Speech Communication 146 (2023) 93–108

Contents lists available at ScienceDirect

Speech Communication

journal homepage: www.elsevier.com/locate/specom

The acquisition of L2 voiced stops by English learners of Spanish and Spanish learners of English

Celia Gorba *, Juli Cebrian

Departament de Filologia Anglesa i Germanıstica, Facultat de Filosofia i Lletres, Universitat Autonoma de Barcelona, Campus UAB, Bellaterra 08193, Spain

ARTICLE INFO

Keywords:
L2 acquisition
Voiced stops
VOT
Prevoicing
L2 experience
Length of residence

ABSTRACT

Previous studies investigating the acquisition of L2 stops have found a positive effect of L2 experience, but few have focused on voiced stops, particularly on prevoicing. This study investigates the acquisition of /b/ and /g/ by two populations, English learners of Spanish and Spanish learners of English. Three groups varying in amount of L2 experience (mainly length of residence, but also L2 use and L2 instruction) were investigated for each L1 population. Participants completed a carrier sentence reading task in their L1 and L2. Results showed that amount of L2 experience had a positive effect on L2 stop production, as the least experienced groups were outperformed by the experienced ones. No clear effect of L2 experience was observed on the L1, as learners did not differ from monolingual controls, but some differences between learner groups emerged. Moreover, overall, L2 learners were able to produce L1 and L2 stops differently, which indicates that their L1 and L2 categories were not merged. Still, the L1-English L2-Spanish speakers produced L2 stops more accurately than the L1-Spanish L2-English groups, suggesting that learning to rely on an existing L1 cue may be easier than learning to use a cue associated with a different L1 category.

1. Introduction

Adult learners of a second language (L2) often produce L2 sounds differently from native speakers, as their first language (L1) may have influenced their L2 speech (Best and Tyler, 2007; Flege, 1995; Kuhl, 1991). Gaining experience in the L2 – which may involve length of residence in an L2 setting (e.g., Flege 1987, 1995, Flege et al. 1997; Gorba 2019; Gorba and Cebrian 2021; Jun and Cowie 1994; Levy and Law 2010), L2 use (e.g., Kartushina and Martin 2019; Piske et al. 2001) or L2 instruction (e.g., Dmitrieva et al. 2020; Nagle 2019) – may help L2 learners improve their pronunciation in the target language. Moreover, the L1 of an L2 learner may also be affected in the process of learning an L2. This may happen as a result of phonetic drift from the L1 towards the L2, when L1 sounds or features become more L2-like, (e.g., Chang 2012, 2013; Dmitrieva et al. 2020; Flege 1987; Herd et al. 2015; Kartushina et al. 2016), or away from the L2, when L1 structures are modified to differentiate them from L2 structures (Flege and Eefting, 1987). Thus, crosslinguistic influence – i.e., the way the first (L1) and second language (L2) of an L2 learner interact – may occur in either direction (e.g., Flege 1987; Major 1992; Sancier and Fowler 1997). The aim of the present study is twofold: (1) to investigate the effect of L2 experience on the production of L2 and L1 voiced stops, and (2) to contrast two groups of learners, whose L1 systems differ in the use of phonetic cues to implement the voicing contrast for initial stops (English learners of Spanish and Spanish learners of English), as explained below.

Testing these two populations is of special theoretical interest given the different distribution of the cues used to signal stop voicing in the two languages. Specifically, on the one hand, Spanish voiced stops are prevoiced, and are contrasted with short-lag VOT (voiceless stops) to signal voicing. On the other hand, in English prevoicing and short-lag VOT do not contrast and are both used to produce voiced stops, short-lag VOT being more common, particularly in initial post-pausal position (e.g., Lisker and Abramson 1964; see Section 1.2.1. for a detailed review of VOT use in the languages under study). Thus, the underlying question is whether the acquisition of L2 voiced stops by these two populations may differ due to the distributional differences of VOT between English and Spanish. Several scenarios can be considered. For instance, one possibility is that learners would benefit from L1 transfer. In the case of the English learners of Spanish, positive transfer would involve using prevoicing in the L2, as voice-lead VOT is a possible, albeit less common, cue to stop voicing in initial position, used in free variation with short-lag VOT. In the case of the Spanish speakers, producing L2

* Corresponding author.
E-mail address: celia.gorba@uab.cat (C. Gorba).

https://doi.org/10.1016/j.specom.2022.12.003
Received 9 December 2021; Received in revised form 17 November 2022; Accepted 23 December 2022
Available online 25 December 2022

stops with prevoicing, the cue used in the L1, could be regarded as accurate, as prevoicing is a possible and unequivocal cue to voiced stop production in English, even if a more target-like speech would be achieved with a reduction in the frequency of use of this cue. A second scenario involves negative transfer and would take place if English learners simply used short-lag VOT predominantly in Spanish voiced stop productions, and Spanish learners used the same amount and frequency of prevoicing as in their L1. The former case would be more problematic than the latter, as short-lag VOT stops signal voiceless stops in Spanish. Furthermore, acquiring the L2 categories for voiced stops may also influence the L1 – i.e., it may result in a phonetic drift towards L2 values. A third possibility is thus that the learners' L1 is affected by the L2 if the learning of the L2 category results in a drift of the L1 values to more L2-like values, for instance, by reducing the amount of prevoicing in L1 Spanish or increasing it in L1 English – i.e., category assimilation in Flege's (2003) terms. The opposite scenario, category dissimilation, is also a possibility. In the case of the English learners of Spanish, this would involve making a greater distinction between the two languages by reducing the frequency of use of prevoicing in the L1, while in the case of the Spanish learners of English, by increasing the amount of prevoicing in the L1. Examining the L1 and L2 production of speakers of both languages and considering the differences between the two phonetic systems thus allows us to assess these different possible learning scenarios.

Another important contribution of the present study is that it investigates the acquisition of voiced stops, which has received little attention in previous studies compared to the acquisition of voiceless stops. Moreover, as discussed above, the inclusion of both Spanish learners of English and English learners of Spanish allows for the evaluation and comparison of the acquisition of two different main cues to voicing: short-lag VOT for Spanish learners of English and predominant use of prevoicing for English learners of Spanish. Finally, the current study also presents a methodological innovation, given that a categorical measure, namely the presence (yes/no) of prevoicing, is used to study voiced stops, whereas most previous research use VOT duration (e.g., Baese-Berk 2019; Casillas 2019; Dmitrieva et al. 2020; Flege and Eefting 1987; Nagle 2019; Zampini 1998, amongst others). The following section will review aspects that are relevant to the present paper, including the different processes involved in crosslinguistic influences and the acquisition of voiced stops.

1.1. Effects of L2 experience on L1 and L2

The manner in which the first and the second language interact has been addressed by the most influential L2 speech models. For example, the Speech Learning Model (SLM, Flege, 1995, 2002, 2007), and its revised version (SLM-r, Flege and Bohn, 2021), posit that crosslinguistic influence is motivated by the fact that the L1 and the L2 coexist in a common phonetic space, which allows bidirectional influence – i.e., from the L1 on the L2 as well as from the L2 on the L1. At initial stages of L2 acquisition, an L2 phone may be assimilated to a close L1 phone – i.e., perceived in terms of a similar L1 phone – hindering the creation of a new L2 category. Given sufficient exposure, an L2 learner may be able to eventually create a separate category for the L2 phone. Early versions of the SLM considered L2 experience, quantified mainly as length of residence, as the main factor affecting crosslinguistic influence. The extent and direction of crosslinguistic influence may in fact be modulated by a number of factors, including, for instance, starting age of learning, L1 and L2 use, formal instruction and motivation (Piske et al., 2001). A wide range of studies have investigated the effect of length of residence in an L2 setting (e.g., Bohn and Flege 1990; Flege 1987; Flege et al., 1997; Gorba 2019; Gorba and Cebrian 2021; Jun and Cowie 1994; Levy and Law 2010; Stevens 2001). Overall, living in an L2 setting has been found to improve the production of L2 sounds. For example, Stevens (2001) found that American learners of Spanish studying in Spain with no previous pronunciation instruction improved their production of L2 voiceless stops to a greater extent than learners studying in the US. Generally, the longer the length of residence, the more target-like L2 productions may be (e.g., Flege 1987; Flege et al. 1997; Lev-Ari and Peperkamp 2013; Levy and Law 2010). Moreover, immersive experience tends to have a greater impact on L2 learning than L2 instruction. For instance, Mora (2008) conducted a longitudinal study in which L1-Spanish/Catalan learners of English were tested before and after receiving formal instruction and, also, after a three-month stay abroad. Results suggested that instruction had only a small effect on L2 VOT productions, but production became more target-like after the stay abroad.

On the other hand, a number of studies have also shown that L2 experience may result in less native-like productions in the L1 due to phonetic drift towards L2-like values (e.g., Bergmann et al., 2016; Chang 2012, 2013; Flege 1987; Harada 2003; Kartushina et al. 2016, see Section 1.2.1). In short, it appears that, overall, living in an L2 setting may accelerate the acquisition of L2 phones and that length of residence modulates the amount of crosslinguistic influence. However, if living in an immersion setting does not go hand in hand with other factors such as amount of L2 use, the effects of length of residence on both the L2 and the L1 may be lower. Flege and Bohn (2021) analyzed the data in Flege (1995) as well as data collected in 2003 from the same population – i.e., native Italian speakers living in Canada – and observed that, even though participants were comparable in terms of length of residence, some improved in their production of English /p, t, k/, whereas others showed the opposite pattern. This difference was attributed to differences in language use, as those who improved used the L2 to a greater extent. The results also showed that the L2 may keep changing over time as a function of L2 use. Thus, the revised version of the SLM (SLM-r, Flege and Bohn, 2021) claims that the effect of length of residence alone is not an appropriate predictor of L2 learning, and points to the importance of the amount of L1 and L2 use. In this regard, the SLM-r proposes a variable, referred to as Full Time Equivalent (FTE), that combines length of residence in the L2 setting and proportion of L2 use to quantify L2 input. Furthermore, previous studies show that L2 use, as well as L2 instruction, can have an effect not only on the L2 but also on the L1 (e.g., Chang 2012; Herd et al. 2015; Nagle 2019), even in a non L2 setting (e.g., Kartushina and Martin 2019). For example, Kartushina and Martin (2019) found that after participating in a two-week study abroad program in the Netherlands – without explicit L2 instruction – using mainly the L2 as a lingua franca, Spanish-Basque learners of English improved their pronunciation of English vowels. In addition, there was evidence of influence from their L2 on their native languages. However, these effects diminished after four weeks back in their home country, highlighting the role of language use in crosslinguistic influence. Extensive L2 instruction may also result in more target-like productions in the L2 and phonetic drift of the L1 towards L2 values even in an L1 setting (e.g., Herd et al. 2015; Nagle 2019). For instance, Herd et al. (2015) found a greater use of negative VOT – i.e., Spanish-like values – in the L1 of advanced L1-English learners of Spanish than in the case of learners who had received a smaller amount of L2 instruction. Another possibility is that L2 experience in the form of L2 instruction may result in changes in the L1 categories in the opposite direction from the L2 counterparts – i.e., category dissimilation. For instance, Huffman and Schuhmann (2016) found that some American learners of Spanish produced fewer instances of prevoiced stops in L1 English after 6 weeks of L2 instruction.

Moreover, combining more than one type of L2 experience may enhance L2 acquisition to a greater extent. Lord (2010) compared the effect of an 8-week immersion program on English learners of Spanish who had previously received L2 instruction and on a similar group with no prior L2 instruction. Results showed that, even though both groups improved, participants with prior instruction outperformed the group with no instruction on the production of L2 obstruents both before and after the immersion setting. This finding indicated that, even though immersion has a beneficial effect on L2 production, it can be boosted with prior L2 instruction. In short, length of residence in an L2 setting modulates the amount and direction of crosslinguistic influence (e.g., Flege 1987; Mora 2008), but other factors, such as L2 use (e.g., Kartushina and Martin 2019) and L2 instruction (e.g., Herd et al. 2015; Nagle 2019) also play an important role. Next, research conducted on the acquisition of L2 stops, which are the focus of the present paper, will be reviewed.

1.2. The acquisition of L2 stops

The acquisition of L2 stops may pose a difficulty for L2 learners given potential phonetic differences between the L1 and L2 phones and dissimilarities regarding the use of the voicing cues. In fact, the L2 and L1 categories of L2 learners may differ from those of monolingual speakers. Particularly, the production of VOT, which has been found to be the main cue for voicing in stops in several languages (Lisker and Abramson, 1964), has been investigated in a number of previous studies assessing L1 and L2 production (e.g., Castañeda 1986; Docherty 1990; Lisker and Abramson 1964). This section reviews relevant studies on the acquisition of L2 stops and ends with a description of the VOT system of both Spanish and in English – i.e., the two languages examined in this study.

1.2.1. Previous studies on the acquisition of L2 stop voicing

The acquisition of stops has received considerable attention in the literature, especially when it comes to voiceless stops (e.g., Caramazza et al. 1973; Casillas 2019; Flege 1990; Fowler et al. 2008; Harada 2003; Hazan and Boulakia 1993; Major 1992; Schuhmann and Huffman 2015; Williams 1977; Zampini 1998). The results generally suggest that L2 production of voiceless stops evolves towards more target-like values with increased L2 experience (e.g., Flege 1987; Gorba and Cebrian 2021; Major 1992; Schuchmann and Huffman 2019; Stevens 2001; Yang et al. 2022), and that the L1 category may also experience changes in the direction of the L2, particularly given sufficient L2 input and use (e.g., Flege 1987; Major 1992; Sancier and Fowler 1997; Herd et al. 2015). For instance, Flege (1987) examined the production of L1 and L2 /t/ by French learners of English – as well as L1-English learners of French – differing in setting and length of residence and found evidence of bidirectional influence: French learners of English living in the US produced /t/ in both languages with VOT values that differed from monolingual speakers' values and were intermediate to the values expected for each language. Flege (1987) also found that two groups of American learners of French in an L1 setting produced significantly longer VOT values for French /t/ than French monolinguals, whereas the Americans living in an immersion setting did not, illustrating a positive effect of residence in an L2 setting on L2 production. In a study that examined the production of voiceless stops by the same group of L1-English speakers as in the current study, Gorba and Cebrian (2021) found that English learners of Spanish with experience living in the L2 setting outperformed those with no experience in an L2 setting in their production of Spanish /p/ and /k/, indicating a positive effect of L2 experience on L2 accuracy. The same result was replicated by Spanish learners of English (Gorba, 2020). However, contrary to Flege (1987), no influence of the L2 on the L1 was found regardless of the amount of L2 experience. The different outcome regarding the effect of the L2 on the L1 in the studies reported above could be explained by the fact that the length of residence of the most experienced group in Gorba and Cebrian (2021) and Gorba (2020) was considerably shorter than that of the group tested in Flege (1987) – about 4 years vs. 12, respectively –, as well as other factors, including a limited use of the L2 – even in the L2 setting – and amount of L2 instruction.

As reported above, other factors, such as formal instruction, have also been found to have a positive effect on the acquisition of voiceless stops (e.g., Casillas 2019; Chang 2012, 2013; Nagle 2019; Schuhmann and Huffmann 2015; 2019; Yang et al. 2022; Zampini 1998). For instance, Zampini (1998) investigated the production of initial bilabial stops in three different sessions by L1-English students enrolled in a 15-week Spanish phonetics course. Results showed that, by the last session, the L2 learners produced Spanish /p/ with VOT values that resembled those of Spanish native speakers and that differed from their own productions of English /p/. Similarly, Casillas (2019) investigated the progress of English learners of Spanish in their production of /p/ throughout a seven-week immersion course where the use of their L1 was prohibited and found that learners performed in a target-like manner at the end of the program, as their VOT values did not differ from those of Spanish-English bilinguals. Chang (2012) studied the production of L1 and L2 stops and vowels of L1-English learners of Korean who were both enrolled in a language course and living in an immersion setting. Results showed evidence of L1 phonetic drift towards L2 values as early as two weeks into the L2 course. It should also be noted that the influence of the L2 on the L1 was greater for beginner learners than for learners who already had some knowledge of the L2 (Chang, 2013). This difference was attributed to a novelty effect – i.e., the fact that novel perceptual experiences may result in a greater salience and heightened encoding of L2 stimuli. In short, L2 learners are capable of improving their production of L2 voiceless stops by both learning to use long-lag VOT (e.g., Spanish and French learners of English) and reducing VOT duration towards short-lag values (e.g., English learners of French/Spanish). Moreover, the L1 may also drift towards L2 values given sufficient exposure to the L2 – e.g., a long residence in the L2 setting – or due to a novelty effect.

So far, the studies reported above have focused on voiceless stops. Research investigating their voiced counterparts seems to suggest overall that voiced stops may pose a greater difficulty for L2 speakers than voiceless stops and, in fact, studies investigating both types of categories show a more target-like production regarding the latter (Gorba, 2020). For example, a number of studies have investigated the acquisition of short-lag voiced stops, especially English stops, by speakers of languages that contrast prevoicing (voiced stops) and short-lag VOT (voiceless stops), such as Spanish, Italian or French (e.g., Caramazza et al. 1973; Flege and Eefting 1987; Hazan and Boulakia 1993; MacKay et al. 2001; Nathan 1987; Williams 1977). These studies have generally found that L2 learners did not produce L2 stops with target-like values. Flege and Eefting (1987) found that L1-Spanish L2-English speakers – of different ages and with different degrees of L2 exposure – produced English voiced stops with short-lag VOT significantly less often than L1-English speakers. Even though they made a small numerical difference in the use of prevoicing in Spanish as opposed to English (6% more frequent in Spanish), this difference did not reach significance. Similarly, MacKay et al. (2001) found that L1-Italian L2-English speakers differing in age of learning and L2 use prevoiced English stops significantly more often than English native speakers. Even though all groups differed from a group of English controls, the early learners used prevoicing numerically less frequently than the late learners. Regarding their L1, the Italian speakers were found to produce prevoiced L1 stops less often than Italian monolinguals did, particularly in the case of early bilinguals with a low L1 use – but it should be noted that the learners and the monolinguals completed different production tasks. Moreover, results showed that the percent use of prevoicing in the L1 and in the L2 were positively correlated for all groups – i.e., the more they used prevoicing in the L1 the more they used it in the L2. Nathan (1987) elicited the production of English stops by L1-Spanish speakers at two different times 18 months apart and found different results for /b/ and /d/ as opposed to /g/. Speakers' L1 and L2 production for /b/ and /d/ did not differ significantly at any testing time, as the L2 learners used prevoicing in both languages. However, in the case of /g/, the learners were found to use prevoicing significantly more frequently in Spanish (100% both times) than in English (80% and 50%, the first and second time tested, respectively). This result could be linked to the fact that maintaining the rate of glottal airflow needed to produce voicing involves a greater articulatory effort for a velar occlusion than for more anterior articulations (e.g., Ohala 1983). Still, this difficulty exists for Spanish velars too. Thus, the results for /g/ may show that learners were already making a difference between the L1 and the L2, as an articulatorily marked gesture is maintained in the L1 but a more unmarked production emerges in the L2, where prevoicing is not crucial.[1]

The studies described above explored the acquisition of the short-lag vs. long-lag VOT distinction. The reverse pattern, that is, the acquisition of prevoicing as a contrastive feature has received comparatively less attention in the literature, although it has been addressed by some recent studies (Baese-Berk, 2019; Casillas, 2019; Dmitrieva et al., 2020; Herd et al., 2015; Huffman and Schuhmann, 2016; Hutchinson and Dmitrieva, 2021; Maye and Gerken, 2000; Nagle, 2019; Schuhmann and Huffman, 2015, 2019; Yang et al., 2022; Zampini, 1998). Maye and Gerken (2000) and Baese-Berk (2019) studied the perception of the contrast between short-lag /t/ and prevoiced /d/ by English speakers. Both studies found that after a short perception training – of one session in the case of Maye and Gerken (2000) and of two and three sessions regarding Baese-Berk (2019) – the English speakers started to discriminate the two phones, particularly if exposed to a bimodal distribution of the VOT continuum – i.e., if learners were mainly presented with the stimuli at the two ends of the continuum. Zampini (1998); Casillas (2019); Nagle (2019) and Schuhmann and Huffmann (2015) examined the production of Spanish voiced stops by English learners. Zampini (1998) found that English speaking learners did not produce Spanish /b/ accurately, as they presented mean positive VOT values across all testing sessions, although, as reported above, they produced /p/ with target-like values. Even though the learners appeared to produce English and Spanish /b/ with somewhat numerically different values – being the English VOT values greater – this difference did not reach significance. Conversely, Casillas (2019) found that adult English learners of Spanish did improve their production of Spanish /b/ throughout the seven-week immersion course described above. In fact, their production of /b/ significantly shifted towards more target-like values – i.e., longer prevoicing – after the third week of the program and continued to improve during the following weeks. However, their VOT use differed from that of English-Spanish bilinguals, as they presented significantly less prevoicing. It should also be noted that, even though there was some improvement in their production of /b/ – i.e., learners presented more target-like VOT at the end of the course –, their improvement with /p/ was greater and occurred earlier in the course. Similarly, Nagle (2019) found that beginning L1-English learners of Spanish improved their pronunciation of voiced stops over the course of two semesters, and that the greatest improvements took place during the first half of the study. Schuhmann and Huffmann (2015) observed that after receiving phonetic instruction, English learners of Spanish produced Spanish stops with more prevoicing. However, the difference between the learners' L1 and L2 stops approximated – but did not reach – statistical significance. As for the L1, although the English learners of Spanish presented numerically lower VOT values at the end of the program than before receiving training, this difference was also not significant. Interestingly, both before and after phonetic training, participants who prevoiced the most in Spanish also did so in English. Baese-Berk (2019) also found that, when trained on both perception and production, the participants improved their production of voice-lead VOT, although their perception improved to a lesser extent than when trained on perception only.

In short, L2 learners may be able to improve their production of voiced stops, although generally to a lesser extent than voiceless stops (Nagle, 2019; Dmitrieva et al., 2020; Gorba, 2020; Schuhmann and Huffman, 2019, but cf. Herd et al., 2015; Hutchinson and Dmitrieva, 2021). Moreover, the acquisition of the voiced categories may be affected by whether the main cue to voicing is available in the L1 as an additional albeit non-predominant cue to the same category (e.g., voice-lead VOT for English speakers), or if the L2 cue signals a different category in the L1 (e.g., short-lag VOT for Spanish and Italian speakers). Thus, when it comes to the acquisition of L2 prevoiced stops on the part of English learners, improvement has been observed with sufficient exposure or instruction, (e.g., Casillas 2019; Nagle 2019; Schuhmann and Huffman 2015). However, Spanish and Italian learners of English appear to change the use of these cues in the direction of native speakers to a lesser extent (e.g., Flege and Eefting 1987; Mackay et al. 2001; Nathan 1987). Still, in both cases, learners rarely achieve target-like values and their productions tend to differ from those of native speakers.

1.2.2. Comparison of the English and Spanish VOT systems

In English, the voicing distinction in stops is generally described as a contrast between short-lag and long-lag VOT, particularly in syllable initial position. For example, Lisker and Abramson (1964) measured the VOT productions of American English speakers and reported that /p/ is produced with 58 ms of VOT, /t/ with 70 ms and /k/ with 80 ms (see Table 1 for the values reported in several studies). As for voiced stops, most participants were found to use short-lag VOT, although some used voice-lead VOT (average short-lag and voice-lead VOT for /b/ were 1 ms and -101 ms for /b/, respectively: 5 ms, and -102 for /d/ and 21 ms and -88 ms for /g/). That is, short-lag VOT and voice-lead VOT do not contrast, as both are used in the production of voiced stops.

The Spanish VOT range is shorter than that of English in terms of the possible realizations of voiced and voiceless stops, which range from short-lag VOT for voiceless stops to voice-lead VOT for voiced stops (see Table 1). For example, Castañeda (1986) found that speakers of Castilian Spanish produced /p/ with a mean VOT of 7 ms, /b/ with -70 ms, /t/ with 10 ms, /d/ with -78, /k/ with 29 ms and /g/ with -174 ms. It should also be noted that a greater variability in terms of VOT duration has been found for voiced stops than for voiceless stops, as the duration of prevoicing in voiced stops can vary from relatively very long to relatively short (Lisker and Abramson, 1964; Castañeda, 1986, see Table 1 for ranges).

Considering these differences between the two languages under study, it can be assumed that the acquisition of L2 voiced stops for English learners of Spanish and for Spanish learners of English will involve different processes. In the case of the former, learners will need to learn that the use of prevoicing is distinctive in Spanish, and that short-lag VOT signals a different category. As for the Spanish learners, the opposite process needs to take place: they will have to learn that short-

Table 1
VOT means in ms for initial stops in English and Spanish reported in some previous studies. Standard deviations (single values) and ranges (range of values) are given within parentheses, as reported by each study.

Language/Study	/p/	/b/	/t/	/d/	/k/	/g/
English						
British English (Docherty, 1990)	42 (10)	15 (8)	65 (13)	21 (6)	62 (14)	27 (11)
American English (Lisker and Abramson, 1964)	58 (20 – 120)	1/-101 (0 – 5 / -130 – -20)	70 (30 – 105)	5/-102 (0 – 25 / -155 – -40)	80 (50 – 135)	21/-88 (0 – 35/ -150 – -60)
Spanish						
Castilian Spanish (Castañeda, 1986)	7 (0 – 24)	-70 (-166 – -24)	10 (5)	-78 (26)	26 (11)	-58 (-132 – -16)
Castilian Spanish (Rosner et al., 2000)	13 (4 – 49)	-92 (-173 – -23)	14 (6.2 – 30.5)	-92 (-188 – -28)	29 (10 –63)	-74 (-165 – -14
Puerto Rican Spanish (Lisker, and Abramson, 1964)	4 (0 – 15)	-138 (-235 – -60)	9 (0 - 15)	-110 (-170 – -75)	29 (15 – 55)	-108 (-165 – 45)

[1] We are grateful to an anonymous reviewer for pointing out this fact.

lag and voice-lead do not contrast and that short-lag VOT, which in Spanish signals voiceless stops, is typically characteristic of voiced stops in English, particularly in stressed initial position, in opposition to long-lag VOT. Note, however, that using prevoicing to produce English voiced stops would not result in intelligibility problems, since, as mentioned above, that is a possible albeit less frequent way to produce them.

1.3. *The present study*

The current study presents two experiments which investigate the acquisition of L2 voiced stops by two populations that are mirror-image in terms of L1 and L2. Specifically, the acquisition of /b/ and /g/ on the part of English learners of Spanish (experiment 1) and Spanish learners of English (experiment 2) varying in L2 experience will be examined. The coronal stop /d/ is not investigated, as it is produced in a different place of articulation in the two languages under study – i.e., dental in Spanish and alveolar in English (Lisker and Abramson, 1964). The learners' production of both L1 and L2 stops will be examined in order to assess if learners produce L2 stops with different values from L1 stops, if the learning of the L2 stops has any effect on the L1 categories, and if L2 experience has an effect on these issues. The main research questions are thus the following:

(1) What is the effect of L2 experience on the production of L2 voiced stops?
(2) Does the learning of L2 voiced stops have an effect on the production of L1 stops? If so, is this effect modulated/affected by L2 experience?

Furthermore, given that two different populations are tested, namely L1-English L2-Spanish speakers and L1-Spanish L2-English speakers, the design of the study allows for comparisons between the two. This comparison is of particular theoretical interest since, even though the two languages under study share a distinction between voiced and voiceless stops, each language makes use of a different distribution of VOT duration to signal the contrast, as illustrated above (see Section 1.2.2)

2. Experiment 1. English learners of Spanish

2.1. *Methodology*

2.1.1. *Participants*

A total of 50 participants completed this study, including a control group of functional monolinguals for English and Spanish, respectively, and two groups of English learners of Spanish. The English control group (EnMono) were nine functional English monolinguals who were enrolled in undergraduate or graduate studies at Queen Mary University of London (QMUL) and were mostly from the London metropolitan area (EnMono)[2] and 10 Spanish functional monolinguals constituted the Spanish control group (SpMono). Table 2 summarizes the main characteristics of each group, including sex, L1, L2, months spent in an L2 setting, location, percent weekly L2 use and years of L2 instruction. The information was retrieved using a linguistic background questionnaire based on the Bilingual Language Profile (BLP, Birdsong et al., 2012), which also included additional questions regarding the recency of the stay in an L2 setting. Percent weekly L2 use was drawn from the participants' self-reported amount of L2 use in different contexts (i.e., at work/school, with family and with friends). The inexperienced English group (EnInexpUK) consisted of 11 L1-English students of Spanish who were enrolled in the first or second year of Hispanic Studies or in the highest levels of Spanish language courses at a language center at QMUL. They had never lived in a Spanish-speaking country, although some reported short holiday visits, and were living in the UK at the time of testing. EnInexpUK reported using Spanish 12.3% on a weekly basis and had been learning Spanish for a mean of 3.8 years. The moderately experienced English group (EnExpUK) were fourth-year students of Hispanic Studies at QMUL and had spent between a term and a year ($M = 9.4$ months) in a Spanish country the previous academic year. They were back in the United Kingdom at the time of testing. EnExpUK had studied Spanish for 7.5 (SD: 5.3) years and reported a weekly use of Spanish of 15%. The most experienced English group in terms of time spent in an L2-speaking setting (EnExpSp) had been living in Spain – where they resided at the time of testing – for a mean of 50.7 months – i.e., 4.2 years. In spite of having resided in an immersion setting for the longest period of time, EnExpSp had received the least amount of L2 instruction ($M = 2.4$). Moreover, their percent weekly L2 use was very similar to that of the English learners in an instructional setting (16.3 %). The Spanish control group was made up of 10 students of History and Spanish Studies at Universitat Autònoma de Barcelona (Spain) who were functional monolingual speakers of Spanish. They all had studied some English at school, as English is a mandatory course in the Spanish primary and secondary education curricula, but all reported not using it.

2.1.2. *Task and procedure*

Participants had to complete a sentence reading task in both languages (see Appendix A for the complete list)[3]. Control groups performed the task only in their L1. The reading task consisted of a list of 32 carrier sentences – for each language – that presented bilabial and velar stops in word-initial position, as well as fillers. All words were disyllabic and carried stress on the penultimate syllable and all target stops were followed by vowel /i/. Each sentence was repeated twice and a total of 20 productions of voiced stops – 10 per segment – were elicited for each participant and language. It was decided to use a carrier sentence reading task in order to be able to control the position in which the target phones appeared, since Spanish stops in medial position – except after nasal – are realized as approximants (Colina 2020; Hualde 2005). The carrier sentence in the English task was 'X is the next word' and the Spanish equivalent – i.e., 'X es la siguiente palabra' – was used in the Spanish condition. Order of completion of the English and Spanish reading tasks was counterbalanced. Furthermore, participants completed an online linguistic background questionnaire based on the Bilingual Language Profile (Birdsong et al., 2012) before the testing session.

The recordings made in an English setting took place in a soundproof room at QMUL and a Neumann TLM 103 microphone, a booth mixer Steinberg UR22 MKII and the software Audacity Team, 2016 were used. Most of the participants tested in Spain were recorded at the speech laboratory at Universitat Autònoma de Barcelona (UAB) in an acoustically treated room using an Audio-Technica AT 2050 microphone and an Alesis Multimix 8 mixer. For participants' convenience, a few recordings were made in a different institution in Barcelona, namely Universitat de

[2] There were some dialectal differences between the English speakers, particularly in the group tested in Spain due to the difficulty recruiting volunteers. Four participants were from Southern England, one from Northern England, one from Scotland, two from Ireland, one from Canada, one from the East Coast of the United States and one from the South of the United States. However, these dialectal differences were not expected to affect the L1 production of voiced stops, and, in fact, no significant differences were observed between groups (see results section).

[3] The data presented in this paper are part of a larger study that tested the perception and production of stops by Spanish learners of English and English learners of Spanish. In the production task, which was a carrier sentence reading task, the production of both voiced and voiceless stops was collected (see Appendix A for the complete list of sentences). Participants also had to complete two identification tests (one for the /p/-/b/ contrast and one for the /k/-/g/ contrast) in both languages (see Gorba 2019 and Gorba and Cebrian 2021).

Table 2
Characteristics of the five groups that participated in Experiment 1: number of participants (N), gender, L1, L2, months in L2 setting, location at the time of testing, % weekly L2 use and years of L2 instruction (inst.). Standard deviations are provided in parentheses.

Group	*N*	Sex	L1	L2	Months in L2 setting	Location	% weekly L2 use	Years of L2 inst.
EnMono	9	4 F; 5 M	English	NA	NA	UK	-	-
EnInexpUK	11	6 F; 5 M	English	Spanish	none / minor	UK	12.3% (8)	3.8 (3.3)
EnExpUK	9	4 F; 6 M	English	Spanish	9.4 (4.2)	UK	15% (6.5)	7.5 (5.3)
EnExpSp	11	6 F; 5 M	English	Spanish	50.7 (27.3)	Spain	16.3% (10.5)	2.4 (2.9)
SpMono	10	4 F; 6 M	Spanish	NA	NA	Spain	-	-

Barcelona, in a soundproof booth using a Shure-SM58 microphone and a Marantz PMD-660 digital recorder. All sound files were recorded in WAV format.

2.1.3. Data analysis

A total of 2780 target words were analyzed in Praat (Boersma and Weenik, 2016). Some productions were disregarded because they were produced with a sound other than a stop – i.e., an approximant or a fricative – or because VOT was difficult to measure due to the quality of the recording at the point where the stop was produced – e.g., due to overlapping noise. The VOT of the remaining words, a total of 2756 (English /b/: 689 words; English /g/: 682 words; Spanish /b/: 696 words; Spanish /g/: 689 words), was measured manually. Short-lag VOT productions were measured from the onset of the burst to the beginning of voicing as indicated by the periodicity of the signal, considering both the spectrogram and the oscillogram. Measurements of voice-lead productions were made from the beginning of the prevoicing until the onset of the burst.

Two different measures were obtained, namely a continuous measure (VOT duration) and a categorical measure (the presence or the absence of prevoicing). Most previous studies investigating voiced stops have used VOT duration as the measure of analysis – with the exception of MacKay et al. (2001), which used percentage of prevoiced productions, and, to our knowledge, none of them used a categorical variable (presence of prevoicing: yes/no). Given the great variability often found in the duration of prevoicing (Castañeda, 1986; Lisker and Abramson, 1964), a decision was made to use the categorical measure in the statistical analysis[4]. Therefore, although descriptive results report both VOT means and percentages of prevoiced productions, the latter is the one used in the statistical analysis throughout the paper. Thus, for each stop produced by each speaker in each language a negative VOT value was considered an instance of prevoicing, whereas a VOT of 0 ms or greater was coded as an instance of short-lag VOT (i.e., as lack of prevoicing). In order to examine differences between groups in their L1 and L2 productions, the results were submitted to a series of generalized linear mixed models (GLMM) with presence of prevoicing (yes/no) as the dependent variable, and group (three learner groups and the monolingual control group) and segment (/b/ and /g/) as independent variables. Separate analyses were carried out for each language – English and Spanish. In addition, between-language differences for each L1 population (i.e., differences in the production of voiced stops in the L1 and in the L2) were also explored by means of a series of GLMMs with language (English and Spanish), group (the three learner groups for each L1 population) and segment (/b/ and /g/), and their interactions, as fixed factors. Different combinations of random intercepts and slopes for participant and word were also examined in the models. The selection of the best model was based on the lowest Akaike Information Criterion (AIC) and on model convergence. In fact, in every case, the results for the different models were very consistent and the levels of significance for each factor and pairwise comparison were very similar across models. SPSS software (IBM Corp, 2017) was used. Finally, correlation tests were conducted to determine whether the use of prevoicing in the two languages was related.

[4] In fact, a parallel statistical analysis involving a series of linear mixed models with VOT duration as the dependent variable was also conducted. The two sets of analyses – on VOT duration or on presence/absence of prevoicing – yielded very similar results, the only difference being that more cases of significant group differences – particularly between groups sharing the same L1 – were revealed in the prevoicing analysis.

2.2. Results

2.2.1. Production of English voiced stops

First, the production of the L1-English groups in English was examined. Tables 3 and 4 present the mean VOT and the proportion and percentage of prevoiced productions for /b/ and /g/, respectively, for each group. Given the bimodal distribution of VOT durations described for English voiced stops (Lisker and Abramson, 1964), the tables also show the mean VOT for short-lag productions and for voice-lead productions separately. As expected, most voiced stops were produced with short-lag VOT, but instances of prevoicing were found in all groups. Regarding /b/, EnInexpUK had the highest average VOT (8 ms), followed by EnMono and EnExpSp (close to 0) and EnExpUK a mean negative VOT (-9 ms). The two groups with the longest length of residence also showed the largest ranges of VOT, and the mean durations of voice-lead VOT were also the greatest (EnExpUK: -90; EnExpSp: -76). A similar pattern of results was found with respect to the percentage of prevoicing, with EnInexpUK having the lowest amount (1%) and EnExpUK the highest (18%).

Similar results were found for /g/ (see Table 4). EnInexpUK presented the greatest VOT value (20 ms), followed by EnMono (14 ms), whereas EnExpSp and EnExpUK had the lowest VOT means (6 and 1 ms, respectively). As for the percentage of prevoiced /g/, EnExpUK presented the greatest amount of prevoiced tokens (26%), whereas EnInexpUK prevoiced the least (5%).

Table 3
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of English /b/ by English learners of Spanish and English monolinguals. Standard deviations are given in parentheses.

Group	/b/ VOT mean	/b/ VOT range	/b/ mean short-lag VOT	/b/ mean voice-lead VOT	Proportion prevoiced
EnMono	1 (25)	-94 – 20	7 (4)	-69 (26)	9/90 (10%)
EnInexpUK	8 (6)	-47 – 24	8 (4)	-47	1/110 (1%)
EnExpUK	-9 (40)	-146 – 23	9 (4)	-90 (31)	16/90 (18%)
EnExpSp	0 (35)	-160 – 19	7 (10)	-76 (36)	17/110 (15%)

The results of a GLMM with presence (yes/no) of prevoicing as the dependent variable, group, segment and the interaction as fixed factors and a random intercept for participant revealed a significant effect of group [$F(3, 799) = 3.521$; $p = .015$], but no effect of segment [$F(1, 799) = 2.437$; $p = .119$], and no interaction [$F(3, 799) = 1.065$; $p = .363$].

Table 4
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of English /g/ by English learners of Spanish and English monolinguals. Standard deviations are given in parentheses.

Group	/g/ VOT mean	/g/ VOT range	/g/ mean short-lag VOT	/g/ mean voice-lead VOT	Proportion prevoiced
EnMono	14 (38)	-255 – 43	23 (7)	-93 (77)	7/90 (8%)
EnInexpUK	20 (25)	-109 – 45	23 (8)	-84 (28)	5/110 (5%)
EnExpUK	1 (48)	-129 – 58	27 (9)	-77 (26)	23/90 (26%)
EnExpSp	6 (41)	-137 – 49	24 (8)	-75 (28)	20/110 (18%)

Fisher's LSD[5] pairwise comparisons exploring group differences showed a significant difference between the groups with experience and the least experienced group, that is, between EnExpSp (mean = 17%, SD = 37) and EnInexpUK (mean = 3%, SD = 16), $p = .039$, and between EnExpUK (mean = 22%, SD = 41) and EnInexpUK, $p = 031$. EnInexpUK was the group that prevoiced the least often, but none of the groups differed significantly from EnMono.

2.2.2. Production of Spanish voiced stops

As for the L2 production of Spanish stops, as expected, the English groups prevoiced less often than the Spanish monolinguals and, consequently, presented greater VOT means (see Tables 5 and 6). Regarding /b/, EnExpUK, with a mean VOT of -48 ms, was the group that presented the closest mean VOT duration to SpMono's (-78 ms), and prevoiced /b/ the most often (64%). EnExpSp presented a mean VOT for /b/ of -33 ms and prevoiced /b/ 56% of time, while EnInexpUK presented the closest mean VOT to 0 ms, with -13 ms, and prevoiced it the least (26%). A similar pattern was followed in the case of /g/. EnExpUK performed the most similarly to SpMono – with a mean VOT of -55 ms and 79% of prevoiced tokens – followed by EnExpSp – with a mean VOT of -29 ms and 62% of prevoiced /g/ –, whereas EnInexpUK presented the closest mean VOT to 0 ms (-6 ms) and prevoiced the least (32% of time).

In this case, the results of a GLMM on amount of prevoicing with group, segment and their interaction as fixed factors and a random intercept for participant also revealed a significant effect of group [$F(3, 799) = 10.916$, $p = .000$], no effect of segment [$F(1, 799) = 3.008$, $p = .083$] and no interaction [$F(3, 799) = 0.449$, $p = .718$]. Pairwise comparisons with a Bonferroni correction indicated that EnInexpUK (mean = 29%, SD = 45) differed significantly from SpMono (mean = 98%, SD = 14), $p < .001$, and EnExpSp (mean = 58%, SD = 49) also differed significantly from SpMono ($p = .011$). By contrast, the difference between SpMono and EnExpUK (mean = 71%, SD = 46) approximated but did not reach significance ($p = .077$). Moreover, EnExpUK were found to use prevoicing in Spanish significantly more often than EnInexpUK ($p = .001$).

Table 5
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of Spanish /b/ by English learners of Spanish and Spanish monolinguals. Standard deviations are given in parentheses.

Group	/b/ VOT mean	/b/ VOT range	/b/ mean short-lag VOT	/b/ mean voice-lead VOT	Proportion prevoiced
SpMono	-78 (29)	-137 – 7	7 (0)	-80 (26)	95/97 (98%)
EnExpSp	-33 (47)	-118 – 119	7 (19)	-70 (26)	61/110 (56%)
EnExpUK	-48 (52)	-203 – 49	6 (10)	-82 (34)	58/90 (64%)
EnInexpUK	-13 (45)	-207 – 22	10 (4)	-78 (47)	28/109 (26%)

Table 6
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of Spanish /g/ by English learners of Spanish and Spanish monolinguals. Standard deviations are given in parentheses.

Group	/g/ VOT mean	/g/ VOT range	/g/ mean short-lag VOT	/g/ mean voice-lead VOT	Proportion prevoiced
SpMono	-77 (29)	-138 – 22	20 (3)	-79 (26)	97/99 (98%)
EnExpSp	-29 (48)	-123 – 51	25 (9)	-61 (30)	68/110 (62%)
EnExpUK	-55 (51)	-175 – 43	22 (10)	-75 (35)	67/85 (79%)
EnInexpUK	-6 (52)	-198 – 48	25 (9)	-72 (43)	34/107 (32%)

2.2.3. L1 vs. L2 production of voiced stops

The production of voiced stops in each of the languages under study was also compared within each group. Figs. 1 and 2 show the distribution of VOT durations for each learner group in each language. All English groups presented either longer positive VOT values or shorter voice- lead VOT values in English than in Spanish for the production of /b/ and /g/, although the difference was greater in some cases than in others. EnInexpUK used a smaller range of values and produced generally shorter voice-lead VOT durations, whereas EnExpUK displays the greater differences between L1 and L2 and shows greater variability in L1.

As described above (see Section 2.1.3), a series of generalized linear mixed models were conducted with presence or absence of prevoicing as the dependent variable. Fig. 3 below presents the mean percentages of use of prevoicing per group and language. The results of a GLMM that included group, language, segment, group x language and segment x language as fixed factors and a random intercept for participant revealed

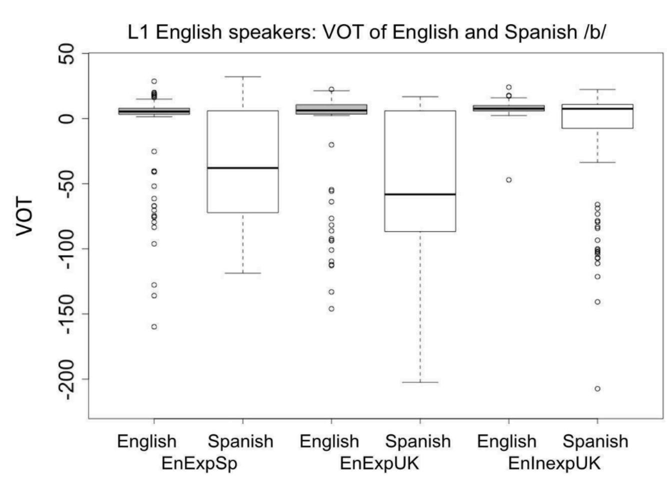

Fig. 1. Boxplot of VOT production (in ms) for English and Spanish /b/ by the L1-English speakers.

[5] Fisher's Least Significant Difference (LSD) test was chosen because, despite the significant effect of group, Bonferroni-adjusted pairwise comparisons did not reach significance and it was relevant to explore what group differences drove the effect of group. Although the LSD does not correct for multiple comparisons, it is deemed appropriate for comparisons involving less than five levels Derek and DeMars (2019). In the remaining of the analyses, Bonferroni-adjusted pairwise comparisons were applied.

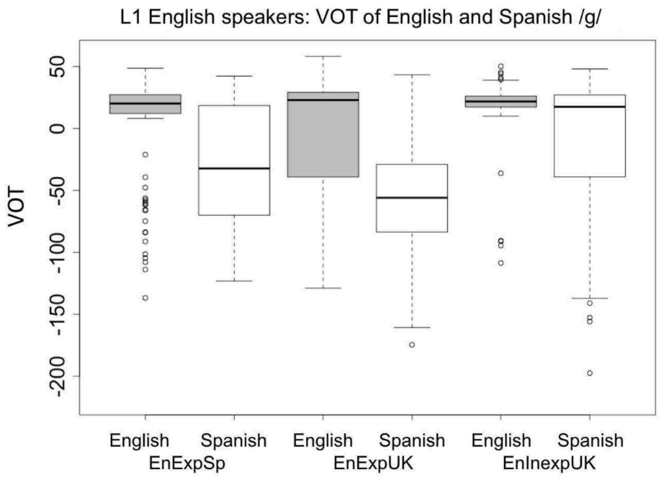

Fig. 2. Boxplot of VOT production (in ms) for English and Spanish /g/ by the L1-English speakers.

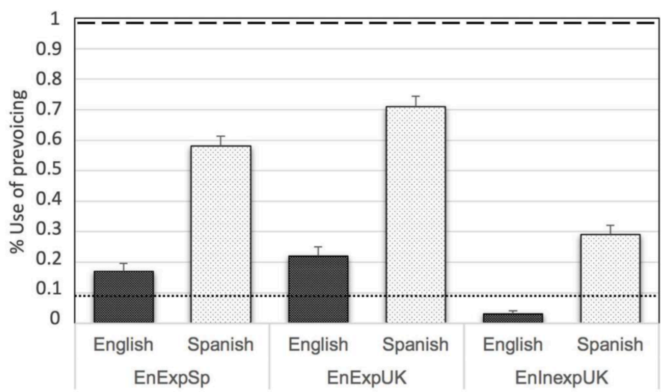

Fig. 3. Mean use of prevoicing in the production of English and Spanish stops by the L1-English speakers. The dashed line and the dotted line show the mean values for SpMono and EnMono, respectively. Percentage use of prevoicing and standard error of the mean are given.

an effect of group [F (2, 1223) = 7.4; p = .001], language [F(1, 1223) = 173.9; p < .001] and segment [F(1, 1223) =10.032; p = .002], but no significant interaction (language x group: ([F(2, 1223) = 0.483; p = .617]; language x segment: [F(1, 1223) = 0.08; p = .777]). Thus, all groups produced voiced stops with more prevoicing in Spanish (mean = 51, SD = 50) than in English (mean = 13, SD = 34), and prevoicing was somewhat more frequent with /g/ (mean = 35, SD = 48) than with /b/ (mean = 29, SD = 45). Regarding the effect of group, Bonferroni-adjusted pairwise comparisons indicated that both EnExpSp (mean = 38%, SD = 48) and EnExpUK (mean = 46%, SD = 50) used prevoicing more frequently (across languages) than EnInexpUK (mean = = 16%, SD = 36), p = .027 and p = .004, respectively)

Finally, in order to explore whether the learners' production in the L1 and the L2 were related, the relationship between the use of prevoicing in L1 and L2 voiced stops was examined by conducting a one-tailed Pearson correlation for each stop. The use of prevoiced stops in the L1 and in the L2 was significantly, albeit moderately, correlated for both /b/ (r = .359, N = 31, p = .024) and /g/ (r = .479, N = = 31; p = .002).

2.3. Interim discussion

Experiment 1 investigated the production of L2 and L1 /b/ and /g/ by English learners of Spanish. Results showed that, as expected, the learners presented greater VOT values in English than in Spanish. In fact, no group differences were observed in the production of L1-English stops, which was comparable to that of monolingual English speakers, whereas group differences emerged in the production of L2-Spanish stops. Regarding the production of L2 voiced stops, both EnInexpUK – the least experienced group – and EnExpSp – the group with the longest length of residence in the target-language setting – were found to differ from SpMono, whereas EnExpUK did not differ from the Spanish controls. In other words, the groups with the least and the most experience living in an L2 setting produced Spanish voiced stops differently from Spanish controls, whereas the group with intermediate experience in an L2 setting – but a greater amount of L2 instruction – presented the most target-like VOT values. Thus, it appears that length of residence in an L2 setting alone does not account for the differences observed between the English groups. Despite residing in an immersion setting, EnExpSp did not report a notably greater amount of L2 use than the other L1-English groups (12% and 15% for EnInexpUK and EnExpUK, respectively, vs. 16% for EnExpSp). EnExpSp's relatively small L2 use may be related to the fact that they used English at work, as they were English teachers, and that most of them had English-speaking partners. This finding is in line with the fact that Flege and Bohn's (2021) SLM-r underscores the importance of language use and proposes measuring L2 experience in terms of a composite measure combining length of residence and amount of L2 use (Full Time Equivalent or FTE), which is calculated by multiplying the number of years spent in an L2 setting by the percentage of weekly L2 uses. Although the current study was not designed to evaluate specifically the effect of FTE, the data in Table 1 above allow us to compare the groups in terms of this variable. EnExpSp and EnExpUK did not differ greatly in FTE of target language input (0.12 vs 0.69, respectively)[6] despite the former's much longer stay in the L2 setting. Moreover, the fact that EnExpUK outperformed EnExpSp in Spanish voiced stop production may be related to amount of L2 instruction and student status. Recall that EnExpUK were the English group that had received the most L2 instruction (7.5 years, vs. 2.4/3.8, see Table 2 above). Furthermore, EnExpUK were in fact enrolled in Spanish classes at the time of testing and at the institution where the testing took place. This may have triggered more careful productions, as students may have interpreted the experiment task as an academic test. By contrast, in spite of having lived in an L2 setting for the longest period of time, EnExpSp had received scarce – sometimes none – formal instruction in Spanish (mean = 2.4 years) and were not studying Spanish at the time of testing. Thus, the current results add to the findings from previous studies showing a beneficial effect of formal instruction on voiced stop production (e.g., Casillas 2019; Lord 2010; Nagle 2019; Schuhmann and Huffmann 2015). It is possible, thus, that a greater amount of L2 instruction had resulted in more target-like productions of Spanish stops on the part of EnExpUK compared to EnExpSp.

With respect to the production of English voiced stops, none of the groups differed from EnMono statistically. This finding is in line with Gorba and Cebrian (2021), a study involving the same population which did not find an effect of the L2 on the L1 production of voiceless stops. The authors argued that a greater amount of L2 experience and L2 use may be necessary for an L1 phonetic drift towards the L2 to occur, an explanation that may also apply to the present study. A novelty effect similar to that found in Chang (2012, 2013) would not apply to the present study, as none of the groups were absolute beginners in an immersion setting. Still, although EnExpUK and EnExpSp did not differ significantly from the English controls in their use of prevoicing in the L1, the two more experienced groups made a greater use of prevoicing than EnInexpUK (see Fig. 3) and this difference reached significance. This may indicate that learning to use more prevoicing in the L2 may carry with it a greater use of that feature in the L1, providing some

[6] Flege and Bohn's (2021) exemplifications of differences in FTE between experienced vs inexperienced groups are notably greater, e.g., 4.3 vs 0.3 (Flege et al., 1997) or 17.2 vs 9.2 (Flege, 1991), respectively.

indication of a possible effect of the L2 on the L1. As for the inexperienced group, even though they used prevoicing numerically less than the English native speakers, there is no evidence for category dissimilation, as this difference was numerically very small (EnMono: /b/ = 10%, /g/ = 8% prevoiced productions; EnInexpUK: /b/ = 1%, /g/ = 5% prevoiced productions) and did not reach significance (c.f., Dmitrieva et al., 2020). The difference between the L1 and L2 production of voiced stops was also assessed. All groups were found to use prevoicing less often in English than in Spanish, although this difference was greatest in the case of EnExpUK and smallest regarding EnInexpUK. That is, all English learners of Spanish were able to distinguish L1 and L2 voiced stops in production – to a greater or lesser extent. A previous study looking at English learners of Spanish, Zampini (1998), also found that L2 learners made a numerical difference between their L1 and L2 production of voiced stops, in this case regarding VOT values, but this difference was not significant. It is possible that methodological differences between the two studies may explain the different result, as Zampini's study examined a smaller amount of production data and involved participants who had no experience living in the L2 setting.

Finally, it was investigated whether there was a relationship between the use of prevoicing in the L1 and in the L2. Given that the use of prevoicing is subject to individual and free variation in English, it was hypothesized that the use voice-lead VOT in their L2 may be related to its use in the L1. The correlations conducted on the use of prevoicing in the L1 and L2 indicated that those English learners that used prevoicing the most frequently in English also did so in Spanish. As a matter of fact, at a group level, EnExpUK used prevoicing the most in both languages. This outcome is in agreement with some previous studies that showed that the use of negative VOT in the L1 and the L2 were correlated, and, thus, that the use of prevoicing in the L1 was overall transferred to the L2 (MacKay et al., 2001; Schuhmann and Huffmann, 2015). That is, even though the English learners of Spanish were able to distinguish L1 and L2 stops by producing shorter voice-lead VOT and longer short-lag VOT in English than in Spanish, the use of prevoicing might have been – positively – transferred from the L1 to the L2 to a certain extent, and, thus, facilitated the production of target-like – prevoiced – voiced stops.

In sum, a positive effect of L2 experience was found on the production of L2 stops, as those learners with some experience in the L2 setting outperformed those with none. However, other factors, including longer L2 instruction on the part of EnExpUK and insufficient amount of L2 use in the case of EnExpSp, may have played a role, as the former outperformed the latter. No clear evidence of L1 phonetic drift towards the L2 were found regardless of amount of L2 experience, as no group differed from the monolingual speakers in their L1 production. Still, the greater use of prevoicing on the part of the two more experienced groups, both in L1 and L2, points to a possible influence of the L2 on the L1. In addition, all groups seemed to make a difference between L1 and L2 stops by using prevoicing more often in the L2. It remains to be seen if the current results are specific to the population under study (L1-English learners of L2- Spanish) or generalizable to other populations. This issue is explored in Experiment 2, which examines the production of L1 and L2 voiced stops by Spanish learners of English.

3. Experiment 2. Spanish learners of English

3.1. Methodology

3.1.1. Participants

This study involved 49 participants (see Table 7), including 29 L1-Spanish L2-English speakers, who were distributed in groups that were parallel to the English learners of Spanish in Experiment 1 in terms of L2 experience, and the same two control groups as in Experiment 1 (SpMono and EnMono).[7] The inexperienced Spanish learners (SpInexpSp) were 10 undergraduate students in their first or second year of English Studies at UAB who had never lived in an immersion setting. They used English 22.7% of time on a weekly basis and had learnt it for 13.2 years. The moderately experienced Spanish learners of English (SpExpSp) were 10 students in their fourth year of English studies in the same institution who had spent 7.4 months in an immersion setting the previous academic year as part of a study-abroad program. SpExpSp reported to use English 26.3% of the time weekly and had been learning it for 13.9 years. Finally, the highly experienced Spanish group (SpExpUK) was made up of nine L1-Spanish speakers who had been living – and still resided – in the UK for 47.9 months – i.e., four years. SpExpUK were the group that used the L2 the most (54.6% weekly) and had learnt it for 13.4 years. Even though the L1-English and the L1-Spanish groups were comparable in terms of length of residence and setting, there were considerable differences regarding amount of L2 instruction and L2 use that stem from the greater presence of English in the Spanish education system than of Spanish in the UK and in the media and entertainment industry.

3.1.2. Tasks, procedure and data analysis

Participants had to complete the same tasks that were presented in Experiment 1 and the data was analyzed following the same procedure.

3.2. Results

3.2.1. Production of Spanish voiced stops

The L1 speakers of Spanish, as expected, produced most voiced stops with voice-lead VOT, although a few instances of short-lag productions were found in all groups (see Tables 8 and 9). All groups showed a similar mean duration of voice-lead VOT, ranging from a mean of -64 and -68 ms for SpInexpSp and SpExpUK, respectively, to a mean of -77 and -80 ms for SpMono and SpExpSp, respectively. Regarding the percentage of prevoiced productions, results were also relatively similar across groups. SpExpSp produced almost exclusively prevoiced stops (99%), whereas SpExpUK had the lowest percent of prevoicing (88%). Regarding /g/, SpExpSp presented the longest mean voice-lead VOT (-80 ms), followed by SpMono and SpExpUK (-77 ms), whereas SpInexpSp presented a somewhat shorter mean (-71 ms). As for the percentage of prevoiced tokens, SpInexpSp prevoiced the most (100%), whereas SpExpUK prevoiced the least (93%). These results were submitted to a series of generalized linear mixed models with presence of prevoicing as the dependent variable, following the same procedure as described for Experiment 1 (see section 2.1.3). A GLMM with group, segment and their interaction as fixed factors and a random intercept for participant revealed no significant main effects and no interaction (group: [$F(3, 766) = 1.737$, $p = .158$]; segment: $F(1, 766) = 0.001$, $p = .98$]; group x segment: $F(3, 766) = \ = 0.165$, $p = .92$]). Thus, on the whole, experience with English did not seem to affect the use of prevoicing in the L1, although SpExpUK was the group to display the numerically lowest percent use of prevoicing, particularly in the case of /b/.

3.2.2. Production of English voiced stops

The same analysis as in the case of Spanish was conducted for the production of English voiced stops. All groups tended to produce voiced stops with prevoicing more often than with short-lag VOT (see Tables 10 and 11). Regarding /b/, SpExpUK presented the shortest mean amount

[7] All L1-Spanish participants were speakers of Castilian Spanish. Most of them were also speakers of Catalan, but the knowledge of this language was not expected to have any effect on the production of VOT, as Spanish and Catalan present identical VOT values (Julià i Muné, 1981).

Table 7
Characteristics of the five groups that participated in Experiment 2, including number of participants (N), gender, L1, L2, months in L2 setting, location at the time of testing, % weekly L2 use and years of L2 instruction (inst.). Standard deviations are provided in parentheses.

Group	*N*	Sex	L1	L2	Months in L2 setting	Location	% weekly L2 use	Years of L2 inst.
EnMono	9	4 F; 5 M	English	NA	NA	UK	-	-
SpExpUK	9	5 F; 4 M	Spanish	English	47.9 (23.3)	UK	54.6% (24.6)	13.4 (2.3)
SpExpSp	10	7 F; 3 M	Spanish	English	7.4 (6)	Spain	26.3% (10)	13.9 (1.9)
SpInexpSp	10	6 F; 4 M	Spanish	English	none / minor	Spain	22.7% (9.7)	13.2 (2.2)
SpMono	10	4 F; 6 M	Spanish	NA	NA	Spain	-	-

Table 8
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of Spanish /b/ by Spanish learners of English and Spanish monolinguals. Standard deviations are given in parentheses.

Group	/b/ VOT mean	/b/ VOT range	/b/ mean short-lag VOT	/b/ mean voice-lead VOT	Proportion prevoiced
SpMono	-78 (29)	-137 – 7	7 (0)	-80 (26)	95/97 (98%)
SpInexpSp	-64 (36)	-137 – 69	10 (2)	-72 (25)	94/100 (94%)
SpExpSp	-80 (26)	-135 – 3	3	-81 (25)	99/100 (99%)
SpExpUK	-68 (37)	-126 – 19	8 (4)	-79 (19)	79/90 (88%)

Table 9
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of Spanish /g/ by Spanish learners of English and Spanish monolinguals. Standard deviations are given in parentheses.

Group	/g/ VOT mean	/g/ VOT range	/g/ mean short-lag VOT	/g/ mean voice-lead VOT	Proportion prevoiced
SpMono	-77 (29)	-138 – 22	20 (3)	-79 (26)	97/99 (98%)
SpInexpSp	-71 (23)	-128 – -28	-	-71 (23)	100/100 (100%)
SpExpSp	-80 (51)	-175 – 11	11	-81 (32)	97/98 (99%)
SpExpUK	-77 (40)	-239 – 25	18 (5)	-84 (31)	84/90 (93%)

Table 10
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of English /b/ by Spanish learners of English and English monolinguals. Standard deviations are given in parentheses.

Group	/b/ VOT mean	/b/ VOT range	/b/ mean short-lag VOT	/b/ mean voice-lead VOT	Proportion prevoiced
EnMono	1 (25)	-94 – 20	7 (4)	-69 (26)	9/90 (10%)
SpExpUK	-43 (49)	-182 – 29	8 (4)	-79 (31)	53/90 (59%)
SpExpSp	-71 (45)	-217 – 16	7 (3)	-83 (35)	85/99 (86%)
SpInexpSp	-66 (44)	-182 – 27	13 (6)	-79 (30)	88/100 (88%)

of voice-lead VOT (-43 ms) – i.e., the closest to EnMono (1 ms) – and used prevoicing the least (59%). SpExpSp and SpInexpSp used prevoicing more often (86% and 88%, respectively) and produced the longest mean negative VOT (-71 ms and -66 ms, respectively). With respect to velar stops, SpExpUK presented the shortest mean voice-lead VOT duration (-24 ms) and used prevoicing less frequently than the other Spanish groups (55% of time, see Table 11). SpExpSp prevoiced

Table 11
Mean VOT and range of all productions, mean VOT of short-lag and voice-lead productions, and proportion and percentage (in parentheses) of prevoicing in the production of English /g/ by Spanish learners of English and English monolinguals. Standard deviations are given in parentheses.

Group	/g/ VOT mean	/g/ VOT range	/g/ mean short-lag VOT	/g/ mean voice-lead VOT	Proportion prevoiced
EnMono	14 (38)	-255 – 43	23 (7)	-93 (77)	7/90 (8%)
SpExpUK	-24 (53)	-132 – 65	27 (11)	-68 (24)	50/90 (55%)
SpExpSp	-49 (41)	-127 – 59	15 (6)	-62 (30)	80/100 (80%)
SpInexpSp	-50(46)	-166 – 46	26 (11)	-66 (31)	83/100 (83%)

/g/ 80% of time and presented a mean voice-lead VOT of -49 ms, and SpInexpSp prevoiced /g/ 83% of time and produced it with the longest mean negative VOT (-50 ms).

In this case, the results of a GLMM on presence (yes/no) of prevoicing as the dependent variable, with group, segment and their interaction as fixed factors and a random intercept for participant revealed a significant effect of group [$F(3, 742) = 20.95$, $p < .001$], but no effect of segment [$F(1, 742) = 0.868$, $p = .352$] and no interaction [$F(3, 742) = 0.008$, $p = .967$]. All three Spanish groups (SpExpUK: mean = 58%, SD = 50, SpExpSp: mean = 85%, SD = 35, SpInexpSp: mean = 86%, SD = 34) used prevoicing significantly more frequently than EnMono (mean = 9%, SD = 29), as confirmed by Bonferroni adjusted pairwise comparisons ($p < .001$ in all three cases). In addition, SpExpUK were also found to use prevoicing less frequently than the other two Spanish groups ($p = .021$ in both cases).

3.3. L1 vs. L2 production of voiced stops

The results obtained for the production of Spanish and English stops by the Spanish learners of English were compared. All Spanish groups presented mean negative VOT values for both English and Spanish voiced stops, although the voice-lead VOT duration tended to be numerically shorter in English than in Spanish for all groups, particularly in the case of the most experienced learners (see Figs. 4 and 5).

As done in experiment 1, the results involving the presence (yes/no) of prevoicing (presence or absence of prevoicing for each production) were submitted to a series of GLMMs (see Fig. 6 for the mean percent use of prevoicing per group and language). In this case, the best fitting model was a GLMM with group, language, segment, group x language and segment x language as fixed effects and random intercepts for participant and word. The statistical model showed a significant effect of group [$F(2, 1140) = 5.928$, $p = .003$] and language [$F(1, 1140) = 37.958$, $p = .000$], but no effect of segment [$F(1, 1140) = 2.142$, $p = .144$]. Thus, prevoicing was more frequently used in Spanish (mean = 96%, SD = 21) than in English (mean = 77%, SD = 42). Bonferroni pairwise comparisons indicated that SpExpUK used prevoicing (across languages) significantly less often (mean = = 74%, SD = 44) than SpExpSp (mean = 92%, SD = 27) and SpInexpSp (mean = 91%, SD = 28), $p = .029$, in both cases. The language x group interaction did not reach significance [$F(2, 1140) = 1.527$ $p = .218$], but the segment x

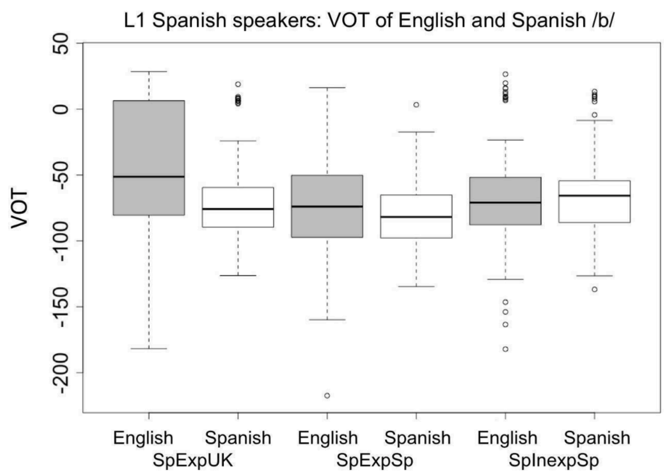

Fig. 4. Boxplot of VOT production (in ms) for English and Spanish /b/ by the L1-Spanish speakers.

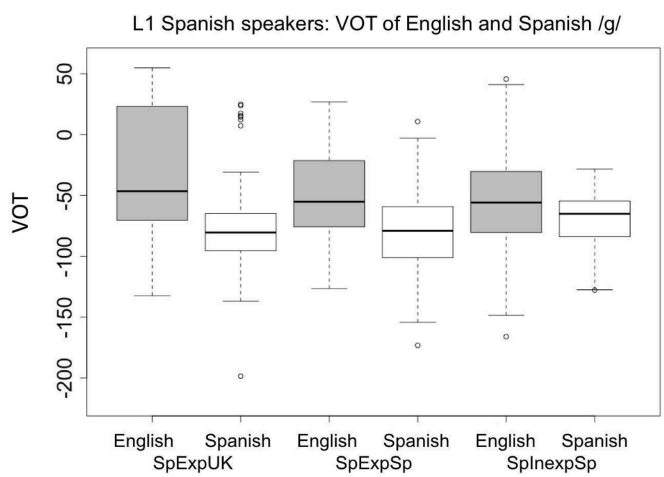

Fig. 5. Boxplot of VOT production (in ms) for English and Spanish /g/ by the L1-Spanish speakers.

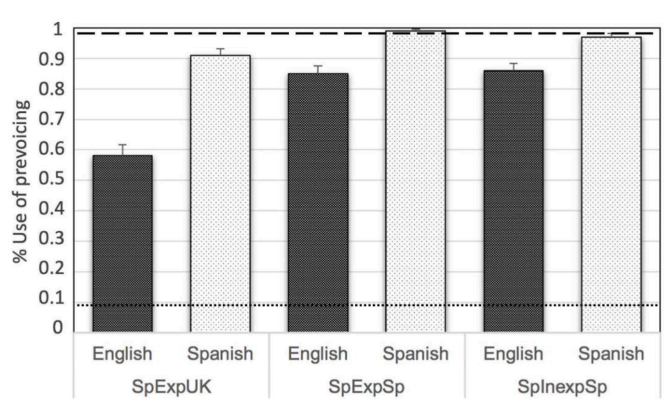

Fig. 6. Mean use of prevoicing in the production of for English and Spanish stops by the L1-Spanish speakers. The dashed line and the dotted line show the mean values for SpMono and EnMono, respectively. Percentage use of prevoicing and standard error of the mean are given.

language did [$F(1, 1140) = 4.073$, $p = .044$]. Bonferroni pairwise comparisons indicated that /b/ and /g/ differed in Spanish (/b/: 93%, SD = 25; /g/: 98%, SD = 15; $p = .038$), but not in English (/b/: mean = 78%, SD =41; /g/ = 76%, SD = 43), possibly because SpExpUK used less prevoicing with Spanish /b/ than with /g/ (88% vs. 93%, respectively). Finally, the Pearson correlation test revealed a significant – though moderate – correlation between L1 and L2 /b/ and /g/ (/b/: r =.316; N = 29, p = .047; /g/: r = .466; N = 29; p = .005).

3.4. Interim discussion

Experiment 2 focused on the production of L2 and L1 /b/ and /g/ by Spanish learners of English varying in amount of L2 experience. Results showed that L2 experience was found to modulate the accuracy of production of L2 voiced stops, since SpExpUK presented the most target-like productions and SpInexpSp were the Spanish group that performed the least accurately. That is, the greater the amount of L2 experience learners had, the more target-like their productions of English voiced stops were. In this case, differences in experience are to be explained by differences in length of residence, as, unlike with the English L1 groups, the Spanish L1 groups did not differ in years of L2 instruction (averaging from 13.2 to 13.9 years, see Table 5 above). In addition, and unlike the L1-English groups, the L1-Spanish groups did differ in amount of L2 use, as the group living in the L2 setting used the L2 the most (54.6%, vs. 26.3% and 22.7% for SpExpSp and SpInexpSp, respectively). Consequently, in this case, differences in FTE years of L2 input (Flege and Bohn, 2021) between the two most experienced groups were also greater (2.18 vs. 0.16) than with the L1-English groups. It should be noted, though, that none of the L1-Spanish groups produced English voiced stops in a target-like manner, as all used prevoicing significantly more than EnMono. Still, SpExpUK, who presented the closest values to EnMono's and prevoiced less often, differed significantly from SpInexpSp and SpExpSp. This finding is in line with previous studies that have found that L2 experience results in a more accurate production in the L2 (e.g., Flege 1987; Flege et al. 1997; Gorba and Cebrian 2021; Lev-Ari and Peperkamp 2013; Levy and Law 2010) and indicates that the positive effect of L2 length of residence on voiceless stop production can be extended to voiced stops.

As for the results obtained for L1 production, the Spanish learners' L1 production of voiced stops was still found to fall within native-like values, as none of the groups differed from Spanish monolinguals. Thus, it seems that living in an L2 setting did not result in phonetic drift towards L2 values. A similar explanation as the one given for the English learners above could apply to the Spanish learners: specifically, it is possible that their length of residence was not long enough for the L1 to change towards L2 values and novelty effects would not apply given their amount of L2 experience. However, SpExpinUK used prevoicing numerically less frequently than the other L1 Spanish-groups, particularly in the case of /b/. This finding may indicate that the production of /b/ in Spanish may be beginning to shift towards more L2-like values in the case of the most experienced group.

Regarding the differences between L1 and L2 production in the Spanish learners, all groups were found to produce voiced stops differently in Spanish and in English – i.e., used prevoicing significantly more often in Spanish than in English. Nevertheless, this difference was numerically small, and the production of short-lag VOT voiced stops (14-15% for SpInexpSp and SpExpSp, 42% for SpExpUK) was still far from native English values (91%). This shows a notable difficulty in using an L2 cue that is associated with a voiceless category in their L1. Still, an effect of experience was observed, since the most experienced group used prevoicing the least. In short, L2 experience, related to longer residence in an L2 setting and greater amount of L2 use, affected the L2 production of Spanish learners of English. A greater amount of experience with English seemed to result in a greater use of short-lag VOT, and a reduction in the use of prevoicing, the predominant L1 cue. No clear effects of the L2 on the L1 were found, possibly due to a relatively small amount of L2 experience. Overall, all groups used prevoicing significantly more frequently in the L1 than in the L2, even though the difference between the two languages was very small, particularly for the less experienced groups.

4. General discussion

This study has examined the production of L1 and L2 voiced stops (specifically /b/ and /g/) by English learners of Spanish (experiment 1) and Spanish learners of English (experiment 2), whose L1s differ in the use and weight of the main voicing cues, namely short-lag VOT and prevoicing. The production of the two populations in the two languages has been compared in terms of the presence or absence of prevoicing (i.e., voice-lead VOT), which is the main cue to stop voicing (in initial post-pausal position) in Spanish, and a possible, but not main, cue in English, where voiced stops are often produced with short-lag VOT. The effect of L2 experience on L2, and L1, production has been assessed by comparing three groups for each population differing in experience (in terms of length of residence in the L2 setting, but also amount of L2 use and L2 instruction). Thus, L1 and L2 production were examined by comparing the production of the L2 learners sharing the same L1 and the corresponding control monolingual group in each experiment. The performance of the learner groups in the two languages was also compared so as to determine whether learners produced L1 and L2 voiced stops differently and to evaluate the role that L2 experience played on the ability to make this difference. So far, the outcomes of each experiment have been discussed separately. This section will bring together the results obtained for both experiments and will attempt to answer the research questions presented in Section 1.3.

The first research question addressed the effect of L2 experience on the production of L2 voiced stops. Results indicated that, overall, none of the groups performed in a target-like manner, as, with the sole exception of EnExpUK, all groups from both populations differed significantly from the target language controls in the use of prevoicing in the production of L2 stops. In the case of the Spanish learners, all groups differed from English monolinguals, but the most experienced group (SpExpUK) performed in a more target like-manner, as they prevoiced the least often. In fact, SpExpUK used prevoicing in the production of English voiced stops significantly less frequently than the other L1-Spanish groups (58% vs. 85%, 86%; EnMono: 9%). Differences in years of residence in the L2 setting, together with the amount of L2 use, can explain SpExpUK's better performance (recall that the L1-Spanish groups hardly differed in years of L2 instruction). In the case of the English learners of Spanish, as mentioned above, EnExpUK – the group with some previous experience living in the L2 setting – presented the most target-like use of prevoicing for Spanish voiced stops (71%, vs. EnExpSp: 58% and EnInexpUK: 29%; SpMono: 98%). As discussed in the interim discussion (see Section 2.3), other factors may have resulted in EnExpUK's more target-like performance than EnExpSp, the group with the longest length of residence. The fact that EnExpUK had received the greatest amount of L2 instruction together with their student status at the time of testing may have resulted in a more accurate production. In fact, previous studies report positive effects of L2 instruction on L2 productions in L1 settings (e.g., Dmitrieva et al. 2020; Nagle 2019). Moreover, the group with longer experience in an L2 setting did not report using Spanish notably more often than the learners in the UK (16% vs 15% and 13% for EnExpSp and ExpInexpUK, respectively), contrary to expectation, which resulted in a relatively small difference between EnExpSp and EnExpUK in terms of years of FTE. Hence, in the absence of actual differences in L2 use or FTE, EnExpUK's better performance in the L2 can be attributed to the role of L2 instruction. Another explanation, as will be discussed below, is related to the possible transfer of the use of prevoicing from the L1 to the L2. In brief, L2 experience has been found to improve L2 performance, but the factors that contribute to experience vary for the two populations. Years of L2 instruction seems to play the main role in the case of the English learners of Spanish, given the absence of real differences in L2 use or FTE, while length of residence and L2 use are the crucial factors in the case of the Spanish learners of English. These differences in experience between the two populations are further discussed below.

The second research question investigated the effect of the L2 on the production of L1 voiced stops and the role of L2 experience. Similar results were obtained for the two populations, as a clear effect of the L2 on the L1 was not observed – none of the groups differed significantly from their L1 controls, that is, no clear influence of the L2 was found on the L1. A greater variation was observed in the case of the English learners of Spanish, both within group and between groups, possibly due to the fact that prevoicing is used natively in English, in free variation with short-lag VOT. The L1-English group that prevoiced the most often in English was EnExpUK (22%), the one with some previous experience in the L2 setting and the greatest amount of L2 instruction, as mentioned above, but this greater use of prevoicing was not significantly different from the English monolingual pattern (9%). Still, EnExpUK and EnExpSp (17%) made a significantly greater use of prevoicing in English than EnInexpUK (3%), showing that an increase in prevoicing in the L2 may carry with it an increment of prevoicing in L1, and illustrating a possible effect of L2 learning on L1 patterns. In the case of the Spanish learners, the only indication that L2 experience may have influenced L1 production is the fact that the most experienced Spanish group (SpExpUK) yielded a lower percent use of prevoicing in Spanish compared to the other L1-Spanish groups (/b/: 88% vs. 94-99%; /g/: 93% vs. 99-100%). In other words, SpExpUK produced the greatest amount of short-lag VOT voiced stops in Spanish (12%), thus showing some influence from the L2 on the L1. Still, given that no significant differences were observed with the Spanish control group, there is not enough evidence to claim that there has been a drift towards L2 values in the production of the L1. The absence of a clear effect of the L2 on the L1 may be explained by the fact that the experienced learners in the present study had been living in an immersion setting for a shorter period than participants in previous studies (e.g., about four years in the present study as opposed to 12 in studies such as Flege (1987)). Therefore, just as in the case of Gorba and Cebrian (2021), it is possible that phonetic drift was not evident because a longer period of L2 experience and a more intensive use of the L2 – in detriment of the L1 – is required for it to occur. Moreover, none of the groups in the present study were absolute beginners, as they had either resided in an L2 setting or received L2 formal instruction – or both –, thus canceling a novelty effect (Chang, 2012, 2013). In this line, other previous studies that have reported changes in the L1 involved very short periods of time using the L2 intensively, and where changes in the L1 diminished a few weeks after returning to the L1 setting, for Spanish learners of English (Kartushina and Martin, 2019). As for English learners, amount of L2 instruction has also been found to be key for L1 drift to occur, both in the direction of the L2 (Herd et al., 2015; Nagle, 2019) and in terms of dissimilation between L1 and L2 categories (Dmitrieva et al., 2020).

By testing both the L1 and the L2, it is possible to determine whether the L2 learners produced the L1 and the L2 with different values and to assess which of the possible scenarios discussed in the introduction may have taken place. The overall results show that all groups produced voiced stops with different values in English and Spanish, and, thus, that the production of L2 stops is distinguished from the production of L1 stops – i.e., all groups produced more prevoiced stops in Spanish than in English. This may indicate that learners are not simply transferring their L1 patterns, particularly the English learners of Spanish, who increased considerably the amount of prevoiced productions in the L2 with respect to the L1. Still, producing L1 and L2 stops differently does not mean that L2 stops were produced in a target-like manner. Only EnExpUK's L2 values were comparable to – that is, not significantly different from – native values, and even in this case the frequency of use was notably lower (71%, vs. 98% for SpMono). This result is similar to the outcome reported by Gorba and Cebrian (2021), who examined the production of voiceless stops by the same groups of English participants as the current study. The results showed that all learners, regardless of their amount of L2 experience, produced English voiceless stops with significantly longer VOT values than Spanish stops. In that case, however, the two experienced groups performed similarly to the Spanish control group, and only the inexperienced group differed significantly from the controls. The more target-like performance observed by Gorba and Cebrian (2021) suggests a somewhat greater difficulty to acquire L2 voiced stops than voiceless stops, in line with what the outcomes of previous studies suggest (see Section 1.2.2. for a review of studies involving voiceless and voiced stops). Regarding the Spanish learners of English, despite significant differences between L1 and L2, the use of prevoicing was predominant in both languages, and close to L1-Spanish values (see Figs. 3 and 6 above), pointing to a clearer evidence of transfer of the L1 pattern onto the L2. Transfer in this case resulted in acceptable productions in English given that, as stated above, prevoicing is also found in native English even if less frequently than short-lag VOT. Only SpExpUK illustrates a greater percent use of short-lag VOT (42%), but still far from the English norm (EnMono: 91% short-lag productions).

The relationship between the use of prevoicing in the L1 and in the L2 was also investigated. Moderate though significant correlations were found for both populations and both phones, in line with some previous results (Hutchinson and Dmitrieva, 2021; MacKay et al., 2001). When it comes to the English learners of Spanish, given that the use of prevoicing is subject to individual and free variation in English, it was hypothesized that the use of voice-lead VOT in their L2 may be related to its use in the L1 due to phonetic transfer of an available L1 cue. The correlations conducted on the use of prevoicing in the L1 and L2 indicated that those English learners that used prevoicing the most frequently in English also did so in Spanish. As a matter of fact, at a group level, EnExpUK used prevoicing the most in both languages. This outcome is in agreement with some previous studies that showed that the use of negative VOT in the L1 and the L2 were correlated, and, thus, that the use of prevoicing in the L1 was transferred to the L2 (MacKay et al., 2001; Schuhmann and Huffmann, 2015). Still, the fact that use of prevoicing in the L1 and the L2 is correlated could also be interpreted as an increase in prevoicing in the L1 resulting from its use in the L2. In any case, the fact that prevoicing is a possible native cue to voiced stops in English may have facilitated the production of target-like – prevoiced – voiced stops in Spanish; hence the overall better L2 performance of the L1-English groups than of the L1-Spanish groups. Still, instances of negative transfer are also evident in the fact that the L1-English speakers produced Spanish stops with short-lag VOT stops 29-71% of the time. As we have seen, this type of negative transfer diminished as a function of L2 experience. With respect to the Spanish learners of English, the common use of prevoicing in both languages shows L1 transfer onto the L2, as discussed above.

On the other hand, little evidence has been found regarding the other two scenarios discussed in the introduction, i.e., involving an effect of the L2 on the L1 either as the result of the L2 pulling the L1 towards more L2-like forms or as the result of category dissimilation between L1 and L2 sounds. With respect to the former, recall that the two most experienced L1-English L2-Spanish groups (EnExpUK and EnExpSp) were found to use prevoicing in English significantly more often than the least experienced group (EnInexSp). This may indicate that learning to use prevoicing predominantly as a cue to stop voicing in Spanish may have carried with it a greater use of prevoicing in the L1. Still, none of the three L1-English groups differed in use of prevoicing from the English controls. Regarding the L1-Spanish groups, some influence of the L2 on the L1 can be detected particularly in the case of the most experienced group, who prevoiced the least in both languages. In the L1, SpExpUK used prevoicing numerically less than the Spanish controls, particularly in the case of /b/ (SpMono: 98% of prevoiced tokens; SpExpUK: 88%). The use of short-lag VOT on the part of SpExpUK (12% of productions) is mainly attributable to English influence, given that short-lag VOT is the most common cue for voiced stops in English, while it is not used to signal voicing in Spanish. Thus, the greater use of short-lag VOT in L1 production, as well as the fact that the use of prevoicing in the two languages was correlated, may also show some influence of the L2 on the L1. Recall, however, that this effect on the L1 was not significant, as no differences emerged in L1 production between any of the Spanish learners of English and the Spanish monolinguals. Finally, no indication of the remaining possibility, i.e., L1-L2 category dissimilation as a result of L2 learning, was found (EnInexpSp's use of prevoicing was numerically lower than EnMono's, 9% vs. 3%, respectively, but as mentioned above this difference was non-significant).

The main research questions have been discussed. Given that two populations were investigated, and different outcomes were observed in some cases, the possible explanations for these disparities will be discussed next. The first difference between the two populations examined has to do with the factors that contributed to L2 experience in each case. As mentioned above, the L1-Spanish group that performed in a more target-like manner was SpExpSp, that is, the group with the longest length of residence in the L2-speaking setting. By contrast, its L1-English counterpart, EnExpSp, was outperformed by EnExpUK, a group with only a previous shorter stay in the L2 country. This illustrates that length of residence did not have the same effect for both populations, despite the fact that the groups were comparable across populations in terms of the number of years spent in the L2 country. This difference between the two populations is related to a related variable, amount of L2 use. In the case of the L1-Spanish groups, SpExpUK and SpExpSp differed both in length of residence and in amount of L2 use. By contrast, the English L1 groups did not differ much in L2 use despite differences in place and length of residence. As a matter of fact, L2 use has been attached great importance in the revised version of the SLM (SLM-r, Flege and Bohn, 2021). Recall that, according to the SLM-r, a determining factor in L2 speech learning is FTE (Full Time Equivalent of years of L2 input), a composite variable that considers both length of residence and L2 use. As discussed above, the English groups differed little in FTE despite the difference in length of residence, while the Spanish groups had more of a difference.

On the other hand, the L1-English groups differed in amount of L2 instruction, as EnExpUK had received on average more years of Spanish instruction than EnExpSp (7.5 vs. 2.4, respectively). L2 formal instruction is a factor that has been reported to improve the production of L2 stops (e.g., Casillas 2019; Chang 2012; Lord 2010; Nagle 2019; Schuhmann and Huffmann, 2015; Zampini 1998). Interestingly, the amount of L2 instruction received by the English learners of Spanish was smaller than that received by the Spanish groups (13.5 years on average across groups) and yet the English learners of Spanish produced L2 voiced stops more similarly to native speakers of the target language than the Spanish learners of English. It seems plausible, thus, that greater differences between the two populations – namely, a greater advantage of the English population over the Spanish population – would have been observed if the L1-English speakers had received a comparable amount of L2 instruction to that of the L1-Spanish speakers.

This difference in what contributes to L2 experience between the two populations (FTE vs. L2 instruction), however, may be an artifact of the samples examined in the current study, and not generalizable to the two populations. For instance, length of residence may have played a greater role with L1-English speakers if the group residing in the L2 setting had consisted of English speakers who made a greater use of the L2, e.g., not English teachers as it was mostly the case in the current study. It should be noted, however, that these differences in L2 use and amount of L2 instruction are in fact a reflection of the linguistic reality of the two populations under study. Regarding language use, English is more present than Spanish in the international media and entertainment industry, resulting in a relatively greater presence of English in an L1-Spanish setting than the reverse. Moreover, whereas Spanish speakers in the UK generally use English at work and in everyday life interactions, the reverse is not always the case for English speakers in Spain, since many use English at work and in their social interactions. In fact, that is mostly the case in this study, as most of the participants in the EnExpSp group were teaching English as a foreign language. In addition, regarding L2 instruction, English is more widely taught as a foreign language in Spain than the other way around. Finding truly comparable groups in terms of L2 use and L2 instruction may be challenging, at least for these specific populations (English learners of Spanish in Spain and Spanish learners of English in the UK).

A second difference between the two populations has to do with the allophonic distribution of the feature under study in the two languages investigated. Recall that in English the initial stop voicing contrast is implemented as a short-lag vs. long-lag VOT contrast, and voice-lead unequivocally signals a voiced stop, although it is used less frequently than short-lag VOT. In Spanish, on the other hand, the contrast is between voice-lead and short-lag VOT. Thus, an English learner of Spanish has to learn to make a greater use of an existing L1 cue, voice-lead, when producing L2-Spanish voiced stops, and to reduce the amount of VOT used in the case of L2 voiceless stops. By contrast, a Spanish learner of English needs to learn that what constitutes a voiceless category in the L1, i.e., short-lag VOT, in fact typically signals the contrasting category in the L2, a voiced stop. Increasing the use of an existing albeit not predominant L1 cue (voice-lead for English L1 speakers) may be easier to achieve than to use an existing L1 cue that signals a different category in the L1 (short-lag VOT for Spanish L1 speakers). This may be due to the fact that while, in the case of the English L1 speakers, learning the L2 system does not affect the phonological value associated with voice-lead VOT (a voiced stop both in the L1 and the L2), in the case of the Spanish L1 speakers, learning the English contrast implies learning to dissociate short-lag VOT from the L1 (voiceless) category and associating it with the contrasting (voiced) category. Hence the greater difficulty on the part of the Spanish learners of English to produce English voiced stops with short-lag VOT than on the part of English learners of Spanish to produce Spanish voiced stops with prevoicing. Furthermore, simply transferring the use of the main L1 cue onto the L2 has different consequences for the two populations. For L1-Spanish L2-English speakers using prevoicing in English does not impede intelligibility, since prevoiced stops can only be interpreted as voiced stops. By contrast, for English L1 Spanish L2 speakers, producing Spanish stops with short-lag VOT can result in lack of intelligibility. Consequently, it is more crucial for the L1-English L1 speakers than for the L1-Spanish speakers to avoid transfer of the L1 main cue.

Nevertheless, stating that the Spanish learners of English did not produce L2 voiced stops accurately on the basis that they used prevoicing more often than English monolingual speakers is arguable. This is because prevoicing is used in English voiced stops in initial position in free variation – that is, voice-lead and short-lag VOT do not contrast – although short-lag VOT is more commonly used (Lisker and Abramson, 1964; Docherty, 1990). Perhaps, it would be more appropriate to state that the Spanish learners generally did not produce L2 voiced stops in the most common manner for native English speakers and were not able to make a clear distinction between the L1 and the L2. On the other hand, the English learners of Spanish possibly made a clearer distinction between the two languages, thanks to the subphonemic status of prevoicing in English, which coincides with that of prevoicing in Spanish – i.e., is used to produce voiced stops. In short, it appears that learning to produce on an existing L1 cue may be easier than learning to use a cue associated with a different L1 category.

This study is not without limitations. As mentioned above, the two populations under study were mirror-image in terms of L1 and L2 and comparable in terms of length of residence in an L2 setting but differed in some respects – i.e., in amount of L2 use and L2 instruction. It is possible that different outcomes would have been observed given more similar groups. Therefore, future research with more comparable L1-English and L1-Spanish speakers in terms of length of residence as well as language use is necessary to fully assess the roles of these variables for these two populations. In fact, future studies may benefit from using compound independent variables – i.e., that combine more than just one factor –, such as it is the case of the SLM-r's FTE, as that may help to control for individual differences between participants. Furthermore, the amount of L2 experience of the participants tested – i.e., the number of years spent in an L2 setting – was considerably smaller in the present study than it was in some previous research (e.g., Flege 1987). As a result, it is possible that certain crosslinguistic processes, such as it is the case of phonetic/phonological changes of the L1 towards the L2, were not clearly observed (Gorba and Cebrian, 2021). Another limitation regarding participants is related to sample size: although around 50 participants were tested in each experiment, each group was made up of between 9 and 11 participants. It should also be noted that the tasks used to retrieve the production of voiced stops – i.e., a carrier sentence reading task in each language – was a formal and controlled task. While this was motivated by allophonic patterns of Spanish and was done to prevent spirantization of the stops (/b/ and /g/ are produced as [β] and [ɣ], respectively, between continuant sounds (e.g., Colina 2020; Hualde 2005)), the data examined in the current paper may not be representative of spontaneous speech. Further research should also collect spontaneous data in order to corroborate that the results of this study reflect the real speech of L2 learners. The acquisition – on the part of English learners of Spanish – and transfer to the L2 – in the case of Spanish learners of English – of the process of spirantization needs also to be investigated in order to get a full picture of the acquisition of voiced stops in these populations in Spanish.

5. Summary and conclusions

This study has investigated the acquisition of L2 voiced stops by two populations, English learners of Spanish and Spanish learners of English, whose L1s base the initial stop voicing contrast on a different use of the same phonetic cues. Moreover, the effect of L2 experience on L1 and L2 production of /b/ and /g/ was analyzed by comparing groups differing in amount and type of L2 learning experience. Differences in the production of the two languages was also assessed by comparing L1 and L2 production. Overall, L2 experience has been found to have a positive effect on L2 voiced stop production, as the inexperienced learners of both populations consistently performed in a less target-like manner than the groups with some L2 experience. Still, L2 speakers' production generally differed from native speech, particularly in the case of the Spanish learners of English, who were found to produce prevoiced stops to a much greater extent than monolingual English speakers. Different factors have been found to contribute to L2 experience for each population of L2 learners. Length of residence together with the amount of language use accounted for between group differences among the Spanish learners of English. In the case of the English learners of Spanish, differences in the amount of time spent in the L2-speaking country did not influence L2 performance, presumably because groups did not differ in the amount of L2 use. In this context, the amount of L2 instruction and the condition of being a student may have been a more determining factor in the case of the English learners of Spanish. No notable effect of L2 experience on the L1 was found, as none of the learner groups differed from their L1 controls. The only indication of an influence of the L2 on the L1 can be found in the English learners of Spanish with some experience in the L2 setting, as they used prevoicing in English significantly more often than the least experienced group and numerically more than the English controls. Overall, longer length of residence in the L2 setting, together with a greater amount of L2 use, may be necessary for L2 effects of the L1 to be more generally observed. Regarding the comparisons between L1 and L2 production, all groups were found to use significantly more prevoicing in Spanish than in English. Despite the significant difference, most Spanish groups were found to use prevoicing predominantly in both languages, pointing to the transfer of the L1 pattern onto the L2. L1-English speakers made a greater difference between the two languages than L1-Spanish L1 speakers did and were overall more successful in that they approximated target language use of prevoicing more closely. The differences between the English and the Spanish learners can be explained by the differences in the use of the voicing cues in the languages under study. The fact that prevoicing is a possible cue, even if not the main one, to initial stop voicing in English may have facilitated a more target-like production of voiced stops on the part of the English learners of Spanish. On the other hand, Spanish learners may have initially assimilated L2 /b/ and /g/ to their corresponding L1 categories, that is, with voice-lead VOT, as short-lag VOT is associated with phonologically voiceless stops in their L1. Recall, however, that prevoicing is possible, although less common, in English voiced stops and, thus, it can be argued that using Spanish values in English is acceptable. Further research is necessary to evaluate the contribution of the experience variables examined here, with more balanced designs. In addition, it remains to be seen if the results of this study would be replicated in the production of more spontaneous speech, as opposed to read speech. These issues are left for future research.

CRediT authorship contribution statement

Celia Gorba: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Juli Cebrian:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Juli Cebrian reports financial support was provided by Spain's Ministry of Science and Innovation.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by Research Grants No. FFI2017-88016-P and PID2021-122396NB-I00 from the Spanish Ministries of Economy and Competitiveness and Science and Innovation.

Appendix A. List of sentences used in the production tasks

A.1. English production task

Read each of the following sentences twice:

1 Kitten is the next word
2 Peeler is the next word
3 Beaches is the next word
4 Watches is the next word
5 Gearbox is the next word
6 Cellphone is the next word
7 Keychain is the next word
8 Peaceful is the next word
9 Tiger is the next word
10 Pieces is the next word
11 Houses is the next word
12 Beefcake is the next word
13 Keenly is the next word
14 Girlfriend is the next word
15 Keeper is the next word
16 Music is the next word
17 Geezer is the next word
18 Razor is the next word
19 Beetle is the next word
20 Flawless is the next word
21 Peanut is the next word
22 Geekfest is the next word
23 Flipflop is the next word
24 Beetroot is the next word
25 Beating is the next word
26 Headphones is the next word
27 Keener is the next word
28 Geeky is the next word
29 Ancient is the next word
30 Peacock is the next word
31 Gearshift is the next word
32 Keyhole is the next word

A.2. Spanish production task

Lee cada una de las siguientes frases dos veces:

1 Monte es la siguiente palabra
2 Pila es la siguiente palabra
3 Hombre es la siguiente palabra
4 Quicio es la siguiente palabra
5 Bizco es la siguiente palabra
6 Guinda es la siguiente palabra
7 Piso es la siguiente palabra
8 Mono es la siguiente palabra
9 Bicho es la siguiente palabra
10 Guiso es la siguiente palabra
11 Rata es la siguiente palabra
12 Birla es la siguiente palabra
13 Ante es la siguiente palabra
14 Quinta es la siguiente palabra
15 Orden es la siguiente palabra
16 Guía es la siguiente palabra
17 Pista es la siguiente palabra
18 Perro es la siguiente palabra
19 Biblia es la siguiente palabra
20 Guiño es la siguiente palabra
21 Uña es la siguiente palabra
22 Quita es la siguiente palabra
23 Uso es la siguiente palabra
24 Bici es la siguiente palabra
25 Móvil es la siguiente palabra
26 Quince es la siguiente palabra
27 Pico es la siguiente palabra
28 Cama es la siguiente palabra
29 Quise es la siguiente palabra
30 Susto es la siguiente palabra
31 Guita es la siguiente palabra
32 Pino es la siguiente palabra

References

Audacity Team., 2016. Audacity(R): free audio editor and recorder [computer application]. Version 2.1.1. Retrieved in February 2016 from https://audacityteam.org/.
Baese-Berk, M.M., 2019. Interactions between speech perception and production during learning of novel phonemic categories. Atten. Percept. 81 (4), 981. http://mendeley.csuc.cat/fitxers/40f99bf6fdfa610abb22465d179a7f43.
Bergmann, C., Nota, A., Sprenger, S.A., Schmid, M.S., 2016. L2 immersion causes non-native-like L1 pronunciation in German attriters. J. Phon. 58, 71–86. https://doi.org/10.1016/j.wocn.2016.07.001.
Best, C. T., Michael D. T., 2007. Nonnative and second-language speech perception: commonalities and complementarities. In Ocke-Schwen Bohn & Murray J. Munro (Eds.), Language Experience in Second Language Speech Learning: In Honor of James Emil Flege, 10389 (2): 13–34. doi:10.1121/1.1332378.
Birdsong, D., Gertken, L.M., Amengual, M., 2012. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/.

Boersma, P., Weenink, D., 2016. Praat: doing phonetics by computer [computer software]. Version 6.0.19. Retrieved February 2016 from http://www.praat.org/.
Bohn, O.S., Flege, J.E., 1990. Interlingual identification and the role of foreign language experience in L2 vowel perception. Appl. Psycholinguist. 11 (3), 303–328. https://doi.org/10.1017/S0142716400008912.
Caramazza, A., Yeni-Komshian, G.H., Zurif, E.B., Carbone, E., 1973. The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. J. Acoust. Soc. Am. 54 (2), 421. http://mendeley.csuc.cat/fitxers/503cef848b96b7b7a540101eabff812.
Casillas, J., 2019. Phonetic category formation is perceptually driven during the early stages of adult L2 development. Lang. Speech 28. https://doi.org/10.1177/0023830919866225.
Castañeda, M.L., 1986. El VOT de las oclusivas sordas y sonoras españolas. Estud. Fon. Exp. 2, 92–110.
Chang, C.B., 2012. Rapid and multifaceted effects of second-language learning on first language speech production. J. Phon. 40 (2), 249–268. https://doi.org/10.1016/j.wocn.2011.10.007.
Chang, C.B., 2013. A novelty effect in phonetic drift of the native language. J. Phon. 41 (6), 520–533. https://doi.org/10.1016/j.wocn.2013.09.006.
Colina, S., 2020. Spirantization in Spanish: the role of the underlying representation. Linguistics 58 (1), 1–35.
Derek, C., DeMars, C., 2019. An updated recommendation for multiple comparisons. Adv. Methods Pract. Psychol. Sci. 2 (1), 26–44.
Docherty, G.J., 1990. An experimental phonetic study of the timing of voicing in English obstruents. Diss. Abstr. Int. 51 (3). http://mendeley.csuc.cat/fitxers/41e02d2e0aa875774b1d4557e8b2ded4.
Dmitrieva, O., Jongman, A., Sereno, J.A., 2020. The effect of instructed second language learning on the acoustic properties of first language speech. Languages 5 (4), 44.
Flege, J.E., 1987. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. J. Phon. 15 (1), 47–65. https://doi.org/10.1016/S0095-4470(19)30537-6.
Flege, J.E., 1990. Production of the English word-final /t/-/d/contrast by native speakers of Mandarin and Spanish. Acoust. Soc. Am. J. 88, 81. https://doi.org/10.1121/1.2029176.
Flege, J.E., 1991. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. J. Acoust. Soc. Am. 89 (1), 395–411. https://doi.org/10.1121/1.400473.
Flege, J.E., Winifred, S., 1995. Second language speech learning: theory, findings, and problems. Speech Perception and Linguistic Experience: Issues in Cross-Language Research. York Press, Timonium, MD, pp. 233–277. https://doi.org/10.1111/j.1600-0404.1995.tb01710.x.
Flege, J.E., Bunneister, P., Piske, T., Rohde, A., 2002. Interactions between the native and second-language phonetic systems. An Integrated View of Language Development: Papers in Honor of Henning Wode. Wissenschaftlicher Verlag, Trier, pp. 217–244.
Flege, J.E., Schiller, Niels O., Meyer, Antje S., 2003. Assessing constraints on second-language segmental production and perception. Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities. de Gruyter Mouton, Berlin, pp. 319–355.
Flege, J.E., Cole, J., Hualde, J.I., 2007. Language contact in bilingualism: phonetic system interactions. Laboratory Phonology 9. De Gruyter Mouton, New York, pp. 353–381.
Flege, J.E., Bohn, O.S., 2021. The Revised Speech Learning Model (SLM-r). In: Wayland, R. (Ed.), Second language speech learning: theoretical and empirical progress. Cambridge University Press, New York, NY, pp. 3–83.
Flege, J.E., Bohn, O.S., Jang, S., 1997. Effects of experience on non-native speakers' production and perception of English vowels. J. Phon. 25 (4), 437–470. https://doi.org/10.1006/JPHO.1997.0052.
Flege, J.E., Eefting, W., 1987. Production and perception of English stops by native Spanish speakers. J. Phon. 15 (1), 67–83. https://doi.org/10.1016/S0095-4470(19)30538-8.
Fowler, C.A., Sramko, V., Ostry, D.J., Rowland, S.A., Hallé, P., 2008. Cross language phonetic influences on the speech of French-English bilinguals. J. Phon. 36 (4), 649–663. https://doi.org/10.1016/j.wocn.2008.04.001.
Gorba, C., 2019. Bidirectional influence on L1 Spanish and L2 English stop perception: the role of L2 experience. J. Acoust. Soc. Am. 145 (6), EL587–EL592. https://doi.org/10.1121/1.5113808.
Gorba, C., 2020. Crosslinguistic influence in L2 speakers: the effect of language experience on L1 and L2 VOT production and perception [Unpublished PhD thesis Universitat Autònoma de Barcelona].
Gorba, C., Cebrian, J., 2021. The role of L2 experience in L1 and L2 perception and production of voiceless stops by English learners of Spanish. J. Phon. 88.
Harada, T., 2003. L2 influence on L1 speech in the production of VOT. In: Proceedings of the 15th International Congress of Phonetic Sciences, pp. 1085–1088.
Hazan, V.L., Boulakia, G., 1993. Perception and production of a voicing contrast by french-english bilinguals. Lang. Speech 36 (1), 17–38. https://doi.org/10.1177/002383099303600102.
Herd, W.J., R.L. Walden, Whitney L. Knight, and Savana N. Alexander. 2015. Phonetic drift in a first language dominant environment. Proceedings of the Paper Presented at the Meetings on Acoustics, Jacksonville, FL, USA, November 2-6; Volume 23, pp. 1-9.
Hualde, J.I., 2005. The Sounds of Spanish with Audio CD. Cambridge University Press.
Huffman, M.K., Schuhmann, K.S., 2016. Effect of Early L2 Learning on L1 Stop Voicing. In: Proceedings of the Paper presented at the Meetings on Acoustics, 23. Acoustical Society of America, Pittsburgh, PA, USAPittsburgh, pp. 1–10. September 9.
Hutchinson, A.E., Dmitrieva, O., 2021. Individual variation in acquisition of voicing contrasts by classroom learners of French. Journal of Second Language Pronunciation 7 (3), 408–434.
Julià i Muné, J., 1981. Estudi Contrastiu del Oclusius de L'anglès i del Català : Un Experiment Acústic, 1. Estudi General: Revista de La Facultat de Lletres de La Universitat de Girona, p. 75, 7.
Jun, S., Cowie, I., 1994. Interference for 'New' and 'Similar' Vowels in Korean Speakers of English, 43. Ohio State University, pp. 117–130. Working Papers.
Kartushina, N., Frauenfelder, U.H., Golestani, N., 2016. How and when does the second language influence the production of native speech sounds: a literature review: L2 influences on L1: a literature review. Lang. Learn. 66 (S2), 155–186. https://doi.org/10.1111/lang.12187.
Kartushina, N., Martin, C.D., 2019. Third-language learning affects bilinguals' production in both their native languages: a longitudinal study of dynamic changes in L1, L2 and L3 vowel production. J. Phon. 77, 100920 https://doi.org/10.1016/j.wocn.2019.100920.
Kuhl, P.K., 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. Perception 50 (2), 93.
Lev-Ari, S., Peperkamp, S., 2013. Low inhibitory skill leads to non-native perception and production in bilinguals' native language. J. Phon. 41 (5), 320–331. https://doi.org/10.1016/j.wocn.2013.06.002.
Levy, E.S., Law, F.F., 2010. Production of french vowels by American-English learners of French: language experience, consonantal context, and the perception-production relationship. J. Acoust. Soc. Am. 128 (3), 1290. https://doi.org/10.1121/1.3466879.
Lisker, L., Abramson, A., 1964. A cross-language study of voicing in initial stops: acoustical measurements. Word 20, 384–422. https://doi.org/10.1080/00437956.1964.11659830.
Lord, G., 2010. The combined effects of immersion and instruction on second language pronunciation. Foreign language annals 43 (3), 488–503.
MacKay, I., Flege, J., Piske, T., Schirru, C., 2001. Category restructuring during second-language (L2) speech acquisition. J. Acoust. Soc. Am. 110, 516–528. https://doi.org/10.1121/1.1377287.
Major, R.C., 1992. Losing English as a first language. Mod. Lang. J. 76 (2), 190. http://mendeley.csuc.cat/fitxers/0af2fb3eae3a42c82adde937d9b0701f.
Maye, J., Gerken, L., 2000. Learning phonemes without minimal pairs. In: Proceedings of the 24th Annual Boston University Conference on Language Development, 2, pp. 522–533.
Mora, J. C., 2008. Learning context effects on the acquisition of a second language phonology. In C. Pérez-Vidal, M. Juan-Garau, A. Bel, (eds.), A Portrait of the Young in the New Multilingual Spain. Clevedon: Multilingual Matters. 241-263.
Nagle, C., 2019. A longitudinal study of voice onset time development in L2 Spanish stops. Appl. Linguist. 40 (1), 86–107. https://doi.org/10.1093/applin/amx011.
Nathan, G.S., 1987. On second-language acquisition of voiced stops. J. Phon. 15 (4), 313–322.
Ohala, J.J., MacNeilage, P.F., 1983. The origin of sound patterns in vocal tract constraints. The Production of Speech. Springer-Verlag, New York, pp. 189–216.
Piske, T., MacKay, I.R.A., Flege, J.E., 2001. Factors affecting degree of foreign accent in an L2: A review. J. Phon. 29 (2), 191–215. https://doi.org/10.1006/jpho.2001.0134.
Sancier, M.L., Fowler, C.A., 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. J. Phon. 25 (4), 421–436. https://doi.org/10.1006/jpho.1997.0051.
Schuhmann, K., Huffman, M. K., 2015. L1 Drift and L2 Category Formation in Second Language Learning.
Schuhmann, K.S., Huffman, M.K., 2019. Development of L2 Spanish VOT before and after a brief pronunciation training session. J. Second Lang. Pronunciation 5 (3), 402–434. https://doi.org/10.1075/jslp.18018.sch.
Stevens, J., 2001. Study abroad learners' acquisition of the Spanish voiceless stops. MIFLC Rev. 10, 137–151.
Williams, L., 1977. The perception of stop consonant voicing by Spanish-English bilinguals. Percept. Psychophys. 21 (4), 289–297. https://doi.org/10.3758/BF03199477.
Yang, Y., Chen, X., Xiao, Q., 2022. Cross-linguistic similarity in L2 speech learning: evidence from the acquisition of Russian stop contrasts by Mandarin speakers. Second Lang. Res. 38 (1), 3–29. https://doi.org/10.1177/0267658319900919.
Zampini, M.L., 1998. The relationship between the production and perception of L2 Spanish Stops. Tex. Pap. Foreign Lang. Educ. 3 (3), 85. https://mendeley.csuc.cat/fitxers/77294836f95fbc4b90c465f296937509.