



L2 Pronunciation accuracy and context: a pilot study on the realization of geminates in Italian as L2 by French learners

Sonia d'Apolito, Barbara Gili Fivela

Università del Salento - CRIL

sonia.dapolito@unisalento.it, barbara.gili@unisalento.it

This paper investigates the interaction between the characteristics of both L1 and L2 phonetic-phonological systems and how the context, in terms of the amount of information available (less vs more information), may influence the accuracy in producing L2 sounds as well as speech fluency. Specifically, it focuses on how French learners of Italian as L2, representing two different competence levels (lower and higher), realize geminates (non-native sounds) in two different contexts (less and more rich). A rich context is expected to induce lower accuracy. Acoustic data of nine subjects (three beginners, three advanced and three natives as control) were collected and analyzed in order to observe: 1) the realization of geminates (duration of the consonant and preceding vowel as an index of accuracy); and 2) the speech fluency (number and duration of disfluencies; speech/articulation rate). Results suggest that learners' productions are affected by L1, above all in the case of beginners, who show a lower degree of accuracy. As regards the accuracy and context interaction, results show that the production of geminates is more accurate (longer duration) in poor than in rich context. Further, a higher number of disfluencies is found in rich than in poor context.

Index Terms: geminates, Italian as L2, context, L2 speech accuracy, L2 speech fluency.

1. Introduction

It is well known that accuracy in L2 speech is affected by the characteristics of L1 [1] and, thus, investigating accuracy in L2 pronunciation offers an overview of the interaction of the L1 and L2 systems. For instance, the production of Italian geminates by French speakers is expected to lack in accuracy, as gemination is not familiar to French. The duration of consonants is a contrastive feature in Italian so that words have a different meaning in case a singleton (C) or a geminate (CC) consonant appears [2]. Moreover, gemination implies a change not only in the duration of the consonant but also in the duration of the preceding vowel, which is shorter when followed by CC and it is longer before C [3]. On the contrary, French does not have gemination, except for the uvular approximant within the opposition between the imperfect tense and the conditional mood (e.g. *pourait vs. pourrait*) [4]. Thus, gemination represents an interesting phenomenon to investigate the interaction between a French-L1 and an Italian-L2 system.

The production of the geminates can be problematic for L2 learners [5,6,7,8] and accuracy in production also has an important impact on comprehension. Thus, studying pronunciation accuracy of L2 non-native sounds gives suggestion on learner's ability in communication. On the one hand, the ability to speak accurately and at comfortable rate

(without long pauses and/or hesitations, that is fluently) allows learners to express freely their own thoughts [9] and to be comprehensible; on the other hand, it allows listeners to comprehend the message without much effort. Nevertheless, in the literature, phonetic and phonological errors are reported to account only for about 22% of L2 speech comprehension [10,11]. However, studies conducted on L2 pronunciation often regard English as L2. Further, the context in which communication takes place may heavily affect the accuracy in producing the intended message as well as its comprehension (see Hyper and Hypo-articulation theory [11]).

Thus, it may be interesting: 1) to observe the interaction between phonetic-phonological systems different from English, such Italian as L2 spoken by French learners; and 2) to investigate the role of context on both production accuracy and speech fluency.

2. Goals and hypotheses

In this study, the singleton and geminate consonants are realized by Italian speakers and by French learners of Italian as L2 (two different competence levels, i.e. low and high proficiency level) in two different contexts in which the amount of information available varies (poor vs. rich context). The contrast (measured in terms of duration) is observed in order to investigate: 1) the interaction between the L1 and L2 phonetic-phonological systems in the production of non-native sounds (geminates) by French speakers, who reached two different competence levels; 2) the interaction between the amount of information available in the context (less vs more) and L2 speech, in terms of: a) the accuracy in producing L2 non-native sounds; and b) speech fluency (disfluencies and speaking/articulation rate).

The hypotheses are that: 1) French learners, due to the influence of their L1, may reduce or substitute the non-native sounds to/with native sounds; moreover, a lower accuracy is expected by beginners [1]; 2a) information available in the context affects accuracy, as lower accuracy (duration) is expected when the context is rich, in that available information may help to disambiguate the message [11]; and 2b) the context affects speech fluency (speaking and articulation rate), as a higher number disfluencies is expected in the rich context, due to the greater attention to context details required by the task.

3. Method

Nine subjects participated in the experiment. Six French learners were recruited among Erasmus students at University of Salento (female; mean age 21.5: they came from Paris (2), Nantes (3) and Nancy (1)). According to the Erasmus test results, they were gathered into two groups: low competence

level (L1; L2; L3) and high competence level (H1; H2; H3). The advanced learners studied Italian language for 4.5 years, while the beginners for 2.5 years; all of them spoke Italian language during their Erasmus period and have never been in Italy before such period. Three native speakers as control (C1, C2, C3) were recruited among students at University of Salento (female; age: 23.6; from Maglie (LE), in Salento).

L2 Italian sounds of interest were /t, d, s, n, l, r/, both as singletons (C) and as geminates (CC). For each phoneme two minimal pairs were found (except for the phoneme /d/ which only shows one minimal pair). All speakers produced the words, including the target sounds, inserted in a sentence elicited in two interactional contexts: 1) poor context, which does not facilitate disambiguation, namely a carrier phrase (Table. 1A); 2) rich context, which does help to disambiguate the words, that is an interaction appropriate to the meaning of the target word (Table. 1B). Minimal pairs were inserted in both initial and final position of the phrase (22 words x 2 contexts x 2 positions x 3 repetitions= 264 phrases).

Table 1: An example of the corpus

| A. Poor context |
|---|
| Cosa hai detto? – Maria ha detto sera/serra di nuovo What did you say? – Mary said evening/greenhouse again |
| B. Rich context |
| Cosa ti va di fare? – Questa sera vorrei andare al cinema What would like to do? – This evening I would like to go to the cinema |

Acoustic segmentation was performed in PRAAT [12] in order to label boundaries of the phrase, word and target segments C1V1C2V2 (where C2= C/CC). Acoustic analysis related to: a) duration and normalized duration (segment duration/word duration) of the target consonant segment (C2) as well as the preceding vowel (V1); and b) speech fluency in terms of speaking (SR) and articulation rate (AR - excluding silent pauses), as well as number and duration of pauses/hesitations. Statistical t-tests were run for comparisons separately for each speaker ($p < 0.05$) in order to observe any individual difference due to different backgrounds.

4. Results

4.1. Geminate vs singleton consonants

Both duration and normalized duration were calculated, though only significant results for normalized duration are reported here. In the poor context, all control speakers show a significantly longer duration in CC than in C for all phonemes. Advanced learners show results similar to controls, while in the case of beginners the duration differs significantly (with longer CCs than Cs) for two out of three subjects: L2, though only as for /s/, and L3 as for /l, n, s, t/. The beginner L1, on the contrary, does never differentiate CC from C.

In the rich context, native speakers confirm results found in the poor context, that is CCs are significantly longer than Cs. Such results are also confirmed in the case of advanced learners, except for /d/ and /n/ by H3 (which, however, shows the same tendency). As for beginners, L1, in line with poor context results, does not show any difference in duration between CCs and Cs. However, L2 shows a significantly longer duration for CCs in the case of /l/, as well as for /s/, and

L3 shows a longer duration in the case of /l, n, s/ and /d/. Results are reported in Figure 1 and in Table 2.

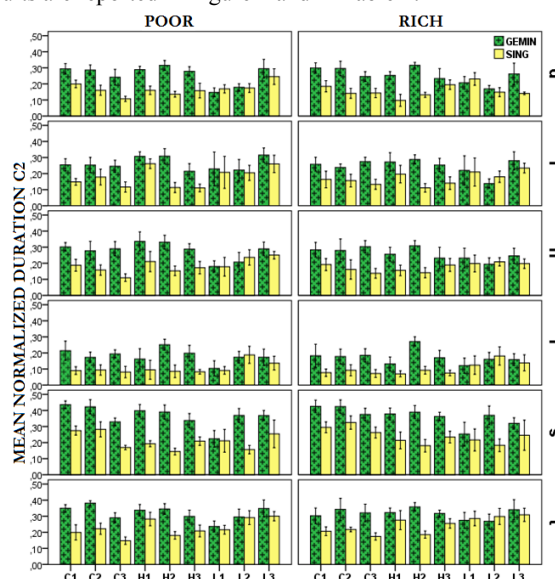


Figure 1: Bar graphs for C/CC normalized duration in poor (left) and rich (right) context for geminate (green) and singleton (yellow) for all speakers

Table 2: T-test results C2 normalized duration

| Spk | t-test | Poor cont. | Rich cont. | Durat. |
|-----|------------|--------------------|--------------------|--------|
| C1 | $p < 0.05$ | /d, l, n, r, s, t/ | /d, l, n, r, s, t/ | +CC |
| C2 | $p < 0.05$ | /d, l, n, r, s, t/ | /d, l, n, r, s, t/ | +CC |
| C3 | $p < 0.05$ | /d, l, n, r, s, t/ | /d, l, n, r, s, t/ | +CC |
| H1 | $p < 0.05$ | /d, l, n, r, s, t/ | /d, l, n, r, s, t/ | +CC |
| H2 | $p < 0.05$ | /d, l, n, r, s, t/ | /d, l, n, r, s, t/ | +CC |
| H3 | $p < 0.05$ | /d, l, n, r, s, t/ | /l, n, r, s, t/ | +CC |
| L1 | n.s. | never | never | - |
| L2 | $p < 0.05$ | /s/ | /l, s/ | +CC |
| L3 | $p < 0.05$ | /l, n, s, t/ | /d, l, n, s/ | +CC |

Statistical tests run with context as a factor show that, as for controls, geminates differ significantly in very few cases (C2: in the case of /t/, which is longer in poor context; C3: in the case of /s/, which is longer in rich context). However, we hypothesized that the context could be crucial for learners. For advanced learners, the test reaches the significance for H1, as for /d/ and /n/, and in the case of H3, as for /n/. For both speakers, the duration is longer in poor context. As for beginners, the test can be performed only for L2 (/s/) and L3 (/l, n, s/) - L1 does not produce geminates. However, only L3 (/n, s/) shows a longer duration in poor than in rich context.

Table 3: T-test results for context as factor

| Spk | t-test | Geminates | Durat. |
|-------|------------|-----------|--------|
| C1 | n.s. | never | - |
| C2 | $p < 0.05$ | t/ | +poor |
| C3 | $p < 0.05$ | /s/ | +rich |
| H1 | $p < 0.05$ | /d, n,/ | +poor |
| H2 | n.s. | never | - |
| H3 | $p < 0.05$ | /n / | +poor |
| L1/L2 | n.s. | never | - |
| L3 | $p < 0.05$ | /n, s/ | +poor |

As regards the duration of the vowel (V1) preceding the geminate/singleton consonant, in the poor context, for control speakers C2 and C3 the duration of V1 is significantly shorter when a geminate follows, for all phonemes. As for speaker C1, the preceding vowel is significantly shorter before a geminate only when /d, r, s/ follows. Similarly to natives, for the advanced learners H1 and H2 the duration of V1 is always shorter when followed by a geminate, for all phonemes apart of /t/ for H1. For the learner H3, V1 is significantly shorter before CC only in the case of /d, n, s/. As for beginners, L1 never differentiates CCs from Cs and, consistently, V1 does not differ; L2 distinguishes the geminate from the singleton only in the case of /s/, and in such case V1 is significantly shorter when followed by a geminate. L3 differentiates CC from C for /l, n, s, t/, but the duration of V1 is significantly shorter before geminate only in the case of /s/.

As for the rich context, the native speaker C3 shows a significantly shorter V1 duration before CC for all phonemes, except /d/; as for C1 the test is significant only in the case of /r, t/; while for C2 only /t/ offers significant results. As for the advanced learner H2, the duration of V1 is always significantly shorter before a geminate, in line with results of C2 duration. Despite the fact that even H2 and H3 always distinguish C from CC, V1 is significantly shorter before geminates only in the case of /n, s/ for H1. As for beginners, the duration of V1 never differs for L1, confirming that C and CC are not distinguished; as for L2, vowel duration is significantly shorter before CC only in the case of /s/, while C2 duration shows significant results for /l/ too (see above); along similar lines, L3 distinguishes CC from C in the case of /d, l, n, s/, but V1 is significantly shorter before CC only for /n/. Results are reported in Figure 2 and in Table 3.

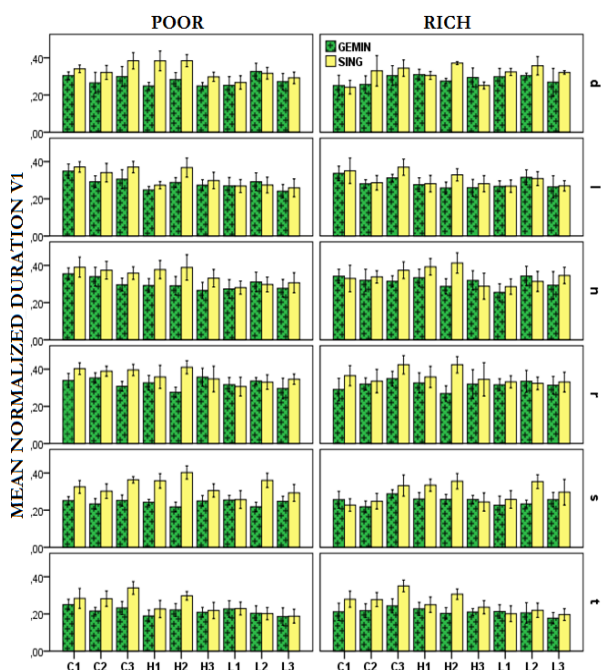


Figure 2: Bar graphs for V1 normalized duration in poor (left) and rich (right) context for geminate (green) and singleton (yellow) for all speakers

Table 3: T-test results V1 normalized duration

| SPK | t-test | Poor cont | Rich cont. | Durat. |
|-----|--------|--------------------|--------------------|--------|
| C1 | p<0.05 | /d, r, s/ | /r, t/ | -VICC |
| C2 | p<0.05 | /d, l, n, r, s, t/ | /t/ | -VICC |
| C3 | p<0.05 | /d, l, n, r, s, t/ | /l, n, r, s, t/ | -VICC |
| H1 | p<0.05 | /d, l, n, s, t/ | /n, s/ | -VICC |
| H2 | p<0.05 | /d, l, n, r, s, t/ | /d, l, n, r, s, t/ | -VICC |
| H3 | p<0.05 | /d, n, s/ | /t/ | -VICC |
| L1 | n.s. | never | never | - |
| L2 | p<0.05 | /s/ | /s/ | -VICC |
| L3 | p<0.05 | /s/ | /n/ | -VICC |

4.2. Speech fluency

Disfluencies realized by speakers are pauses, repetitions of syllables or words and lengthening of segments. Table 4 shows how disfluencies are distributed among speakers in both contexts. Generally, it is observed that both context and competence level affect speech fluency. As for context, learners realize a higher number of disfluencies in rich than in poor context (except for H2). As for the competence level, beginners realize a higher of disfluencies than advanced learners, and this is particular evident in rich context. Moreover, in poor context also the word containing the target consonant (geminate) seems to play a role, since the advanced learners H2, the beginners L1 and L3 and the native C3 realize above all pauses in conjunction with the word containing a geminate. On the contrary, in rich context the realization of disfluencies is mainly linked to the richness of the context.

Table 4: Realization of disfluencies and their duration in poor and in rich context for all speakers

| Poor context | | | | |
|--------------|----|-----------|-------------|-------------|
| SPK | N | Pause | Repetitions | Lengthening |
| C1 | 1 | 255ms | | |
| C2 | 2 | | 254ms | |
| C3 | 16 | 133ms | | |
| H1 | 1 | 255ms | | |
| H2 | 36 | 166ms(35) | | 270ms(1) |
| H3 | 5 | 259ms | | |
| L1 | 28 | 278ms(20) | 631ms(8) | 383ms(5) |
| L2 | 1 | | 401ms | |
| L3 | 11 | 253ms(10) | 651ms(1) | |
| Rich context | | | | |
| C1 | 2 | 194ms(1) | 230ms(1) | |
| C2 | 3 | 130ms(3) | 373ms(2) | |
| C3 | 4 | 150ms(3) | 450ms(1) | |
| H1 | 10 | 295ms(10) | 632ms(5) | |
| H2 | 12 | 217ms(8) | 296ms(4) | |
| H3 | 26 | 259ms(19) | 360ms(10) | 253ms(7) |
| L1 | 74 | 413ms(28) | 781ms(37) | 360ms(30) |
| L2 | 7 | 566ms(7) | 416ms(3) | |
| L3 | 50 | 435ms(32) | 850ms(19) | 338ms(20) |

Results of t-test for both Speech and Articulation Rate (respectively, SR and AR) are shown in Table 5 and in Figure 3. As shown, all control speakers speak significantly faster in rich context than in poor one. As for the advanced learners, a significant difference is found for H2 for both SR and AR, while in the case of H3 for SR only. The learner H2 speaks

faster in rich context, on the contrary the learner H3 speaks faster in poor context. All beginners speak differently according to the context. Greater fluency is shown by both L1 and L3 in poor context, while L2 shows it in rich context.

Table 5: *T*-test results for SR and AR for all speakers

| SPK | Speaking rate | Articulation rate |
|-----|---------------|-------------------|
| C1 | p<0.05 | p<0.05 |
| C2 | p<0.05 | p<0.05 |
| C3 | p<0.05 | p<0.05 |
| H1 | n.s | n.s |
| H2 | p<0.05 | p<0.05 |
| H3 | p<0.05 | n.s. |
| L1 | p<0.05 | p<0.05 |
| L2 | p<0.05 | p<0.05 |
| L3 | p<0.05 | p<0.05 |

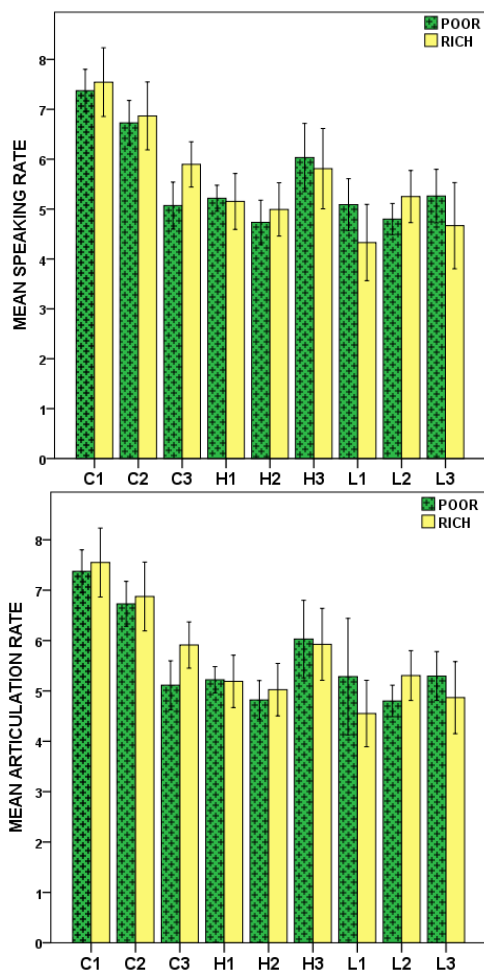


Figure 3: Bar graphs for SR (above) and AR (below) in poor (green) and rich (yellow) context for all speakers

5. Discussion and conclusions

This acoustic study focuses on the realization of geminates in Italian as L2 by French learners who reached two different competence levels (low vs high proficiency). The goal was to observe: 1) how the L1 and L2 phonetic-phonological systems

interacts as for non-native sounds like geminates, by also taking a look at how the competence level affects the accuracy in producing geminates. Further, the study takes into account the context in terms of amount of information available (less vs more information) in order to observe: 2) the interaction between context and L2 speech both in terms of: a) the accuracy in producing L2 sounds (duration); and b) speech fluency (number of disfluencies and speaking and articulation rate). We hypothesized that the production of geminates was influenced by the L1 and also depended on the competence level, that is a lower accuracy was expected by learners and above all by beginners. As for the context, we expected a lower accuracy and a higher disfluency in the case of rich rather than poor context.

In line with our first hypothesis, the production of French learners is affected by L1 phonetics and phonology and by the competence level. Advanced learners show results to similar control speakers', producing longer geminates than singletons in the case of all phonemes. In addition, even the preceding vowel is shortened before a geminate. On the contrary, the beginners' production varies from a very low to a higher degree of accuracy: as for the lower accuracy, see learner L1, who shows no distinction at all, and L2 who distinguishes CC from C in few cases (/s/ in both poor and rich context and /l/ in rich context only); as for the higher degree of accuracy, think of L3, who differentiates CC from C in the case of /l, n, s/ in both poor and rich context, and in the case of /t/ or /d/ depending on the context. However, when they differ, CCs are longer than Cs and the vowel is shorter when followed by a geminate.

As regards the interaction between context and accuracy, results show that geminates are longer in poor than in rich context. However, it is rather the preceding vowel that seems to change with reference to the richness of the context. In particular, controls seem to shorten the vowel preceding geminates more in poor than in rich context and in more cases than learners do, especially as for beginners. This seem to support the hypothesis that speakers pay more attention to produce accurately in poor context (if not by lengthening the geminate consonant, by shortening the preceding vowel) because no other elements within the phrase can facilitate disambiguation. Finally, as expected, a higher number of disfluencies is related to the richness of the context, and beginners realize more disfluencies than other speakers. Speech and articulation rate show that all control speakers, one advanced (H2) and one beginner (L2) speak faster in rich than in poor context, while one advanced learner (H3) and two beginners (L1 and L3) speak faster in poor than in rich context, since in rich context they realize a high number of disfluencies.

In conclusion, these preliminary results suggest that the accuracy in production of non-native sounds is not only due to the influence of mother tongue phonetics and phonology, but it is also influenced by the amount of the information available in the context. Moreover, context affects speech fluency too.

However, further work is needed to confirm these findings. Other acoustic measurements will be taken into account, such as fricative CoG and formant values for nasals and laterals; further, other statistical tests will be performed, such as mixed models. Overall, the work would benefit from a larger set of experimental subjects.

6. References

- [1] J. E. Flege, O. S. Bohn, & D. Meador D., "Native Italian speakers' production and perception of English vowels", *Journal of Acoustical Society of America*, 106, 2973-2987, 1999.
- [2] A. Esposito, M. G. Di Benedetto, "Acoustical and perceptual study of gemination in Italian stops", *Journal of the Acoustical Society of America*, 106 (4), 2051-2062, 1999.
- [3] P. M. Bertinetto, *Strutture prosodiche dell'italiano*, Firenze, Accademia della Crusca, 1981.
- [4] A. Battye, M. Hintze, P. Rowlett, *The French language today: A linguistic introduction*, London: Routledge, 2003.
- [5] M. Han, "The timing control of geminate and single stop consonants in Japanese: a challenge for nonnative speakers", *Phonetica* 49, 102 – 127, 1992.
- [6] J. Mah, and J. Archibald, "Acquisition of L2 Length Contrasts", in Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002), ed. Juana M. Liceras et al., 208-212, 2003.
- [7] B. Kabak, T. Reckziegel, and B. Braun., "Timing of second language geminates and singletons" , in Proceedings of the 17th International Congress of the Phonetic Sciences: 994–97, 2011.
- [8] P. Sorianello, Italian geminate consonants in L2 acquisition, Consonant gemination in first and second language acquisition in Costamagna L. & C. Celata (eds.), Pisa, Pacini Editore, pp. 25-46, 2017.
- [9] S. Zhang, "The role of input, interaction and output in the development of oral fluency", *English language teaching*, vol. 2 (4), 91-100, 2009.
- [10] L. Smith and C. Nelson, "World Englishes and issues of intelligibility" *The handbook of world Englishes*, edited by Kachru B., Kachru Y., Nelson, C., Blackwell Publishing Ltd, 428-445, 2006.
- [11] B. Lindblom, "Explaining phonetic variation: a sketch of the H&H theory", *Speech production & speech modeling*, in Hardcastle & Marchal (Eds.) Dordrecht, 403-439, 1990.
- [12] P. Boersma and D. Weenink, Praat: doing phonetics by computers, 2008.