Estudios de Fonética Experimental
Journal of Experimental Phonetics
ISSN: 1575-5533 · ISSN-e: 2385-3573

Phonetic variation in Italian L2: An acoustic analysis of sibilant fricatives in the speech of L1 Spanish learners

Martina Rossi [1] 0000-0001-5970-3366
Chiara Meluzzi [2] 0000-0002-2291-006X
José María Lahoz-Bengoechea [3] 0000-0002-4654-6776

[1] Christian-Albrechts-Universität zu Kiel (Germany)
[2] Università degli Studi di Milano (Italy)
[3] Universidad Complutense de Madrid (Spain)

DOI: 10.1344/efe-2023-32-67-86

Corresponding address: mrossi@isfas.uni-kiel.de

Received: 15/03/2023 Accepted: 19/05/2023 Published: 07/06/2023

Rossi, M., Meluzzi, C., & Lahoz-Bengoechea, J.M. (2023). Phonetic variation in Italian L2: An acoustic analysis of sibilant fricatives in the speech of L1 Spanish learners. *Estudios de Fonética Experimental*, *32*, 67–86. https://doi.org/10.1344/efe-2023-32-67-86

Abstract

The present study investigates the acoustic features of Italian sibilant fricatives (/s/, /z/ and /ʃ/) in the speech of L1 Spanish learners. Segmental duration, degree of voicing and place of articulation of learners' productions are analysed alongside those of a control group of L1 Italian speakers to investigate the fine-grained phonetic differences between native and non-native pronunciation. Results from a quantitative analysis suggest that factors such as the degree of typological markedness of the phoneme and the influence of the L1 in perception and production affect L2 pronunciation to a different extent for each of the target sounds.

Keywords

Italian L2; Phonetic variation; Spanish-speaking learners; Sibilant fricatives; Phonological acquisition

Variació fonètica en italià com a L2: Una anàlisi acústica de les fricatives sibilants en la parla d'aprenents amb l'espanyol com a L1

RESUM

En aquest estudi s'investiguen les característiques acústiques de les fricatives sibilants de l'italià (/s/, /z/ i /ʃ/) per part de parlants d'espanyol com a L1. S'analitzen la durada dels segments, el grau de sonoritat i el punt d'articulació. Amb l'objectiu d'investigar les diferències fonètiques específiques entre la pronunciació nativa i la no nativa, els sons produïts pels aprenents es comparen amb els d'un grup de control format per parlants d'italià com a L1. Els resultats quantitatius suggereixen que factors com el grau de marcatge tipològic de cada fonema i la influència de la llengua materna en la percepció i en la producció afecten la pronunciació de l'italià com a L2 de manera diversa en cada un dels sons estudiats.

MOTS CLAU

Italià L2; Variació fonètica; Aprenents castellanoparlants; Fricatives sibilants; Adquisició fonològica

1. Introduction

This paper reports the results of a cross-sectional study aimed at analyzing the acoustic characteristics of Italian as a Second Language (L2) in a group of Spanish-speaking learners by focusing on the phonetic implementation of Italian sibilant sounds, i.e., fricatives (s, z, ʃ) and affricates (t͡s, d͡z, t͡ʃ, d͡ʒ). While these sounds have shown to be a source of difficulties for learners (Giannini, 2003; Costamagna, 2003), they still have not been extensively addressed through fine-grained acoustic studies, especially for Italian. This article exposes the results of the analysis of sibilant fricatives in the speech of upper-intermediate and advanced L2 learners. Our main objectives are to provide a fine-grained description of sibilant fricatives' acoustic characteristics in the Italian L2, to observe how Spanish-speaking learners' productions differ phonetically from those of Italian control subjects and, ultimately, to investigate which factors (e.g., typological markedness, aerodynamic complexity, L1 influence, task, extra-linguistic factors) may motivate the phonetic configuration of sibilant fricatives in their interlanguages.

After an overview of the main theoretical approaches to the acquisition of L2 pronunciation (Section 2), we will describe the phonetic features of sibilant sounds and their distribution (Section 3). After that, the methodology and the analysis design (Section 4) will be addressed, followed by the description of the main results obtained through both qualitative and quantitative methods (Section 5). In the discussion (Section 6) we will compare our results with the previously formulated hypotheses, theoretical assumptions and the results from previous research, before concluding with a brief overview of the investigation and the future perspectives of our work (Section 7).

2. The acquisition of L2 pronunciation: theories and approaches

The adults' pronunciation of the L2 typically differs from that of the native speakers (Colantoni et al., 2015). In their speech, in fact, certain non-native acoustic features and patterns of variation may never assume a proper target-like configuration. This is due to several linguistic factors, such as the characteristics of the phonological inventory of the learner's L1, those of the L2's inventory, and the degree of typological markedness of the target phonemes (Major, 2001).

Especially in the earlier stages of the acquisition, the L1 has a very strong influence on the interlanguage (Major, 2001), functioning as a phonological filter (Trubetzkoy, 1939) which causes cross-linguistic influence phenomena that constitute what is commonly referred to as "foreign accent". The interlanguages of learners with the same L1 tend to display several common acoustic characteristics, which may function as a cue for the native listener to identify the L1 or the country of origin of the speaker (Costamagna, 2008).

The deeply-rooted experience with the L1 also affects L2 perception. As postulated by Flege's (1995, 2003) Speech Learning Model (SLM), the adult learner, being used to L1 perceptual categories, possesses a decreased ability to identify the phonetic differences between L1 and L2 sounds, especially when two instances are similar. In fact, the SLM predicts that when L1 and L2 share a same phoneme, but its phonetic implementation differs in not very salient ways, an accented pronunciation is more likely to be expected: L2 instances will be equated to those of the L1 and produced following the acoustic and articulatory routines of the native language. On the other hand, accented production is less common with new L2 sounds, which cannot be confused with existing L1 instances, and appear therefore to be learned more easily than similar sounds.

Another factor that has been observed to play a role in L2 phonological acquisition is typological markedness. Marked, less natural sounds tend to create a "long-term learning obstacle" (Costamagna, 2008, p. 138), especially when the phone does not exist in the L1 system, as postulated by Eckman's (1977, 1991) Markedness Differential Hypothesis. In the Ontogeny and Phylogeny Model (OPM),

Major (2001) also claims that for very marked and very infrequent phenomena, little to no acquisition may take place, so that determinate sounds may never appear in the interlanguage of the adult learner or may appear in a non-native-like configuration.

Experimental research on the phonological acquisition of Italian L2 also gave some insight on the action of the above factors on the acquisition process. For example, in her analysis of the acquisition of Italian affricates by a group of Spanish, Albanese and Igbo learners, Sorianello (2019) observes that the production of post-alveolar affricates tends to cause less problems than the production of dental affricates, which are typologically rarer and less natural. In their acoustic study on the production and perception of Italian vowels and consonants by Galician learners, Romito et al. (2016) observe a strong influence of L1 along difficulties in the production of very marked segments, such as voiced sibilant affricates. Finally, in their research on the devoicing of voiced obstruents of Italian by Swiss German learners, Schmid & Wachter (2015) conclude that this tendency is to be allocated both to the phonology of the L1 of the subjects and to general linguistic principles related to the complexity of sounds, such as voiced sibilants, i.e. the segments most affected by the devoicing phenomenon.

3. Sibilant fricatives: phonetic properties and distribution

Fricatives (and affricates) are classified as sibilants when they are characterized by a high intensity noise, especially in the high-frequency regions of the spectrogram (Balise & Diehl, 1994; Ladefoged & Maddieson, 1997). Typological studies on the distribution of obstruents in language inventories have highlighted the general infrequency of voiced segments in opposition to their voiceless counterpart, with the asymmetry for fricatives being the "most extreme" (Żygis et al., 2012, p. 301). This distribution reflects the claim that voiced fricatives are the marked elements of the opposition voiceless/voiced (e.g., Trubetzkoy, 1939; Greenberg, 1966; Żygis et al., 2012). The asymmetry can be explained by the phonetic properties of these sounds and by the aerodynamic conditions required to produce an audible friction. In fact, in order to produce voiced fricatives, two conflicting conditions must be fulfilled, i.e., a high oral pressure must be maintained to obtain sufficient air velocity to produce the high-intensity frication noise, and, at the same time, a relatively low oral pressure is required to ensure voicing (Żygis, 2008). For voiced sibilants in particular, the vibration of the vocal folds reduces the transglottal airflow volume enough "that the airflow through the constriction downstream is too low to produce sibilants' characteristic high intensity noise" (Kingston, 1993, p. 75). These adverse conditions make voiced sibilant sounds harder to produce, but also harder to be perceived by the listener, since they tend to be significantly less intense and shorter in duration than voiceless sibilants (Ohala & Solé, 2010).

As Maddieson (1984) stated "generally, the existence of a voiced fricative in the inventory implies the presence of a voiceless counterpart in the inventory" (p. 47). In the Italian inventory of sibilant fricatives there are two voiceless fricatives, one for the dento-alveolar place of articulation (/s/) and one for the post-alveolar place of articulation (/ʃ/). While the latter does not have a phonological voiced counterpart ([ʒ] appears only as an allophone of the affricate /d͡ʒ/ in the regional Italian of Tuscany and in a few loanwords: Bertinetto, 2010), voiced /z/ is present. Despite being in a complementary distributional relationship in the standard pronunciation (but with a significant degree of variation among regional varieties: Schmid, 1999; Canepari, 2005), and the opposition between voiced and voiceless being neither functional nor productive, both /s/ and /z/ are traditionally considered two phonemes of Italian, since there actually exist a few minimal pairs formed by their voicing opposition (e.g. /ˈkjɛse/, 'asked' ~ /ˈkjɛze/, 'churches'); however, this opposition pertains basically to high stylistic varieties (De Dominicis, 1999; Schmid, 1999). As far as /ʃ/ is concerned, in Italian phonology it is traditionally considered to be one of the *rafforzato* consonants, i.e., inherently long or "intrinsic geminate" (Loporcaro &

Bertinetto, 2005, p. 134), along with /ʎ, ɲ, t͡s, d͡z/. Evolved from Latin consonant clusters and geminates, these sounds share some tendencies with 'canonical' geminates, i.e., high segment duration and, when occurring in the intervocalic context, shortened duration of the preceding vowel (Sorianello, 2002). For this reason, intervocalic /ʃ/ is often transcribed as /ʃʃ/ or /ʃː/ (e.g.: Canepari, 1999).

Evolved from a medieval inventory which included both voiced and voiceless sibilant fricatives, i.e., /s, z, ʃ, ʒ/, modern Castilian Spanish only maintained the voiceless phoneme /s/ (Lapesa, 1942; Fradejas Rueda, 1997; the high dialectal and intra-speaker variability has been frequently addressed by Spanish phoneticians, e.g. Univaso et al., 2014; Del Saz, 2023). The voiced allophone [z] may appear in free variation before a voiced consonant in syllable codas unless in very careful speech (Schwegler et al., 2010; Harris, 1969) or, more rarely, even between vowels, especially in fast and informal speech, as attested by Torreblanca (1983; 1986) for a few areas in Central Spain. As for the post-alveolar/palatal place of articulation, Spanish inventory does not include the fricative, but it does have the sibilant voiceless affricate /t͡ʃ/, present also in Italian but generally produced in Spanish with a more posterior place of articulation (Mazzotta, 1984, and also Regan 2020: 60, for Spanish varieties in Andalucía).

Thus, as far as sibilant fricative sounds are concerned, Spanish speakers learning Italian will encounter a new sound, i.e., /ʃ/, a new contrast, i.e., /s ~ z/, and a sound which is maximally similar to the already known /s/. As mentioned before (cfr. Section 2), the situation for the acquisition of similar sounds is not as straightforward and uncomplicated as it may seem, given the fact that the L1 implementation of the phoneme may have different phonetic features than that of the same sound in the L2, and the transfer of those features may contribute to the acoustic configuration of the foreign accent. This is the case for the /s/ sound, which in Spanish has a phonetic implementation different from Italian. Besides appearing as voiceless in all phonotactic contexts, Italian /s/ is overall pronounced with more tension and with a more anterior place of articulation than in Castilian Spanish /s/ (Mazzotta, 1984; Hernandez, 2009). Italian /s/ is stated to be articulated with the tongue tip towards the alveoli of the upper incisors (Albano Leoni & Maturi, 2009, p. 57), even if great variability among regional varieties of Italian has also been detected, albeit not much explored in acoustic terms (e.g., Canepari, 1979). In fact, in Italian phonology /s/ has been classified as an alveolar (Schmid, 1999), an alveo-dental (Canepari, 1979) and as dental fricative (De Dominicis, 1999). Castilian Spanish /s/ is described as an apico-alveolar fricative, articulated with the tip of the tongue towards the alveoli (Hidalgo Navarro & Quilis Merlín, 2012). Moreover, in the articulation of /s/ the tongue tends to assume a slightly concave shape in Castilian Spanish, while in Italian (but also in French and German) it is more convex, which confer to the Spanish /s/ a lower timbre (Navarro, 1957). Given these fine-grained acoustic differences of this sound in Italian and Spanish, it has been observed that the pronunciation of /s/ by Spanish-speaking learners contributes to signal a 'Spanish accent' to Italian native listeners (Devís, 2005).

Based on these premises and on the theoretical assumptions outlined in the previous section, we formulated the following hypotheses regarding the sibilant fricatives object of analysis:

a) Being the difference from the L2 not very salient, we hypothesize that Italian /s/ will be produced with the same acoustic correlates and articulatory routines of the learners' native language, so presumably with a more posterior place of articulation and less articulatory tension than Italian natives.
b) Voiced /z/ is a marked sound and it is hard to be correctly perceived and appropriately produced. As previous research shows, there are two possible strategies that speakers may apply "when faced with the demands of producing voicing and frication noise" (Żygis et al., 2012, p. 311) at the same time: subjects may choose to maintain frication rather than voicing, in which case a voiceless fricative will be

produced; on the other hand, they may opt for maintaining voicing over friction, in which case an approximant will be the result (Żygis et al., 2012; Ohala & Solé, 2010). For Spanish learners of Italian, it might be expected that the first strategy will be the predominant one, since learners may perceive Italian /z/ as the already known voiceless /s/, also considering the distribution of the two phonemes in the target and the high degree of regional variation in the distribution of voiced/voiceless in spoken Italian. Moreover, Italian /s/ and /z/ are expressed by the same grapheme in the orthography, <s>, which also represents /s/ in Spanish. We do not exclude that /z/ may appear in our production data, but we expect it to have a very low frequency and an unsystematic distribution.

c) Finally, learners are not expected to encounter big difficulties for the post-alveolar fricative /ʃ/, which can be considered as a new phoneme for them. However, since Spanish-speaking learners of Italian tend to use the post-alveolar /tʃ/ as a reference to produce this new sound (Schmid, 2004; Maturi, 2014), we may observe a more posterior place of articulation than natives in their productions of the Italian /ʃ/.

4. Methods

The research outlined in this work is aimed at answering the following three main research questions: What are the fine-grained acoustic characteristics of sibilant sounds in the speech of Spanish-speaking learners of Italian? How do they differ from those of L1 speakers? What are the linguistic (but also the extra-linguistic) factors that influence pronunciation? To answer these questions and to test the hypotheses outlined above (Section 3), we planned a cross-sectional study aimed at eliciting the production of Italian sibilants by Spanish-speaking learners and L1 Italian control subjects.

4.1. Tasks

Data were collected through individual experimental sessions in which the subject had to carry out three production tasks in front of a Zoom H1 audio recorder (without an external microphone). The tasks of the experimental session were set up to elicit different speech styles (spontaneous and read) and speech rates (normal and fast), while ensuring the occurrence of sibilant segments in a high number of phonetic contexts. First, the subjects took part in a brief semi-structured interview (Speaking task); the second and third tasks consisted in the reading of three sentence lists (one for fricative segments, the other two for affricate segments), first at a ‘normal’ speech rate (Reading task), then at a ‘fast’ speech rate (Fast Reading task). The sentences were composed of real Italian words. Each sentence contained either one or two words with the target sounds. The variables controlled to develop the sentence list for fricative segments were: the phonological status of the sound as voiced or voiceless; point of articulation (alveolar, post-alveolar); position (word initial, word internal); context (alveolar fricatives: intervocalic, geminate, post-sonorant consonants, pre-stop consonants, pre-consonant clusters; post-alveolar fricatives: intervocalic, post-sonorant consonants); following vowel (i, e, a, o, u); stress (stressed syllable, unstressed syllable). Here’s some examples of sentences from the sentence list, with the IPA transcription of the target words (following Canepari, 1999):

a) Non oso [ˈozo] aprire la cassa [ˈkasːa] (‘I dare not open the box’).
b) Sono lisci [ˈliʃːi] come seta (‘These are smooth as silk’).
c) Lo scranno [ˈskranːo] del senatore (‘The senator’s seat’).
d) Non sei conscio [ˈkɔnʃo] del rischio (‘You are not aware of the risk’).
e) Ho un casale [kaˈzale] in campagna (‘I own a farmhouse in the countryside’).
f) Bisogna fasciare [ˈfaʃːare] il piede (‘The foot must be bandaged’).

After the production tasks, learners were invited to fill out a Language Background questionnaire in Spanish aimed at collecting their personal information (age, gender, other L1s), their level of proficiency in Italian, the amount of formal instruction received, the amount of time spent in Italy working or studying, and the amount of use and exposure.

4.2. Participants

All the students of Italian that took part in the experiments were volunteers. Learners were recruited among the students of Italian at the Complutense University of Madrid. The call for participants was open to learners with every level of proficiency, but for the purposes of this study, and because of its cross-sectional approach, only upper-intermediate and advanced learners' data were analyzed, since they were able to produce all the target items in the entirety of the controlled contexts. This way, we have been able to carry out a complete description of the phonetic characteristics of the target sounds, along with an extensive comparison with the native controls' patterns of variation. Further analyses of our data may possibly focus on the configurations assumed by sibilant sounds at different levels of learners' proficiency.

The subjects analyzed for the experimental group are 5 upper-intermediate and advanced Spanish-speaking learners of Italian, 21-24 years old, all female (due to the availability of the subjects). An overview of each learners' characteristics (participants are anonymized through their initials) is reported in Table 1.

They are self-assessed either as upper-intermediate or advanced in the Language Background questionnaires, where they also reported the level of the last official language examination they successfully passed: based on the CEFR, their level goes from B2 to C2 (Common European Framework of Reference; Council of Europe, 2001). Their years of formal instruction in Italian span from 2 years to the 12 years of the subject CCB, who has been learning Italian since elementary school. Three subjects (BAM, NPV and JAG) spent periods of study/work in Italy. Besides Spanish as their L1, two subjects reported to have native-like proficiency of Basque and Catalan, respectively (this issue will be addressed further in the Discussion).

For the control group, only female speakers have been considered, in order to exclude the acoustic variation that would have come with the gender variable (see Fuchs & Toda, 2010): this group thus includes two L1 Italian female subjects from Northern Italy, in particular from the Milan area, aged 21 and 24. The complete corpus consists of 02h06m of recordings, with a total of 2203 sibilant tokens, 852 of which are sibilant fricatives.

4.3. Data extraction and statistical analysis

The speech data collected were digitized at a sampling rate of 44.1 kHz, 16-bit, mono format. The annotation and the acoustic analysis of the audio files was carried out using Praat (Boersma & Weenink, 1992–2021; Boersma, 2001). The annotation followed protocol developed for affricates by Meluzzi (2014; 2020), with few alterations to be adapted to fricatives. It was carried out manually by the first author and double-checked by the second and third authors. In order to provide a comprehensive description of their configuration and their pattern of variation, our study of sibilant fricatives has been structured around three main acoustic parameters, i.e., Voice, Duration, Place of articulation (Meluzzi, 2014; 2020). For the analysis of these parameters, the acoustic variables extracted from Praat have been the following:

	Age	Instruction	Level	Stay in Italy
BAM	22	3 years	C1	<1 year
CCB	21	12 years	C2	NA
NPV	23	2 years	B2	<1 year
JAG	24	2 year	B2	2 years
CGR	23	5 years	C1	NA

Table 1. Age, years of formal L2 instruction, CEFR proficiency level and length of their stay in Italy for each subject in the learners' group.

a) Periodicity: through a visual assessment of the spectrogram, if the signal was periodic for the whole segment the token was annotated as voiced; if periodicity in the lower frequencies was absent the sound was tagged as voiceless; if the signal was periodic for just a part of the segment, it was also annotated as voiceless. These labels, observed together with a second set of tags indicating the phonological status of the fricative as voiced/voiceless, helped us differentiate between the segments that were produced as voiceless according to the phonological rules of the language, and those that were expected as voiced but failed to be fully produced as such, in spite of displaying spectral periodicity for the initial portion of the segment (which is accounted for by the label for the Voice Decay Time, see below). As will be further discussed in the next sections, this distinction was particularly useful to contrast the fully voiced phones produced by the natives and the realizations of the learners, who seem to have acquired the underlying voiced fricatives but were not always able to maintain voicing for the whole duration of the segment due to the conflicting aerodynamic conditions required (see Section 3).

b) Duration of the segment, in milliseconds: this parameter was extracted using Praat from the segment interval, annotated manually. The left boundary marking the beginning of the phone was placed at the spectral offset of the preceding segment, while the right boundary was placed at the disappearance of the fricative's F2, when visible, or at the offset of the high-frequency intense spectral noise.

c) pVDT and pVOT: Voice Decay Time (VDT) indicates the amount of time between the start of the fricative segment and the decay of the periodicity, while Voice Onset Time (VOT) refers to the amount of time between the end of the segment and the beginning of the successive periodicity. In order to normalize these measures, the values of the VDT and the VOT were calculated as a proportion of the duration in milliseconds of the VDT/VOT and the total duration in milliseconds of the segment (Olmo-López, 2017).

d) Center of Gravity (CoG): extracted automatically from Praat, this measure in Hertz gives information about where the energy is concentrated on the spectrum; former acoustic studies on fricatives proved that CoG is very sensitive to fine-grained variations in the place of articulation (Jongman et al., 2000; Harrington, 2010; Fuchs & Toda, 2010). The CoG correlates negatively with the length of the oral cavity, so the more anterior the place of articulation and the narrower the front cavity, the higher the Center of Gravity will be (Jongman et al., 2000; Fuchs & Toda, 2010; Cruselles et al., 2017).

The independent variables were the subject, the L1 of the subject (Spanish or Italian), the phonological status of the token as a voiced or voiceless fricative, its position, its context, the following vowel, the presence or absence of the stress in the syllable where the segment occurred, the task (Reading, Fast Reading, Speaking). Excel was used for the qualitative analysis, while the statistical analysis was carried out with the software R (R Core Team, 2020) using linear mixed-effect models with the package lmerTest (Kuznetsova et al., 2017). In each

model, the subject was included as a random factor, while all the aforementioned independent (categorical) variables were adopted in the models as fixed effects. The subject was coded such that the reference was the grand-mean; the rest of the variables were dummy-coded, taking the following levels as reference: Italian (L1), voiceless (phonological status), word-initial (position), intervocalic (context), 'a' (following vowel), unstressed (syllable), and reading (task).

5. Results

In this section we provide an overview of the main results obtained through the qualitative and quantitative analysis of alveolar and post-alveolar fricatives following the parameters of Voice, Duration and Place of articulation.

As for the Voice parameters, a first qualitative exploration of the data revealed a very low percentage of voiced alveolar fricative occurrences in the learners' sub-corpus (2%). The great majority of alveolar fricatives were realized as the voiceless [s], as it was expected (Hypothesis II). Even if /z/ occurrences were too few to effectively obtain a generalization from our data, it appeared that the distribution of the voice feature was not systematic, nor did it follow the distribution rules of Italian or the pattern of variation displayed by the natives. While Italian controls produced voiced /z/ in the intervocalic context and when preceding a voiced consonant in word-internal position (94% of the total occurrences of [z] for L1 Italian speakers), learners' majority of voiced alveolar fricatives appeared before a voiced consonant in word-initial position (60% of the total occurrences of [z]), especially in the Fast-Reading task (42%). It must be noted that, before a voiced consonant, [s] was the preferential choice, especially in word-internal position, and, in some cases, also the following consonant, or part of it, was devoiced, as the high values of the pVOT for this context shows: .54 vs. .13 for Italian controls. All the few instances of intervocalic sonorization appeared in the tasks aimed at eliciting a "less careful" speech, i.e., the Fast Reading and the Speaking task.

The question at this point was the following: do learners perceive a phonological difference between /s/ and /z/? If we hypothesize a positive answer, it could have been the case that learners tried to produce a voiced fricative but failed to maintain the periodicity throughout the whole segment because of the conflicting aerodynamic conditions required for production. To test this possibility, we looked at the variation of the pVDT, modeling it with the expected realization as voiced/voiceless of the sound in Italian, based on the phonotactic context. If learners did try to produce a voiced sound, but eventually failed to do so, then the value of the pVDT might be systematically higher in the contexts where voicing was required. Qualitatively, we observed that the pVDT appeared to be on average slightly longer for voiceless realizations of segments expected as voiced (0.11 for /z/ vs. 0.09 for /s/; see Figure 1) and the statistical model tested significant: $F(1, 884.16) = 7.9407$, $p < .01$.

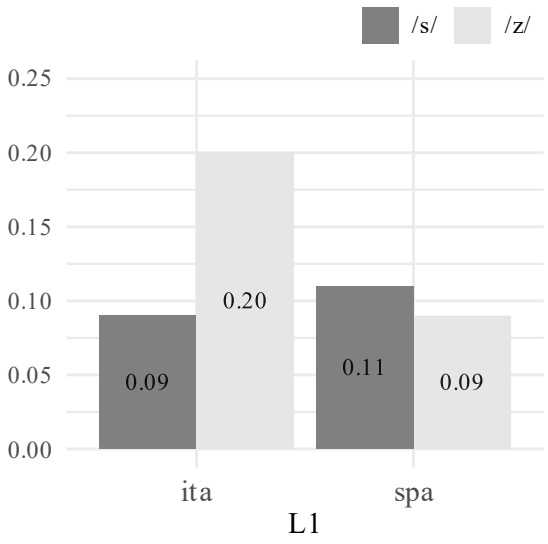

Figure 1. Average pVDT extension in voiceless realizations of phonological /s/ and /z/ for control subjects and learners.

Duration was one other feature that appeared to differentiate the realizations of sounds expected as /z/ and those expected as /s/. Voiced fricatives are generally shorter than their voiceless counterpart and, in our data, we observed that learners did produce the segments expected as /z/ as slightly shorter than those expected as /s/: 104 msec vs. 115 msec (excluding geminates). This difference

appeared to be statistically significant: $F(1, 691.14) = 7.1545$, $p < .01$.

Observing the variation of the Duration parameter, we noticed that Spanish-speaking learners' instances of alveolar fricatives appear to be on average shorter than those produced by the native controls, which could be a correlate of the lesser tension that characterizes Spanish with respect to Italian. In fact, Italian voiceless [s] (excluding geminates) has the average duration of 128 msec vs. 115 msec for learners; Italian voiced [z] has the average duration of 85 msec vs. 40 msec for learners. Observing the pattern variation of Duration in each phonotactic context, focusing only on the Reading task, and excluding the few voiced occurrences, we observed a shorter duration of the learners' productions for each level, except the intervocalic context (see Table 2). The effect of syllable stress on the duration of [s] in the Reading task was statistically significant: $F(1, 667.94) = 14.0970$, $p < .001$. In the L1 group, the stressed syllable showed a 43 msec increase in [s] duration compared to the unstressed syllable, while in the learners' group, it exhibited an increase of 11.5 msec.

	Learners	**L1**
Geminate	174	217
Intervocalic	118	114
Post-Sonorant (/n/)	113	125
Post-Sonorant (/l/)	133	164
Post-Sonorant (/r/)	119	142
Pre-Voiced	87	90
Pre-Voiceless	90	128

Table 2. Average Duration of [s] in msec in each phonotactic context for Italian controls and Spanish-speaking learners of Italian in the Reading task.

	Speaking	**Reading**	**Fast Reading**
Geminate	147	182	134
Intervocalic	120	122	109
Post-Sonorant (/n/)	117	126	88
Post-Sonorant (/l/)	NA	142	114
Post-Sonorant (/r/)	112	126	99
Pre-Voiced	65	93	79
Pre-Voiceless	102	102	80

Table 3. Average Duration of [s] in msec in each phonotactic context for Spanish-speaking learners of Italian in the Speaking, Reading and Fast Reading tasks.

Gemination is executed by learners by increasing the consonant duration by +56 msec, compared to the intervocalic singleton, even if for the Italians the difference is higher, with the consonant duration being increased by +103 msec compared to the intervocalic singleton. Even if the segment duration is higher for Italian controls, the variation based on the context appears to be very similar between the two groups, with the geminate context displaying the longer segments, followed by the post-sonorant contexts and the intervocalic one, with the shorter durations appearing for the pre-voiced consonant context. In fact, in the statistical models, the effect of the context on duration is significant: $F(7, 685.03) = 60.9194$, $p < .001$; while the effect of the L1 is not: $F(1, 5.09) = 1.1547$, $p = .33$.

Furthermore, the duration of alveolar fricatives in our learners' sub-corpus seems to have been influenced also by the task. Segment duration in the

Speaking task is shorter than in the Reading task, but it is in the Fast-Reading task that we observe the average lower values (see Table 3). While the Task variable has a significant effect on the variation of the Duration parameter considering both the natives' and the learners' productions ($F(2, 690.20) = 19.7899$, $p < .001$), it does not reach statistical significance when modeled in interaction with L1 ($F(1, 689.35) = 2.2587$, $p = .13$). Thus, for Spanish learners, the task had no significant effect on alveolar fricatives' duration, even if some trends emerge.

The Place of articulation of alveolar fricatives was observed through the values in Hertz assumed by the Center of Gravity. The CoG of the voiced alveolar fricatives was extracted from Praat applying a high pass filter to the spectrogram (from 1000 Hz), so that the signal in the lower frequencies, given by the vibration of the vocal folds, would not interfere with the actual Center of Gravity in the higher frequencies. As we expected (Hypothesis I), the place of articulation appears to be more advanced in native Italians' productions, with higher values of the CoG, while for learners it seems to be more posterior, with lower values of the CoG (see Table 4).

	Learners	**L1**
Geminate	7046	9179
Intervocalic	6629	8401
Post-Sonorant (/n/)	6959	8924
Post-Sonorant (/l/)	6474	8672
Post-Sonorant (/r/)	6749	8139
Pre-Voiced	6439	8877
Pre-Voiceless	6787	9044

Table 4. Average CoG in Hz of alveolar fricatives in each phonotactic context for Italian controls and Spanish-speaking learners of Italian.

In general, the Place of articulation for L1 Spanish subjects when producing the alveolar sibilant fricative in Italian is situated in the 7000–6400 Hz range, so in the alveolar area, closer to the palate than to the teeth, while the native Italian CoG for our control subjects stretches from 9200 Hz to 8100 Hz, much closer to the dental region than to the palate (see Figure 2). The L1 appears in fact to have a statistically significant effect on the variation of the CoG in our corpus: $F(1, 5.09) = 12.5116$, $p < .02$.

As far as the palate-alveolar sibilant fricative /ʃ/ is concerned, there have been no phenomena to be observed for the Voice parameter: all the segments were realized as voiceless, as expected, since in the phonological inventory of Italian the voiced /ʒ/ is not present. As mentioned above (cf. Section 3), it appears only in a few loanwords, which were not included in the experimental materials for this research.

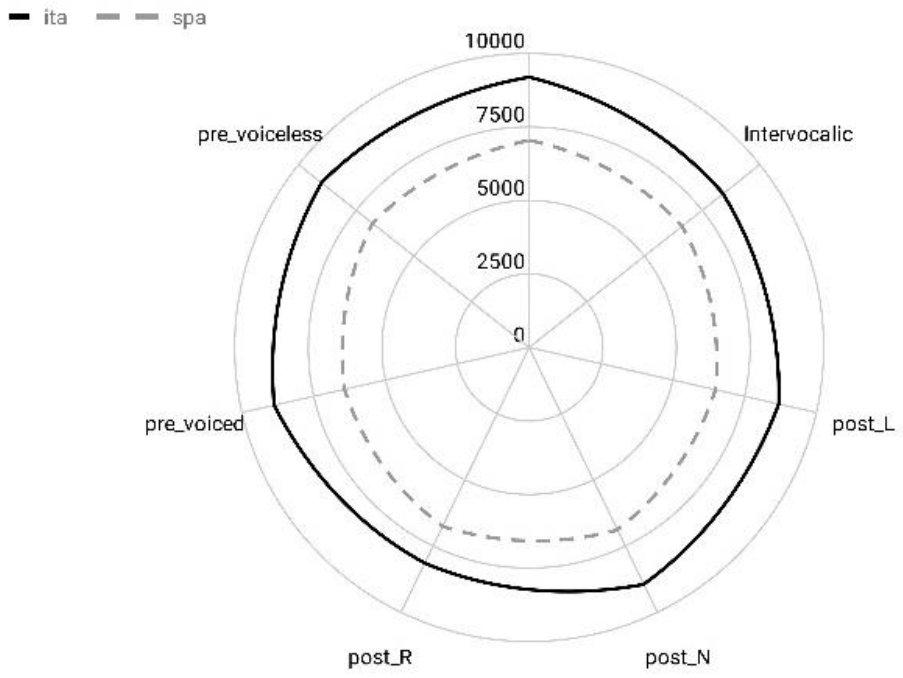

Figure 2. Representation of the CoG range (in Hz) of alveolar fricatives in each phonotactic context for Italian controls and Spanish-speaking learners of Italian.

In terms of Duration, the post-alveolar fricatives produced by the Spanish-speaking learners appeared once again to be shorter than the segments produced by the native controls in the intervocalic context (see Figure 3), but this seemingly different behavior of the two groups does not reach statistical significance: $F(1, 5.4227) = 0.6165$, $p = .46$.

The average duration of the post-alveolar fricative in the intervocalic context for the Reading task as produced by our L1 Italian subjects reaches 203 msec, while the segments produced by L2 subjects have an average duration of 161 msec, which for both groups is shorter than the geminate but longer than the singleton /s/ (cf. Table 2), which is in line with the literature on *rafforzato* consonants (Sorianello, 2002; Loporcaro & Bertinetto, 2005). There are very few extensive experimental acoustic studies of Italian fricatives; however, as far as /ʃ/ is concerned, our results correspond to those of Endo & Bertinetto (1999), who have shown that, in a Northern variety of Italian, the duration values of the post-alveolar fricatives rank below those of real geminates, but above those of singletons, as it happens in our results, both for Italian controls and for learners.

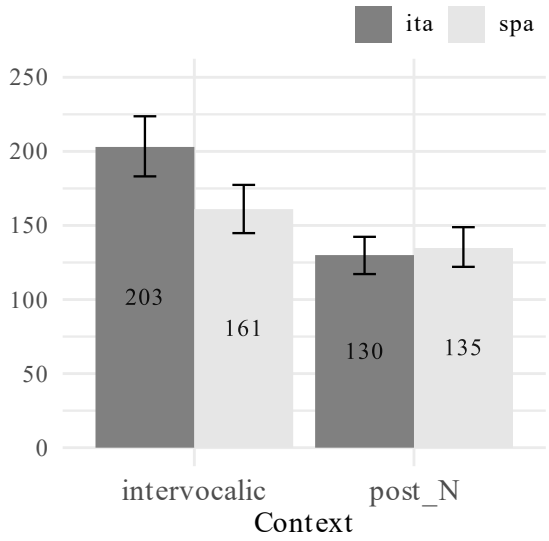

Figure 3. Average duration of [ʃ] in msec in each phonotactic context for Italian controls and Spanish-speaking learners of Italian in the Reading task.

Focusing only on the Reading task, we observed an average duration of 159 msec in word-initial position for Spanish learners (vs. 167 msec for L1 Italian) and 150 msec in word-internal position (vs. 159 msec for L1 Italian). While the position effect on consonant duration does not reach the significance level ($F(1, 145.07) = 3.6721$, $p = .06$),

the effect of the phonotactic context appears to be significant: $F(1, 145.09) = 22.704$, $p < .001$.

Still in the Reading task, the duration of the post-alveolar fricative seems to be also influenced by the presence or absence of stress: in the stressed syllable, the duration of the consonant is increased by +7 msec in both groups (L1: 160 msec vs. 167 msec; L2: 150 msec vs. 157 msec). However, the effect of stress is not significant: $F(1, 145.16) = 2.7257$, $p = .11$. On the other hand, the Duration parameter appears instead to be significantly influenced by the effect of the Task, like was the case for alveolar fricatives: $F(2, 147.28) = 11.434$, $p < .001$.

As shown in Table 5, the longer duration values for the Intervocalic context appear for the Reading task (161 msec) while the shorter duration values are observed for the Speaking task (112 msec), with intermediate values for the Fast-Reading task (136 msec). A comprehensive comparison is not possible for the post-sonorant /ʃ/, since it did not appear in any of the learners' Speaking tasks. However, the fricative appears to be on average shorter in the Fast-Reading task than in the Reading task.

	Speaking	**Reading**	**Fast reading**
Intervocalic	112	161	136
Post-Son. (/n/)	NA	135	108

Table 5. Average Duration of [ʃ] in msec in each phonotactic context for Spanish-speaking learners of Italian in the Speaking, Reading and Fast Reading tasks.

Finally, the analysis of the Place of articulation of post-alveolar fricatives revealed once again a slightly more posterior point of constriction in learners' pronunciation of the segment (see Table 6), but this difference from the native speakers did not turn out to be statistically significant: $F(1, 4.9127) = 1.2838$, $p = .31$. In fact, the difference between the two groups, however present, is rather small: the average CoG for learners is 4285 Hz, while for native controls it is 4795 Hz. The context, however, appears to have a significant effect on the variation of this place of articulation ($F(1, 144.69) = 4.8388$, $p < .05$), with the values of the CoG observed at around 4200 Hz for learners and 4700 Hz for Italians for both levels of the variable.

	Learners	**L1**
Intervocalic	4344	4794
Post-Son. (/n/)	4088	4797

Table 6. Average CoG in Hz of post-alveolar fricatives in each phonotactic context for Italian controls and Spanish-speaking learners of Italian.

In order to get a clearer picture of the range of the CoG, we observed how it varies as a function of the following vowel: consistently with what was reported before, the learners' average CoG for post-alveolar fricatives extends from 4200 Hz to 4400 Hz before higher vowels, while the controls' average CoG extends from 4600 Hz to 4900 Hz, with the learners' place of articulation slightly closer to the palatal region (see Figure 4).

6. Discussion

The acoustic characteristics of sibilant fricatives (and affricates) were analyzed by focusing on three main parameters of description, i.e., Voice, Duration and Place of articulation. As for the Voice parameter, we observed a very small number of occurrences of voiced alveolar fricatives. Our Hypothesis II (cf. Section 3) was confirmed: the phone appeared in the productions of L2 speakers, though with a very low frequency and an unsystematic distribution. It appears that the less marked, already known [s] sound was the preferred segment in every phonotactic context considered by the study. Also, since both phonemes are represented by the same grapheme in Italian (<s>), which in Spanish orthography represents /s/, a possible explanation from the data may be that the 5 advanced Spanish learners of Italian we analyzed might not perceive any difference between [s] and [z] and might connect both instances of the target to the phoneme /s/. However, from our analysis, two other characteristics that differentiated the realization of /s/ and /z/ by learners have emerged, i.e., duration and pVDT: the variation of both

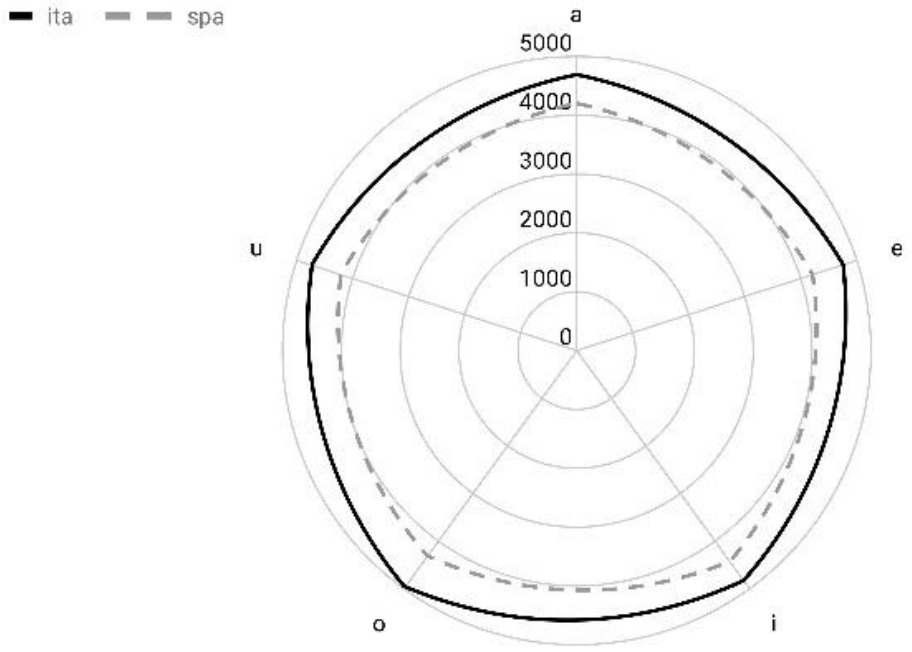

Figure 4. Representation of the CoG range (in Hz) of post-alveolar fricatives for each following vowel for Italian controls and Spanish-speaking learners of Italian.

parameters in function of the expected phonological realization as voiced/voiceless reached the significance threshold in the statistical analysis. In general, [s] occurrences in the positions and contexts where /z/ was expected had a shorter average duration and a higher value of pVDT than [s] occurrences where /s/ was expected. A possible explanation for this tendency could be that learners tend to rely on different cues, i.e., duration and pVDT instead of voice, to acoustically implement the Italian /s ~ z/ contrast. A similar conclusion was reached by Eckman et al. (2014, but cf. also 2015) in their analysis of the acquisition of the same contrast in English by adult native speakers of Spanish. In their investigation, 4 of their 14 subjects, who were transcribed by annotators as not producing the target contrast, displayed a different percent of voicing overlapping the fricative noise between /s/ and /z/. Their hypothesis is that learners were producing a covert contrast, i.e., "in their acquisition of a target-language phonemic distinction, some second-language learners may implement the TL contrast acoustically in a way that is not perceived by native speakers of the TL" (Eckman et al., 2015, p. 7). Most importantly, they claim that the production of these covert contrasts might be one of the necessary stages of all phonological acquisition.

The case of the two bilingual speakers of Basque and Catalan (subjects JAG and NPV respectively) is interesting because their pronunciation of voiced fricative does not show any difference from that of the monolingual subjects. With a native proficiency of languages whose phonological inventories is even richer than Italian in terms of voiced sibilants, it could have been expected that they would have been advantaged in the production of such sounds in the L3 (Antoniou et al., 2015; Kopečková, 2016). Conversely, they displayed the same difficulties as monolingual learners, as far as voiced sibilants are concerned. Some studies have already observed that bilingual speakers are not advantaged in the production of difficult L3 sounds, even if a sound with the same or similar features is present in their native repertoires (Gallardo del Puerto, 2008; Davine et al., 1971; Lambert & McNamara, 1969; Werker, 1986). Recalling the distinction made by Cummins (1979, 1980, 1983, 1984) between Cognitive Academic Language Proficiency (CALP) and Basic Interpersonal Communication Skills (BICS), Gallardo del Puerto (2008) claims that the production of non-native sounds does not benefit so

much from the cognitive advantage offered by bilingualism, since pronunciation, along with oral fluency and listening comprehension, is one of "those linguistic areas that do not rest so much with cognitive development" (Gallardo del Puerto, 2008, p. 12), as opposed to all those skills related to language literacy, such as reading, writing, lexicon and morphology. Moreover, there is also an important physical component to pronunciation, i.e., articulation, for which L3 learners will face the same difficulties as L2 learners (Gallardo del Puerto, 2008).

Concerning the voicing, our Hypothesis III (cf. Section 3) was confirmed: without any voiced counterpart in the Italian inventory, and with a less marked status and a low degree of articulatory difficulty, post-alveolar fricatives were correctly produced by learners as voiceless, in each phonotactic context and experimental task.

As for our second research question, we observed some differences in the production of sibilant fricatives between Italian controls and L1 Spanish learners in terms of Duration and Place of Articulation. Both alveolar and post-alveolar fricatives appeared to be on average shorter in the speech of L2 subjects in comparison to the productions of L1 speakers. The transfer of the duration features in the pronunciation of alveolar fricatives between Spanish and Italian as L2s was already noted by Hernandez (2009, p. 65): in particular, he observed that in the speech of Italian learners of Spanish, the [s] tends to sound overly prominent because it is produced with higher articulatory tension, and thus has longer duration than it is expected in the Target Language (cf. also Canepari, 1979, p. 269), reflecting the phonetic characteristics of the [s] in their L1. In our data, the opposite situation was observed: in Italian, L1 Spanish learners produced shorter sibilant fricatives than the native controls. Learners appear to maintain the L1 articulatory settings when pronouncing L2 sounds. This tendency is also visible in the variation of the Place of articulation and its comparison with that of L1 speakers, especially to produce [s]. In fact, while the post-alveolar fricatives of learners are just slightly more posterior than those of the natives, this tendency is especially visible for the alveolar fricative, where the learners' place of articulation is more posterior than the L1 speakers', mirroring the more posterior production of the [s] in Spanish; however, a major difference could lay in the tongue shape more than in the place of articulation. The outcome of our analysis was indeed expected (cf. Hypothesis I, Section 3): beside the fact that articulatory routines of the L1 are "highly automatized and highly resistant to change" (Zybert, 1997, p. 117), this difference between the Spanish [s] and the Italian [s] is not very salient, thus likely to go unnoticed without proper phonetic instruction (Birdsong, 2007; Bongaerts, 1999), and moreover, it does not lead to the production of a non-native sound that would impair intelligibility.

As for the third research question, the influence of several factors on learners' production of sibilant fricatives was observed. First, the data we obtained on voiced sibilant fricatives lead us to believe that such sounds are challenging for learners, both in terms of perception and production, because of their degree of typological markedness. In line with previous studies on phonological acquisition (cf. Section 2), our results thus seem to support the claim that typological markedness is a good predicting factor of the difficulties that a learner will face in the acquisition of pronunciation in an L2, with less natural, more marked sounds being harder to be perceived and produced in a native-like way (Major, 2001; Sorianello, 2019).

Moreover, in line with the SLM (Flege, 1995; 2003), the subjects' L1 appeared to function as a filter influencing the pronunciation of the target consonants, especially the Italian /s/, which can be considered "similar" to the Spanish voiceless alveolar fricative phoneme. Specifically, through a fine-grained analysis of the spectral characteristics of this segment, we were able to observe the concrete effects of this influence, with the voiceless alveolar fricative being produced by learners with less tension and with a more posterior place of articulation, thus sounding shorter, and more grave and palatal in their speech, which is, as mentioned

in Section 3, one of the features that signal the Spanish accent of an L2 speaker to the Italian native listener (Devís, 2005).

Also, the phonetic environment in which the segments appeared within the word turned out to be a significant predictor of the characteristics assumed by target sounds, at least in terms of duration. Interestingly, both groups showed a very similar pattern of variation for the duration of /s/ and /ʃ/ between the controlled contexts. Finally, we observed a few differences in consonant duration attributable to the effect Task variable, which may suggest some minor stylistic adjustments between the Reading Task, which obviously elicits a more careful and more articulated speech, and the Speaking Task, which elicits a more spontaneous, less careful speech; but also, variation conditioned by the speech rate between the Reading and the Fast-Reading task, for both target sounds.

7. Conclusion

This research provided an extensive description of the phonetic variation of Italian sibilant fricatives in the speech of L2 Spanish-speaking learners, giving a further insight on the factors that influence pronunciation in the interlanguage. This was achieved by addressing the configuration of several fine-grained acoustic features and their variation both through a descriptive and an inferential approach. By outlining the acoustic and articulatory characteristics of what is commonly referred to as foreign accent, investigations such as this represent an important resource for language instruction, since the implementation of focused phonetic adjustments may help learners develop a native-like pronunciation of very complex and marked phonemes, too. Further developments of the research should focus a larger sample of participants: in particular, both male and female learners should be considered, in order to analyze the possible differences between the two genders in the acoustic configuration and the phonetic variation of sibilants in the interlanguage. Moreover, the implementation of perception tasks is necessary for the purpose of giving further aid to our final considerations regarding the priority of perception on production for voiced sibilants, or the implementation of covert contrasts. Finally, the analysis of speakers of Italian L2 with different levels of proficiency could shed some light on how both production and perception of fricatives develop with instruction and exposure.

Author contributions

- Conceptualization: MR, CM
- Methodology: MR, CM, JML
- Investigation: MR
- Resources: MR, CM, JML
- Formal analysis: MR
- Visualization: MR
- Data interpretation: MR, CM, JML
- Writing - original draft: MR
- Writing - review & editing: MR, CM, JML

This paper has been jointly conceived and written by all the authors. For the requirements of the Italian Academia, M. Rossi is responsible for sections 4.2, 5 and 6; C. Meluzzi for sections 1, 3 and 4.1; J.M. Lahoz-Bengoechea for sections 3, 4.3, 7.

References

Albano Leoni, F., & Maturi, P. (2009). *Manuale di fonetica* (3rd ed.). Carocci.

Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, *18*(4), 683–695. https://doi.org/10.1017/S1366728914000777

Balise, R. R., & Diehl, R. L. (1994). Some distributional facts about fricatives and a perceptual explanation. *Phonetica*, *51*(1–3), 99–110. https://doi.org/10.1159/000261962

Bertinetto, P. M. (2010). Fonetica italiana. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore*, *9*(1), 1–30.

Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In O. S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech*

learning (pp. 99–116). John Benjamins. https://doi.org/10.1075/lllt.17.12bir

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9–10), 341–345. https://hdl.handle.net/11245/1.200596

Boersma, P., & Weenink, D. (1992–2021). *Praat: Doing phonetics by computer* (Version 6.1.39) [Computer program]. https://www.praat.org/

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133–159). Lawrence Erlbaum. http://www.jstor.org/stable/44280132

Canepari, L. (1979). *Italiano standard e pronunce regionali*. Cleup.

Canepari, L. (1999). *Dizionario di pronuncia italiana*. Zanichelli.

Canepari, L. (2005). *A Handbook of Phonetics*. LINCOM.

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech: Theory and practice*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087636

Costamagna, L. (2003). Affricates in Italian as L2: The role of psycho-attitudinal parameters. In L. Costamagna, & S. Giannini (Eds.), *La fonologia dell'interlingua: Principi e metodi di analisi* (pp. 95–130). Franco Angeli.

Costamagna, L. (2008). The acquisition of Italian L2 affricates: The case of a Brazilian learner. In A. S. Rauber, M. A. Watkins, & B. O. Baptista (Eds.), *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech* (pp. 138–148). Federal University of Santa Catarina.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Syndicate of the University of Cambridge.

Cruselles, P., Herrero, R., & Jiménez, J. (2017). Propietats acústiques i percepció de contrastos mínims en les sibilants del valencià. *Estudios de Fonética Experimental*, *26*, 93–132. https://raco.cat/index.php/EFE/article/view/331098

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*(2), 221–251. https://doi.org/10.2307/1169960

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, *14*(2), 175–187. https://doi.org/10.2307/3586312

Cummins, J. (1983). Language proficiency, biliteracy and French immersion. *Canadian Journal of Education*, *8*(2), 117–138. https://doi.org/10.2307/1494722

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Multilingual Matters.

Davine, M., Tucker, R., & Lambert, W. E. (1971). The perception of phoneme sequences by monolingual and bilingual elementary school children. *Canadian Journal of Behavioural Sciences*, *3*(1), 72–76. https://psycnet.apa.org/doi/10.1037/h0082251

De Dominicis, A. (1999). *Fonologia comparata delle principali lingue europee moderne*. CLUEB.

Del Saz, M. (2023). (In)complete neutralization in Western Andalusian Spanish. *Estudios de Fonética Experimental*, *32*, 21-42.

Devís, A. (2005). *Percezione dell'accento straniero* [Conference presentation]. VIII Convegno AISC 2005, Napoli, Italy.

Eckman, F. R. (1977). Markedness and the Contrastive Analysis Hypothesis. *Language Learning*, *27*(2), 315–330. https://doi.org/10.1111/j.1467-1770.1977.tb00124.x

Eckman, F. R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, *13*(1), 23–41. https://doi.org/10.1017/S0272263100009700

Eckman, F. R., Iverson, G., & Song J. Y. (2014). Covert contrast in the acquisition of second language phonology. In A. W. Farris-Trimble, & J. A. Barlow (Eds.), *Perspectives on phonological theory and development: In honor*

of Daniel A. Dinnsen (pp. 25–48). John Benjamins. https://doi.org/10.1075/lald.56.05eck

Eckman, F. R., Iverson, G., & Song J. Y. (2015). Overt and covert contrast in L2 phonology. *Journal of Second Language Pronunciation*, *1*(2), 254–278. https://doi.org/10.1075%2Fjslp.1.2.06eck

Endo, R., & Bertinetto, P. M. (1999). Caratteristiche prosodiche delle così dette "rafforzate" italiane. In R. Delmonte, & A. Bristot (Eds.), *Aspetti computazionali in fonetica linguistica e didattica delle lingue: Modelli e algoritmi* (pp. 243–255). Unipress.

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. 233–276). York Press.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. O. Schiller, & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355). Mouton de Gruyter. https://doi.org/10.1515/9783110895094.319

Fradejas Rueda, J. M. (1997). *Fonología histórica del español.* Visor Libros.

Fuchs, S., & Toda, M. (2010). Do differences in male versus female /s/ reflect biological or sociophonetic factors? In S. Fuchs, M. Toda, & M. Żygis (Eds.), *Turbulent sounds: An interdisciplinary guide* (pp. 281–302). De Gruyter. https://doi.org/10.1515/9783110226584.281

Gallardo del Puerto, F. (2007). Is L3 Phonological Competence Affected by the Learner's Level of Bilingualism? *International Journal of Multilingualism*, *4*(1), 1–16. https://doi.org/10.2167/ijm042.0

Giannini, S. (2003). Typological comparison and interlanguage phonology: Maps or gaps between typology and language learning of sound systems? In A. Giacalone Ramat (Ed.), *Typology and second language acquisition* (pp. 403–440). De Gruyter. https://doi.org/10.1515/9783110891249.403

Greenberg, J. H. (1966). *Universals of Language*. MIT Press.

Harrington, J. (2010). *Phonetic analysis of speech corpora.* Wiley-Blackwell.

Harris, J. (1969). *Spanish Phonology.* MIT Press.

Hernandez, I. A. (2009). Tensión y cantidad en la corrección fonética con italófonos: Una propuesta para el aula de ELE. *RESLA: Revista Española de Lingüística Aplicada*, *22*, 59–76.

Hidalgo Navarro, A., & Quilis Merlín, M. (2012). *La voz del lenguaje: Fonética y fonología del español.* Prosopopeya.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252–1263. https://doi.org/10.1121/1.1288413

Kingston, J. (1993). Size, structure and markedness in phonological inventories. *University of Massachusetts Occasional Papers in Linguistics*, *16*(1), 69–123.

Kopečková, R. (2018). Exploring metalinguistic awareness in L3 phonological acquisition: The case of young instructed learners of Spanish in Germany. *Language Awareness*, *27*(1–2), 153–166. https://doi.org/10.1080/09658416.2018.1432629

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Ladefoged, P., & Maddieson, I. (1996). *The Sounds of the World's Languages.* Blackwell.

Lambert, W. E., & Macnamara, J. (1969). Some cognitive consequences of following a first-grade curriculum in a second language. *Journal of Educational Psychology*, *60*(2), 86–96. https://psycnet.apa.org/doi/10.1037/h0027092

Lapesa, R. (1942). *Historia de la lengua española.* Gredos.

Loporcaro, M., & Bertinetto, P. M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence. *Journal of the International Phonetic Association*, *35*(2), 131–151. https://doi.org/10.1017/S0025100305002148

Maddieson, J. (1984). *Patterns of sounds.* Cambridge University Press.

Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Lawrence Erlbaum Associates.

Maturi, P. (2014). *I suoni delle lingue, i suoni dell'italiano: Nuova introduzione alla fonetica.* Il Mulino.

Mazzotta, G. (1984). *Italiano, francese, spagnolo: Sistemi fonologici a confronto.* Minerva Italica.

Meluzzi, C. (2014). *Le affricate dentali nell'italiano di Bolzano. Un approccio sociofonetico.* [Doctoral dissertation]. Università di Pavia/Bolzano, Italy.

Meluzzi, C. (2020). *Sociofonetica di una varietà di koinè: Le affricate dentali nell'italiano di Bolzano.* Franco Angeli.

Navarro, T. (1957). *Manual de pronunciación española.* Consejo Superior de Investigaciones Científicas.

Ohala, J. J., & Solé, M. J. (2010). Turbulence and phonology. In S. Fuchs, M. Toda, & M. Żygis (Eds.), *Turbulent sounds: An interdisciplinary guide* (pp. 37–102). De Gruyter. https://doi.org/10.1515/9783110226584.37

Olmo-López, A. (2017). *Análisis acústico de la oposición entre oclusivas tensas y laxas del chino mandarín* [Master's Thesis]. Universidad Internacional Menéndez Pelayo, Spain.

R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* https://www.r-project.org/

Regan, B. (2020). El [ʃ]oquero: /tʃ/ variation in Huelva Capital and surrounding town. *Estudios de Fonética Experimental, 29*, 55-90.

Romito, L., Frontera, M., & Ciardullo, M. A. (2016). Studio acustico-percettivo di contrasti fonemici in Italiano L2/LS. In R. Savy, & I. Alfano (Eds.), *Studi AISV 2: La fonetica nell'apprendimento delle lingue* (pp. 23–44). Officinaventuno. https://doi.org/10.17469/O2102AISV000002

Schmid, S. (1999). *Fonetica e fonologia dell'italiano.* Paravia scriptorium.

Schmid, S. (2004). Dallo spagnolo all'italiano: Elementi di analisi contrastiva nella prospettiva dell'acquisizione. In C. Ghezzi, F. Guerini, & P. Molinelli (Eds.), *Italiano e lingue immigrate a confronto. Riflessioni per la pratica didattica* (pp. 197–219). Guerra Edizioni. https://doi.org/10.5167/uzh-112601

Schmid, S., & Wachter, S. (2015). Le ostruenti sonore nella pronuncia dell'italiano di apprendenti svizzero-tedeschi. In M. Vayra, C. Avesani, & F. Tamburini (Eds.), *Il farsi e disfarsi del linguaggio: Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio* (pp. 203–217). Officinaventuno. https://doi.org/10.17469/O2101AISV000013

Schwegler, A., Kempff, J., & Ameal-Guerra, A. (2010). *Fonética y fonología españolas: Teoría y práctica.* Wiley.

Sorianello, P. (2002). I suoni fricativi dell'italiano fiorentino. *Quaderni del Laboratorio di Linguistica*, *3*, 26–39.

Sorianello, P. (2019). Le consonanti affricate: strategie di acquisizione in italiano L2. *Italiano LinguaDue*, *11*(1), 68–88. https://doi.org/10.13130/2037-3597/11845

Torreblanca, M. (1983). Las sibilantes sonoras del Oeste de España: ¿Arcaísmos o innovaciones fonéticas? *Revista de Filología Románica*, *1*, 61–108.

Torreblanca, M. (1986). La sonorización de /s/ y /θ/ en el noreste toledano. *Lingüística Española Actual*, *8*(1), 5–19.

Trubetzkoy, N. S. (1939). *Grundzüge der Phonologie.* Vandenhoeck & Ruprecht.

Univaso, P., Martínez Soler, M. & Gurlekian, J.A. (2014). Variabilidad intra- e inter-hablante de la fricativa sibilante /s/ en el español de Argentina. *Estudios de Fonética Experimental*, *23*, 95-124.

Werker, J. (1986). The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics*, *7*(2), 466–472. https://psycnet.apa.org/doi/10.1017/S0142716400007360

Zybert, J. (1997). Acquisition of L2 phonetic features. *Papers and Studies in Contrastive Linguistics*, *32*, 103–117.Żygis, M. (2008). On the avoidance of voiced sibilant affricates. *ZAS Papers in Linguistics*, *49*, 23–45. https://doi.org/10.21248/zaspil.49.2008.362

Żygis, M., Fuchs, S., & Koenig, L. L. (2012). Phonetic explanations for the infrequency of voiced sibilant affricates across languages. *Laboratory Phonology*, *3*(2), 299–336. https://doi.org/10.1515/lp-2012-0016