants and shorter vowels before geminate consonants) is entirely ignored by Chinese learners.

In relation to the third research question that addresses the role played by the learning experience in Chinese learners' acquisition of Italian consonant quantity distinction, the results are contrary to our H5. That is, Chinese learners' increased learning experience does not seem to enhance either their perception or production of Italian consonant quantity contrast. Previous studies focusing on L2

vowel/consonant quantity contrast acquisition have reported similar results (e.g., for perception, see [Hayes, 2002](#); for production, see [Harada, 2006](#); [Hirata, 2017](#); [Lee & Mok, 2018](#); [Toda, 1997](#)). Why is this? Some plausible reasons can be found in the literature.

[Luo et al. \(2020\)](#) argue that the majority of L2 quantity contrast acquisition may occur at the initial stage of L2 acquisition, of which the cut-off point should be 6–12 months according to [Best and Tyler \(2007\)](#). Possibly due to this, in L2 quantity contrast acquisition, learners with language learning experience always outperform those without experience (e.g., [Hayes-Harb & Masuda, 2008](#); [Kabak et al., 2011](#); [Tsukada & Hajek, 2019a](#)). However, once the initial stage is passed and the quantity categories are established, for most learners an increased learning experience benefits no longer L2 quantity contrast acquisition. Alternatively, considering that learners with little learning experience can achieve accuracy rates far above chance level in perception, and can make short-long differences in production, it is possible that they do not encounter intelligibility issues when communicating with native speakers. Thus, the urgency to further improve the L2 quantity contrast pronunciation becomes low, and the inaccuracy problems persist even when their learning experience increases ([Hirata, 2017](#); [Lee & Mok, 2018](#)).

As an alternative explanation, we argue that Chinese learners' scarce improvement in L2 Italian consonant length contrast acquisition may be caused by their insufficient L2 perceptual input. The L2LP model ([Escudero, 2005](#)) indicates that L2 learners modify their L2 perception grammar through increased exposure to the L2 perceptual input. The Chinese learners in this study were learning L2 Italian exclusively in China. Therefore, their daily-life communication is predominantly in Mandarin Chinese; and their authentic L2 Italian perceptual input is limited to lessons given by native Italian-speaking teachers that last only a few hours per week. Thus, the increase in the present Chinese learners' learning experience is not paralleled by a substantial increase in their exposure to L2 perceptual input, causing them to fail to get closer to a native level of proficiency in perception. So, the Chinese learners' scarce improvement in perception causes, in turn, their standstill in production.

This argument seems compatible with the L1 consonant quantity contrast acquisition process. That is, the perceptual input plays a vital role in consonant length distinction acquisition also for L1 learners. For example, both Finnish and Japanese have singleton/geminate consonant contrast in their phonology, though geminate consonants occur more frequently in Finnish than in Japanese. As such, in Finnish children's books, geminate consonants represent a total of 9.9% of all the consonant occurrences, while this proportion is 5.1% in Japanese children's books ([Aoyama, 2000](#)). More importantly, in Finnish and Japanese mothers' speech to children, the proportions of geminate consonants in all consonant occurrences are respectively 13.2% and 6.5% ([Kunnari et al., 2001](#)). As a result, Finnish children are exposed to geminate consonant input almost twice as frequently as Japanese children. Possibly due to this input difference, L1 Finnish children acquire consonant length distinction in production more rapidly than L1 Japanese children ([Aoyama, 2000](#); [Kunnari et al., 2001](#)).

There is another issue worth considering. Recall that the actual implementations of consonant length contrast in Italian and Japanese differ to a considerable extent (see §1.1 and §1.2 for a comparison). Firstly, consonant length is contrastive for a wider number of phonemes in Italian than in Japanese; and secondly, the duration interplay between singleton/geminate consonants and preconsonantal vowels in Italian is the opposite to that in Japanese. However, despite these differences, Chinese learners' consonant length contrast acquisition patterns for these two languages show some similarities. Specifically, in perception, Chinese learners have accuracy rates well above chance level but significantly lower than the ceiling. In production, Chinese learners produce geminate consonants that are longer than singletons, though their singletons tend to be longer than native speakers'. Besides, Chinese learners ignore the duration interplay between geminate/singleton consonants and preconsonantal vowels, and produce pre-singleton and pre-geminate vowels with similar durations. This may indicate that Chinese learners follow a similar acquisition pattern in dealing with consonant quantity contrasts in different L2s. However, at the present stage this is a purely speculative conclusion and would need to be addressed with further investigations.

Several crucial issues are left for future research. For instance, in the present perceptual identification test, some target minimal pairs consisted of words with unbalanced lexical frequency, and it is unclear whether this imbalance has affected the participants' perception. Therefore, to consolidate the current conclusions, future perception experiments using nonwords contrasting in consonant length as stimuli may help. Besides, it has been shown that the manners of articulation of consonants (i.e., sonorant vs. obstruent; [Dmitrieva, 2018](#)), the voicing feature of obstruent consonants ([Tsukada & Hajek, 2019b](#)), the length differentials of singleton/geminate consonant contrast ([Hayes, 2002](#)), and the sonority differences between consonants and post-consonantal vowels ([Motohashi-Siago & Hardison, 2009](#)) could all affect cross-linguistic perceptual accuracy of consonant quantity contrast. Therefore, it would also be interesting to know whether/how these factors interact with learning experience to affect Chinese learners' perceptual discrimination between Italian singleton and geminate consonants.

Also, some pedagogical implications can be drawn from this study. Specifically, the Chinese learners in the present study fail to fully master the duration differences creating Italian consonant length contrast (especially the duration interplay between Italian consonants and preconsonantal vowels) and to overcome on their own the acquisition standstill mainly because (i) they are relatively "quantity-insensitive" and (ii) have limited L2 Italian perceptual input - as they have learned Italian exclusively in China. Therefore, one possible way to improve Chinese learners' acquisition of Italian consonant length contrast is to increase their L2 Italian input through creating an immersion context (e.g., study-aboard context) for them. Also, considering that computer-assisted (e.g., [Hirata, 2004](#); [Motohashi-Siago & Hardison, 2009](#)) and gesture-aided (e.g., [Li et al., 2020](#)) teaching methods, by visualizing L2 segmental duration differences through either acoustic waveforms or hand movements, have been shown to be effective in helping "quantity-insensitive" learners to enhance their L2 durational contrast acquisition, it is highly recommended to incorporate these two methods in Chinese learners' L2 Italian pronunciation training classrooms.

Funding

The first author is supported by a grant from the China Scholarship Council [grant number 201908210329].

Author statement

Qiang Feng: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Funding acquisition. **M. Grazia Busà:** Conceptualization, Methodology, Investigation, Resources, Writing - Review & Editing, Supervision.

Declaration of competing interest

None.

Acknowledgments

The authors wish to express sincere gratitude to the participants involved in this study. Many thanks also to the editor and the two anonymous reviewers for their helpful comments and valuable suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.system.2022.102938>.

Appendix A. Word stimuli for the perception and production (in bold) experiments

Target words	Distractors
<i>papa</i> 'pope' vs. <i>pappa</i> 'pulp'	<i>passo</i> vs. <i>basso</i>
<i>dita</i> 'fingers' vs. <i>ditta</i> 'company'	<i>pollo</i> vs. <i>bollo</i>
<i>Luca</i> 'person's name' vs. <i>Lucca</i> 'city's name'	<i>tetto</i> vs. <i>detto</i>
<i>pena</i> 'penalty' vs. <i>penna</i> 'pen'	<i>quanto</i> vs. <i>quando</i>
<i>pala</i> 'shovel' vs. <i>palla</i> 'ball'	<i>cara</i> vs. <i>gara</i>
<i>cade</i> 'falls' vs. <i>cadde</i> 'fell'	<i>eco</i> vs. <i>ego</i>
<i>face</i> 'torch' vs. <i>facce</i> 'faces'	<i>pasta</i> vs. <i>basta</i>
<i>bevi</i> 'drink' vs. <i>bevvi</i> 'drank'	<i>panca</i> vs. <i>banca</i>
<i>agio</i> 'ease' vs. <i>aggio</i> 'premium'	<i>pari</i> vs. <i>Bari</i>
<i>tuffo</i> 'tuff' vs. <i>tuffo</i> 'dip'	<i>noto</i> vs. <i>nodo</i>

References

- Altmann, H., Berger, I., & Braun, B. (2012). Asymmetries in the perception of non-native consonantal and vocalic length contrasts. *Second Language Research*, 28(4), 387–413. <https://doi.org/10.1177/0267658312456544>
- Aoyama, K. (2000). *A psycholinguistic perspective on Finnish and Japanese prosody: Perception, production and child acquisition of consonantal quantity distinctions*. University of Hawai'i at Manoa.
- Bassetti, B., Sokolović-Perović, M., Mairano, P., & Cerni, T. (2018). Orthography-induced length contrasts in the second language phonological systems of L2 speakers of English: Evidence from minimal pairs. *Language and Speech*, 61(4), 577–597. <https://doi.org/10.1177/0023830918780141>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Best, C. T., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro, & O. S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–34). John Benjamins.
- Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer (6.1.16)* [Computer software] <http://www.praat.org/>.
- Brown, C. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14(2), 136–193. <https://doi.org/10.1191/026765898669508401>
- Canepari, L., & Giovannelli, B. (2008). La buona pronuncia italiana del terzo millennio. In *Manualetto d'italiano neutro con sonori, esercizi e test (4a edizione)*. Aracne.
- Cavanaugh, J. R. (2005). Accent matters: Material consequences of sounding local in northern Italy. *Language & Communication*, 25(2), 127–148. <https://doi.org/10.1016/j.langcom.2005.02.002>
- Di Benedetto, M. G., & De Nardis, L. (2021a). Consonant gemination in Italian: The affricate and fricative case. *Speech Communication*, 134, 86–108. <https://doi.org/10.1016/j.specom.2021.07.005>
- Di Benedetto, M. G., & De Nardis, L. (2021b). Consonant gemination in Italian: The nasal and liquid case. *Speech Communication*, 133, 62–80. <https://doi.org/10.1016/j.specom.2021.07.006>
- Dmitrieva, O. (2018). The role of perception in the typology of geminate consonants: Effects of manner of articulation, segmental environment, position, and stress. *Language and Speech*, 61(1), 43–70. <https://doi.org/10.1177/0023830917696113>
- Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford: OUP.
- Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Netherlands Graduate School of Linguistics.

- Esposito, A., & Di Benedetto, M. G. (1999). Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America*, 106(4), 2051–2062. <https://doi.org/10.1121/1.428056>
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.
- Fox, J., & Weisberg, S. (2019). *An {R} Companion to applied regression (third)*. Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- Gottfried, T. L., & Beddor, P. S. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech*, 31(1), 57–75. <https://doi.org/10.1177/002383098803100103>
- Han, M. S. (1992). The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. *Phonetica*, 49(2), 102–127. <https://doi.org/10.1159/000261906>
- Han, M. S. (1994). Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America*, 96(1), 73–82. <https://doi.org/10.1121/1.410376>
- Harada, T. (2006). The acquisition of single and geminate stops by English-speaking children in a Japanese immersion program. *Studies in Second Language Acquisition*, 28(4). <https://doi.org/10.1017/S0272263106060281>
- Hayes, R. L. (2002). The perception of novel phoneme contrasts in a Second Language: A developmental study of native speakers of English learning Japanese singleton and geminate consonant contrasts. *Coyote Papers*, 12, 28–41.
- Hayes-Harb, R., & Masuda, K. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research*, 24(1), 5–33. <https://doi.org/10.1177/0267658307082980>
- Hirata, Y. (2004). Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning*, 17(3–4), 357–376. <https://doi.org/10.1080/0958822042000319629>
- Hirata, Y. (2017). Second language learners' production of geminate consonants in Japanese. In H. Kubozono (Ed.), *The phonetics and phonology of geminate consonants* (pp. 163–183). Oxford University Press.
- Idemaru, K., & Guion, S. G. (2008). Acoustic covariants of length contrast in Japanese stops. *Journal of the International Phonetic Association*, 38(2), 167–186. <https://doi.org/10.1017/S0025100308003459>
- Kabak, B., Reckziegel, T., & Braun, B. (2011). Timing of second language singletons and geminates. In W.-S. Lee, & E. Lee (Eds.), *Proceedings of the 17th international congress of the phonetic sciences (ICPhS XVII)* (pp. 994–997). City University of Hong Kong.
- Kawahara, S. (2015). The phonetics of sokuon, or geminate obstruents. In H. Kubozono (Ed.), *Handbook of Japanese phonetics and phonology* (pp. 43–78). De Gruyter Mouton.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59(5), 1208–1221. <https://doi.org/10.1121/1.380986>
- Kramer, M. (2009). *The phonology of Italian*. Oxford University Press on Demand.
- Kubozono, H. (2013). Introduction to the special issue on Japanese geminate obstruents. *Journal of East Asian Linguistics*, 22(4), 303–306.
- Kunnari, S., Nakai, S., & Vihman, M. M. (2001). Cross-linguistic evidence for the acquisition of geminates. *Psychology of Language and Communication*, 5(2), 13–24.
- Lee, S. J., Kurita, G., & Perkins, J. (2018). Identifying the locus of L2 pronunciation: An exploratory study of geminate production by Chinese L2 learners of Japanese. *Foreign Language Education Research*, 22, 69–86.
- Lee, A., & Mok, P. (2018). Acquisition of Japanese quantity contrasts by L1 Cantonese speakers. *Second Language Research*, 34(4), 419–448. <https://doi.org/10.1177/0267658317739056>
- Lee, W.-S., & Zee, E. (2008). Prosodic characteristics of the neutral tone in Beijing Mandarin. *Journal of Chinese Linguistics*, 36(1), 1–29.
- Lehiste, I. (1970). *Suprasegmentals*. Massachusetts Inst. of Technology.
- Lenne, M. (2002). *Calculate segment durations.praat*. https://github.com/lenne/spect/blob/1.0.0/scripts/calculate_segment_durations.praat.
- Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-Squares means (1.5.3 [Computer software]* <https://CRAN.R-project.org/package=emmeans>.
- Li, P., Baills, F., & Prieto, P. (2020). Observing and producing durational hand gestures facilitates the pronunciation of novel vowel-length contrasts. *Studies in Second Language Acquisition*, 42(5), 1015–1039. <https://doi.org/10.1017/S0272263120000054>
- Lin, T. (1985). Tantaou Beijinghua qingyin xingzhi de chubu shiyan [On neutral tone in Beijing Mandarin]. In S. Hu (Ed.), *Beijing yuyin shiyanlu [Working papers in experimental phonetics]* (pp. 1–26). Beijing Daxue chubanshe [Peking University Press].
- Lin, M., & Yan, J. (1980). Beijinghua qingsheng de shengxue xingzhi [Acoustic characteristics of neutral tone in Beijing Mandarin]. *Fangyan*, 3, 166–178.
- Lu, S., Chen, L., & Yamada, J. (2016). Taiwanese learners' durational patterns for Japanese obstruent geminates/singletons and adjacent vowels. *Perceptual & Motor Skills*, 122(1), 67–75. <https://doi.org/10.1177/0031512516630018>
- Luo, J., Li, V. G., & Mok, P. P. K. (2020). The perception of Cantonese vowel length contrast by Mandarin speakers. *Language and Speech*, 63(3), 635–659. <https://doi.org/10.1177/0023830919879471>
- Maddieson, I. (1985). Phonetic cues to syllabification. In V. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of peter ladefoged* (pp. 203–221). Academic Press.
- Mairano, P., & De Iacovo, V. (2020). Gemination in northern versus central and southern varieties of Italian: A corpus-based investigation. *Language and Speech*, 63(3), 608–634. <https://doi.org/10.1177/0023830919875481>
- Masuko, Y., & Kiritani, S. (1990). Perception of Mora sounds in Japanese by non-native speakers of Japanese. In Acoustical Society of Japan (Ed.), *Proceedings: 1990 international conference on spoken language processing* (pp. 1333–1336). Acoustical Society of Japan.
- Meng, Y., Kotzor, S., Xu, C., Wynne, H. S., & Lahiri, A. (2021). Mismatch negativity (MMN) as an index of asymmetric processing of consonant duration in fake Mandarin geminates. *Neuropsychologia*, 163, 1–8. <https://doi.org/10.1016/j.neuropsychologia.2021.108063>
- Minagawa-Kawai, Y., & Kiritani, S. (1998). *Non-native productions of Japanese single stops that are too long for one mora unit*. Proceedings of the 5th International Conference on Spoken Language Processing.
- Minagawa, Y., & Kiritani, S. (1996). Discrimination of the single and geminate stop contrast in Japanese by five different language groups. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 30, 23–28.
- Motohashi-Siogo, M., & Hardison, D. M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language, Learning and Technology*, 13(2), 29–47.
- Payne, E. M. (2005). Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association*, 35(2), 153–181. <https://doi.org/10.1017/S0025100305002240>
- Peterson, R. A., & Cavanaugh, J. E. (2020). Ordered quantile normalization: A semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, 47(13–15), 2312–2327. <https://doi.org/10.1080/02664763.2019.1630372>
- Pickett, E. R., Blumstein, S. E., & Burton, M. W. (1999). Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica*, 56(3–4), 135–157. <https://doi.org/10.1159/000028448>
- R Core Team. (2020). *R: A language and environment for statistical computing* [Computer software] <https://www.R-project.org/>.
- Tanaka, S. (2017). The relation between L2 perception and L1 phonology in Japanese loanwords: An analysis of geminates in loanwords from Italian. In H. Kubozono (Ed.), *The phonetics and phonology of geminate consonants* (pp. 321–339). Oxford University Press.
- Toda, T. (1997). Strategies for producing mora timing by non-native speakers of Japanese. In T. Toda (Ed.), *Acquisition of Japanese as a Second Language* (pp. 157–197). John Benjamins Publishing.
- Toda, T. (2003). Acquisition of special morae in Japanese as a second language. *Journal of the Phonetic Society of Japan*, 7(2), 70–83.
- Tsukada, K., Cox, F., Hajek, J., & Hirata, Y. (2018). Non-native Japanese learners' perception of consonant length in Japanese and Italian. *Second Language Research*, 34(2), 179–200. <https://doi.org/10.1177/0267658317719494>
- Tsukada, K., & Hajek, J. (2019a). 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, & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences* (pp. 82–86). Australasian Speech Science and Technology Association Inc.

- Tsukada, K., & Hajek, J. (2019b). Misperception of Italian singleton and geminate obstruents by native Japanese speakers. *Journal of the Phonetic Society of Japan*, 23, 148–156.
- Tsukada, K., & Hajek, J. (2020). Perception of consonant length in familiar and unfamiliar languages by native speakers of Mandarin, Italian and Japanese. *Proceedings of the 10th International Conference on Speech Prosody*, 66–70. <https://doi.org/10.21437/SpeechProsody.2020-14>
- Tsukada, K., Hirata, Y., & Roengpitya, R. (2014). Cross-language perception of Japanese vowel length contrasts: Comparison of listeners from different first language backgrounds. *Journal of Speech, Language, and Hearing Research*, 57(3), 805–814. https://doi.org/10.1044/2014_JSLHR-S-12-0416
- Yamakawa, K., Amano, S., & Kondo, M. (2021). Mispronunciation of Japanese singleton and geminate stops by Korean and Taiwanese Mandarin speakers. *Acoustical Science and Technology*, 42(2), 73–82. <https://doi.org/10.1250/ast.42.73>