Subtle Changes in L1 Stops of Late Salento Italian-French Bilinguals: An Acoustic Study Using AutoVOT Adapted for Italian and French

Marie Fongaro[1(✉)], Barbara Gili Fivela[2], Maud Pélissier[3], and Gabriel Hévr[4]

[1] University of South Bohemia in České Budějovice, České Budějovice, Czech Republic
mfongaro@ff.jcu.cz

[2] University of Salento, Lecce, Italy
barbara.gilifivela@unisalento.it

[3] Université Paris Cité, ALTAE, 75013 Paris, France
maud.pelissier@u-paris.fr

[4] Czech Technical University, Prague, Czech Republic
hevrgabr@fit.cvut.cz

Abstract. This study examines the influence of the second language (L2) on the native language (L1) of Italians from Salento (Southern Italy) who have started to learn French, i.e., their L2, mostly in adolescence and have moved to the Paris region as adults; they have been living in that region by different amounts of time and using French on a daily basis. Unlike French, Italian shows both geminate and singleton stops; the contrast mainly lies in stop length, even though the duration of the preceding vowel also changes. Because of this difference between languages, we investigated, at the phonetic level, the L1 singleton and geminate stops of 15 late Salento Italian-French bilinguals by comparing their L1 speech production with that of 15 controls, who are Italians born and living in Salento. Normalised acoustic duration of the stops in word-internal intervocalic position as well as that of the vowels preceding these stops was analysed. Results on stops show that bilinguals and controls significantly differ in the normalised duration of some, but not all, investigated stops; for vowels, there is no significant difference. Another key point of this paper is the description of the version of AutoVOT adapted for French and Italian, which was developed and used to obtain an accurate normalized duration of stops and vowels in this study.

Keywords: L2 influence on L1 stops · Late bilinguals · AutoVOT for Italian and French

1 Introduction

In today's world, many people move abroad as adults. Living in their non-native language (L1) country, they use a second language (L2) on a daily basis and,

A. Karpov and G. Gosztolya (Eds.): SPECOM 2025, LNAI 16187, pp. 241–255, 2026.
https://doi.org/10.1007/978-3-032-07956-5_17

if they acquired the L2 after the age of six, they are considered to be 'late bilinguals' according to the terminology of [1]. Cross-Linguistic Influence (CLI), here the influence of one speaker's language on another speaker's language [2], has often been studied in late bilingual L1 and L2 speech at the phonetic level, over-focussing on (1) the influence of L1 on L2 of bilinguals (for an overview, see [3]), and (2) the L1 speech of late bilinguals whose L1 or L2 was English (see [4]). Moreover, the L2 influence on L1 has been studied more in some phonetic segmental and suprasegmental features than in others (for an overview, see, e.g., [5]). For instance, very few authors have addressed the L2-L1 influence with a focus on the length of stop consonants as this contrasts phonologically in few languages only. This study fills these gaps by dealing with stop consonants in a not yet explored L1 speech (here Italian) of late bilinguals from Salento (geographical area in South of Italy) who have lived in the Paris region for varying lengths of time: they moved to the Paris region as adults and started to learn French, their L2, mainly in adolescence (hereafter late Salento Italian-French bilinguals – B). We compare the L1 speech of B with that produced by controls who are Italians born and living in Salento (hereafter C).

1.1 Differences Between Italian and French Stop Consonants

Italian and French stops fundamentally differ as in Italian, the stops (/p/, /t/, /k/, /b/, /d/, /g/) may be found as both singleton (e.g., /fato/ 'fate') and geminate (e.g., /fatto/ 'fact'). Geminates are usually analysed as bifonematic and affiliated to different syllables, but the consonant duration within the word is one of the main correlates differentiating minimal pairs. On the contrary, French stops (/p/, /t/, /k/, /b/, /d/, /gg/) do not show similar duration differences. To illustrate, the Italian word 'se/t/e' means 'thirst', while 'se/tt/e' means 'seven'; the orthographic difference between the words consists in the use of single or double consonant (i.e., *sete vs sette*). Also in French, there are single and double consonants, but they do not differ in pronunciation; for instance, the French words *date*, which means 'date' with the sense of a day of the year as specified by a number, and *datte*, which means 'date' with the sense of fruit, are both pronounced as /dat/. In Italian, the geminates occur in word-internal intervocalic position and they primarily differ from singletons in their duration but also the duration of the preceding vowel changes; the geminates are longer than singletons (their duration was found to be about twice the singleton consonant duration) and the vowel preceding geminates is shorter than the one preceding singletons [6,7]. This difference in singleton and geminate duration and in the duration of the preceding vowel also exists in Salento dialects and the variety of Italian spoken in Salento that both B and C have spoken and have been exposed to [8].

1.2 The L2 Influence on L1 Singleton and Geminate Consonants in SLM, SLM-R and L2LP Models

To date, the L2 influence on L1 singleton and geminate consonants have been investigated only in three studies: The first was conducted by [9] on the perception of singleton/geminate consonant contrast by immigrants from Lucca (Italy) living in California. Contrary to American English, the singleton *vs* geminate contrast is present in the Lucchese dialect even if degemination occurs in Lucchese dialect depending on a speaker's education, family context, social class and others. The authors tested the contrast perception between selected singleton and geminate consonants (/r/, /s/, /t/ *vs* /rr/, /ss/, /tt/) on three groups of subjects: 8 Lucchese immigrants living in California from 28 to 54 years (i.e., first-generation immigrants), 7 s-generation immigrants (i.e., children of the first generation), and 16 Italian controls living in Lucca. The results of two perception experiments (one including real words, another including non-words) showed that first-generation immigrants maintained the ability to discriminate the consonant contrast in the non-words but not in the real words. This ability wasn't found neither in real and nor non-words in second generation immigrants.

The second and third study were acoustic studies. In the second study [10], the authors examined the duration of the closure portion in singleton and geminate stops produced in a word repetition task by first and second generation of Palestinian Arabic and Italian immigrants living in the US and the respective Palestinian Arabic and Italian controls. The results showed a general significance of the factors GROUP (first-, second-generation or control speakers), VOICING (voiced/-less stop) and CONSONANT STATUS (singleton/geminate) on the duration of the stop closures in both languages. In Italian, the closure duration of stops was often similar in the production of first and second generation speakers and shorter than in the production of controls. The third study was conducted by [11] on the duration of singleton and geminate consonants in a word-naming task by four groups of speakers: first-, intermediate- and second-generation of Iranian immigrants (Farsi speakers) living in Toronto, and Farsi controls. The intermediate-generation was composed of children of Iranian immigrant families that arrived at an age between 5 and 14 in an English-speaking country, while the second generation consisted of children of Iranian immigrants who were either born there or arrived there before the age of 5. Results showed that geminate duration shortens across successive generations.

Researchers have widely used the Speech Learning Model (SLM, [12]) and its revised version (SLM-r, [13]) to predict CLI, including L2-L1 influence in the speech of bilinguals (see, e.g., [14]). The core suppositions of SLM and SLM-r for studies on L2-L1 influence lies in a conception of L1 and L2 phonetic systems of bilinguals as existing in a common phonetic space possibly inducing the interaction between the L1 and L2 sounds. This interaction may lead to non native-like L2 speech production as well as to less native-like L1 speech production because of changes in the L1 of a speaker (see, e.g., [15,16]). However, whether or not and to what extent the changes in L1 and L2 of a speaker occur often varies considerably between individuals, as it depends on many factors described in SLM-r (for instance, endogenous factors related to an individual speaker, the degree of perceived phonetic dissimilarity of L2 sound from the closest L1 sound, the quantity and quality of L2 input the speaker received for the given L2 sound in 'meaningful conversations'). Similarly, there is a supposition in the Second Language Linguistic Perception model (L2LP, [17]) that to maintain the optimal L1 and L2 perception and production, speakers must be exposed to rich L1 and L2 input, otherwise their L2 will affect their L1. Hence, the results of the three studies presented above ([9–11]) go in the direction of what is assumed by the models: We may assume that the less the speakers are exposed to singleton-geminate contrast, the bigger is a decline in their ability to maintain this contrast, as it is the case across successive generation of immigrants, because the L1 input that immigrants receive decreases with each generation.

2 Goal and Hypotheses

The goal of this work is to investigate, in the context of the L2-L1 influence research, the possible changes in L1 stops of B. In particular, the analysis focuses on the duration of (1) singleton and geminate Italian stops and (2) vowels that precedes these stops in the word (hereafter 'preceding vowels') in the L1 speech production of B and compares it with the L1 speech production of C. In view of what discussed in the introductory section, and supposing B are exposed to less rich L1 input than C as living in their not L1-country, we hypothesise that:

1. L1 geminate stops of B will differ in their duration from those of C as a consequence of the possible influence of the L2 of B on their L1.
2. L1 preceding vowels of B will differ in their duration from those of C as a consequence of a possible influence of the L2 of B on their L1.
3. The amount of time B have lived in France (hereafter Length of Residence – LOR) will affect the duration of L1 geminate stops and L1 preceding vowels of B.

3 Method

3.1 Speakers, Speech Recording and Material

For the purpose of the study, we recorded the L1 of 15 B (9M, 6F; mean age = 41.13 y.o.; SD = 10.39) who have lived in France for various amount of time (LOR [year]: 1, 4, 6, 8, 8, 10, 12, 14, 14, 15, 18, 22, 24, 27, 33), and the L1 of 15 C (9M, 6F) matched as closely as possible to B for age, sex and education level. All speakers were asked to complete a short questionnaire concerning details about their LOR, language(s) use and background, the places they lived in and the L2 acquisition. Both B and C had Italian parents and declared to speak only Italian and/or a Salento dialect until the age of seven. The B were born and/or grew up in the geographical area of Salento, and moved to France as adults. The C were born and/or grew up in Salento and have lived there for all or most of their lives. The C declared using mainly Italian and/or Salento dialects in their everyday life; the B declared using both French and Italian and, in some cases, also Salento dialects. One B reported using English at work. All speakers reported no speech disorder.

The majority of B was recorded in Paris while C and a few B were recorded in Lecce (Salento). These few B were recorded during a summer vacation, a few days after their arrival in Salento, to avoid a long immersion into the Salento Italian variety, i.e., ensuring a linguistic context comparable to that of the recordings made in Paris. The recording took place in a quiet room, using a Neumann KMS 105 microphone and a Focusrite scarlett solo 3rd Gen USB sound card. Since the L2-L1 influence was found to be more obvious in spontaneous than in reading-elicited speech (see, e.g., [4,18]), we opted for the elicitation of a sort of spontaneous speech: we created a set of pictures evoking mostly two-syllable and a few three-syllable Italian words (hereafter 'target words') with Italian stops (/p/, /pp/, /t/, /tt/, /k/, /kk/, /b/, /bb/, /d/, /dd/, /g/, /gg/) placed in intervocalic word-intermediate stress-controlled position (hereafter 'target stops'). The singleton and geminate stops forming a pair (e.g., /t/ and /tt/) were placed in the target words in similar vocalic contexts (e.g., for the pair involving /t/ and /tt/, we had the target Italian words 'se/t/e' and 'se/tt/e' meaning 'thirst' and 'seven' respectively, and 're/t/e' and 're/tt/e' meaning 'net' and 'tuition' respectively). Each target stop occurred in two target words.

The recording included two main phases. During the first phase, a target word X was shown on the PC screen by the experimenter and the speaker had to produce it in the carrier sentence 'I say X'. During the second phase, more pictures were presented on the PC screen and, for each target word X, the speaker had to describe the way the experimenter was moving it on the screen using the carrier sentence 'I put X next to Y. I moved X'. Hence, a single speaker produced each target word four times, that is, s/he produced it the first time during the first phase of the recording (i.e., the carrier sentence 'I say X'), the second, third and fourth time, s/he produced it during the second phase of the recording (i.e., the carrier sentence 'I put X next to Y. I moved X'). A single target word occurred first as X, then as Y. We obtained a total of 30 recordings (one per speaker; 15 recordings from B and 15 recordings from C). For the analysis, we included all target words contained in these recordings except for some few not well-pronounced target words (e.g., a speaker produced a hesitation inside the target word). Thus, we had 2859 target words for the analysis; we analysed 2859 target stops and 2859 preceding vowels contained in these words (one target stop and one preceding vowel per target word if pronounced, see Table 1 for more details).

3.2 Acoustic Analysis

First of all, it was necessary to determine the acoustic parameter to be used in the analysis to verify our three hypothesis. The authors widely used the ratio of consonant to word duration to acoustically examine the difference between Italian geminate and singleton consonants (see, e.g. [6,19]), but to examine this difference, they focused only on Italian geminate and singleton consonants in words with the same number of syllables (e.g., two-syllable Italian words with singleton and geminate consonants: 'sete' vs 'sette', 'papa' vs 'pappa'). Consequently, this ratio allows us to accurately distinguish the Italian singleton and geminate consonants when these occur in words with the same number of syllables, as it is robust to the variation of speech rate between speakers [20]. To

Table 1. Number (n) of segments analysed in the current study by target stop and group of speakers, B: bilinguals and C: controls (for preceding vowels, n is equal to n for target stops).

Stop	n for C	n for B
b	119	119
bb	120	116
d	120	120
dd	120	120
g	119	119
gg	120	119
p	120	120
pp	118	115
t	119	119
tt	119	120
k	120	119
kk	120	119

illustrate this fact by an example, imagine two speakers. The second speaker speaks twice as slowly as the first. Both of them produce the Italian word 'sete'. The duration of the Italian word 'sete' produced by the first speaker is 200 ms, and the duration of the consonant /t/ in the word 'sete' is 50 ms for this speaker. The duration of the Italian word 'sete' produced by the second speaker is 400 ms, and the duration of the consonant /t/ in this word is 100 ms for this speaker. Thus, the ratio is the same for both speakers, i.e., 0.25. However, if we focus on Italian singleton and geminate consonants in words with different number of syllables, in order for this ratio to remain a parameter for the correct distinction between Italian singleton and geminate consonants, this ratio must be multiplied by the number of syllables of a word. To illustrate, imagine another speaker in addition to the two speakers from the previous example. This third speaker speaks as fast as the first, but produces the three-syllable pseudo-word 'setele'. The duration of the consonant /t/ in this pseudo-word for the third speaker is 50 ms, i.e., the same as the duration of /t/ for the first speaker, and the duration of this pseudo-word is 300 ms, because it is a three-syllable word. If we calculate only the ratio, we get 0.25 for the first speaker and 0.17 for the third speaker, which leads us to the erroneous conclusion that the /t/ of the third speaker is shorter than the /t/ of the first speaker. However, if we multiply the ratio by the number of syllables of a word, i.e. 2 for 'sete' and 3 for 'setele', we get a value of 0.5 for both speakers, which means that their /t/ is of identical duration, which is correct. Since the ratio of consonant to word duration is a well-established and methodologically approved parameter for distinguishing Italian singletons and geminates, we decided to focus on it in our analysis, knowing that since the target stops and preceding vowels (henceforth 'target segments') occur in both two-syllable and three-syllable target words in the current study, it is necessary to multiply the ratio by the number of syllables of a target word.

Therefore, for the current study, we decided to use the term *normalised duration* to speak about the parameter to be examined in the analysis in this study. In view of the above, we calculate it as the ratio of the duration of a target segment (stop or preceding vowel) to target word duration multiplied by the number of syllables of a target word. Additionally, we multiplied it by 1000 to obtain its value in *ms* and by the average of durations of all target words (see 1).

$$\text{Normalised dur. } [ms] = \frac{\text{Segment dur.}}{\text{Word dur.}} \cdot \text{Mean words dur.} \cdot \text{Nb of syll.} \cdot 1000 \quad (1)$$

To calculate the *normalised duration*, we needed to know the exact duration of target segments and target words, and the number of syllables of target words. Therefore, the recordings were manually orthographically transcribed, semi-automatically segmented and labelled into word and phone tiers using BAS Web Services [21,22]. We added additional tiers into TextGrids by scripts in PRAAT [23] and in R [24] using the package *rPraat* [25], so the number of syllables of a target words was indicated into the TextGrid and the target stops and preceding vowels were labelled.

Since the calculation of *normalized duration* requires the exact duration of target segments and target words, we manually corrected the placement of target segments and target word boundaries in PRAAT, with the exception of the placement of the left boundary of target words beginning with voiceless stops, of which there were several in this study (e.g., we had Italian words *papa*, *pappa*, *tubi*, *tocca*). The placement of the left boundary of these target words (i.e., the target words beginning with voiceless stops) was very inconsistent after semi-automatic segmentation and labelling by BAS Web Services (see Fig. 1 for illustration). This was probably because the duration of the closure of voiceless word-initial stop cannot be consistently identified in the cases where a silent pause precedes the word. To correct the placement of the left boundary in the target words beginning with voiceless stops, we used an adapted version of the AutoVOT software program developed by the fourth author of the current study (see below for the description) together with PRAAT scripts.

3.3 AutoVOT for Italian and French

AutoVOT is a tool which was created in 2014 by [26] and originally developed for English to automatically detect the voice onset time (VOT) of stop consonants in speech recordings. It uses machine learning models trained on annotated data to predict VOT boundaries in a phonetic segment and generates new tiers into Praat TextGrids with the corresponding time-aligned labels. The tool is shipped with pre-trained models mostly tuned for American English, which limits its accuracy for other languages, such as Italian and French (both spoken by our B), which differ from English in terms of stop articulation. To address this limitation, the fourth author of the current study adapted the existing AutoVOT to Italian and French, retraining the models.

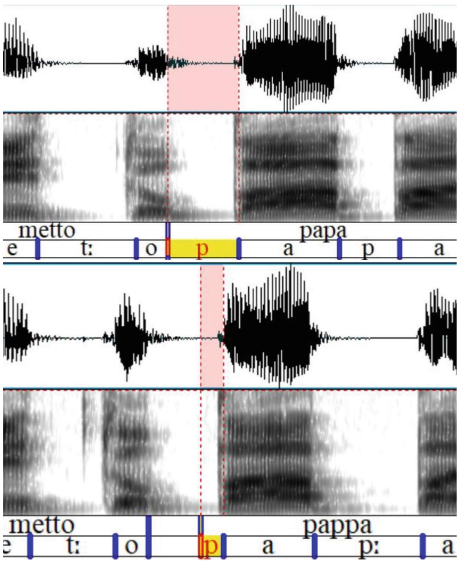

Fig. 1. Examples of placement of the left boundary of the word-initial voiceless stops in target words after semi-automatic segmentation and labelling.

The adapted version was created by first compiling a single dataset containing both Italian and French recordings with manually annotated VOT intervals for both voiceless French stops and voiceless Italian singleton and geminate stops. The dataset was composed of a part of recordings in Italian of three C speakers and one B speaker (recorded as described above and used in the current study for the analysis). Additionally, the dataset included a part of recordings in French of one B speaker and two French native speakers (similar recording procedure to the one described above). Therefore, the fourth author of the current study used the dataset containing at all 410 manually annotated VOT intervals of voiceless French and Italian stops to train new models that better reflect the phonetic properties of these languages. We chose to train new models on the recordings of both languages, French and Italian, as B have spoken both. The training process followed the same methodology used in the original AutoVOT, exploiting feature extraction and classification algorithms available in the AutoVOT training pipeline.

The adapted tool was incorporated into our segmentation workflow in PRAAT. We used it to automatically segment and label the VOT of word-initial voiceless stops in our recordings of B and C (i.e., recordings in Italian only). The VOT segmentation and labelling by the adapted tool was 99% accurate for Italian. Hence, manual correction was needed only for very few data points (note, we do not know the accuracy of this adapted tool for French as the speech production in French is not a focus of the current paper). Then, on the basis of the left boundary of the VOT interval placed by the adapted tool, using a

Praat script, we automatically moved, in the phone tier of the TextGrids, the left boundary of the word-initial voiceless stops so that the duration of closure portion of all word-initial voiceless stops corresponded to the standard closure duration, i.e., 40 ms. Next, using a Praat script, we automatically aligned the left boundary of the target words beginning with voiceless stops to the phone tier using a Praat script. This resulted in a uniform closure portion duration of word-initial voiceless stops allowing to measure an accurate duration of target words beginning with a voiceless stop. Please note that in the current study, the analysis does not concern the VOT but the *normalised duration*. The boundaries of VOT interval were inserted into TextGrids only for target words beginning with a voiceless consonant, so that the exact duration of these target words could be measured.

Figure 2 shows an example of a Praat TextGrid processed with AutoVOT before and after the training of the tool. The adapted version of AutoVOT, along with instructions and the pre-trained model for Italian and French is publicly available in the GitHub repository of the fourth author of the current study [27].

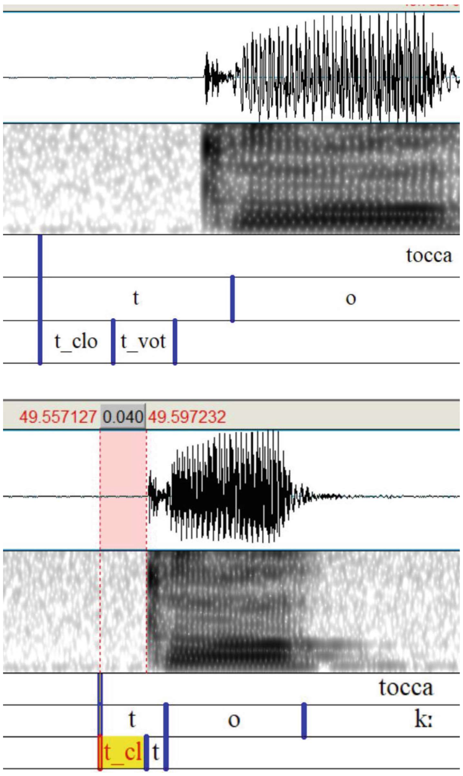

Fig. 2. Example of a TextGrid processed by AutoVOT before (top) and after (bottom) the training of the tool.

Having the boundaries of both target words and target segments correctly placed, we automatically measured their duration in seconds in PRAAT by a script.

3.4 Statistical Analysis

The statistical analyses ($\alpha = 0.05$, conf. interval = 95%) were carried out in R where the *normalised duration* of target segments was computed. For the statistical analysis, we used the R packages *lme4* [28], *dplyr* [29], and *ggplot2* [29].

To test Hypothesis 1 on the normalised duration of target stops, we built a linear mixed-effects model. As fixed effects, we entered GROUP (B *vs* C speakers), CONSONANT STATUS (singleton *vs* geminate), and PAIR of stops (p-consonants, t-consonants, k-consonants, b-consonants, d-consonants and g-consonants) with an interaction term into the model. As random effects, we had intercepts for SPEAKERS and WORDS. The random effect SPEAKERS had a random slope CONSONANT STATUS; other random slopes were not included as it would result in model convergence issues. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality. The comparison of estimated means was carried out with *emmeans* package [30].

Hypothesis 2, concerning the normalised duration of the preceding vowels, was tested as Hypothesis 1. Consequently, the levels of fixed effect CONSONANT STATUS changed to 'vowel before singleton' *vs* 'vowel before geminate', and the levels of PAIR to the levels indicating which pair of consonants the vowel precedes.

To test Hypothesis 3, we first computed the average of *normalised duration* of target segments for each speaker and target stop or preceding vowel separately. Then, we built a set of linear regressions, one for each target stop and preceding vowel, with the independent variables LOR (value in years) and *normalised duration* as dependent variable.

4 Results

The analysis of stops showed that B and C significantly differ in the normalised duration of /b/ ($\beta = 29.93\,\text{ms}, SE = 8.01, t = 3.74, p = 0.0003$), /k/ ($\beta = -21.81\,\text{ms}, SE = 8.01\,\text{ms}, t = -2.725, p = 0.0079$), and /gg/ ($\beta = 33.80\,\text{ms}, SE = 10.00\,\text{ms}, t = 51.90, p = 0.0014$): the /b/ as well as /gg/ of bilinguals were significantly shorter that those of controls while their /k/ was significantly longer than that of controls (see Fig. 3). The analysis of the preceding vowels did not show any significant result (see Fig. 4). The results of linear regressions did not show any significant relationship between *LOR* of B and *normalised duration* of any target stop nor preceding vowel.

5 Discussion and Conclusion

In this study, we predicted the L1 geminate stops and the preceding vowels produced by B (bilinguals) and C (controls) to differ in their duration (Hypothesis

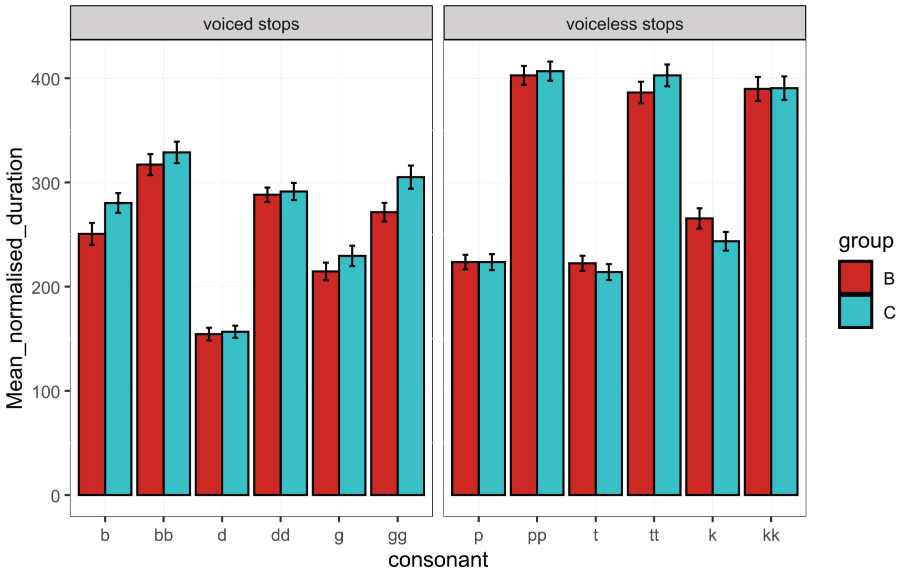

Fig. 3. Average *normalised duration* of target stops produced by bilinguals (B) and controls (C). Error bar: 95% conf. interval.

1 and Hypothesis 2 respectively). As for Hypothesis 1, the current study shows the significant difference between the B and C *normalised duration* only for the geminate /gg/. Thus, Hypothesis 1 was confirmed only for one of six Italian geminate stops. Concerning Hypothesis 2, we found the L1 preceding vowels in the production of B and C speakers (i.e., the vowels preceding geminates) do not significantly differ in their duration. Hypothesis 2 was therefore not confirmed at all. We also predicted LOR to affect the duration of geminates of B speakers (Hypothesis 3). However, this hypothesis was not confirmed by the analysis.

We may argue that the significant result on /gg/ of B, that is, its shorter *normalised duration* in B compared to C speech, might be interpreted as a change in L1 of B that would be induced by their less rich L1 input. The B undergo this L1 change because, as they do not live in their L1 country, they are probably exposed to less rich L1 input than C. This interpretation goes in the direction of L2LP, SLM and SLM-r assumption that to maintain the L1, a rich L1 input must be received, if not leading to changes in L1 induced by the effect of L2 on L1. Nevertheless, as we found the significant result only for the duration of /gg/ of B, and not for the duration of other geminates, we have to admit that our B and C differ in the duration of importantly less numerous geminates than in the previous research (i.e., studies of [10] and [11]) where bilinguals significantly differ from controls in all investigated geminates. In contrast, the duration of the preceding vowels of B and C analysed in this study unfortunately cannot be confronted to the results of [10] and [11] as these authors did not examine the duration of the preceding vowels.

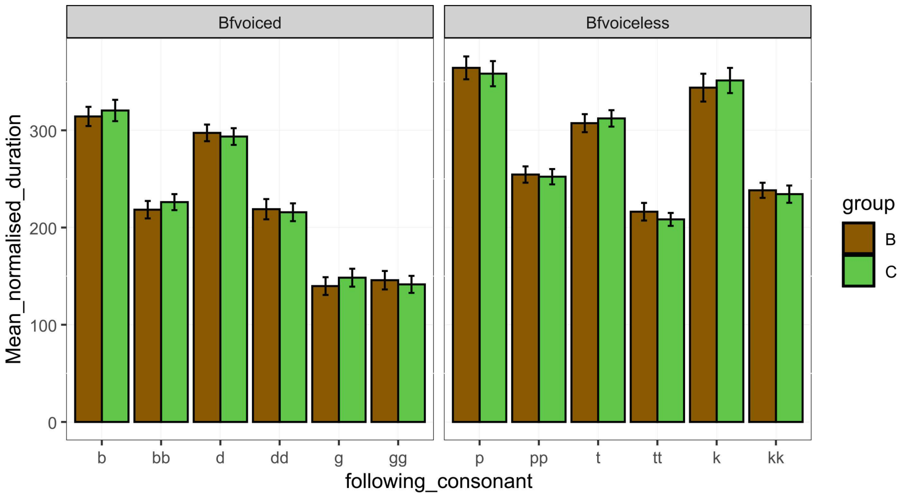

Fig. 4. Average *normalised duration* of the preceding vowels produced by bilinguals (B) and controls (C). Vowels preceding voiced stops: 'Bfvoiced', vowels preceding voiceless stops: 'Bfvoiceless'. Error bar: 95% conf. interval.

The disparity between our results on geminate durations and those of [10] and [11] might be due to the LOR of our bilinguals: in our study, we included the B of very various LOR, with minimum of 1 and maximum of 33 years, while in [10,11], the bilinguals were of higher LOR (e.g., the LOR in [11] was from 19 to 40 years). It can be speculated that we would have found more significant differences between the *normalised duration* of geminates of B and C if additional B with very high LOR (e.g. LOR of more than 35 years) were included or if only B with high LOR were analysed in this study. However, this speculation contrasts with the fact that no significant relationship between LOR and the duration of germinates of B was found in this study. Therefore, we believe that LOR can play a role in changes in L1 in B only when considered together with other factors, such as the use of L1 and L2 by B. With regard to geminate production, this study deserves to be expanded to analyse the inter-speaker differences taking these factors into account.

Moreover, this study showed that B and C significantly differ in the *normalised duration* of /b/ and /k/; this result was not predicted by our hypotheses as they did not concern the singleton consonants. Singleton duration was analysed in [10]: the authors found all singletons of bilinguals to be of shorter duration than those of controls. Similarly, in our study the singleton /b/ of bilinguals was significantly shorter than that of controls. However, we found the contrary for /k/ of bilinguals which was significantly longer than that of controls. While the result on the duration of /b/ may be quite easily interpreted and discussed, that concerning /k/ is difficult to interpret. Concerning the shorter duration of /b/, an influence of L2 may be in place, as voiced bilabial stops are particularly lengthened in Southern varieties of Italian [31]; such lengthening could influence the segment duration in word internal position as well, and the shortening in B production may then be due to the French influence. Regarding /k/, on the contrary, as far as we know, there is no interlingual comparison of its duration in Italian and French that would allow us to speculate whether or not the duration value of the /k/ of B shifts towards the one of the French /k/. To rectify this, we plan to extend this study by examining the duration of French stops, including /k/, produced by French controls to determine if there is any significant difference between the duration of French stops and Italian singleton stops, which will allow us to more easily interpret this result.

To conclude, this paper showed changes in some but not all L1 stops of late Salento Italian-French bilinguals by comparing their L1 speech production with that of Italian controls living in Salento: We could only partially compare our results with those of a few existing studies on changes in L1 geminates of bilinguals. To better compare the results across studies, it would be useful to enlarge this study by investigating changes in L1 geminates in the speech of second- and third-generation of Italian immigrants from Salento living in France.

Acknowledgments. This study was supported by the Ministry of Education, Youth and Sports of the Czech Republic: Operational Programme Johannes Amos Comenius (OP JAC), no. CZ.02.01.01./00/22_010/0008126. Co-founded by the European Union.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Grosjean, F.: Bilingualism: A Short Introduction, pp. 5–25. Wiley, Hoboken (2013)
2. Jarvis, S., Pavlenko, A.: Crosslinguistic Influence in Language and Cognition. Routledge, New York (2008)
3. Hansen Edwards, J.G., Zampini, M.L.: Phonology and Second Language Acquisition. John Benjamins, Amsterdam (2008)
4. Hevrova, M.: Phonetic attrition and cross-linguistic influence in L1 speech of late Czech-French bilinguals. Ph.D. thesis, Charles University & University Jean-Jaurès, Prague & Toulouse (2021)
5. Leeuw, E.: Phonetic attrition. In: Schmid, M.S., Köpke, B. (eds.) The Oxford Handbook of Language Attrition, pp. 204–217. Oxford University Press, Oxford (2019)
6. Gili Fivela, B., Zmarich, C.: Italian geminates under speech rate and focalization changes: kinematic, acoustic, and perception data. In: INTERSPEECH 2005, pp. 2897–2900 (2005). https://doi.org/10.21437/Interspeech.2005-765

7. Bertinetto, P.: Strutture prosodiche dell'Italiano. Accademia della Crusca, Firenze (1981)
8. Loporcaro, M.: On the analysis of geminates in Standard Italian and Italian dialects. In: Hurch, B., Rhodes, R.A. (eds.) Natural Phonology: The State of the Art, pp. 153–187. de Gruyter, New York (1996). https://doi.org/10.5167/uzh-221893
9. Celata, C., Cancila, J.: Phonological attrition and the perception of geminate consonants in the Lucchese community of San Francisco. Int. J. Biling. **14**(2), 185–209 (2010)
10. Ciccone, M., Hanini, R., Sciannantena, M.: A cross-linguistic examination of geminate consonant attrition. In: Pavlova, A. (ed.) ESSLLI and WeSSLLI 2020, pp. 173–184. Brandeis University (2020)
11. Rafat, Y., Mohaghegh, M., Stevenson, R.: Geminate attrition across three generations of Farsi-English bilinguals living in Canada: an acoustic study. Ilha do Desterro **70**, 151–168 (2017). https://doi.org/10.5007/2175-8026.2017v70n3p151
12. Flege, J.E.: Second language speech learning: theory, findings, and problems. In: Strange, W. (ed.) Speech Perception and Linguistic Experience: Issues in Cross-Language Research, pp. 233–277. York Press, Timonium (1995)
13. Flege, J.E., Bohn, O.S.: The revised speech learning model (SLM-r). In: Wayland, R. (ed.) Second Language Speech Learning: Theoretical and Empirical Progress, pp. 3–83. Cambridge University Press, Cambridge (2021)
14. Lang, B., Davidson, L.: Effects of exposure and vowel space distribution on phonetic drift: evidence from American English learners of French. Lang. Speech **62**(1), 30–60 (2019)
15. Sancier, M.L., Fowler, C.A.: Gestural drift in a bilingual speaker of Brazilian Portuguese and English. J. Phon. **25**(4), 421–436 (1997)
16. Mayr, R., Sánchez, D., Mennen, I.: Does teaching your native language abroad increase L1 attrition of speech? The case of Spaniards in the United Kingdom. Languages **5**(4), 1–41 (2020)
17. Leussen, J.W., Escudero, P.: Learning to perceive and recognize a second language: the L2LP model revised. Front. Psychol. **6**, 1000 (2015)
18. Major, R.C.: Losing English as a first language. Mod. Lang. J. **76**(2), 190–208 (1992)
19. D'Apolito, S., Gili Fivela, B.: 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 (2019)
20. Hermes, A., Tilsen, S., Ridouane, R.: Cross-linguistic timing contrast in geminates: a rate-independent perspective. In: ISSP 2020, pp. 52–55 (2020). https://doi.org/10.13140/RG.2.2.18040.26889
21. Kisler, T., Reichel, U.D., Schiel, F.: Multilingual processing of speech via web services. Comput. Speech Lang. **45**, 326–347 (2017)
22. Schiel, F.: Automatic phonetic transcription of non-prompted speech. In: Ohala, J.J., Hasegawa, Y., Ohala, M., Granville, D., Bailey, A.C. (eds.) ICPhS 1999, pp. 607–610. San Francisco, CA, USA (1999)
23. Boersma, P., Weenink, D.: Praat: doing phonetics by computer (2025), [Computer program]. Version 6.4.27
24. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2019). https://www.R-project.org/

25. Bořil, T., Skarnitzl, R.: Tools rPraat and mPraat. In: Sojka, P., Horák, A., Kopeček, I., Pala, K. (eds.) TSD 2016. LNCS (LNAI), vol. 9924, pp. 367–374. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45510-5_42
26. Keshet, J., Sonderegger, M., Knowles, T.: AutoVOT: a tool for automatic measurement of voice onset time using discriminative structured prediction (2014). https://github.com/mlml/autovot/, version 0.94, retrieved May 2025
27. Hévr, G.: Autovot for Italian and French: Adapted version (2025). https://github.com/GabTux/autovot, gitHub repository, retrieved May 2025
28. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using LME4. J. Stat. Softw. **67**(1) (2015)
29. Wickham, H.: ggplot2: Elegant Graphics for Data Analysis. Springer, New York (2016)
30. Lenth, R.V.: emmeans: estimated marginal means, aka least-squares means (2021), R package version 1.5.4
31. Romano, A.: Indici acustici di alcune geminate iniziali salentine. In: Marotta, G., Nocchi, N. (eds.) GFS 2003, pp. 233–241. ETS, Pisa (2003)