THE CLASSIFICATION OF NATIVE AND NON-NATIVE PRODUCTIONS OF ITALIAN CONSONANT LENGTH CONTRAST

Qiang Feng, M. Grazia Busà

Università degli Studi di Padova
qiang.feng@phd.unipd.it, mariagrazia.busa@unipd.it

ABSTRACT

It has been suggested that the duration ratio between Italian consonants and preconsonantal vowels (C/pre-V ratio) is a potentially good candidate for automatic classification of Italian singleton vs. geminate consonants produced by native speakers. The optimal C/pre-V ratio threshold value for discriminating between these two categories of consonants is about 1. This triggers the question of whether this threshold also applies to the classification of non-native productions of Italian consonant length contrast. To answer this question, in this study we classified Italian singleton vs. geminate consonants produced by Chinese speakers and native Italian speakers using two approaches, namely (a) relying on native listeners' perceptual identification and (b) using the above threshold. A comparison of the classification results shows significant differences between the two approaches, suggesting that, for both native and non-native productions, using the C/pre-V ratio threshold value of 1 cannot always accurately classify Italian singleton vs. geminate consonants.

Keywords: Consonant length, singleton, geminate, Italian, Chinese.

1. INTRODUCTION

Italian consonants (i.e., /b, d, g, p, t, k, tʃ, dʒ, f, v, s, m, n, l, r/) can contrast in phonemic length. Some examples are *pena* 'penalty' vs. *penna* 'pen', *papa* 'pope' vs. *pappa* 'mush', etc. Previous investigations on the status of Italian consonant length contrast [1]–[3] show that (i) Italian geminate consonants are longer than their singleton counterparts, and (ii) Italian pre-geminate vowels are shorter than the corresponding pre-singleton ones because of an anticipatory durational compensation. Take the minimal pair of *pena* vs. *penna* as an example: the "nn" in *penna* is longer than the "n" in *pena*, and the "e" in *penna* is shorter than the "e" in *pena*.

However, what noted above seems to hold only within a given speech rate. This is because Italian consonant length is found to vary with speaking rates, and a singleton consonant produced at a slow speaking rate may be as long as a corresponding geminate produced at a fast speech rate [4]. Therefore, consonant duration may not be a suitable criterion for the classification of Italian singleton vs. geminate consonants across speech rates. By contrast, according to [4], the duration ratio between Italian consonants and preconsonantal vowels (i.e., C/pre-V ratio) may serve as a better candidate for the classification of Italian consonant length contrast across speech rates. This was confirmed by [1], [2], who showed that the optimal C/pre-V ratio threshold value for discriminating between Italian singleton and geminate consonants is about 1. Specifically, if the C/pre-V ratio is greater than 1, a geminate consonant will be detected; otherwise, it will be a singleton.

This triggers an interesting question. That is, could the C/pre-V ratio threshold value of 1 that is optimal for discriminating between Italian singleton and geminate consonants produced by native Italian speakers also apply to non-native productions?

This question is raised for two reasons. First, according to the cue-weighting theory, in determining the perceptual identity of a sound, some acoustic dimensions play a greater role than others [5]. In our case, the C/pre-V ratio works better than the absolute consonant duration in classifying native Italian speakers' singleton vs. geminate consonants. Thus, it is reasonable to speculate that Italian listeners, in perceptually identifying native productions of Italian consonant length contrast, may pay more attention to the duration interplay between Italian consonants and preconsonantal vowels than to the absolute consonant duration. However, we do not know whether Italian listeners also apply this cue-weighting strategy to their perception of non-native productions of Italian consonant length contrast. Moreover, as shown in previous studies [6]–[12], Italian singleton and geminate consonants produced by non-native speakers often differ from those produced by native Italian speakers. Therefore, it seems interesting to examine whether the ratio threshold above is also valid for the classification of non-native productions of Italian consonant length contrast.

Second, empirical studies on L2 speech acquisition often rely on the perception of native speakers to assess non-native productions. For example, to know whether an Italian consonant produced by a non-native speaker is a singleton or geminate, the most reliable way is to find out how native Italian speakers identify it perceptually.

However, the recruitment of native speakers and the development of perceptual identification tasks can be rather effortful and time-consuming. Thus, if using the above threshold could represent an alternative, more economic approach for easily and accurately classifying Italian singleton vs. geminate consonants produced by non-native speakers, future research might become less burdensome. For this reason, we wanted to test whether the above threshold can be considered valid for the classification of non-native productions of Italian consonant length contrast.

In this study, we will use both native Italian speakers' perceptual identification and the above-mentioned C/pre-V ratio threshold to classify the production of Italian singleton vs. geminate consonants by Chinese speakers and native Italian speakers. Through a comparison of the classification results of these two approaches, we aim to assess whether the C/pre-V ratio threshold value of 1 can be used to accurately classify non-native (as well as native) productions of Italian consonant length contrast.

2. METHOD

Three perceptual identification tasks were developed using the ExperimentMFC 7 in Praat [13]. The tokens to be identified in these tasks were as described in [6]. In short, they were 5 Italian minimal pairs contrasting in consonant length produced by 30 Chinese speakers and 10 native Italian speakers, both in isolation and in carrier sentences. In total, 1600 tokens were elicited (40 tokens × 40 speakers = 1600).

2.1. Listeners

Three native Italian-speaking undergraduates (one 23-year-old female and two 24-year-old males) were recruited as listeners. They were from the Veneto region in the North-East of Italy, the Abruzzo region in central Italy, and the Puglia region in the South of Italy, respectively. The reason for recruiting listeners from three different regions in Italy was to minimize the dialectal influence on the perceptual identification tasks.

2.2. Stimuli

We divided the above 1600 tokens into 3 parts. Each part served as the stimuli for one identification task. Table 1 shows the stimulus composition of each task. As the table shows, each task consisted in the identification of a set of tokens produced by 12 Chinese speakers and 4 Italian speakers; 12 of them were different across the three tasks and the remaining 4 (in bold) were identical. So, each task had 640 stimuli (40 tokens × 16 speakers = 640), 480 of which (40 tokens × 12 speakers = 480) were different across tasks and the remaining 160 (40 tokens × 4 speakers = 160) were identical.

Task	Stimuli: tokens produced by
1	CS-01, CS-02, CS-03, CS-04, CS-05, CS-06, CS-07, CS-08, CS-09, **CS-28, CS-29, CS-30,** IS-01, IS-02, IS-03, **IS-10**
2	CS-10, CS-11, CS-12, CS-13, CS-14, CS-15, CS-16, CS-17, CS-18, **CS-28, CS-29, CS-30,** IS-04, IS-05, IS-06, **IS-10**
3	CS-19, CS-29, CS-21, CS-22, CS-23, CS-24, CS-25, CS-26, CS-27, **CS-28, CS-29, CS-30,** IS-07, IS-08, IS-09, **IS-10**

Table 1: Stimulus composition of the identification tasks (CS=Chinese speaker, IS=Italian speaker).

At a perceptual examination, 43 tokens were found to contain a conspicuous misreading of the target words (e.g., *papa* 'pope' misread as *papà* 'father') and therefore were discarded. Thus, rather than 640 the actual number of stimuli in Task 1, 2, and 3 was 624, 622, and 625 respectively; and the actual number of identical stimuli in the three tasks was 157 (rather than 160).

2.3. Procedure

The listeners were randomly assigned to one of the three identification tasks. They completed the tasks with headphones on in front of a computer in the Language and Communication Lab at the University of Padova in Italy.

Each task consisted of two parts. In the first part, the tokens produced in isolation were presented in random order. Two words (i.e., a minimal pair) were shown on the computer screen as options, along with an "oops" button. The listeners were asked to click on the option that corresponded to the word heard. If they accidentally clicked on the undesired option, they could go back to the word heard previously by clicking on the "oops" button. Each word appeared only once and could be listened to up to three times when it appeared. After every 30 words there was a break.

The second part was structured similarly to the first one, except that the stimuli were produced in carrier sentences. The listeners' task was to determine which of the two options on the computer screen appeared in the sentences heard.

The tasks were self-paced, and each lasted about 25 minutes. At the end of the tasks, the listeners' responses that contained target geminate consonants (e.g., *penna*, *pappa*) were extracted as "geminate", and those containing target singletons (e.g., *pena*, *papa*) were extracted as "singleton".

2.4. Data preparation

In the first place, we had to determine the responses to the 157 stimuli that were identical across the three identification tasks. This is because all three listeners had identified them, and thus we had three different sets of responses to the stimuli. When the listeners' responses were divergent, the determination was made following the "majority rule". For example, if a stimulus was classified as "singleton" by two listeners and as "geminate" by one listener, the response would be counted as "singleton". In total, 35 stimuli required the application of the "majority rule". Afterwards, we merged the listeners' responses to the 157 stimuli with their responses to the other stimuli. In this way, we obtained the first set of classification results, which was based on the native Italian listeners' perceptual identification.

Secondly, we measured the durations of the target consonants and preconsonantal vowels in each token. Based on the duration values, we calculated the C/pre-V ratios for all tokens. The C/pre-V ratios of 27 tokens could not be calculated because the target consonants and/or preconsonantal vowels were unmeasurable. The tokens with a C/pre-V ratio smaller than or equal to 1 were classified as "singleton", and those with a C/pre-V ratio greater than 1 were classified as "geminate". In this way, we obtained the second set of classification results, which was based on the tokens' C/pre-V ratios.

2.5. Analyses

We first calculated the interrater reliability using the irr package 0.84.1 [14] in R 3.6.3 [15] based on the three listeners' responses to the 157 stimuli that were identical across the three identification tasks.

Then, we calculated the overlap rate between the two sets of classification results. By overlap we refer to the cases in which the two approaches classified one token in the same way. For example, if a token was classified as "geminate" by both approaches, there was an overlap between the two sets of classification results.

Afterwards, we calculated the Chinese speakers' and native Italian speakers' production accuracy rates. Specifically, if the speakers' target consonants were produced as intended (e.g., a "*pena*" was classified as "singleton"), they would be noted as correct. Otherwise (e.g., a "*pena*" was classified as "geminate"), they would be noted as incorrect.

Finally, we applied a generalized linear mixed-effects model (GLMM) with a binomial link function to the classification results (i.e., singleton vs. geminate) using the lme4 package 1.1.26 [16] in R 3.6.3 [15]. The GLMM had Approach (two levels: relying on native Italian listeners' identification ["by identification" for short] vs. using the C/pre-V ratio threshold ["by ratio" for short]), Group (two levels: native vs. non-native), Consonant type (two levels: singleton vs. geminate), and the interaction between them as fixed factors, and Subject and Item as random intercepts. The main effects of the fixed factors were assessed with the Type II Wald chi-squared tests using the car package 3.0.10 [17]. Post-hoc pairwise comparisons with FDR (false discovery rate) correction were implemented using the emmeans package 1.5.3 [18].

3. RESULTS

The interrater reliability was moderately high (Fleiss $k = 0.7$), indicating that though the three native Italian listeners were from different regions of Italy, they generally agreed with each other in identifying Italian singleton/geminate consonant contrast perceptually. Thus, the first set of classification results obtained on the basis of the native listeners' perceptual identification is reliable.

The overlap rate between the two sets of classification results reached 89.65% and 77.07% for the tokens produced by the native Italian speakers and the Chinese speakers, respectively. This indicates that there was more agreement between the two approaches in classifying native productions than in classifying non-native ones.

The Chinese speakers' and native Italian speakers' production accuracy rates for Italian singleton and geminate consonants are shown in Table 2. As can be seen, the production accuracy rates calculated by the two classification approaches reveal some common tendencies. That is, the Chinese speakers had lower production accuracy for either Italian singleton or geminate consonants as compared to the native Italian speakers; and the Chinese speakers' production accuracy was higher for Italian singleton consonants than for geminates. However, it should also be noted that the accuracy rates calculated by the two classification approaches are not exactly the same. Specifically, the accuracy rates calculated on the basis of the Italian listeners' identification were almost always higher than those calculated on the basis of the C/pre-V ratio threshold.

		By identification	By ratio
Non-native	Singleton	66.67%	71.50%
	Geminate	46.63%	38.51%
Native	Singleton	95.48%	88.94%
	Geminate	96.00%	89.50%

Table 2. Chinese speakers' (non-native group) and Italian speakers' (native group) production accuracy rates for Italian singleton and geminate consonants

Table 3 shows the results of the GLMM applied to the classification results of the native and non-native productions of Italian consonant length contrast. As can be seen, the GLMM yielded a significant main effect on the interaction between Approach, Group and Consonant type. Since we were particularly interested in whether there were any differences between the two classification approaches, we conducted post-hoc pairwise comparisons. The results, in Table 4, show that in the classification of native and non-native productions of both Italian singleton and geminate consonants, the two approaches performed significantly differently.

Fixed effects	χ^2	*df*	*p*
Approach (A)	13.63	1	0.0002*
Group (G)	3.85	1	0.0498*
Consonant type (C)	6.10	1	0.0135*
A×G	1.23	1	0.2669
A×C	1.45	1	0.2288
G×C	232.58	1	<.0001*
A×G×C	8.97	1	0.0027*

Table 3: Results of the GLMM applied to the classification results of the native and non-native productions of Italian consonant length contrast.

By identification vs. By ratio					
		Est.	*SE*	*z*	*p*
Non-native	Singleton	0.37	0.15	2.53	0.0115*
	Geminate	0.41	0.14	2.94	0.0032*
Native	Singleton	-0.91	0.42	-2.17	0.0303*
	Geminate	1.06	0.44	2.39	0.0171*

Table 4: Results of the post-hoc pairwise comparisons of the two classification approaches.

4. DISCUSSION AND CONCLUSIONS

In this study, we classified native and non-native productions of Italian consonant length contrast using two approaches, namely (a) relying on native listeners' perceptual identification and (b) using the consonant to preconsonantal vowel duration ratio (C/pre-V ratio) threshold (i.e., 1). The results are as follows.

Regarding native productions, there was a large degree of overlap between the classification results obtained with the two approaches, suggesting that the C/pre-V ratio does serve as an important acoustic cue in Italian listeners' perceptual categorization of Italian consonant length contrast. Nonetheless, the differences between the two approaches were significant, indicating that the (b) approach cannot always accurately classify Italian consonant length contrast produced by native Italian speakers. This is in line with [1], [2], who showed that when the C/pre-V ratio threshold value of 1 is applied to the classification of native Italian speakers' singleton vs. geminate consonants, certain error rates exist. We argue that this may be because the C/pre-V ratio is an important but not the only acoustic cue that could determine the perceptual identification of Italian consonant length category. In some cases, other acoustic cues (e.g., absolute consonant duration) may determine the perceived category membership.

As for the classification of non-native productions, the differences between the two approaches were also significant. Moreover, the differences were more evident than in the classification of native productions. This suggests that for Italian listeners the C/pre-V ratio plays a less important role in their perceptual identification of non-native productions than in their perceptual identification of native ones. We argue that this is because the Chinese speakers' productions of Italian consonant length contrast differed greatly from the native Italian speakers'. Specifically, the Chinese speakers completely ignored the duration interplay between Italian consonants and preconsonantal vowels in production [6]. Maybe for this reason, native Italian listeners are forced to rely more on other acoustic cues (e.g., absolute consonant duration) in perceptually classifying non-native productions than in perceptually categorizing native ones. In other words, Italian listeners apply somewhat different cue-weighting strategies to their perception of non-native productions of Italian consonant length contrast.

To conclude, for both native and non-native productions, using the C/pre-V ratio threshold cannot always accurately classify Italian singleton vs. geminate consonants. Therefore, for future research aiming to classify Italian consonant length contrast produced by non-native speakers, this approach cannot replace native speakers' perceptual identification. However, this study only classified Italian singleton vs. geminate consonants produced by Chinese speakers. To consolidate the present conclusions, future studies including non-native productions by speakers of other languages are needed.

5. REFERENCES

[1] M. G. Di Benedetto and L. De Nardis, "Consonant gemination in Italian: The nasal and liquid case," *Speech Communication*, vol. 133, pp. 62–80, 2021, doi: https://doi.org/10.1016/j.specom.2021.07.006.

[2] M. G. Di Benedetto and L. De Nardis, "Consonant gemination in Italian: The affricate and fricative case," *Speech Communication*, vol. 134, pp. 86–108, Nov. 2021, doi: 10.1016/j.specom.2021.07.005.

[3] 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, no. 4, pp. 2051–2062, Oct. 1999, doi: 10.1121/1.428056.
[4] E. R. Pickett, S. E. Blumstein, and M. W. Burton, “Effects of Speaking Rate on the Singleton/Geminate Consonant Contrast in Italian,” *Phonetica*, vol. 56, no. 3–4, pp. 135–157, Dec. 1999, doi: 10.1159/000028448.
[5] L. L. Holt and A. J. Lotto, “Cue weighting in auditory categorization: Implications for first and second language acquisition,” *The Journal of the Acoustical Society of America*, vol. 119, no. 5, pp. 3059–3071, 2006.
[6] Q. Feng and M. G. Busà, “Mandarin Chinese-speaking learners’ acquisition of Italian consonant length contrast,” *System*, vol. 111, p. 102938, Dec. 2022, doi: 10.1016/j.system.2022.102938.
[7] C. Celata and L. Costamagna, “Timing delle geminate nell’italiano L2 di apprendenti estoni,” *Quaderni del laboratorio di linguistica*, vol. 10, no. 1, pp. 1–27, 2011.
[8] B. Kabak, T. Reckziegel, and B. Braun, “Timing of second language singletons and geminates,” in *Proceedings of the 17th International Congress of the Phonetic Sciences (ICPhS XVII)*, W.-S. Lee and E. Lee, Eds., Hong Kong: City University of Hong Kong, 2011, pp. 994–997.
[9] S. D’Apolito and B. G. Fivela, “L2 Pronunciation Accuracy and Context: A Pilot Study on the Realization of Geminates in Italian as L2 by French Learners.,” in *INTERSPEECH*, 2019, pp. 1706–1710.
[10] 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, no. 2, pp. 3–29, 2014.
[11] Q. Feng and M. G. Busà, “How Do L1 Mandarin Chinese Learners Produce L2 Italian Singleton and Geminate Consonant Contrasts?,” *Pronunciation in Second Language Learning and Teaching Proceedings*, vol. 12, no. 1, 2022.
[12] V. Galatà, G. Meneguzzi, L. Conter, and C. Zmarich, “Primi dati sull’acquisizione fonetico-fonologica dell’italiano L2 in prescolari rumeni,” in *La voce nelle applicazioni*, A. Paoloni and M. Falcone, Eds., Roma: Bulzoni Editore, 2012, pp. 35–50.
[13] P. Boersma and D. Weenink, “Praat: doing phonetics by computer.” 2020. [Online]. Available: http://www.praat.org/
[14] M. Gamer, J. Lemon, and I. Singh, “irr: Various Coefficients of Interrater Reliability and Agreement.” 2019. [Online]. Available: https://CRAN.R-project.org/package=irr
[15] R Core Team, “R: A language and environment for statistical computing.” R Foundation for Statistical Computing, Vienna, Austria, 2020. [Online]. Available: https://www.R-project.org/
[16] D. Bates, M. Maechler, B. Bolker, and S. Walker, “Fitting Linear Mixed-Effects Models Using lme4,” *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015, doi: 10.18637/jss.v067.i01.
[17] J. Fox and S. Weisberg, “An {R} Companion to Applied Regression.” Sage, Thousand Oaks CA, 2019. [Online]. Available: https://socialsciences.mcmaster.ca/jfox/Books/Companion/
[18] R. Lenth, “emmeans: Estimated Marginal Means, aka Least-Squares Means.” 2020. [Online]. Available: https://CRAN.R-project.org/package=emmeans