Open Research Online

The Open University's repository of research publications
and other research outputs

The acquisition of voicing assimilation by advanced Hungarian learners of Spanish

Journal Item

For guidance on citations see FAQs.

oro.open.ac.uk

The acquisition of voicing assimilation
by advanced Hungarian learners of Spanish

Zsuzsanna Bárkányi
The Open University

This study examines the effect of explicit phonological instruction on the acquisition of variable /s/ voicing in Spanish by advanced Hungarian learners. Hungarian and Spanish have very similar, yet not identical, voicing assimilation (VA) systems; the most important difference lies in the pre-sonorant context as sonorant consonants trigger voicing assimilation in Spanish but not in Hungarian. Data were collected in acoustic experiments from 7 native speakers of Northern Peninsular Spanish and 12 Hungarian university students who were advanced learners of Spanish. The latter group was tested twice: before and after a three-month *Spanish phonetics and phonology* course. Our data reveal that this amount of instruction is not enough for L2 speakers to overcome their first language VA system, which might be attributed, in part, to the variable allophonic nature of the process and the similarity between the two languages' VA systems.

Keywords: voicing assimilation, pre-sonorant voicing, Spanish, foreign language acquisition, phonological instruction

1. Introduction

Voicing and phonological processes related to voicing like voicing assimilation (VA) have received little attention in studies dealing with second language (L2) and foreign language (FL) acquisition. The only aspect of voicing which has been thoroughly studied is voice onset time (VOT). Research on L2 acquisition has shown that speakers transfer VOT values from their mother tongue into the L2. Suomi (1980) is one of the earliest examples who reports on the production of English stops by native speakers of Finnish. Flege (1991) observed that native Spanish speakers who were early learners of English produced native-like English VOT values in initial voiceless stops, while late learners produced longer VOT values than in Spanish but shorter ones than native English speakers.

Hungarian and Spanish are both so called true voice languages: voiceless stops are produced with a short voicing lag (voicing begins shortly after the release of the stop), and voiced stops are produced with vocal fold vibration throughout the closure phase. This means that Hungarian learners of Spanish do not have to adjust the phonetic realisation of voiced and voiceless stops in L2/FL Spanish.

We do not know exactly yet which phonological properties of the first language (L1) are the most likely to be transferred into the L2. In general, there are not many studies dealing with assimilatory processes in L2 systems and only a few of these tackle voicing (see also 2.3). Those few studies that focus on the L2/FL acquisition of VA mostly investigate the acquisition of VA in English as a Second Language or by native English speakers of English. Simon (2010) examines the productivity of voicing and devoicing rules in Dutch-English interlanguage. She observes L1 transfer of (de)voicing processes into L2 English and claims that the extent of this transfer might be suppressed or reduced as a result of explicit learning. Skranitzl and Šturm (2017) study VA in L2 English by speakers of Czech and Slovak and demonstrate that there is a clear effect of first language transfer in the implementation of VA. Their claim is based on the fact that pre-sonorant voicing is clearly observable in the productions of Slovak speakers, while in the productions of Czech speakers it is not. The study also demonstrates that the effect of L1 transfer is correlated with the strength of their foreign accent.

The only study that tackles VA in Spanish is Schmidt (2014), who examines /s/ voicing in L2 Spanish by native English speakers. The author observes that only a limited number of advanced L2 learners show contextual voicing effects in Spanish, which means that even advanced learners preserve the voicing patterns of English. The focus of the present study is very similar to Schmidt's (2014) in that it examines to what degree regressive voicing assimilation (RVA) of Spanish /s/ is attested in the speech of advanced L2/FL learners. On the other hand, it differs from the aforementioned study in several ways. Firstly, we investigate the productions of native speakers of Hungarian, a language that belongs to the same typological group with regard to voicing as Spanish (both are true-voicing languages with RVA, unlike English). Secondly, in Schmidt (2014) all voiced consonants are treated as belonging to the same category (i.e., voiced consonants) and are expected to behave in the same way with regard to VA, while this study examines the similarities and differences in the voicing propensity of voiced obstruents compared to sonorant consonants in the Spanish interlanguage of advanced learners whose native language is Hungarian.

Thus the focus of the present study is, on the one hand, to find out whether explicit instruction and awareness raising (a three-month *Spanish phonetics and phonology* seminar) have a beneficial effect on the acquisition of the variable /s/ voicing in Spanish, which clearly is a feature that contributes to more native-like pronunciation but probably has little impact on comprehensibility. On the other

hand, it aims to deepen our understanding of the mechanisms governing phonological transfer from L1 into L2/FL. More specifically, we focus on the acquisition of a post-lexical assimilatory process, voicing assimilation, between two languages that have very similar, yet not identical, VA systems.

2. Background

2.1   Regressive voicing assimilation in Hungarian

Hungarian is an RVA language where adjacent obstruents must agree in their voicing feature, that is: voiced obstruents voice preceding voiceless obstruents (2a); voiceless obstruents devoice preceding voiced obstruents (2b); and VA is right-to-left iterative (2c). As shown in Table 1, Hungarian has a symmetrical obstruent system with contrastive voiced-voiceless pairs at each place of articulation, so /s/ and /z/ contrast word-initially (1a), word-finally (1b), and within the word (1c). Please, note that /s/ in Hungarian is spelt as 'sz' and /z/ is spelt as 'z'.

Table 1. The obstruent inventory of Hungarian

	Stop	Fricative	Affricate
Bilabial	p – b		
Labiodental		f – v	
Alveolar	t – d	s – z	ts – dz
Postalveolar		ʃ – ʒ	tʃ – dʒ
Palatal	c – ɟ		
Velar	k – g		

(1)   a.   *szár* 'stem' vs. *zár* 'lock'
       b.   *mész* 'whitewash' vs. *méz* 'honey'
       c.   másznak 'climb.pl.3.pres' vs. máznak 'gloss.dat'

(2)   a.   /tb/ → [db]: *hát-ba* 'back.ill'; *két barát* 'two friends'
            /ʃb/ → [ʒb]: *has-ba* 'stomach.ill'; *hús bolt* 'meat shop'
       b.   /bt/ → [pt]: láb-tól 'foot.abl'; láb torna 'foot exercise'
            /zt/ → [st]: *víz-től* 'water.abl'; *víz torony* 'water tower'
       c.   /gdh/ → [kth]: smaragd-hoz 'emerald.allat'

Unlike in many surrounding languages (e.g., German, Slovak), word-final obstruents do not devoice in Hungarian (3). Sonorant consonants do not participate in RVA, so obstruents maintain the voicing contrast before sonorants both within the word (4a) and across a word-boundary (4b).

(3)   *láb-ak* [b] 'foot.PL' ~ *láb* [b] 'foot'
       láp-ok [p] 'marshland.PL' ~ *láp* [p] 'marshland'
       méz-ek [z] 'honey.PL' ~ *méz* [z] 'honey'
       mesz-ek [s] 'limestone.PL' ~ *mész* [s] 'limestone'

(4)   a.   *plakát* [pl] 'poster', *blöki* [bl] 'doggy', *sróf* [ʃr] 'screw', *zrí* [zr] 'fuss'
       b.   /tm/ → [tm] (*[dm]): *két mag* 'two seeds'
             /sl/ → [sl] (*[zl]): *kész leves* 'ready soup'

According to the traditional generative literature, RVA in Hungarian is categorical, exceptionless, and completely neutralising (Vago, 1980; Zsigri, 1994; Siptár & Törkenczy, 2000). This means that voiceless and devoiced, or contextually voiced and underlyingly voiced, segments cannot be distinguished on the basis of their phonetic and phonological behaviour. More recent acoustic phonetic studies partially contradict this. The first notably different approach to RVA in Hungarian is presented in Jansen (2004), who concludes that RVA leads to incomplete neutralisation. He found residual traces of the underlying contrast between /k/ and /g/, and /ʃ/ and /ʒ/ before voiced obstruents. Gráczi (2010), testing nonsense words also claims that there is no complete neutralisation. She claims that consonant duration and vowel – consonant duration ratio does show vowel shortening effects in Hungarian similar to pre-fortis clipping languages (e.g., Harris, 1994; Wells, 2000).

Bárkányi and G. Kiss (2015) study VA in Hungarian (and Slovak) with a special focus on pre-sonorant voicing and show that obstruents in pre-sonorant position preserve their laryngeal contrast, while in pre-obstruent position the process is mostly categorical with sporadic significant differences between voiced and voiceless obstruents. Bárkányi and G. Kiss (forthcoming) study word-final obstruent clusters in VA contexts and conclude that RVA in Hungarian is categorical, but partially contrast-preserving, and stops and fricatives are not affected in the same way by the process. Fricatives are more prone to resist voicing before a voiced obstruent and are more likely to be devoiced utterance-finally than stops. Partial contrast preservation is achieved by duration-related cues that prevent complete laryngeal neutralisation. The perceptual relevance of the observed partial contrast preservation, however, is not well understood yet. The only study trying to establish the cut-off value between voiced and voiceless obstruents in Hungarian is by Bárkányi and Mády (2012), who examined the perception of utterance-final /s/ vs. /z/ using synthesised speech. There are no studies on the cut-off value of voicing between voiced and voiceless consonants in potentially assimilating environments in Hungarian, which means that the observed acoustic differences may or may not be perceived by speakers of Hungarian, therefore, further research will have to clarify whether RVA in Hungarian is partially or completely neutralising from a perceptual point of view.

2.2 /s/ voicing in Spanish

Spanish /s/ voicing presents a special case within RVA languages since there is no phonological voicing contrast in the language at this point of the segment inventory. The Central-Northern Peninsular dialect has two voiceless sibilant fricatives, an interdental /θ/ and an apico-alveolar /ṣ/. All the other dialects have only one sibilant fricative /s/ which has a wide range of dialectal and individual realisations, from apical to laminal, interdental, etc. (Quilis, 1993). A number of studies dealing with VA in different languages focus on the question of whether it is a low-level coarticulatory phonetic process or a feature-changing phonological process which neutralises the underlying laryngeal specifications of the obstruents in pre-obstruent (or any given) context. It is obvious that in the case of Spanish we cannot speak of laryngeal neutralisation given that there is no contrasting voiced sibilant in the language. Furthermore, Spanish clearly shows a preference for open syllables – coda obstruents are fragile and there is high variability in their realisations (Hualde, 2005) – therefore /s/ voicing in Spanish can only be studied in dialects where syllable-final /s/ rarely undergoes aspiration and deletion. In these dialects, when /s/ is followed by a voiced consonant, a voiced obstruent (5a), sonorant (5b), or semivowel (5c) within the same word or across a word-boundary, /s/ gets partially or fully voiced (e.g., Hualde, 2005).

(5)   a.   *esbelto* [zβ] 'slim', *es bueno* [zβ] 'it's good'
       b.   *isla* [zl] 'island', *es largo* [zl] 'it's long'
       c.   *deshielo* [zj] 'thaw', *los hielos* [zj] 'the ices'

The majority of phonologists who have studied /s/-voicing in Spanish were seeking an answer to whether this process in Spanish is gradient or categorical. Romero (1999) investigated the effects of VA on gestural coordination in one person's speech with the help of electromagnetic articulometry. The author concluded that assimilation in Spanish is the result of gestural blending and shows a gradation of voicing that goes from most voiced for the intervocalic voiced stops, to least voiced for the /s/+voiceless stop sequence. The environment (whether /s/ and a following stop were within the same word or across a word boundary) did not turn out to be significant, while the place of articulation did: voicing was consistently higher in labials than in dentals or velars in single word-initial voiced stops. Interestingly, the single voiceless category (i.e. a single voiceless obstruent) showed levels of voicing that are higher than those of the voiced cluster. According to Romero (1999), this might have two explanations: either Spanish voiceless stops are not completely voiceless in colloquial speech, or there is actually no assimilation of voicing to the expected degree. The author concludes that VA in Spanish is not a categorical phenomenon but a gradual process that does not affect all sequences equally.

Schmidt and Willis (2011) report an acoustic experiment on Mexican Spanish /s/ voicing. In this study, data was divided into voiced vs. voiceless realisations: tokens of 60–100% of voicing were labelled as *voiced*, while tokens in which the majority of the segment was voiceless or contained up to 40 ms voicing were labelled as *voiceless*. The authors note that there is considerable left-edge voicing, that is, a continuation of the voicing of the preceding vowel into the following sibilant. In intervocalic position, this was 0–34 ms whereas in /s/ before a voiceless consonant it went up to 37 ms. The number of voiced realisations in the non-voicing context was 5% for /s/ + voiceless consonant and 9% for intervocalic /s/, while it was 63% in the voicing context. This means that 37% of the /s/ in the voicing context was realised voiceless, and one tenth of intervocalic sibilants was actually voiced. Note that Schmidt and Willis (2011), similarly to Schmidt (2014), did not break up the voicing context into pre-sonorant vs. pre-voiced-obstruent environments. They found more voicing in the phrase internal context than phrase final, but the duration of the sibilant itself also varied according to the prosodic environment: /s/ in the onset of the syllable was significantly longer than in the coda. This study also found a high rate of individual variation, and observed that male speakers tend to voice more than female speakers, especially in phrase-final position. The authors conclude that the process is far from categorical.

García (2013) collected read and spontaneous speech data from 15 native speakers of Spanish from different dialect regions and concluded that /s/ voicing was highly variable (both across subjects and within subjects) and was influenced by speech rate but not speech formality. Interestingly, the most variation between subjects was found during normal speech rate. García argues that the process is the result of gestural blending/overlap rather than a phonological rule. This claim is supported and illustrated with spectrograms where the fricative begins without any vibration of the vocal cords and then becomes voiced approximately halfway through. The author claims that voicing occurs in anticipation of the following voiced segment. García's observation is quite surprising as most acoustic studies that found partial /s/ voicing in Spanish (including the present one) observe that, in the vast majority of cases, voicing is preservative, that is, the voicing tail of the preceding vowel continues into the fricative and then dies out, giving evidence for articulatory inertia rather than anticipation. Schmidt and Willis (2011) explicitly state that they started data analysis with the assumption that there would be both left-edge and right-edge voicing but that right-edge voicing was so infrequent that it was eliminated from the calculation (see also Campos-Astorkiza, 2015).

Campos-Astorkiza (2012) also claims that VA in Spanish is a result of gestural blending and that it is a gradual process. Her theory predicts that sonorant consonants voice the preceding segment more than voiced obstruents do. The first study that explicitly compares the voicing propensity of sonorants and voiced

stops (independently of whether they are realised as stops or as approximants) is Bárkányi (2014), which demonstrates that this prediction is not borne out, but rather that sonorants tend to voice less than voiced obstruents. Campos-Astorkiza (2015) confirms Bárkányi's results and claims that the manner of articulation of the following consonant (voiceless obstruents vs. voiced obstruents vs. sonorants) has a significant effect on the percentage of voicing in the fricative. She explains the low degree of RVA in pre-sonorant position following Recasens and Mira (2012) by the conflicting articulatory and aerodynamic requirements for nasals and laterals versus /s/. The author also studied the role of stress with regard to /s/ voicing in Spanish in a series of acoustic experiments. Campos-Astorkiza (2014) did not find a difference in /s/ voicing between stressed and unstressed syllables, neither did she find an effect of word-boundaries, that is, the same amount of voicing could be observed within and across words. The presence of an intonational phrase boundary, however, did have a significant effect of on /s/ voicing by limiting it, although not categorically. Campos-Astorkiza (2015) established four categories when examining the effects of stress on /s/ voicing and claims that stress has no overall effect, but pair-wise comparisons reveal that the post-stress context (*tapas duritas*) displays more voicing than the other conditions. The author concludes that /s/ voicing in Spanish is gradient and often incomplete rather than categorical.

A study that partly contradicts the aforementioned papers is Bárkányi's (2014), who also found significant inter- and intra-speaker variation in /s/ voicing in Spanish, but she claims that gradient data is a result of categorical, but optional, VA in Spanish (see Strycharczuk, 2012 for a similar claim with regard to intervocalic /s/ voicing in Quito Spanish). Bárkányi's conclusions are based on an experiment with seven native speakers of Northern Peninsular Spanish. In the classification of the data, she followed Schmidt and Willis (2011) who found that 14–15 ms voicing is simply coarticulatory and does not indicate a planned gesture and that there can be up to a 37 ms long voice tail where we expect a completely voiceless realization. Bárkányi (2014) divided the /s/ realizations into three groups: tokens that contain up to 15 ms voicing which does not exceed 40% of the segment were labelled *voiceless*; *partially voiced* segments were those which have a 16–35 ms long voiced part and corresponds to at least 40% of the duration of the segment (the average duration of /s/ in her study is 78 ms.); and instances that contain a voiced portion more than 35 ms long were labelled as *voiced*. She found that native speakers' intended articulatory gestures seemed to be categorical in almost 90% of the cases: /s/ was realised as either voiced or voiceless according to this classification, and it was only 12.9% of pre-sonorant /s/ and 13.4% of /s/ before a voiced obstruent that were indeed realised as partially voiced. The author found a significant difference in the likelihood of voicing before sonorants as opposed to before voiced obstruents; while Spanish speakers voiced the sibilant before voiced obstruents in almost three

quarters of the cases, they only did so in about half of the cases in pre-sonorant position. Thus inter- and intra-speaker variability stems from the optional character of the process which overall produces a gradient effect.

2.3   Pronunciation instruction and L2 acquisition

When learning a new language, our mother tongue undoubtedly exerts an influence on our pronunciation, which is perceived by native speakers as a foreign accent. The divergent pronunciation might be due to segmental errors as well as incorrect suprasegmental patterns or the combination of these. Odlin (1989), based on Moulton (1962), describes four types of segmental errors: phonemic errors, phonetic errors, allophonic errors, and distributional errors. Odlin argues that the term *transfer* over *inference* is preferable when describing L2 production because in this way positive and negative transfer can be contrasted. Negative transfer is the influence of the native language (or some other previously learned language) which leads to inaccuracies in the acquisition or use of the target language. In contrast, the influence of the native language can also be helpful when the similarities between the native and the target language facilitate L2 production. This latter case is referred to as positive transfer. There is ample evidence that phonological transfer occurs on the segmental level (see Archibald & Young-Sholten, 2003 for an overview of L2 acquisition of segmental phonology), on the suprasegmental level (White & Mattys, 2007 on the transfer of L1 rhythm into the L2), as well as on the level of phonotactic patterns (Davidson & Wilson, 2016 and the references therein). Since the present article studies RVA in the interlanguage of advanced learners, it is the L2 acquisition of allophonic features that is of special interest to us.

Cebrian (2000) studied final devoicing and VA in the English interlanguage of Catalan learners. He found that both devoicing and VA are transferred into English, but devoicing to a much greater extent than VA. He concludes that "the kinds of L1 rules that cause interference are low level, exceptionless, sensitive to context, and applicable across the board in the L1" (p. 7), that is, post-lexical rules are more likely to be transferred from L1 to L2 than lexical phonological rules. This is supported by Kissling (2013), who found in an experiment with 95 L1 English participants learning L2 Spanish that instruction did not affect all target phones in equal measure and that the pronunciation of voiced approximants [β ð ɣ], that is, the allophonic variants of the voiced stops /b d g/, did not improve. Díaz-Campos (2004) also claims that Spanish approximants are resistant to instruction and are late acquired. Face and Menke (2009), on the other hand, found that 4th-year Spanish majors in their last semester of studies produced /b d g/ as approximants in over 50% of the tokens.

Although pronunciation has been of interest for language learners, language teachers, and the general public for centuries, very little research focused on L2

pronunciation until the second half of the 20th century. Up until the turn of the 21st century, the goal of pronunciation instruction was to achieve a native-like pronunciation. Jenkins (2000) is one of the first to draw attention to adequate goal setting, which is not the acquisition of a close-to-native pronunciation, but rather intelligibility and comprehensibility in the target language. Experiment-based studies examining the effect of formal instruction in pronunciation have yielded contradictory results. It is often difficult to compare different studies as there are differences in the experimental methods (e.g., word list reading versus spontaneous speech tasks), in the instructional techniques (e.g., awareness raising, listen and repeat exercises, self-paced computer assisted practice, etc.), as well as the learning contexts and duration. According to Thomson and Derwing (2015), 82% of the studies on pronunciation instruction reported some kind of significant improvement. Elliot (1995) found that pronunciation when not formally taught remained relatively stable in spite of the high degree of input the learners processed. It is accepted by a number of scholars that explicit phonetic instruction has a beneficial effect on comprehensibility but does not really contribute to achieving a more native-like pronunciation (e.g., Saito, 2011; Lafford & Salaberry, 2003). Learners, on the other hand, aspire to have a native or native-like pronunciation. Derwing and Munro (2009) also observe that accentedness might have only a minor impact on comprehensibility and intelligibility, but learners are still concerned with reducing the accentedness of their speech.

The evaluation of pronunciation instruction is not straightforward either. Derwing, Munro, and Wiebe (1998) evaluated three types of pronunciation instruction in the speech of 48 adult ESL students at the intermediate level. The three groups involved segmental accuracy instruction, general speaking habits and prosody instruction (global group), and no specific pronunciation instruction over a 12-week course. Forty-eight English native speakers evaluated for accentedness, comprehensibility, and fluency in two tasks: sentence production and a continuous narrative. The comprehensibility results of sentence production show that both the segmental and the global group improved significantly, but the control group did not. As for accentedness, all three groups improved, but the segmental group was significantly better than the other two. In the narrative data, on the other hand, it was the global group that showed improvement in comprehensibility and fluency, but none of the groups improved in accentedness.

The second most studied language with regard to pronunciation instruction is L2 Spanish. Most studies focus on the acquisition of voiceless stops, voiced approximants, and rhotics by L1 English speakers (see for instance Lord, 2005; González López & Counselman, 2013; Kissling, 2013). Elliot (1997) in a study with 66 undergraduate students of intermediate Spanish also addressed assimilation processes, such as assimilation in the point of articulation (*un peso* [um'peso]) and

RVA (*mismo* [ˈmizmo]). He concluded that students had not mastered the assimilation processes by the end of the semester, however they did experience significant improvement in the pronunciation of liquids and stop phonemes.

2.4   From L1 Hungarian to L2/FL Spanish pronunciation

We have seen that the main differences between the two languages with regard to VA is that (i) both /s/ and /z/ are phonemic in Hungarian and contrast in all prosodic positions (1), while Spanish has only one voiceless sibilant fricative with a wide range of realisations; (ii) RVA in Hungarian is categorical, RVA in Spanish is gradient/optional; and that (iii) sonorant consonants voice in Spanish, while they do not participate in RVA in Hungarian. We will use the terms L2 and FL indistinctively in the present paper as all of our learners speak English in addition to Spanish, but some only at a very basic level while others have a far better proficiency in English than in Spanish, so for some learners Spanish is an L2, whereas for others it is an L3. Out of the 12 participants, six were learning other languages at the time of the recordings, but they all reported to be more proficient in Spanish than in these additional languages.

To summarise, if Hungarian learners of Spanish are moving towards a native-like pronunciation, they do have to learn the direction of VA in Spanish, but they do have to learn to apply RVA in a new phonetic context as sonorant consonants also voice the preceding voiceless obstruent. They also have to learn that the application of RVA is optional in Spanish as opposed to its categorical nature in Hungarian.

Thus the aim of the present paper is not to examine the relationship of (foreign) accents and RVA in Spanish (this might be the topic of future research), but to contribute to our understanding on the transfer (and acquisition) of automatic allophonic processes from L1 to L2/FL.

The specific research questions we aim to answer are the following:

1. To what extent do advanced Hungarian speakers of L2/FL Spanish acquire /s/ voicing in Spanish? Can we observe the effects of negative and/or positive transfer from L1?
2. Does explicit instruction have an effect on the acquisition of the subtle subphonemic process of /s/ voicing in Spanish?

3.　Methodology

3.1　Participants

12 second-year BA students (10 female, 2 male) aged 20 to 24 from the Department of Hispanic Studies at the Eötvös Loránd University Budapest participated in the experiment. Their demographic and linguistic profile is provided in Table 2. Eight of them reported to have the intention to join the teacher training MA later on. They had all been learning Spanish for at least 4 years, but most of them started in the first year of secondary school (i.e., at the age of 14). Students at the end of the first year of their university studies have to pass a proficiency exam which corresponds to level B2-C1 of the Common European Framework of Reference for Languages. None of the participants spent more than three months in a native environment. All the students in the learner group of this experiment learned Spanish in classroom settings, took the same proficiency exam (albeit with different results) the year before, all attended the same *Spanish phonetics and phonology* class, and none of them spent more than three months in a Spanish-speaking country.

The productions of the learner group are compared to a group of native speakers (henceforth the control group). The control group consisted of 7 speakers (4 female and 3 male) of Northern(-Central) Peninsular Spanish, a dialect where syllable-final /s/ does not undergo weakening. They were students or professors at the University of Oviedo, unaware of the purpose of the experiment. Their ages ranged between 22–41 years. None of the participants reported any speaking or hearing disorder and none of them received any compensation for their participation. Although gender and dialectal differences might influence /s/ voicing in Spanish – see 2.2 –, the detailed discussion of these issues is outside the scope of this paper as it focuses on L2 learners' pronunciation.

Table 2. Demographic and linguistic attributes of the learner group

Participant	Gender	Age	Languages spoken (apart from Spanish and English)	Years of studying Spanish
S7	F	20		6
S8	M	21	Sanskrit	4
S9	F	20		4
S10	F	20		6
S11	F	24	French	5
S12	F	20		6
S13	F	21		7
S15	F	22	Catalan	5
S16	F	20	German	6
S17	F	20	German, Catalan	6
S18	F	20		6
S19	M	21	Italian, Catalan	7

3.2 Stimuli

/s/ was tested in the following positions:

– absolute final position, that is, utterance-finally (*autobús* # 'bus')
– before a voiced obstruent (e.g., *esbelta* 'slim.FEM.', *las botas* 'the boots');
– before a sonorant consonant (e.g., *eslogan* 'slogan', *las loterías* 'the lotteries');
– before a voiceless obstruent (e.g., *peste* 'plague', *las totalmente* 'the totally').

Stimuli were embedded in neutral carrier sentences (the complete list of test sentences can be found in Appendix 1). /s/ was tested both within the word and across a word-boundary (except for the utterance-final position), but within the same syntactic phrase so that no intonational phrase boundary intervened between the target segment and the trigger. We did not control for the place of stress – which might play a role in voicing – as our focus was on the acquisition of VA in Spanish by L1 Hungarian learners and the effect of explicit phonological instruction on this learning.

3.3 Instruction period

Students were recorded at the beginning and at the end of a three-month *Spanish phonetics and phonology* seminar. They had one 90-minute session a week for 12 weeks in groups of 16–20 people. The primary purpose of the course was not pronunciation practice but rather awareness raising, a contrastive analysis of the phonological systems of Hungarian and Spanish, part of which involved pronunciation practice. Students were introduced to the basic phonetic and phonological concepts and discussed topics in Spanish phonology and dialectology during the course. Explanations of grapheme-phoneme correspondences were given and the differences in the articulation of analogous Spanish-Hungarian sounds and the phonological environments in which these sounds are produced were highlighted. The places and manners of articulation were explained in detail and illustrated with animated diagrams of the vocal tract (for the latter, the website of Spanish Sounds developed by the University of Iowa was used.)

As for the teaching of RVA in Spanish, the Presentation, Practice, Production model was adopted. In the presentation phase, learners contrasted the Hungarian and the Spanish laryngeal systems, and explored the differences and similarities with regard to RVA in the two languages. The practice phase involved both written and oral exercises. For instance, learners had to identify in a list of words which instances of /s/ were (or could be) pronounced voiced. The list also contained words which allowed for negative transfer from L1 or from English, including *zebra* 'zebra' where the orthographic 'z' in Hungarian (*zebra*) is pronounced as [z] while in

Spanish it should be [s] or [Θ] depending on the dialect; *presente* 'present' where the orthographic 's' in English is pronounced as a voiced sibilant but should be [s] in Spanish; *eslogan* 'slogan' where the /s/ should be voiced in Spanish, but in the Hungarian cognate *szlogen* it is pronounced as [s] with no pre-sonorant voicing, etc. Learners also had several transcription exercises in which they had to give the phonetic transcription of short texts or very short (1 or 2-sentence) audios. Students also performed a number of 'listen-and-repeat' exercises. The practice exercises were written by the author or taken from Schwegler et al. (2010).

3.4   Testing procedure

Subjects read the test sentences and fillers from a monitor screen in a randomized order, which was generated by Speech Recorder (Draxler & Jänsch, 2012). In the first session, that is, before the phonetics seminar, each test sentence was read five times, but the first reading was considered as the familiarisation phase and was not analysed. We applied the same method in the case of the control group. All the participants were naive as to the aims of the experiment; learners were told that the only aim was for them to see how a phonetic experiment runs and to help in a research, while native speakers were told that the recordings would be used for research purposes. Recordings were made in a sound-proof cabin in the Research Institute for Linguistics of the Hungarian Academy of Sciences, and in a sound-attenuated room at the University of Oviedo, in both cases with a Sony ECM-MS907 microphone connected to a laptop through an M-Audio Mobile Pre USB preamplifier external sound card. The material was recorded at a 44,100 Hz sampling rate, and was resampled at 22,050 Hz for the acoustic measurements. Only the learner group participated in the second, post-instruction session which was a bit shorter than the first one as only three readings were taken into account. By this time the students had figured out that their pronunciation was under scrutiny, but they still did not know the exact nature of the experiment. It was made clear that the recordings do not form part of their course assessment.

There were 17 test sentences with one sentence including two test conditions, thus 18 items were read 4 times by 12 students which gives 864 recordings for the learner group in the first session, and 648 recordings for the learner group in the second session as there were only three repetitions; and 18 test items, 4 readings by 7 speakers gives 504 test items for the control group. Four items had to be removed from the first session of the learner group and one from the control group due to an error or the deletion of the coda /s/, therefore the following number of test items could be analysed: (i) 860 observations from the learner group's first session, (ii) 648 observations from the learner group's second session, and (iii) 503 observations from the control group.

The acoustic analysis was carried out in Praat (version 5.3.12; Boersma & Weenink, 2012). The spectrograms were segmented manually by the author and double-checked by a research assistant. The following measurements were carried out on the basis of the boundaries inserted:

a. the absolute length of the voiced interval in the sibilant (in ms)
b. voicing ratio, the ratio of the voiced part compared to the total length of the sibilant (% voiced part).

Voicing was measured manually, based on the visual inspection of the spectrograms and oscillograms during the whole duration of the frication noise (i.e., presence of the voicing bar on the spectrogram and regular periodic patterns in the waveform, as seen in Strycharczuk, 2012; Schmidt & Willis, 2011), and a low pass filter with a cut-off frequency of 500 Hz was used to securely determine the exact portion of the voicing oscillation.

3.5   Statistical analysis

The statistical analyses were carried out using R (R Development Core Team, 2008), version 3.3.1. The acoustic correlates of voicing were analysed with linear mixed-effects models, using the lme4 package (Bates et al., 2015). For each pho-netic outcome variable, a mixed-effects model was fitted with a random intercept for subjects. The fixed predictor was the 'time' factor (a. native speaker, b. learner before, c. learner after). We employed planned orthogonal contrast coding (for the method, see Field, Miles, & Field, 2012, among others): 1. native speaker vs. learner; 2. learner before instruction vs. learner after instruction. The base or null model always consisted of the fixed predictor plus the random intercept for subject. This base model was then compared to another one that also contained the time predictor as a random slope. If the last two models differed significantly, it would suggest that the subjects' response to the time factor was very different. This, how-ever, did not turn out to be the case for our data, so the random-slope model was not retained. We report the results of the best-fitting model by including the beta coefficient for the given contrast (indicated with b here), the t-statistic, together with its degrees of freedom, and a calculated p-probability of the t-statistic. The significance level was set at $\alpha = 0.05$. The lme4 package does not provide a p-value for the t tests, and therefore we used the *lmerTest* package (Kuznetsova, Brockhoff, & Bojesen-Christensen, 2016) to acquire probabilities. The t test of this package uses 'Satterthwaite approximations' to degrees of freedom. The effect size measure used in the paper is Pearson's correlation coefficient r (see Field et al., 2012, p. 457).

4.   Results

Before comparing learners' productions with the control group in detail, we give a general overview of our findings related to /s/ voicing. Figure 1 (in which *af* stands for absolute final position, *unv* indicates that /s/ is followed by a voiceless consonant, and *voi* indicates that /s/ is followed by a voiced consonant) shows the ratio of voicing, in the /s/, in different contexts by all speakers. Learners' data from sessions 1 and 2 are merged in this figure. Note that 100% corresponds to a fully voiced realisation, while the 0% voiced part describes a completely voiceless fricative with no vocal fold vibration.
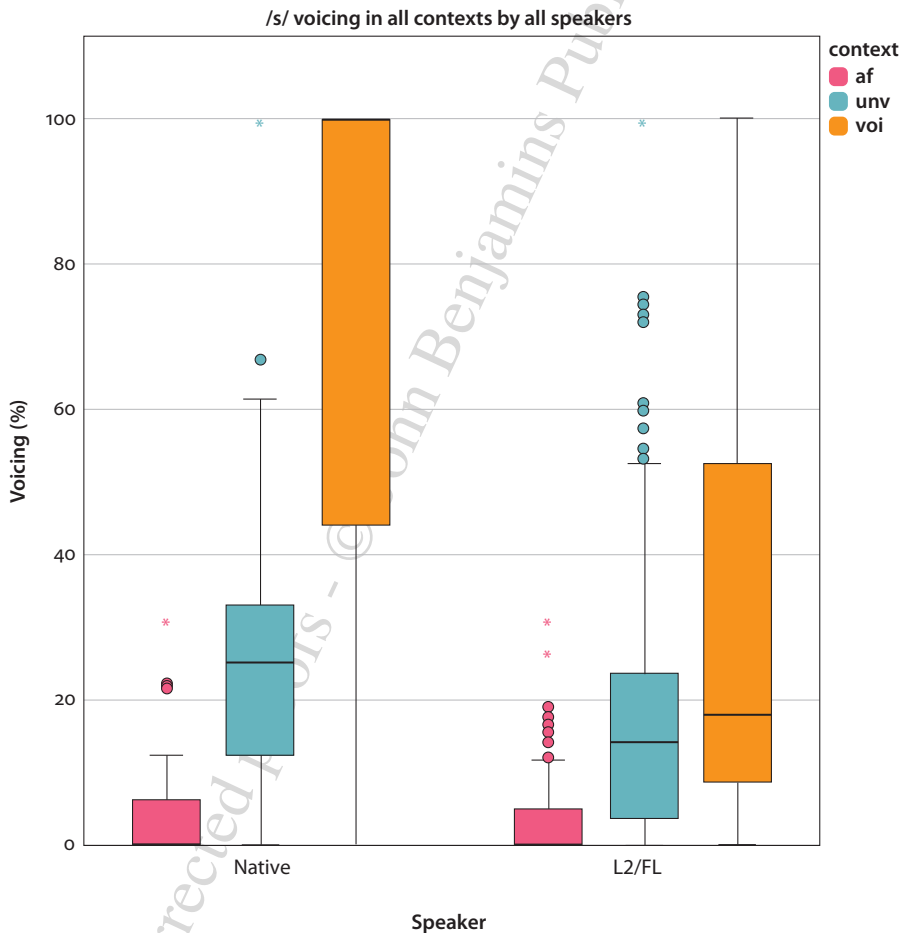
Figure 1. Boxplots with outliers and extreme cases showing the unvoiced part of the /s/ in four contexts by native speakers and L2/FL learners

It is evident at first sight – and is in accordance with our expectations – that it is the voicing environment where learners' and native speakers' data differ the most. In the other two environments, learners and native speakers tend to voice or more precisely unvoice /s/ to a similar extent. In our analysis we will disregard the position of /s/ within the phrase, that is, data from word-internal and word-final phrase-internal contexts will be treated together. As mentioned earlier, Romero (1999) found that the voicing of /s/ is not affected by whether the /s/ + voiced consonant cluster occurs within the word or across word boundaries. In our data, the voicing ratio for /s/ was 33.26%, $SD = 35.24\%$ within the word and 35.58%, $SD = 37.71\%$ across word boundary for all phonetic environments and for all speakers.

The results presented further on in this section combine learners' data from sessions 1 and 2. We merged the two sets of data because in none of the contexts and practically for none of the parameters examined (voicing ratio, voicing duration, duration of the fricative) did we find a statistically significant difference between the two sessions. Table 3 presents the descriptive and inferential statistical results for the pre- and post-instruction sessions. The only parameter that turned out to be statistically significant was utterance-final fricative duration, however, with a weak effect size: $r = 0.21$. We do not have an explanation for why learners produced longer fricatives in the post-instruction session. (The average duration for the native group was 132ms, $SD = 29$ms.)

This means that productions from before the phonetics course started and after the 12-week course are identical with regard to RVA. Language learners have not changed their voicing patterns in the pre-sonorant context as a result of phonological instructions despite the fact that it was in the focus of a number of exercises as explained in 3.3 above. The following two figures (Figure 2 and 3) are examples of complete voicing by a native speaker in the word e**sn**obismo and no voicing by a learner in session 2. In the remainder of the section, we examine each voicing context separately.

Table 3. Statistical results of pre- and post-instruction /s/ voicing of the learner group

| Context	Parameter	Mean		SD		Beta coefficient	t-statistic	p-value
		Pre-instr	Post-instr	Pre-instr	Post-instr			
Utterance-final	Voicing ratio (%)	4.14	2.03	7.41	3.97	1.05	$t(91.71) = 1.58$	0.117
	Voicing duration (ms)	7.24	3.23	12.36	6.22	1.71	$t(91.92) = 1.69$	0.092
	Fricative duration (ms)	154.04	163.02	35.4	31.97	−4.45	$t(92.11) = -2.09$	0.039
Pre-voiceless obstruent	Voicing ratio (%)	17.34	15.33	18.46	13.32	1.006	$t(428.2) = 1.32$	0.187
	Voicing duration (ms)	9.09	9.11	7.10	7.46	0.078	$t(428) = 0.21$	0.831
	Fricative duration (ms)	61.42	63.01	19.97	19.02	−0.581	$t(428.5) = -0.76$	0.448
Pre-voiced obstruent	Voicing ratio (%)	81.88	78.51	31.10	30.47	1.58	$t(425.2) = 1.20$	0.229
	Voicing duration (ms)	41.89	41.03	18.98	18.23	0.23	$t(425.1) = 0.16$	0.874
	Fricative duration (ms)	52.22	53.87	13.21	13.78	1.00	$t(425.1) = -1.49$	0.135
Pre-sonorant	Voicing ratio (%)	15.64	13.08	15.52	14.45	1.29	$t(986.6) = 1.92$	0.055
	Voicing duration (ms)	10.77	9.41	9.39	8.31	0.69	$t(986.7) = 1.51$	0.132
	Fricative duration (ms)	72.79	73.89	19.91	20.89	0.62	$t(987.2) = -0.99$	0.321
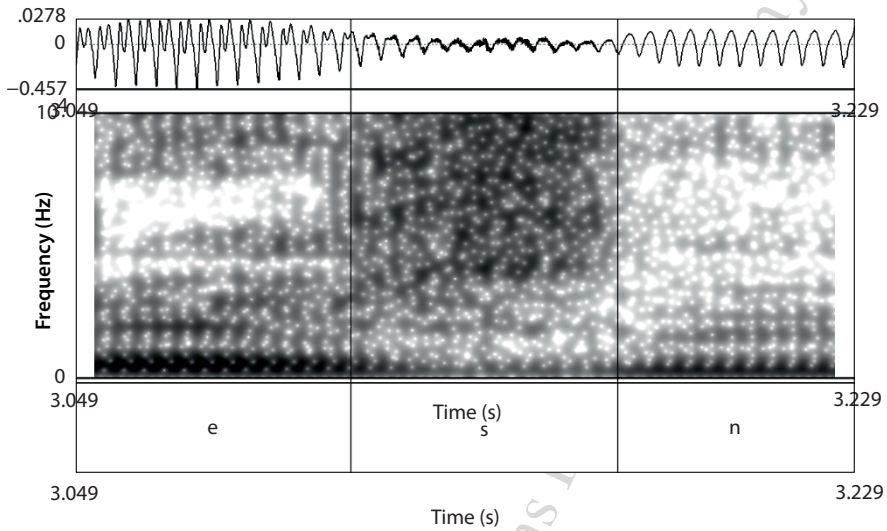
Figure 2. Spectrogram showing the pronunciation of the word *esnobismo* "snobbery" with fully voiced /s/ pronounced by native speaker S4
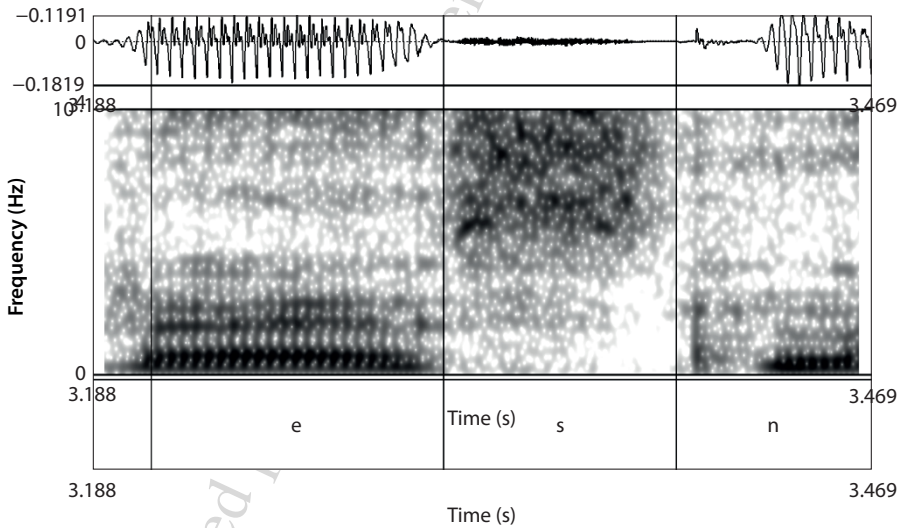

Figure 3. Spectrogram showing the pronunciation of the word *esnobismo* "snobbery" with completely voiceless /s/ pronounced by learner S10

4.1   Expected voiceless realisations

As for the utterance-final position, both the learner group and the control group pro-
nounced the sibilant completely voiceless with mean 3.24%, *SD* = 6.23% and mean
4.19%, *SD* = 8.25% voiced part, respectively (*b* = 0.36, *t*(19.48) = 0.546, *p* = 0.591),
and with a slight voicing tail coming from the previous vowel and dying out in the
fricative which is 5.35 ms on average for native speakers and 5.06 ms for the learner
group. The length of the utterance-final fricative is longer for the learner group
(157.95 ms *SD* = 33.43 ms) than for the control group (132.4 ms *SD* = 28.69 ms on
average), but the difference is not big enough to be statistically significant (*b* = −8.73,
t(17.64) = −1.97, *p* = 0.065). This is probably due to the fact that native speakers
speak at a higher speed (pronounce more syllables per second) than learners.

  In the devoicing context, i.e. when syllable-final /s/ is followed by a voiceless stop
/p t k/ within the same word or across a word-boundary, the target segment was pro-
nounced mostly voiceless (according to the criteria described in Section 3.4) by both
groups (native group mean = 24.22%, *SD* = 17.43%, learner group mean = 16.45%,
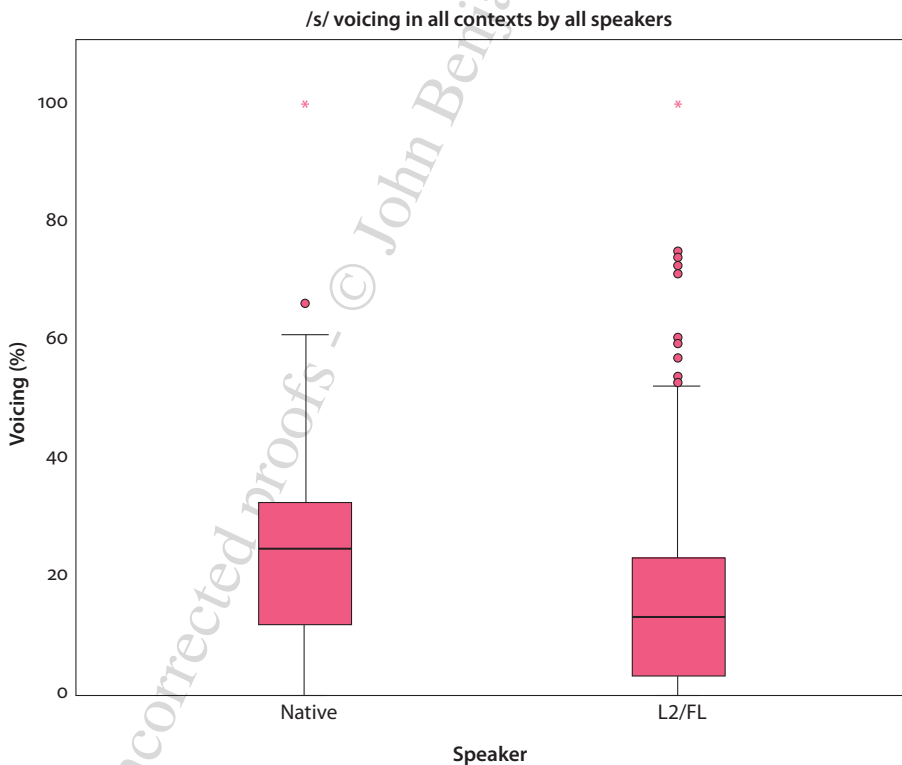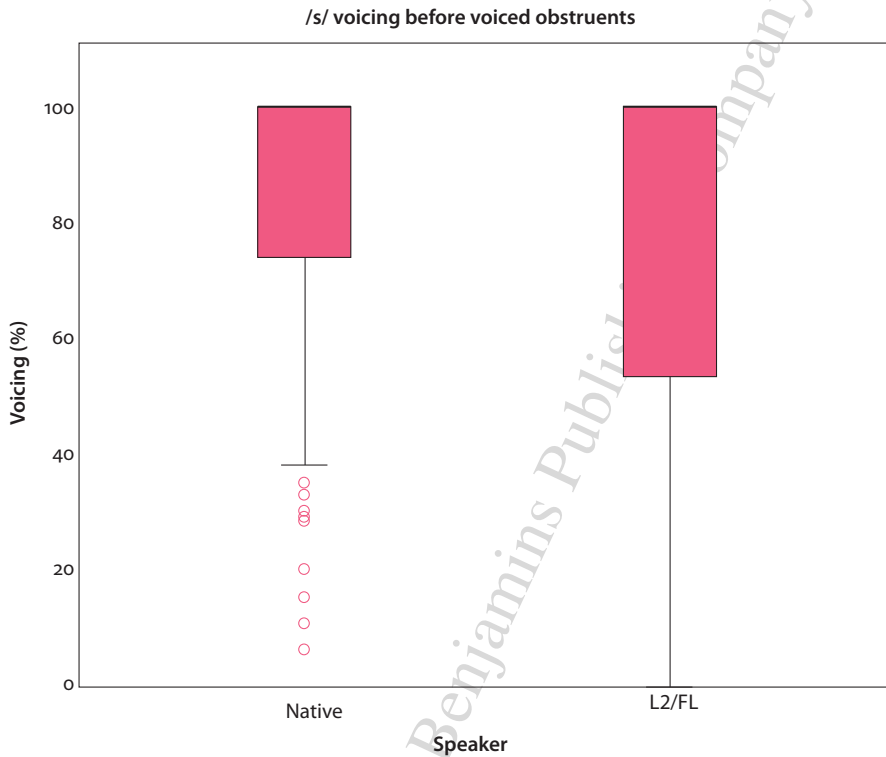*SD* = 16.48%) – see Figure 4. This difference in the voicing ratio is not statistically

Figure 4. Boxplots with outliers and extreme cases showing the voiced ratio in /s/ before
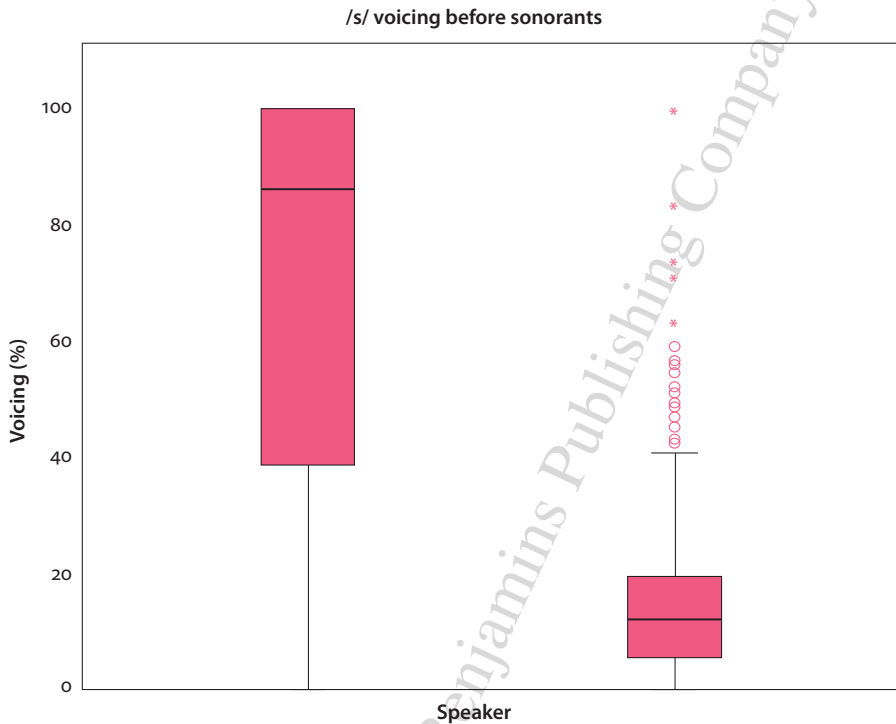voiceless stops in the pronunciation of native speakers and L2/FL learners

Figure 5. Boxplots with outliers and extreme cases showing the voiced part of the /s/ before voiced stops in the pronunciation of native speakers and L2/FL learners

significant ($b = 2.63$, $t(17.7) = 1.64$, $p = 0.118$), and neither is the difference in the fricative duration. The target segment was pronounced shorter by native speakers than by learners, (native group mean = 53.01 ms, $SD = 9.12$ ms and 62.9 ms, $SD = 19.5$ ms for the learner group; $b = -3.06$, $t(17.9) = -1.746$, $p = 0.098$). Native speakers allow a longer voice tail from the preceding vowel than learners, but the difference is not statistically significant (native group mean = 12.88ms, $SD = 9.83$ms, learner group mean = 8.64ms, $SD = 6.85$m; $b = 1.41$, $t(17.4) = 2.02$, $p = 0.595$). We can claim that /s/ in utterance-final position as well as before a voiceless obstruent is pronounced voiceless by both groups of participants.

4.2   Voicing context

As adumbrated in Figure 1, there is a considerable difference in /s/ voicing before a voiced consonant in the productions of the native and the learner group: native group mean of voiced portion = 75.16%, $SD = 31.59$%, learner group mean = 43.69%, $SD = 37.02$%. These figures, however, conceal the different behaviour of voiced

Figure 6. Boxplots with outliers and extreme cases showing the voiced ratio in /s/ before sonorant consonants (/m n l /) in the pronunciation of native speakers and L2/FL learners

obstruents and sonorant consonants as triggers of /s/ voicing, and – as far as we are aware – this is a topic that has not been dealt with in any study on L2/FL Spanish yet. Examining the two data sets separately, we can observe that when /s/ is followed by an intended voiced stop – independently of the fact that /b, d, g/ in our data are mostly realised as approximants with varying degrees of constriction, rather than as true stops – there is no statistically significant difference in the voicing ratio of the fricative between the learner and the native groups (Figure 5). In both groups there are a few fairly or completely voiceless realisations, but in general /s/ is fully voiced in this context (native group mean of voiced portion = 84.93%, *SD* = 25.63%, learner group mean = 80.43%, *SD* = 30.82%; *b* = 1.613, *t*(17.7) = 0.542, *p* = 0.594).

As expected, the most substantial difference lies in the pre-sonorant position (Figure 6). Apart from a handful of voiceless realisations, native speakers produce /s/ in this context voiced. However, it is the other way round for learners: apart from very few voiced realisations, they produce /s/ in this context completely voiceless (native group mean of voiced portion = 70.51%, *SD* = 32.98%, learner group mean = 14.55%, *SD* = 15.11%; *b* = 18.69, *t*(16.9) = 9.327, *p* < 0.001, effect size: *r* = 0.92). We will discuss pre-sonorant voicing in Section 5.2 below in more detail.

5.   Discussion

5-1   Coarticulatory /s/ voicing by native Spanish speakers

A number of studies have shown that the voiceless stops /p t k/ in Spanish are often phonetically voiced in intervocalic position (e.g. Lewis, 2001; Trujillo, 1980, among others); these studies mostly focus on the varieties belonging to the meridional dialectal block (also called *terrabajense*, e.g., Andalusian, Caribbean, etc.). Torreira and Ernestus (2012) demonstrate that intervocalic voicing affects not only voiceless stops but also fricatives in Spanish. Examining over two thousand vowel-fricative-vowel sequences randomly extracted from the Nijmegen Corpus of Casual Spanish, the authors observe that intervocalic voicing affected 34% of /s/ realisations to some extent. This finding is in line with the general lenition tendency of Spanish voiceless obstruents in intervocalic position. In the present study, we do not examine intervocalic /s/ voicing, but it is noteworthy that syllable-final /s/ followed by a voiceless stop tended to be more voiced in native speakers' speech than in learners' productions. The difference is not statistically significant (see 4.1), but it is evident that native speakers allowed a longer voice-tail from the preceding vowel (despite the fact that they produced a shorter fricative on average), which means that they did not counteract coarticulatory voicing (Figure 7) as much as the learner group did.
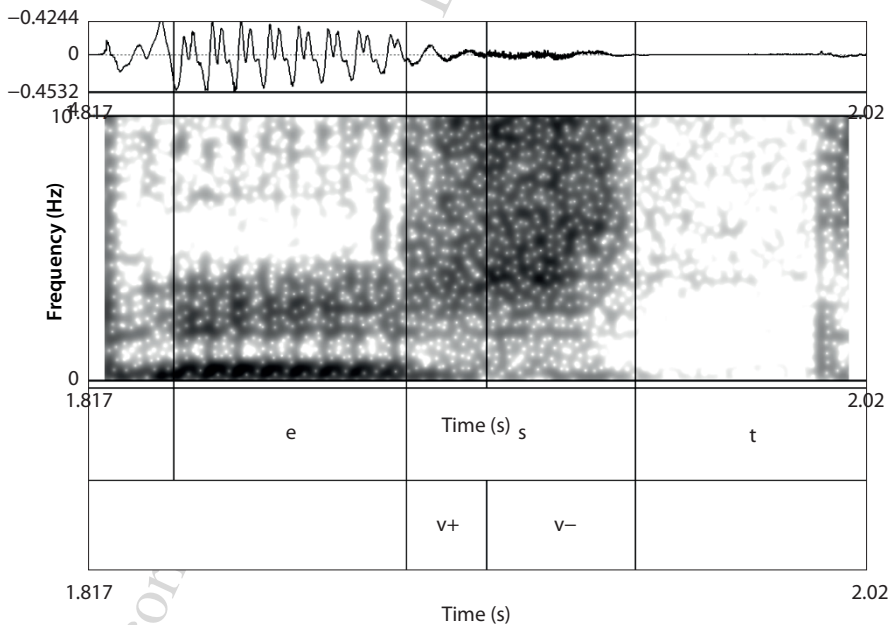

Figure 7. Spectrogram showing the pronunciation of *peste* "plague" with a partially voiced /s/ in the pronunciation of native speaker S1

We explain this difference with the phonemic nature of /s/ vs. /z/ and /s/+stop vs. /z/+stop in Hungarian as opposed to the allophonic status of [s] vs. [z] in Spanish. Recent studies show that voiced fricatives and voiced fricative + stop clusters in Hungarian are not voiced throughout the whole fricative or fricative + stop interval (see Gráczi, 2010; Bárkányi & G. Kiss, forthcoming). Intervocalic /z/ is voiced in 70% of the fricative interval while utterance-final /z/ is voiced only in about 20% according to the literature. This is understandable in light of the aerodynamic requirements of fricative voicing. In order to initiate voicing, the vocal folds must be slightly adducted and there must be sufficient air flowing through the glottis. Aperiodic turbulent noise requires a large volume velocity as well as a narrow constriction in the supraglottal vocal tract. As a result, the vocal folds are to be widely abducted and supraglottal air pressure must exceed subglottal pressure. Voicing, on the other hand, requires the folds to be closely adducted, subglottal air pressure to be greater than supraglottal pressure, and the supraglottal vocal tract to be relatively open. The contradictory targets of voiced fricatives thus indicate that aperiodic turbulent noise and passive voicing cannot be maintained simultaneously: an abducted glottis and decrease in the transglottal pressure differential both remove the basic conditions for vocal fold oscillation. As Ohala (1983) writes, "for the sake of continued voicing the oral pressure should be low, but for the sake of frication the oral pressure should be high" (p. 201). Meeting both of these requirements may be difficult.

Since perceptual cut-off values with regard to voicing in assimilatory contexts have not been specified for Hungarian obstruents so far, we cannot say whether a 25% voiced portion is enough for native speakers of Hungarian to keep the voiced and voiceless clusters safely apart. Bárkányi and Mády (2012) examined the perception of utterance-final /s/ vs. /z/ in the minimal pair *mész* [meːs] 'whitewash' and *méz* [meːz] 'honey' using synthesised speech. Subjects heard the test words in isolation and had to respond in a forced-choice test. The inflection point turned out to be at 30% voicing (*SD* = 8%), that is, if less than 70% of the fricative interval is voiceless the segment is more likely to be perceived as voiced, or to put it the other way, if a final fricative is over 30% voiced, it is more likely to be categorised as voiced, which is not far from the 24.22% voicing observed in native Spanish speakers' speech before voiceless obstruents in our study. As there is no functional load on the [s] – [z] difference in Spanish, L1 Hungarian speakers categorise variably voiced Spanish /s/ as a voiceless sibilant which in their mother tongue tends to be realised with little or no vocal fold vibration. This might explain why native speakers of Hungarian allow less coarticulatory voicing than speakers of Spanish.

5.2 Learning pre-sonorant voicing

Pre-sonorant voicing, that is, cases where a word-final voiceless obstruent assimilates in voicing to a following sonorant in the next word, has raised recurrent interest among phonologists mostly due to the fact that the trigger of voicing assimilation is a segment which is not contrastively specified for voicing as sonorant consonants very rarely contrast voiced-voiceless pairs in a language. Phonetically considered, sonorants may be suitable triggers of RVA as they are phonetically (modally) voiced and rather resistant to devoicing. Yet, typologically, pre-sonorant voicing is rather less frequent than pre-obstruent voicing. As we have seen above, Hungarian is a typical RVA language where sonorant consonants do not participate in this assimilatory process. There are some interesting restrictions that seem to apply to pre-sonorant voicing, which do not apply to pre-obstruent voicing assimilation; as Bárkányi and G. Kiss (2015) observe, pre-sonorant voicing typically occurs in languages which display final devoicing (this, however, does not mean that in all languages with word-final devoicing there is pre-sonorant voicing as well). Pre-sonorant voicing is also generally restricted to the word-final (or syllable-final) position. As we pointed out earlier, this is exactly the point where the VA systems of Hungarian and Spanish differ: in Hungarian, there is no word-final devoicing and no pre-sonorant voicing.

We found no evidence for pre-sonorant voicing of syllable- and word-final /s/ in the Spanish interlanguage of the group of advanced learners in any of the sessions either before or after they received explicit instruction (see Figure 6 above). This means that the learners did not approximate the native voicing patterns. There are several factors that might have obstructed their learning process. First, the two laryngeal and VA systems are so similar (i.e., both are true voice languages, both display RVA among obstruents) that the difference in pre-sonorant voicing might have remained unnoticed. Second, the phonemic status of voiced and voiceless sibilants is different in the two languages. As /s/ and /z/ are contrastive in pre-sonorant position in Hungarian, rendering a number of minimal or quasi-minimal pairs (6) learners tend to keep the two categories safely apart. As Schmidt (2014, p. 101) notes for L1 English learners of Spanish, "L2 learners may be less willing to produce the voiced form as for them it could potentially create a difference in meaning". L1 Hungarian learners are in the same situation; the voiced realisation of the sibilant, according to their L1 voicing system, could create a semantic difference.

(6) *késznek* [sn] 'ready.DAT' vs. *kéznek* [zn] 'hand.DAT'
csempésznek [sn] 'smuggle.PL.3.PRES' vs. *csempéznek* [zn] 'put tiles.PL.3.PRES'

The learning process is further complicated by orthography. As Schmidt (2014) observes, learners of Spanish are usually explicitly taught at the very beginning of their studies that there is no [z] sound in Spanish and that orthographic 'z' corresponds

to [s] or [Ө]. Moreover, the variable nature of voicing in Spanish shown in 2.2 might not serve as a clear and sufficient input for L2/FL learners who based on their L1 expect categorical voiced or voiceless realisations. Recall that native speakers themselves tend to voice less before sonorant consonants than before voiced obstruents (4.2) which might further enforce the false impression that the two RVA systems are identical. Students who participated in our research had received four language practice courses in the previous year, two of which had been taught by a native speaker of Northern-Central Peninsular Spanish. The *Spanish Phonetics and Phonology* class they took during the time of the research described in this paper also served as input for /s/ voicing. However, as about half of the Spanish-speaking world aspirates or deletes coda /s/ (Hammond, 2001), some of the learners might have had limited exposure to RVA in Spanish, which further favours the transfer of their L1 laryngeal system.

Out of the nine pre-sonorant voicing contexts in our experiment (*esnobismo 1, esnobismo 2, eslogan, esmaltadas, mismo, isla, las noches, las loterías, las motos*), the first three might have exacerbated direct negative or positive transfer effects compared to the target pronunciation [eznobizmo] and [ezlogan]. These two items in the test are Spanish-Hungarian cognates: *sznobizmus* /ˈsnobizmuʃ/ 'snobbery' as well as *szlogen* /ˈslogɛn/ begin with voiceless sibilant (this is reflected in the orthography as well), while there is a voiced phone /z/ in the middle of *sznobizmus* [ˈsnobizmuʃ] 'snobbery'. We cannot tease apart whether the lack of voicing in the first two contexts – word-initial /s/ + sonorant – is due to a word-specific negative transfer induced by the cognate word or a general phonological negative transfer, that is, if learners failed to acquire pre-sonorant voicing. Had we found that these two words tended to be less voiced than other similar test items, we could argue that there is a lexical inhibitory effect. This, however, did not turn out to be the case; /s/ in these two words was realised as voiceless as in the other pre-sonorant contexts. Only 4 out of 198 /s/ + sonorant loci were realised voiced (i.e., with over 40% of voicing during the fricative). If we look more closely at these 4 cases, we can see that 3 of them were in the context of *esnobismo 2* in the production of Speaker 8 in both sessions and in Speaker 9 in the first session (before the phonetics seminar), this means that the sporadic /s/ voicing in learners' speech was observed in those test items the Hungarian cognate of which contains the voiced sibilant /z/. Thus these are clearly cases of (word-specific) transfer from L1. Although the systematic study of the relation between foreign accent and pre-sonorant voicing awaits further research, it is interesting to note that it was precisely Speaker 8 who had a marked foreign accent in Spanish. The case of Speaker 9 shows as if the variable nature of the /s/ voicing input had taught this speaker that it is safer not to voice /s/ in pre-sonorant position at all rather than to voice or partially voice it.

As has been mentioned earlier, studies on the effect of instruction on L2 pronunciation have yielded contradictory results. Our data shows that instruction, or at least the amount of formal phonetic, phonological, and pronunciation instruction the students in this research study received, does not enhance the acquisition of the subtle and variable sub-phonemic feature of pre-sonorant voicing in Spanish – while it might have had beneficial effects on other aspects of L2 acquisition (e.g., listening comprehension) not addressed in this paper. The learner group of the present experiment has not changed their voicing pattern as a result of a 12-week Spanish phonetics and phonology seminar. This lack of modification does not prevent or hinder comprehensibility as the RVA systems of the two languages are close enough so that students can get away with their L1 VA system. To conclude, we would like to emphasise the need for realistic goal setting in L2 teaching and learning. While it is worthwhile studying the phonological system of the target language in detail as it can enhance learners' comprehensibility, fluency, listening comprehension, etc., "eradicating traces of foreign accent is widely viewed by applied linguists as an unsuitable goal for L2 pronunciation instruction" (Isaacs, 2014, p. 2). It is neither realistic nor desirable for adult L2 learners to achieve a pronunciation that is completely indistinguishable from native speakers' pronunciation, since accent and identity are closely related (Gatbonton & Trofimovich, 2008). Furthermore, as Derwing and Munro (2009) explain, L2/FL learners do not have to sound like a native speaker in order to be able to successfully carry out their academic or professional tasks.

6.   Conclusions

This article investigated the influence of the following consonant on the phonetic voicing of word-final /s/ in Spanish as produced by native speakers and L2/FL learners. The latter group was tested before and after receiving explicit phonetic and phonological instruction. Our data show that a three-month course does not have an enhancing effect on learners' productions with regard to VA. The Spanish interlanguage of native Hungarian speakers provides clear evidence of L1 interference (or negative transfer), namely that learners transfer the VA rules from Hungarian and do not voice /s/ before a sonorant consonant in Spanish. Consistent word-specific (positive) transfer, that is, pre-sonorant voicing, was observed only in one learner's speech. The relationship between the lack of pre-sonorant voicing, word-specific RVA transfer, and foreign accent awaits further research.

Acknowledgements

References

Archibald, J., & Young-Scholten, M. (2003). The second language segment revisited. *Second Language Research* 19, 163–7.

Bárkányi, Z. (2014). Reflexiones sobre la asimilación de sonoridad de la /s/. In Z. Fábián, I. Szijj, I. Szilágyi, & B. Déri (Eds.), *Köszöntő kötet Giampaolo Salvi 60. születésnapjára. Saggi di linguistica neolatina per i 60 anni di Giampaolo Salvi* (pp. 25–36). Budapest: Eötvös Loránd Tudományegyetem.

Bárkányi, Z., & G. Kiss, Z. (2015). Why do sonorants not voice in Hungarian? And why do they voice in Slovak? In K. É. Kiss, B. Surányi & É. Dékány (Eds.), *Approaches to Hungarian* 14 (pp. 65–94). Amsterdam: John Benjamins. https://doi.org/10.1075/atoh.14.03bar

Bárkányi, Z., & G. Kiss, Z. (forthcoming). Neutralisation and contrast preservation: Voicing assimilation in Hungarian three-consonant clusters. *Linguistic Variation*.

Bárkányi, Z., & Mády, K. (2012, January). *The perception of voicing in fricatives.* Paper presented at the 9th Old World Conference in Phonology, Berlin, Germany.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boersma, P., & Weenink, D. (2005). Praat: Doing phonetics by computer. (*Version 5.3.12*). Retrieved from http://www.praat.org/.

Campos-Astorkiza, R. (2012, May). *Voicing assimilation as gestural blending.* Paper presented at the 20th Manchester Phonology Meeting, Manchester, UK.

Campos-Astorkiza, R. (2014). Sibilant voicing assimilation in Peninsular Spanish as gestural blending. In M. -H. Côte & E. Mathieu (Eds.), *Variation within and across Romance Languages* (pp. 17–38). Amsterdam: John Benjamins.

Campos-Astorkiza, R. (2015). Segmental and prosodic conditionings on gradient voicing assimilation in Spanish. In R. Klassen, J. M. Liceras, & E. Valenzuela (Eds.), *Hispanic Linguistics at the Crossroads: Theoretical Linguistics, Language Acquisition and Language Contact. Proceedings of the Hispanic Linguistics Symposium 2013* (pp. 127–144). Amsterdam: John Benjamins. https://doi.org/10.1075/ihll.4.07cam

Cebrian, J. (2000). Transferability and productivity of L1 rules in Catalan–English interlanguage. *Studies in Second Language Acquisition* 22, 1–26. https://doi.org/10.1017/S0272263100001017

Davidson, L., & Wilson, C. (2016). Processing non-native consonant clusters in the classroom: Perception and production of phonetic detail. *Second Language Research*, 32(4), 471–502.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching* 42, 476–490. https://doi.org/10.1017/S026144480800551X

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393–410. https://doi.org/10.1111/0023-8333.00047

Díaz-Campos, M. (2014). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition* 26, 249–273.

Draxler, C., & Jänsch, K. (2004). Speech recorder (*version 2.2.1*). Retrieved from http://www.bas.unimuenchen.de/Bas/software/speechrecorder/.

Elliot, A. R. (1995). Foreign language phonology: Field independence, attitude, and success of formal instruction in Spanish pronunciation. *Modern Language Journal* 79, 530–542. https://doi.org/10.1111/j.1540-4781.1995.tb05456.x

Elliot, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 80(1), 95–108. https://doi.org/10.2307/345983

Face, T. L., & Menke, M. R. (2009). Acquisition of the Spanish voiced spirants by second language learners. In J. Collentine, M. Garcia, B. Lafford, & F. Marcos Marin (Eds.), *Selected proceedings of the 11th Hispanic Linguistic Symposium*. (pp. 39–52). Somerville, MA: Cascadilla Proceedings Project.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: Sage.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89(1), 395–411. https://doi.org/10.1121/1.400473

García, A. (2013). Allophonic variation in h Spanish sibilant fricatives. PhD dissertation, University of Wisconsin-Milwaukee.

Gatbonton, E., & Trofimovich, P. (2008). The ethnic group affiliation and L2 proficiency link: Empirical evidence. *Language Awareness* 17, 229–249. https://doi.org/10.2167/la464.0

Gráczi, T. E. (2010). A spiránsok zöngésségi oppozíciójának néhány jellemzője [Some characteristics of the voicing contrast of fricatives]. *Beszédkutatás 2010* (pp. 42–56). Budapest: Nyelvtudományi Intézet.

González López, V., & Counselman, D. (2013). L2 Acquisition and category formation of Spanish voiceless stops by monolingual English novice learners. In J. Cabrelli Amaro et al. (Eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium* (pp. 118–127). Somerville, MA: Cascadilla Proceedings Project.

Hammond, R. M. (2001). *The sounds of Spanish*. Somerville, MA: Cascadilla Press.

Harris, J. (1994). *English sound structure*. Cambridge, MA & Oxford: Blackwell.

Hualde, J. (2005). *The sounds of Spanish*. Cambridge: Cambridge University Press.

Isaacs, T. (2014). Assessing pronunciation. In A. J. Kunnan (Ed.), *The companion to language assessment* (Vol. 1, pp. 140–155). Hoboken, NJ: Wiley-Blackwell.

Jansen, W. (2004). Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch. PhD dissertation, Rijksuniversiteit Groningen.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Kissling, E. M. (2013). Is explicit phonetics instruction beneficial for FL learners? *Latin American, Latino and Iberian Studies Faculty Publications* 8, 1–67.

Kuznetsova, A., Brockhoff, P. B., & Bojesen-Christensen, R. H. (2016). lmerTest: Tests in linear mixed effects models. *R package (version 2.0–32)*. Retrieved from http://CRAN.R-project.org/package=lmerTest

Lafford, B., & Salaberry, R. (2003). (Eds.), *Studies in Spanish second language acquisition: State of the science*. Washington, DC: Georgetown University Press.

Lewis, A. (2001). Weakening of intervocalic /p, t, k/ in two Spanish dialects: Toward the quantification of lenition processes. PhD dissertation. Urbana-Champaign, IL: University of Illinois.

Lord, G. 2005. (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 88(3), 557–567. https://doi.org/10.2307/20063159

Moulton, W. (1962). Toward a classification of pronunciation errors. *Modern Language Journal* 46, 101–109. https://doi.org/10.1111/j.1540-4781.1962.tb01773.x

Odlin, T. (1989). *Language transfer: Crosslinguistic influence in language learning*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524537

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). New York: Springer-Verlag. https://doi.org/10.1007/978-1-4613-8202-7_9

Quilis, A. (1993). *Tratado de fonética y fonología españolas*. Madrid: Gredos, Biblioteca Románica.

R Development Core Team. (2008). R: A language and environment for statistical computing. *Foundation for Statistical Computing*. Vienna, Austria. Retrieved from http://www.R-project. org.

Recasens, D., & Mira, M. (2012). Voicing assimilation in Catalan two-consonant clusters. *Journal of Phonetics*, 40(5), 639–654. https://doi.org/10.1016/j.wocn.2012.06.001

Romero, J. (1999). The effect of voicing assimilation on gestural coordination. *International Congress of Phonetic Sciences 99, San Francisco*, 1793–1796.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: An instructed SLA approach to L2 phonology. *Language Awareness* 20, 45–59. https://doi.org/10.1080/09658416.2010.540326

Schmidt, L. B. (2014).Contextual variation in L2 Spanish: Voicing assimilation in advanced learner speech. *Studies in Hispanic and Lusophone Linguistics*, 7(1), 79–113. https://doi.org/10.1515/shll-2014-1159

Schmidt, L. B., & Willis, E. W. (2011). Systematic investigation of voicing assimilation of Spanish /s/ in Mexico City. In S. M. Alvord (Ed.), *Selected proceedings of the 5th conference on Laboratory Approaches to Romance Phonology* (pp. 1–20). Somerville, MA: Cascadilla Proceedings Project.

Schwegler, A., Kempff, J., & Ameal-Guerra, A. (2010). *Fonética y fonología españolas*. 4th edition. Hoboken, N. J.: Wiley.

Simon, E. (2010). Phonological transfer of voicing and devoicing rules. Evidence from L1 Dutch and L2 English conversational speech. *Language Sciences* 32, 63–86. https://doi.org/10.1016/j.langsci.2008.10.001

Siptár, P., & Törkenczy, M. (2000). *The Phonology of Hungarian*. Oxford: Oxford University Press.

Skranitzl, R., & Šturm, P. (2017). Voicing assimilation in Czech and Slovak speakers of English: Interactions of segmental context, language and strength of foreign accent. *Language and Speech* 60, 427–453.

Strycharczuk, P. (2012). Phonetics–phonology interactions in pre-sonorant voicing. PhD dissertation, University of Manchester.

Suomi, K. (1980). Voicing in English and Finnish stops. A typological comparison with an IL study of the two languages in contrast. PhD dissertation. Publications of the Department of Finnish and General Linguistics 10, University of Turku: Turku.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36(3), 326–344. https://doi.org/10.1093/applin/amu076

Torreira, F., & Ernestus, M. (2012). Weakening of intervocalic /s/ in the Nijmegen Corpus of Casual Spanish. *Phonetica* 69, 124–148. https://doi.org/10.1159/000343635

Trujillo, R. (1980). Sonorización de sordas en Canarias. *Anuario de Letras Lingüística y Filología* 18, 247–267.

University of Iowa. *Fonética: Los sonidos del español*. Accessed at http://soundsofspeech.uiowa. edu/spanish/spanish.html

Vago, R. (1980). *The sound pattern of Hungarian*. Washington DC: Georgetown University Press.

Wells, J. C. (2000). *Longman pronunciation dictionary*. Harlow: Longman/Pearson Education.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: first language and second language studies. *Journal of Phonetics*, 35(4), 501–522. https://doi.org/10.1016/j.wocn.2007.02.003

Zsigri, Gy. (1994). Magyar mássalhangzószabályok. [Rules for Hungarian consonants]. PhD dissertation, József Attila University, Szeged. (Published by JATEPress, Szeged, 2006.).

Appendix 1

La gente siempre va en **autobús**.
Desde ayer nos está llamando.
Ya no hay **peste** en Europa.
Llegó la chica **esbelta** de pelo largo.
Que nos **espere** fuera de la casa.
Va al gimnasio solo por **esnobismo**.
Es el mejor **eslogan** del año.
El robo de las joyas **esmaltadas** nos asusta.
Él **mismo** nos acompañó hasta allí.
Dibujó el mapa **isla** por isla.
Las botas están debajo de la cama.
Las potencias nucleares evitan la confrontación.
Las totalmente rubias siempre están de moda.
Habla sobre **las dotes** del futuro vendedor.
Por **las noches** abrimos las persianas.
Se gana mucho en **las loterías**.
En Chile **las motos** son artículos de lujo.

Resumen

En este estudio se examina el efecto de la instrucción fonológica explícita sobre la adquisición de la sonorización de /s/ en español por aprendices húngaros de nivel avanzado. El húngaro y el español tienen sistemas de asimilación de sonoridad muy parecidos pero no idénticos. La diferencia más destacable reside en que las consonantes resonantes causan asimilación de sonoridad en el castellano pero no en el húngaro. Nuestros datos proceden de experimentos acústicos con la participación de dos grupos: 7 hablantes nativos de español del norte peninsular y 12 estudiantes universitarios húngaros, aprendices avanzados de ELE. El segundo grupo fue evaluado dos veces: antes y después de un curso de *Fonética y fonología españolas* de tres meses. Nuestros datos muestran que esta cantidad de instrucción no es suficiente para que los alumnos de ELE superen su sistema de sonoridad nativo. Esto se puede deber en parte a la naturaleza variable alofónica del proceso y a la similitud entre los dos sistemas de sonoridad.

Palabras clave: asimilación de sonoridad, sonorización ante resonantes, español, adquisición de lengua extranjera, instrucción fonológica

Author's address

Zsuzsanna Bárkányi
85 Blandford Rd.
BR34NG
UK

susana.barkanyi@open.ac.uk

Biographical notes

Dr Zsuzsanna Bárkányi is a Lecturer in Spanish at The Open University. Previously, she worked as a Lecturer in Spanish Linguistics at Eötvös Loránd University, Budapest. Zsuzsanna holds an MA with teaching credentials in Spanish and English Language and Literature, an MA in Theoretical Linguistics and a PhD in Spanish Linguistics. Her fields of expertise are the interface of phonetics and phonology and L2 pronunciation.

Author queries

Please provide a complete reference for the citation '(Díaz-Campos (2004))' in this article.

Please provide a citation for the reference id "CIT0665 (Boersma, P., & Weenink, D. (2005)), CIT0673 (Díaz-Campos, M. (2014)), CIT0674 (Draxler, C., & Jänsch, K. (2004)" since citation is missing in the article.