Bilinguals produce language-specific voice onset time in two true-voicing languages

The case of Basque-Spanish early bilinguals

Christoforos Souganidis,[1,2] Nicola Molinaro,[1,3] and Antje Stoehr[1]
[1] BCBL, Basque Center on Cognition, Brain and Language |
[2] Universidad del País Vasco/Euskal Herriko Unibertsitatea UPV/EHU |
[3] Ikerbasque, Basque Foundation for Science

It is well established that early bilinguals who speak languages that differ in the phonetic implementation of the voicing contrast have language-specific voicing systems. This study investigates voicing separation in bilinguals who speak two true-voicing languages, Basque and Spanish. We also describe the voicing system in Gipuzkoan Spanish and Gipuzkoan Basque, which is the closest dialect to Standard Basque and it has not yet been investigated experimentally. Twenty Basque-Spanish early bilingual speakers of Gipuzkoan dialects participated in two picture naming tasks. We described their voicing system by measuring voice onset time (VOT) in both Gipuzkoan Basque and Spanish, and used linear mixed-effects models to investigate between-language production differences. Our results show for the first time that adult early bilinguals who speak two true-voicing languages produce language-specific VOT in 'voiced' plosives. This finding demonstrates that bilinguals' phonetic systems during production are more fine-grained than previously assumed, and contributes to a deeper understanding of granularity in early bilingual phonetic systems.

Keywords: voice onset time production, language-specific production categories, early bilinguals, Basque-Spanish bilinguals, picture naming task

1. Introduction

Bilinguals who acquire both languages at a very young age face the intriguing task of establishing two phonological systems in parallel. This task involves sounds that are unique to one of the two languages and also sounds that are present in

https://doi.org/10.1075/lab.21081.sou | Published online: 31 October 2022
Linguistic Approaches to Bilingualism 14:3 (2024), pp. 370–399. ISSN 1879-9264 | E-ISSN 1879-9272

both languages. The latter may still have subtle phonetic differences between the two languages. For example, plosives are present in the phonologies of virtually all languages (Maddieson, 1984), but their phonetic implementation may differ in terms of voice onset time (VOT). This makes VOT a useful tool to investigate cross-linguistic interactions in bilingual speech production.

VOT is a primary cue to voicing, which distinguishes 'voiced'[1] and 'voiceless' plosives, such as the Spanish labial plosives /b/ and /p/ (Abramson & Lisker, 1973; Williams, 1977a). VOT is the interval between the release of the occlusion of plosives and the onset of voicing (Lisker & Abramson, 1964). The VOT continuum can be separated into three phonetic categories: prevoicing/voicing lead, short-lag, and aspiration/long-lag (see Figure 1). In most languages with a two-way distinction between 'voiced' and 'voiceless' plosives the key contrast is either between prevoicing and short-lag (hereafter, true-voicing languages, e.g., Spanish, Basque) or short-lag and aspiration (hereafter, aspirating languages, e.g., English; Lisker & Abramson, 1964).

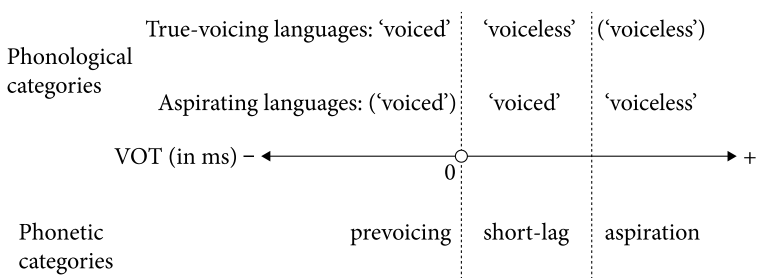

Figure 1. Phonological and phonetic categories for true-voicing and aspirating languages, on the VOT continuum. Parentheses indicate parts of the VOT continuum that speakers occasionally produce, although it is not contrastive for their native language

1.1 VOT production in bilinguals

When researchers first started to investigate VOT production in bilinguals, they aimed to determine whether bilinguals produced language-specific VOT in true-voicing and aspirating languages (e.g., Antoniou et al., 2010; Caramazza et al., 1973; Kang & Guion, 2006) and whether their production diverged from monolinguals (Flege & Eefting, 1987b; Fowler et al., 2008; Kupisch & Lleó, 2017; Sundara et al., 2006). Many of these studies were situated within the framework of the Speech Learning Model (SLM; Flege, 1995). The SLM and its revised ver-

1. Hereafter, single quotation marks are used to refer to phonologically voiced/voiceless plosives.

sion (SLM-r; Flege & Bohn, 2021) provide a theoretical framework for bilingual speech production. The SLM/SLM-r posit that there is a bidirectional influence between a bilingual's languages, which coexist in a shared phonetic space. In line with this framework, a new voicing category will only be established if a new L2 sound is sufficiently different from the corresponding L1 category. If L1 and L2 sounds are largely similar, formation of a new category might be blocked, potentially leading to a merged L1-L2 category. Even if a new category is formed for an L2 sound, its production may not be monolingual-like. It is now well-established that bilinguals who speak both an aspirating language (e.g., English) and a true-voicing language (e.g., Spanish) produce language-specific VOT in 'voiceless' plosives, which may nevertheless diverge from the VOT production of monolinguals (e.g., Flege & Eefting, 1987b), highlighting the fact that bilinguals differentiate their languages even though their production might differ from monolinguals. Within the SLM/SLM-r framework, these findings can be attributed to the formation of a new category due to phonetic differences from the L1. Surprisingly, however, bilinguals who produce monolingual-like 'voiceless' plosives seem to produce 'voiced' plosives mainly with prevoicing, even in the aspirating language (e.g., Hazan & Boulakia, 1993; Sundara et al., 2006). In the SLM/SLM-r framework, this is considered to result from equivalence classification of L2 sounds that are not perceived as sufficiently different from the L1 sounds. Those L1 and L2 sounds are considered to be merged in a composite L1-L2 category. However, recent findings suggest that some bilinguals acquire language-specific 'voiced' plosives even if they do not acquire language-specific 'voiceless' plosives (e.g., Osborne & Simonet, 2021).

Early acquisition of a second language is associated with native-like performance according to the SLM and this has been confirmed by many studies (e.g., Flege et al., 1995, 1999). Yet, minority language speakers do not always exhibit monolingual-like pronunciation (Chang et al., 2011; Kupisch, Barton, et al., 2014; Kupisch, Lein, et al., 2014; Oh et al., 2003), for instance due to restricted minority language input from a restricted number of speakers. Bilinguals may follow individual patterns in phonetic category formation, either forming separate categories for sounds in their L2 with distinct VOT productions or merging them with an L1 category, which would result in indistinguishable VOT production in both languages (e.g., Flege & Eefting, 1987a; Kang & Guion, 2006). Unsurprisingly, most of the bilingual VOT literature has focused on language pairs such as English and Spanish that have different phonetic implementations of the voicing contrast. However, even languages with similar voicing systems might have small VOT differences. For example, VOT differences have been found across a variety of true-voicing languages. For instance, Puerto Rican Spanish has longer prevoicing than Dutch (Lisker & Abramson, 1964) and Canadian French has longer short-lag

VOT than Peninsular Spanish (Llama & Cardoso, 2018). Language pairs with similar voicing contrasts offer a unique opportunity to investigate whether bilinguals can exploit fine-grained phonetic differences to produce language-specific VOT.

Empirical evidence for language-specific productions of VOT in languages with the same phonetic voicing contrast comes from Mayr and Montanari (2014). They investigated the VOT production of two siblings aged 6;8 and 8;1 (years;months), who were simultaneous trilingual speakers of English, an aspirating language, and Italian and Spanish, both true-voicing languages. Surprisingly, the siblings had different VOT production patterns in Italian and Spanish. The authors attributed this to the influence of English on Italian but not on Spanish. Language-specific VOT production has also been documented in children who speak Mandarin and English, both aspirating languages (Yang, 2021; Qi et al., 2012). These children produced distinct VOT in Mandarin and English short-lag and aspirated plosives. These studies provide the first important evidence for fine-grained organization of two languages with similar voicing systems (i.e., true-voicing languages or aspirating languages). Therefore, even when two languages employ the same voicing distinction, bilingual children might still produce slightly different VOT in these two languages.

Early bilingual adults' VOT production in languages with similar phonetic and phonological voicing systems has not yet been investigated. Here, we investigate how bilinguals who speak Gipuzkoan Spanish (a Peninsular Spanish dialect spoken in northern Spain) and Gipuzkoan Basque (a Central Basque dialect spoken in the same region), both true-voicing languages, accommodate possible fine-grained differences in their phonological system. Below, we outline the Spanish and Basque voicing systems before we move on to the current study.

1.2 VOT in Spanish and Basque

Spanish is a Romance language with official status in 21 countries, which had nearly 400 million native speakers spread around the world by the end of the 20th century (Stewart, 1999, p. 3). Thus, it is no surprise that there is considerable dialectal variation in Spanish, not only between countries but also within varieties spoken in the same country. Phonetic and phonological differences have been reported for variants of Latin American Spanish (e.g., Moreno & Mariño, 1998), as well as variants of Peninsular Spanish in Spain (e.g., Torreira, 2012).

Since researchers first started to investigate VOT production, Spanish has received considerable attention (Lisker & Abramson, 1964; Williams, 1977b). Early studies tested speakers of Latin American varieties of Spanish from various countries. Williams (1977b) explored VOT production of speakers from Guatemala, Peru, and Venezuela but found no significant differences between dialects. Rosner, López-Bascuas et al. (2000) replicated Williams' study (1977b) using the same methodology to test Peninsular Spanish speakers.[2] They predominantly found shorter prevoicing in Peninsular Spanish than in the Latin American varieties reported by Williams, with values falling outside the 99% confidence intervals. Table 1 displays an overview of those studies on VOT production in Spanish plosives in word-initial stressed position.

Table 1. By-plosive mean VOT (in ms) in word-initial stressed position, different varieties of Spanish

Study	Variety	Vowels	/p/	/t/	/k/	/b/	/d/	/g/
Williams (1977b)	Guatemalan	/a/, /o/	10	10	26	−120	−109	−101
Williams (1977b)	Venezuelan	/a/, /o/	14	21	33	−95	−79	−64
Williams (1977b)	Peruvian	/a/, /o/	15	16	30	−102	−110	−98
Rosner et al. (2000)	Peninsular	/a/, /o/	13	14	27	−92	−92	−74

In contrast to Spanish, Basque is an understudied language that does not belong to the Indo-European language family. It is predominantly spoken in the Basque Autonomous Community (BAC) in northern Spain, where the majority of Basque speakers live (almost 1 million speakers; Basque Government, 2016), although there are also speakers in Navarre (also in Spain) and some regions in southern France. In the BAC, Spanish and Basque are both official languages. Basque can be considered the minority language, but it is compulsory in education, as is Spanish, and receives broad institutional and social support. Due to this support, it enjoys a higher status than most minority languages spoken elsewhere (Cenoz, 2008, p. 108). Crucially, the Basque language plays an important role in Basque identity (Azurmendi et al., 2008). In sum, bilingual speakers in the BAC receive rich input both in Basque and Spanish.

While it is appealing to study bilingualism in such an intrinsically bilingual community, two methodological issues may hamper modern studies of Basque: the lack of monolingual speakers (Hualde, 2015) and its dialectal variability especially in phonetics and phonology (Hualde, 1991; Hualde & Ortiz de Urbina, 2003). In this study, we focus on Gipuzkoan Basque, the dialect that serves as the basis for Standard Basque (Hualde & Ortiz de Urbina, 2003).

Despite the dialectal variations observed in both Spanish and Basque, they are claimed to be similar in terms of phonetics and phonology (Hualde, 2015). 'Voiced' plosives are prevoiced in both Basque and Spanish. However, we would

2. No information about the specific dialects is provided.

expect small differences between Basque and Peninsular Spanish VOT, given there is research showing 'voiced' plosives have similar prevoicing durations in Basque (Etxebarria, 1987) and Puerto Rican Spanish (Hurch, 1988), while Peninsular Spanish has shorter prevoicing durations than Puerto Rican and other Latin American Spanish (Castañeda Vicente, 1986; Rosner et al., 2000). Based on these previously-observed differences, we suspect fine-grained differences in prevoicing duration between Basque and Peninsular Spanish.

Basque and Spanish 'voiceless' plosives reportedly have short lag VOT (Hualde, 1991) but might still have possible differences. Basque 'voiceless' plosives used to be aspirated, mostly in monosyllabic words (e.g., *ke* [kʰe], 'smoke'), and loanwords (e.g., *pentsa* [pʰents̺a], 'think'; Hualde, 2018). These historically aspirated words are now produced with short-lag VOT in modern Standard Basque and most of its dialects. Aspirated plosives can still be found in at least two varieties of Basque spoken in France: Zuberoan (Gaminde et al., 2002; Mounole, 2004) and Mixean (Egurtzegi & Carignan, 2020) both employ a three-way distinction between 'voiced', unaspirated 'voiceless' and aspirated 'voiceless' plosives, although the VOT of aspirated plosives in Mixean is less extreme than that of Zuberoan. To our knowledge, apart from these three studies, there is no experimental study describing VOT production in Standard Basque or any of its dialects spoken in Spain.

1.3 The current study

The purpose of this study is two-fold: first, to investigate VOT production in Basque-Spanish early bilingual adults, and second, to describe the voicing system of Gipuzkoan Basque and Spanish.

Gipuzkoa is considered the most bilingual region of the BAC in Spain, with the highest percentage of Basque speakers (Basque Government, 2016). Consequently, recruiting participants from Gipuzkoa made it highly likely that they had received rich language input in both Basque and Spanish. Since the Basque language is important for the Basque identity (Azurmendi et al., 2008), we speculate that our participants have a greater intrinsic need for language separation to signal their difference from the Spanish identity, through linguistic means. Importantly, monolingual controls are not necessary for the current study, which aims at investigating VOT differences *within* bilingual speakers (see Hopp & Schmid, 2013; Kroll et al., 2006).

Although the SLM and SLM-r do not explicitly focus on the behavior of early bilinguals, it has been proposed that the SLM could be applied to early bilinguals' speech production (Watson, 2007). Thus, the SLM/SLM-r provide an adequate theoretical framework to address the research question of the present study, as all our participants were early bilinguals. Hereafter, we take the SLM-r perspective, which focuses on individual differences in L2 sound learning and the importance of input (disregarding traditionally-claimed age of acquisition-related differences), although some of the premises of the SLM-r are identical to those of the SLM. Specifically, our research question was: Do Basque-Spanish bilingual adults produce differential VOT in Spanish and Basque 'voiced' and/or 'voiceless' plosives?

We expect to find longer VOT for 'voiceless' plosives in Basque than Spanish (Prediction 1). However, we do not expect to find strongly aspirated plosives in Basque. This is because Basque-Spanish adolescents reportedly have difficulty producing aspiration in English, making it unlikely that they aspirate in Basque (Martínez Adrián et al., 2013). Nevertheless, since aspirated word-initial 'voiceless' plosives were historically part of Basque, it is likely that in modern Standard Basque, word-initial aspirated plosives merged with unaspirated 'voiceless' plosives. This merged category could comprise intermediate values, which fall either in the short-lag or between the short-lag and the aspiration parts of the VOT continuum, and would thus be longer than in Spanish. In this scenario we would expect longer VOT for 'voiceless' plosives in Basque than Spanish.

Given the VOT production differences between Basque (Etxebarria, 1987) and Peninsular Spanish (Castañeda Vicente, 1986; Rosner et al, 2000), we expect Basque-Spanish bilinguals to produce language-specific VOT in 'voiced' plosives, with shorter prevoicing in Spanish than Basque (Prediction 2). However, according to SLM-r, the two languages coexist in a shared phonetic space, so we would not expect complete separation but rather some degree of overlap. Moreover, we would not necessarily expect between-language differences to follow the same directions in all speakers, as Basque lacks monolingual norms (Hualde, 2015): The SLM-r acknowledges individual differences in production of both L1 and L2 sounds, as it has been shown that different patterns are used by different speakers of the same two languages (e.g., the different patterns of 'voiced' plosives of the siblings in Mayr & Montanari, 2014).

2. Methods

2.1 Participants

Twenty Basque-Spanish bilinguals (10 women/10 men; $M_{age} = 23.9$ years old; $SD_{age} = 4.59$; range = 19–34 years old) participated in the study. They did not report any speech, hearing, or reading disorders, and all had at least graduated from high school (8 high school graduates, 4 with technical training, 6 with university, and 2 with postgraduate studies). All of them were speakers of Gipuzkoan dialects of Basque (self-reported), who also spoke Spanish and English, and resided in Gipuzkoa in the BAC of Spain at the time of testing. All participants were early bilinguals, highly proficient in both their languages, who had begun acquiring Spanish before the age of four years and Basque before the age of three years, according to self-report. They also self-reported the L1 of their parents: Spanish for both parents ($n=2$), Basque for both parents ($n=6$), and Spanish for one parent and Basque for the other parent ($n=12$). Participants' language proficiency in Spanish, Basque, and English was assessed using the Basque, English, and Spanish Test (BEST; de Bruin et al., 2017). The BEST includes participants' vocabulary knowledge (score: 0 to 65) and an interview rating (1 to 5).[3] Both parts of the test were conducted by native Basque-Spanish bilingual experimenters, specifically trained for this purpose. We recruited participants with high proficiency in both Spanish and Basque. Our recruitment criterion for Basque, the minority language, was high proficiency, captured through an interview rating >= 4 (corresponding to high fluency and coherence, adequate lexical resources, correct grammar, and no mispronunciations) and a vocabulary knowledge score > 55. Spanish scores were near ceiling (interview = 5, vocabulary knowledge > 63). English is mandatory in secondary education in the BAC, which explains why all participants had at least some knowledge of English. In order to minimize possible interference from other languages, we selected only participants who had not lived abroad. For more information on participant profiles, see Table 2.

Participants were recruited through the participant pool of the Basque Center on Cognition, Brain and Language (BCBL) in Donostia San-Sebastián, Spain. Prior to testing, participants read and signed a consent form that was approved by the BCBL's Ethics Committee, based on the Declaration of Helsinki. Participants received monetary compensation for their time, according to BCBL regulations.

2.2 Materials and design

Participants completed two picture naming tasks (PNT), one in Basque and one in Spanish. Each consisted of 60 target words comprising 10 target words

3. Structured interviews were conducted by professional research assistants who had received standardized training for these interviews. The scores can be interpreted as follows: 5: native speaker competence. 4: speakers are highly fluent, able to talk about a wide range of topics, with some repetition or self-correction. 3: speakers are fluent, able to speak at length using a wide range of vocabulary, and generally easy to understand, although they make some mistakes in complex speech. 2: speakers have limited fluency, frequently unable to convey basic meaning and using limited vocabulary. 1: speakers cannot produce basic sentence forms and sometimes no communication is possible.

Table 2. Participants' age of acquisition, BEST vocabulary knowledge scores and interview ratings for Spanish, Basque, and English

Measure	Language	Mean (SD)	Range
AoA (years)	Spanish	0.7 (1.26)	0–3
	Basque	0.1 (0.45)	0–2
	English	6.5 (1.82)	4–11
Vocabulary knowledge (0–65)	Spanish	64.75 (0.44)	64–65
	Basque	62.3 (3.06)	56–65
	English	48.65 (10.37)	22–65
Interview (1–5)	Spanish	5 (0)	5–5
	Basque	4.85 (0.37)	4–5
	English	3.35 (0.81)	1–4

for each plosive: /b/, /d/, /g/, /p/, /t/, /k/. All target words were plosive-vowel initial, stressed on the first syllable and most were disyllabic (see Appendix 1). Some monosyllabic words were included mainly due to the unavailability of disyllabic Basque words that met our selection criteria, and due to the restriction that words had to be easy to depict (see Appendix 2). Spanish and Basque words were matched for height of vocalic context (high vs. mid or low) since high vowels are associated with longer VOT (e.g., Berry & Moyle, 2011; Esposito, 2002), but they could not be matched for syllable count. Participants' speech was elicited through drawings. Seventy-four (24 Basque, 50 Spanish) were taken from the MultiPic database (Duñabeitia et al., 2018); the remaining 46 were novel drawings (10 Spanish, 36 Basque), either open content drawings found online or drawings created by us.

2.3 Procedure and apparatus

All participants were tested individually in a sound attenuating room at the BCBL facilities in Donostia-San Sebastián, Spain. They were seated at a distance of approximately 45 cm from a computer screen on which the stimuli were presented visually. Tasks were programmed in OpenSesame version 3.2.4 (Mathôt et al., 2012). For the speech recordings, a Sennheiser ME65 microphone was positioned approximately 25 cm in front of the participants, pointing to their vocal folds. Participants were asked to name the drawings shown to them, using a single word. Instructions were provided in the target language both orally by the experimenter and in writing on the screen. If participants did not produce a target word, the experimenter would provide hints in the target language. If participants made a noise, such as coughing, during the production of a target word, they were asked to repeat the word. The order of drawings was randomized for each participant, and each drawing was presented once.

First, participants completed the Basque PNT, then the Spanish PNT. Since one of our aims was to describe the VOT production system of Basque, the Basque task was presented first. Furthermore, it is common procedure to elicit speech first in the minority language or L2 in phonetic research (e.g., Cheng, 2020; Schuhmann & Huffman, 2019). The language of interaction with the participants during each task matched the language tested in the task, as this has been shown to elicit monolingual-like VOT (Antoniou et al., 2010). Between the two tasks, participants were given the option to take a break. Before proceeding to the Spanish PNT, the experimenter changed the language of interaction to Spanish, asking questions about the task they had completed that were unrelated to the research question. The total duration of the experimental session was between 15 and 20 minutes.

2.4 VOT measurements

VOT of the target words in both languages was measured by the first author, using Praat (Boersma & Weenink, 2013). VOT measurements were obtained via visual inspection of the spectrogram and the waveform. For 'voiced' plosives, prevoicing, that is, negative VOT, was measured as the interval between the zero-crossing before the onset of the first periodic waveform component (Francis et al., 2003) and the first release of the closure (i.e., *burst*, Lisker & Abramson, 1964) preceding the higher formant of the vowel following the plosive. For 'voiceless' plosives, VOT was measured as the interval between the burst and the onset of the following vowel. Multiple bursts were identified in many instances of /k/, in which case, the first burst was taken as the starting point of measurement (Mayr & Siddika, 2018).

2.5 Data analyses and model description

We performed linear mixed-effects models (LMEMs) using the lme4 package (Bates et al., 2015) in R (R Core team, 2020) via RStudio v1.2.5042 (RStudio Team, 2020), with the optimx optimizer (Nash, 2014). To obtain *p*-values from the models, we used the lmerTest package (Kuznetsova et al., 2017) which uses Satterthwaite approximations to calculate degrees of freedom (df). To check for collinearity, normality of residuals and random effects, homoscedasticity, and homogeneity of variance, we used the performance package (Lüdecke et al., 2021).

Pairwise comparisons (Bonferroni-corrected, Satterthwaite method for df) were conducted for the main effect of predictors with three levels, after calculating estimated marginal means with the emmeans package (Lenth et al., 2020). Model fit (marginal and conditional R^2) was obtained using the MuMIn package (Bartoń, 2019). For plots we used the ggplot package (Wickham, 2011).

Data analyses included 96.83% of the data (2324 observations). Exclusions included 18 observations in which participants did not produce the target word; 34 'voiced plosives' produced with positive VOT; and 24 Basque words without stress-initial production (due to the variable stress pattern in Basque).

For our analyses we fitted separate LMEMs for 'voiced' and 'voiceless' plosives. Both models had the same fixed and random effects structure and dependent variable – VOT measured in ms. As explained below, the models contained all the predictors of interest as well as other predictors that have been systematically reported to affect VOT production as fixed effects (see below for more information on these factors). To ensure that the coefficients of the fixed effects represented main and not simple effects, we used sum coding, where the first level mentioned in parentheses of each predictor was coded as "−1".

The predictor of interest was *Language* (Basque, Spanish). To account for additional variance, we included the following factors: gender, place of articulation (PoA), vocalic context, and the number of syllables in the word. We included *Gender* (Man, Woman), since longer VOTs have been reported in both 'voiced' and 'voiceless' plosives for women (e.g., Robb et al., 2005; Ryalls et al., 1997; Swartz, 1992; but see Herd, 2020) and *PoA* (Coronal, Dorsal, Labial), where coronal PoA refers to /t/ and /d/, dorsal to /k/ and /g/, and labial to /p/ and /b/, because VOT tends to increase the further away the PoA is from the lips, (e.g., Lisker & Abramson, 1964; Nearey & Rochet, 1994; Volaitis & Miller, 1992; for a physiological explanation see Cho & Ladefoged, 1999). *Vowel Frontness* (Nonfront, Front) and *Vowel Height* (Nonhigh, High) were also included because vocalic context following the plosive alters positive VOT production (e.g., Berry & Moyle, 2011; Esposito, 2002; Nearey & Rochet, 1994; Yeni-Komshian et al., 1977). Additionally, we included *Number of Syllables* (1, 2) as a predictor because this can affect VOT (Flege et al., 1998). Finally, to account for possible interference from English (for an overview see Cabrelli, in press), we included the continuous control variable *English score* (z-scores), which was calculated as the mean percentage of each participant's English vocabulary and interview scores ([%vocabulary score+%interview score]/2).

Random effects in the models included maximal structure (Barr, Levy, Scheepers & Tily, 2013): random intercepts for target word and participant, by-target word random slope for *English score*, and by-participant random slopes for *Language*, *PoA*, *Vowel Frontness*, *Vowel Height*, and *Number of Syllables*.

3. Results

Before proceeding to the results of the LMEMs, we present the descriptive statistics related to our research question (Table 3). Participants' VOT production in Basque and Spanish differ more for 'voiced' (Figure 2) than for 'voiceless' plosives (Figure 3). Although the difference was small, 'voiced' plosives were on average produced with longer prevoicing in Spanish than Basque.

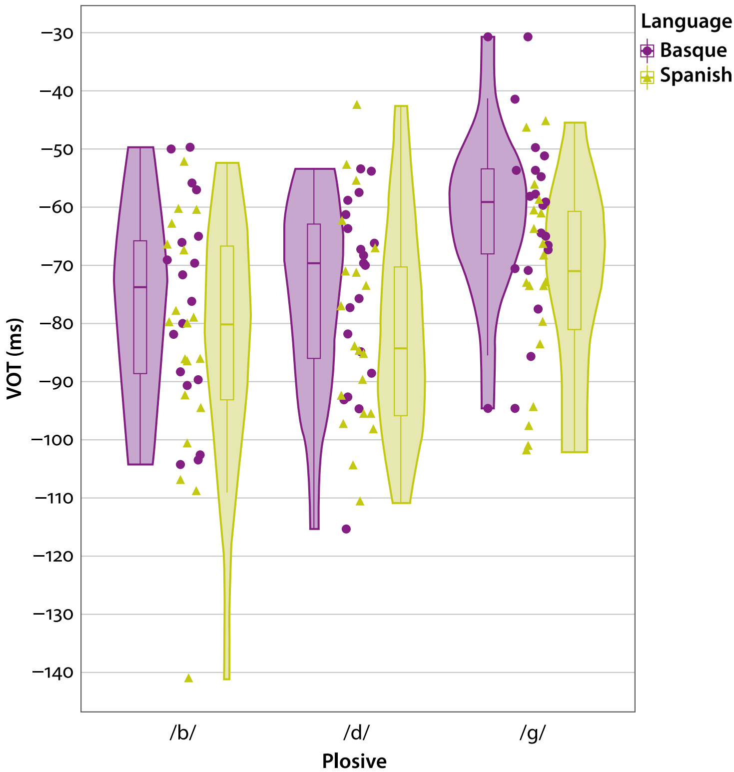

Figure 2. Box and violin plots for VOT production of 'voiced' plosives per language (each dot/triangle represent one participant)

Investigation of individual VOT productions revealed that 16 participants produced longer prevoicing duration in Spanish than in Basque, while the remaining 4 produced longer prevoicing duration in Basque than in Spanish (see https://osf.io/2yxfv).

Statistical analyses for 'voiced' plosives (see Appendix 3 for the model output) showed that bilinguals produced language-specific VOT. The main effect of *Language* showed that 'voiced' plosives were produced with longer prevoicing in

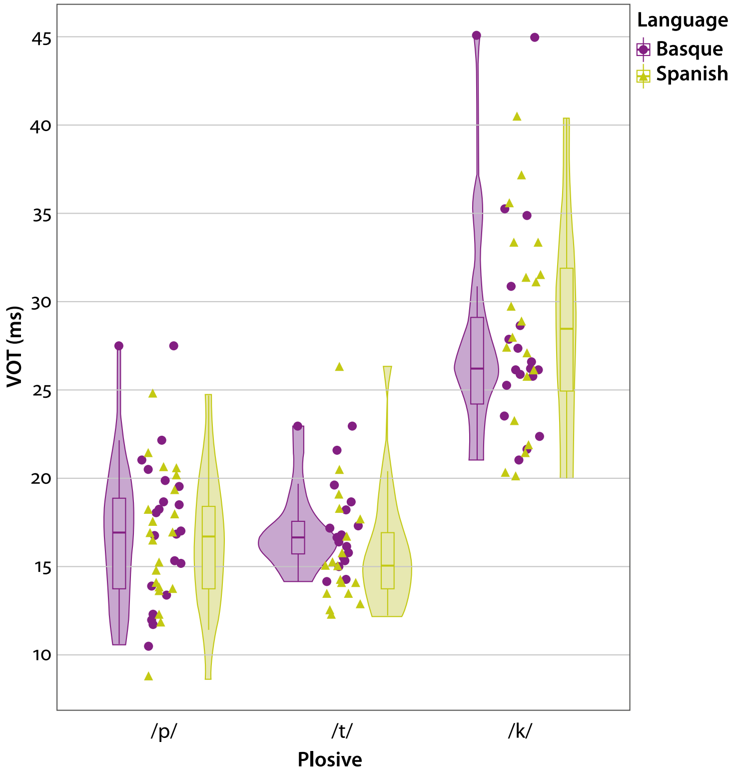

Figure 3. Box and violin plots for VOT production of 'voiceless' plosives per language (each dot/triangle represent one participant)

Table 3. Mean VOT, SD, and number of tokens per plosive and per language

Plosive	Language	Mean	SD	N	Range
/b/	Spanish	−83	32	196	−179 – −19
	Basque	−76	26	186	−154 – −22
/d/	Spanish	−81	34	197	−280 – −16
	Basque	−75	27	183	−151 – −17
/g/	Spanish	−72	30	189	−265 – −18
	Basque	−61	24	189	−153 – −14
/p/	Spanish	16	8	200	3 – 43
	Basque	17	8	198	3 – 52
/t/	Spanish	16	5	200	7 – 37
	Basque	17	6	188	7 – 39
/k/	Spanish	29	9	200	12 – 63
	Basque	28	9	198	11 – 59

Spanish than Basque ($\beta = -5.353$, $SE = 1.670$, $t(36.2) = -3.204$, $p = .003$). No main effects of *Gender, Vowel Height, Number of Syllables* or *English score* were found. Pairwise comparisons revealed a main effect of *PoA*, with shorter prevoicing for dorsals than either coronals ($\beta = 9.964$, $SE = 2.959$, $t(36.5) = 3.367$, $p = .005$) or labials ($\beta = 10.785$, $SE = 3.368$, $t(35.5) = 3.202$, $p = .009$), but not detectably different prevoicing durations for labials and coronals ($\beta = -0.821$, $SE = 3.246$, $t(37.8) = -0.253$, $p = 1$). The main effect of *Vowel Frontness* ($\beta = -3.488$, $SE = 1.569$, $t(43.6) = -2.223$, $p = .031$) showed that plosives followed by a front vowel were produced with shorter prevoicing than plosives followed by a nonfront vowel.

The LMEM for 'voiceless' plosives (see Appendix 4 for the model output) did not detect differences in VOT production between Spanish and Basque, nor any significant effect of *Gender*, *Number of Syllables* or *English score*. Pairwise comparisons showed a main effect of *PoA*, with longer VOT for dorsals than either coronals ($\beta = 12.361$, $SE = 1.333$, $t(52.2) = 9.271$, $p < .001$) or labials ($\beta = 14.668$, $SE = 1.517$, $t(56.4) = 9.666$, $p < .001$) but not detectably different VOT durations for labials and coronals ($\beta = -2.306$, $SE = 1.177$, $t(54.1) = -1.961$, $p = .165$). The main effect of *Vowel Frontness* showed that plosives followed by a front vowel were produced with shorter VOT than plosives followed by a nonfront vowel ($\beta = -1.784$, $SE = 0.505$, $t(57.8) = -3.530$, $p < .001$). Finally, the main effect of *Vowel Height* showed that plosives followed by a high vowel were produced with longer VOT than plosives followed by a nonhigh vowel ($\beta = 2.657$, $SE = 0.575$, $t(55.2) = 4.620$, $p < .001$).

4. Discussion

The goals of this study were to investigate differences in Basque-Spanish early bilingual adults' VOT production and to describe the voicing system of Gipuzkoan Basque. Previous studies that investigated aspirating and true-voicing language pairs found that bilingual speakers keep their languages phonetically separate (Antoniou et al., 2010; Flege & Eefting, 1987b; Fowler et al., 2008; Kupisch & Lleó, 2017; Stoehr et al., 2017; Sundara et al., 2006). The present study revealed that even bilinguals speaking two true-voicing languages produced distinct VOT in each language. In this section we first describe the voicing system of Gipuzkoan Basque and Spanish, and then discuss bilinguals' VOT production. Finally, we consider the granularity of voicing systems in bilingual speakers of two true-voicing languages.

4.1 The voicing system of Gipuzkoan Basque and Spanish

In this study we present the first systematic investigation of the voicing system in Gipuzkoan Basque, the dialect which forms the basis for Standard Basque (Hualde & Ortiz de Urbina, 2003). Table 4 summarizes the by-plosive VOT values we obtained in the present study for Gipuzkoan Basque and Spanish, together with values reported in previous studies for other dialects. We found that VOT for 'voiceless' plosives in Basque and Spanish were not detectably different. Previous research has reported similar VOT for Basque /p/ and /k/ but slightly longer VOT for Basque /t/ compared to the values we observed (Gaminde et al., 2002; Mounole, 2004); and shorter VOT in Spanish for all 'voiceless' plosives than the values we observed (Castañeda Vicente, 1986; Rosner et al., 2000; Table 4). However, all previous studies used word list reading which may have led to hyperarticulation (Lisker & Abramson, 1967). Taking this into account, we hypothesize that previously reported Basque VOT is similar to the measurements we obtained through picture naming in Gipuzkoan Basque. If so, Gipuzkoan Basque and Spanish have numerically similar VOT to unaspirated 'voiceless' plosives in Basque varieties spoken in France, and longer VOT than reported for monolingual Spanish speakers. For 'voiced' plosives, we see the opposite pattern: prevoicing durations in both languages are more similar to those previously reported for Spanish than for Basque. Also, against our expectations, we observed longer prevoicing for all 'voiced' plosives in *Spanish* than Basque. As can be seen in Table 4, previous research has reported the opposite pattern with numerically longer prevoicing in Basque than in Spanish.

Table 4. Mean VOT by-plosive in word-initial position in Basque and Spanish

Language	Study	/p/	/t/	/k/	/b/	/d/	/g/
Basque							
Zuberoan	Mounole (2004)	13	19	27	–	–	–
Zuberoan	Gaminde et al. (2002)	20	24	27	−102	−105	−101
	Mean	16.5	21.5	27	−102	−105	−101
Gipuzkoan	This study	17	17	28	−76	−75	−61
Spanish							
Peninsular	Rosner et al. (2000)	13	14	27	−92	−92	−74
Peninsular	Castañeda Vicente (1986)	6	9	25	−72	−79	−57
	Mean	9.5	11.5	26	−82	−85.5	−65.5
Gipuzkoan	This study	16	16	29	−83	−81	−72

Our results on the effects of place of articulation for Gipuzkoan Basque and Spanish are in line with previous findings: the closer the place of articulation is to the back of the mouth, the longer the (positive) VOT and the shorter the prevoicing duration (e.g., Cho & Ladefoged, 1999; Nearey & Rochet, 1994; Volaitis & Miller, 1992). We found that /k/ had significantly longer VOT than either /t/ or /p/, and /g/ had significantly shorter prevoicing than either /d/ or /b/.

The role of vocalic context was also confirmed by our results. Both 'voiced' and 'voiceless' plosives followed by front vowels /i/ and /e/ had shorter VOT than plosives followed by nonfront vowels /a/, /o/, and /u/. The same tendency for shorter positive VOT for 'voiceless' plosives followed by front vowels compared to nonfront vowels was found in the data of Troya Déniz (2005) for Gran Canarian Spanish. Finally, we found longer positive VOT for 'voiceless' plosives followed by high compared to nonhigh vowels, which is in line with previous studies (Berry & Moyle, 2011; Esposito, 2002; Nearey & Rochet, 1994; Yeni-Komshian et al., 1977).

4.2 VOT production in Basque-Spanish bilinguals

Basque-Spanish bilinguals produced language-specific VOT for 'voiced' plosives in Spanish and Basque. Although the differences were small, we found prevoicing was significantly longer in Spanish than Basque at the group-level. This suggests that bilingual speakers of two true-voicing languages can maintain two separate phonetic categories for 'voiced' plosives, even though those categories belong to the same part of the VOT continuum (i.e., prevoicing). Similar results were recently found by Osborne and Simonet (2021) in speakers of a true-voicing language (Portuguese), who were learners of an aspirating language (English). Those learners distinguished 'voiced' plosives in a nonmonolingual-like manner, through distinct prevoicing durations in the two languages. The differences in prevoicing duration here and in Osborne and Simonet (2021) are small and might not be behaviorally detectable by listeners. However, at the neural level, VOT differences as small as 5 ms show distinct amplitudes of the N1 component during electroencephalography (Toscano et al., 2010). Consequently, such small duration differences appear to be sufficient for bilinguals to form language-specific categories.

Conversely, we did not find support for distinct VOT production of 'voiceless' plosives in Basque and Spanish. This null result suggests that the bilinguals either maintained two separate categories with statistically undetectable VOT differences or had one merged category for 'voiceless' plosives. A merged category could result from a trade-off between the need to separate the two languages and the principle of cognitive economy. The short-lag VOT range is narrower than the prevoicing range (a wide range of negative VOT). It would, therefore, require more precision to produce language-specific plosives in the narrow short-lag range, requiring greater cognitive resources from the articulatory network. For this reason, the most economical way to distinguish Basque and Spanish voicing categories in production would be through language-specific 'voiced' plosives, since the prevoicing part of the VOT continuum is less restricted, providing more space for category formation.

Our results are in line with the SLM-r (Flege & Bohn, 2021). According to the SLM-r, L2 input is crucial: More experience with an L2 helps speakers gradually discern the two closest L1-L2 phonetic categories so as to form a new phonetic category. The early bilinguals tested in the present study had received rich input in both their majority language Spanish and, most importantly, their minority language Basque. Indeed, our results support these hypotheses: the early bilingual speakers tested in the present study produced language-specific VOT for 'voiced' plosives, suggesting they had formed a new phonetic category. Although prevoicing is present in both languages, we argue that early bilinguals living in a bilingual community such as the BAC enjoy such rich language input that they can create and maintain distinct phonetic categories, even if those categories have fine-grained differences. However, future research should assess if even late Basque-Spanish bilinguals show similar language differentiation dependent on the amount of exposure to both languages rather than the age of acquisition.

The SLM-r posits that L2 sounds that are not sufficiently dissimilar to L1 sounds can be assimilated resulting in a composite L1-L2 category. This may have been the case in production of 'voiceless' plosives where we did not find support for differences between the two languages. Both the Basque and Spanish VOT values observed in our study were similar or shorter than the VOT values previously reported for Basque 'voiceless' plosive productions in bilingual Basque-French speakers (Gaminde et al., 2002; Mounole, 2004) but higher than those previously reported for Peninsular Spanish (Castañeda Vicente, 1986; Rosner et al., 2000). These previous studies, however, used different methodologies and sets of stimuli, making direct comparisons with our results impossible. Specifically, previous studies elicited production through word-list reading. Reading can result in hyperarticulation, which may lead to increased VOT durations. This could explain why previous studies reported somewhat longer VOT in Basque and Spanish than the current study, in which production was elicited by a more naturalistic picture naming task.

Importantly, four bilinguals in the current study produced longer prevoicing in *Basque* than Spanish. This is in line with the SLM-r hypothesis of individual differences in L1 categories: Some speakers might use the same acoustic cue (i.e., VOT) in a different way than other speakers of the same language (e.g., Allen, Miller & DeSteno, 2003). Nevertheless, those bilinguals separated their 'voiced' categories in Basque and Spanish like their peers, but exhibited the opposite pattern, with longer prevoicing durations in Basque than in Spanish. This means that some bilinguals might not rely only on statistical distributions of their two languages for category formation as predicted by the SLM-r, especially in the case of Basque where monolingual distributions are not available. Instead, this language differentiation might be created by bilinguals to enhance contrast between their languages.

Additionally, socio-phonetic factors may have contributed to the results of the present study. Although we tried to control for a number of factors known to affect speech production such as gender and dialect, other factors were not controlled (e.g., attitude towards the bilinguals' languages; Law et al., 2021). Basque is a minority language, and a good command of Basque is important for Basque identity (Azurmendi et al., 2008). Therefore, we should not dismiss the possible effects of social leveling, a process whereby speakers diverge from standard phonetic norms in the majority language to indicate group affiliation, in this case diverging from Spanish to signal their linguistic and social Basque identity. This may suggest that early bilinguals produce language-specific 'voiced' plosives, and that the formation of these separate categories could have been augmented due to production differences aimed at social leveling.

Finally, three possible limitations should be considered. First, the uneven distribution of monosyllabic and disyllabic words may have influenced the results even in the absence of a significant effect of the number of syllables in the statistical analyses. Second, the variable levels of English proficiency of our participants should be acknowledged. Nevertheless, since both Basque and Spanish are L1s, we would expect that any possible influence of English, the L3, would equally impact both languages (e.g., Cabrelli, in press), altering the VOT values reported here, but leaving the observed cross-linguistic difference between Spanish and Basque intact. Lastly, differences in speech rate between Spanish and Basque are an important factor as faster speech might cause shorter prevoicing durations in true-voicing languages (e.g., Kessinger & Blumstein, 1997). To check for differences in speech rate, we ran an LMEM on the syllabic rate of the 'voiced' productions (see Appendix 5). Six out of 20 participants showed significant differences in speech rate between Spanish and Basque. Importantly, four out of these six participants had a significantly faster syllabic rate in Spanish than Basque. Faster rate in Spanish would result in shorter prevoicing, which is the opposite of what we observed (i.e., longer prevoicing in Spanish than Basque). We argue that in the absence of speech rate differences, we might have found larger cross-linguistic effects, and we acknowledge that future VOT studies should account for speech rate.

4.3 Voicing granularity in bilinguals

We have seen that bilingual speakers of two true-voicing languages can produce language-specific VOT for 'voiced' plosives. Although VOT is the most salient cue to distinguish 'voiced' from 'voiceless' plosives for true-voicing languages such as Spanish (e.g., Abramson & Lisker, 1973), we argue that bilinguals can attune their voicing systems to produce fine-grained VOT differences in order to effectively maintain separate categories, given: (a) the need for language separation, and (b) that the voicing category through which they separate their languages allows for sufficient variability, rather than a severely restricted phonetic space (i.e., prevoicing contains all negative VOT values, while short-lag contains a restricted range of positive VOT values, usually falling between a range of a few tenths of milliseconds).

A bilingual speaker of two languages with similar voicing systems (i.e., true-voicing languages) might need to separate their two languages, for example, for social reasons. This could lead the bilingual to exploit every phonetic input available in order to construct at least one dissimilar category. This dissimilar category would not necessarily need to be faithful to the prototypical productions of monolinguals in the target language, particularly in the case of Basque, where there are no monolingual norms. We speculate that this might be a product of the bilingual's need for language separation, rather than to achieve monolingual-like productions in both languages. If the two languages have distinct distributions in the same section of the VOT continuum for the same phonological category, the language with which the speaker has greater experience could act as a reference. The new dissimilar category would be then constructed and progressively shaped by experience with both languages as proposed by the SLM-r (Flege & Bohn, 2021). If the two languages do not have distinct typological VOT distributions (i.e., both have a contrast between prevoicing and short-lag), but there is a need for language separation within the bilingual mind, we speculate that productions in the two languages might diverge. This would force the VOT distribution of one language to shift away from the corresponding category in the other language, leading to separate language-specific categories. In order to test this idea, future longitudinal studies should investigate VOT differences between languages and the degree of hyperarticulation via speech rate. Regardless of the existence of distinct productions, the less restricted part of the VOT continuum (i.e., prevoicing) appears to be preferred for the construction of a new category since it allows for more variability. These hypotheses lead us to speculate that bilingual speakers of two aspirating languages will show a similar pattern, with two separate 'voiceless' categories for each language produced with aspiration, as has already been shown for Mandarin-English bilingual children (Yang, 2021). The coarse division of the VOT continuum into three parts (prevoicing, short-lag, and aspiration) may not reflect the fine-grained differences of bilingual speakers of typologically similar languages in term of voicing, and/or for bilinguals, who show considerable interspeaker variation in VOT (e.g., Kupisch & Lleó, 2017). When investigating such populations, a more detailed description of the specific VOT values should be reported to ensure that fine phonetic cross-linguistic differences are not missed.

5. Conclusions

Early bilinguals have to accommodate two voicing systems which can be either dissimilar (e.g., Spanish and English) or similar (e.g., Basque and Spanish). This study provides the first evidence that even early bilingual adults who speak two true-voicing languages, in this case Basque and Spanish, produce language-specific VOT. Whereas previous research found that bilingual speakers of a true-voicing and an aspirating language differentiated 'voiceless', but not 'voiced' plosives in their two languages, we found the opposite pattern for Basque and Spanish: Language-specific VOT for 'voiced' plosives, with different prevoicing durations in Basque and Spanish but no evidence for a difference in 'voiceless' productions. Our results suggest that early bilinguals who are highly proficient in their minority language can dissociate their two languages. Our results detailing bilinguals' language-specific VOT production in two true-voicing languages offer a new perspective for phonetic research in bilingualism, suggesting that bilingual phonetic systems are more nuanced than previously thought. The fact that bilinguals' phonological categories overlap does not automatically imply that their voicing systems have become "lazy". On the contrary, they seem to be engaged in a constant "phonetic struggle" to achieve language differentiation, exploiting even the finest acoustic cues available.

Funding

This research is supported by the Basque Government through the BERC 2022-2025 program and by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation [CEX2020-001010-S]; the Spanish Ministry of Economy and Competitiveness [grant BES-2017-082500, to C.S.]; the Agencia Estatal de Investigación (AEI), the Fondo Europeo de Desarrollo Regional (FEDER), the Spanish Ministry of Science, Innovation and Universities [grant RTI2018-096311-B-I00, to N.M.]; and the European Union's Horizon 2020 research and innovation programme [Marie Skłodowska-Curie grant agreement No. 843533, to A.S.].

Acknowledgements

We would like to thank Ane Zurutuza for her help selecting the stimulus materials, Irene Arrieta, Irati Aperribay, Amaia Cano, David Carcedo, Ainhoa Eguiguren, Manex Lete, Iker Oteiza, and Ane Zurutuza, for administering the present study, and Magda Altman and Caroline Handley for proofreading this article. We are grateful to Efthymia Kapnoula and Clara D. Martin for their insightful comments on earlier versions of this manuscript, and to the members of the Basque Center on Cognition, Brain and Language for valuable feedback on this study.

Data availability statement

The data that support the findings of this study are openly available in OSF at https://osf.io/2yxfv

References

doi Abramson, A. S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, *1*(1), 1–8.

doi Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, *113*(1), 544–552.

doi Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, *38*(4), 640–653.

doi Azurmendi, M. J., Larrañaga, N., & Apalategi, J. (2008). Bilingualism, identity and citizenship in the Basque Country. In M. Niño-Murcia & J. Rothman (Eds.), *Bilingualism and identity* (pp. 35–62). John Benjamins.

doi Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255–278.

Bartoń, K. (2019). MuMIn: Multi-model inference. *R Package Version 1.43.6*.

Basque Government. (2016). *VI. Inkesta Soziolinguistikoa Euskal Autonomia Erkidegoa [6. The Sociolinguistic Survey of the Basque Autonomous Community]*. Vitoria-Gasteiz. Retrieved from http://www.euskara.euskadi.eus/contenidos/nota_prensa/vi_inkesta_sozioling_2016/eu_def/adjuntos/vi_inkesta_eu.pdf

doi Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

doi Berry, J., & Moyle, M. (2011). Covariation among vowel height effects on acoustic measures. *The Journal of the Acoustical Society of America*, *130*(5), EL365–EL371.

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program].

Cabrelli, J. (in press). Language attrition and L3/L*n*. In J. Cabrelli, A. Chaouch Orozco, J. González Alonso, E. Puig Mayenco, S. Pereira Soares, & J. Rothman (Eds.), *The Cambridge handbook of third language acquisition*. Cambridge: Cambridge University Press.

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. *The Journal of the Acoustical Society of America*, *54*(2), 421–428.

Castañeda Vicente, M.L. (1986). El V.O.T. de las Oclusivas Sordas y Sonoras Españolas. *Estudios de Fonética Experimental*, *2*(2), 91–110.

Cenoz, J. (2008). The status of Basque in the Basque country. In M. Barni & G. Extra (Eds.), *Mapping Linguistic Diversity in Multicultural Contexts* (pp. 93–114). Mouton de Gruyter.

Chang, C.B., Yao, Y., Haynes, E.F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America*, *129*(6), 3964–3980.

Cheng, A. (2020). Cross-linguistic fo differences in bilingual speakers of English and Korean. *The Journal of the Acoustical Society of America*, *147*(2), EL67–EL73.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207–229.

de Bruin, A., Carreiras, M., & Duñabeitia, J.A. (2017). The BEST dataset of language proficiency. *Frontiers in Psychology*, *8*.

Duñabeitia, J.A., Crepaldi, D., Meyer, A.S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816.

Egurtzegi, A., & Carignan, C. (2020). An acoustic description of Mixean Basque. *The Journal of the Acoustical Society of America*, *147*(4), 2791–2802.

Esposito, A. (2002). On Vowel Height and Consonantal Voicing Effects: Data from Italian. *Phonetica*, *59*(4), 197–231.

Etxebarria, M. (1987). Iniciación a la fonética acústica. *"Euskal Filología". Cursos de Verano UPV/EHU*, 475–514.

Flege, J.E. (1995) Second-language speech learning: theory, findings, and problem. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J.E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning* (pp. 3–83). Cambridge: Cambridge University Press.

Flege, J.E., & Eefting, W. (1987a). Cross-language switching in stop consonant perception and production by Dutch speakers of english. *Speech Communication*, *6*(3), 185–202.

Flege, J.E., & Eefting, W. (1987b). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*(1), 67–83.

Flege, J.E., Frieda, E.M., Walley, A.C., & Randazza, L.A. (1998). Lexical factors and segmental accuracy in second language speech production. *Studies in Second Language Acquisition*, *20*(2), 155–187.

Flege, J.E., Munro, M.J., & MacKay, I.R.A. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, *97*(5), 3125–3134.

Flege, J.E., Yeni-komshian, G.H., & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, *41*, 78–104.

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. *Journal of Phonetics*, *36*(4), 649–663.

Francis, A.L., Ciocca, V., & Ching Yu, J.M. (2003). Accuracy and variability of acoustic measures of voicing onset. *The Journal of the Acoustical Society of America*, *113*(2), 1025–1032.

Gaminde, I., Hualde, J.I., & Salaberria, J. (2002). Zubereraren herskariak: azterketa akustikoa. *Lapurdum*, (7), 221–236.

Hazan, V.L., & Boulakia, G. (1993). Perception and production of a voicing contrast by french-english bilinguals. *Language and Speech*, *36*(1), 17–38.

Herd, W. (2020). Sociophonetic voice onset time variation in Mississippi English. *The Journal of the Acoustical Society of America*, *147*(1), 596–605.

Hopp, H., & Schmid, M.S. (2013). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, *34*(2), 361–394.

Hualde, J.I. (1991). *Basque Phonology*. Routledge.

Hualde, J.I. (2015). Basque as an extinct language. In B. Fernández & P. Salaburu (Eds.), *Ibon Sarasola, gorazarre: homenatge, homenaje* (pp. 319–326). University of the Basque Country.

Hualde, J.I. (2018). Aspiration in Basque. *Papers in Historical Phonology*, *3*, 1–27.

Hualde, J.I., & Ortiz de Urbina, J. (2003). *A Grammar of Basque*. Mouton de Gruyter.

Hurch, B. (1988). Is Basque a syllable-timed language?. *International Journal of Basque Linguistics and Philology*, *22*(1), 813–825.

Kang, K.-H., & Guion, S.G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *The Journal of the Acoustical Society of America*, *119*(3), 1672–1683.

Kessinger, R.H., & Blumstein, S.E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of phonetics*, *25*(2), 143–168.

Kroll, J.F., Bobb, S.C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism*, *9*(2), 119–135.

Kupisch, T., Barton, D., Hailer, K., Klaschik, E., Stangen, I., Lein, T., & van de Weijer, J. (2014). Foreign Accent in Adult Simultaneous Bilinguals. *Heritage Language Journal*, *11*(2), 123–150.

Kupisch, T., Lein, T., Barton, D., Schröder, D.J., Stangen, I., & Stoehr, A. (2014). Acquisition outcomes across domains in adult simultaneous bilinguals with French as weaker and stronger language. *Journal of French Language Studies*, *24*(3), 347–376.

Kupisch, T., & Lleó, C. (2017). Voice onset time in German-Italian simultaneous bilinguals: Evidence on crosslanguage influence and markedness. In M.S. Yavaş, M. Kehoe & W. Cardoso (Eds.), *Romance-Germanic Bilingual Phonology* (pp. 79–98). Equinox.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(*13*).

doi Law, W. L., Dmitrieva, O., & Francis, A. L. (2021). Language attitudes modulate phonetic interactions between languages in bilingual speakers in diglossic settings. *Linguistic Approaches to Bilingualism*, *11*(3), 289–322.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). Package ' emmeans.' *CRAN Repository*.

doi Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD*, *20*(3), 384–422.

doi Lisker, L., & Abramson, A. S. (1967). Some Effects of Context On Voice Onset Time in English Stops. *Language and Speech*, *10*(1), 1–28.

doi Llama, R., & Cardoso, W. (2018). Revisiting (Non-)Native Influence in VOT Production: Insights from Advanced L3 Spanish. *Languages*, *3*(3), 30.

doi Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). "performance: An R Package for Assessment, Comparison and Testing of Statistical Models." *Journal of Open Source Software*, *6*(60), 3139.

doi Maddieson, I. (1984). The size and structure of phonological inventories. In I. Maddieson (Ed.), *Patterns of Sounds* (pp. 5–24). Cambridge University Press.

Martínez Adrián, M., Gallardo del Puerto, F., & Gutiérrez Mangado, J. (2013). Phonetic and syntactic transfer effects in the English interlanguage of Basque/Spanish bilinguals. *Vigo International Journal of Applied Linguistics*, *10*(1), 51–83.

doi Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324.

doi Mayr, R., & Montanari, S. (2014). Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast. *Journal of Child Language*, *42*(5), 1006–1035.

doi Mayr, R., & Siddika, A. (2018). Inter-generational transmission in a minority language setting: Stop consonant production by Bangladeshi heritage children and adults. *International Journal of Bilingualism*, *22*(3), 255–284.

doi Moreno, A., & Mariño, J. B. (1998). Spanish dialects: phonetic transcription. *ICSLP 1998. Proceedings of the 5th International Conference on Spoken Language Processing*, (Icslp 98), 5–8. Retrieved from http://www.shlrc.mq.edu.au/proceedings/icslp98/PDF/AUTHOR/SL980598.PDF.

Mounole, C. (2004). Zubererazko herskarien azterketa akustikoa. *International Journal of Basque Linguistics and Philology*, *38*(1), 207–248.

doi Nash, J. C. (2014). On Best Practice Optimization Methods in R. *Journal of Statistical Software*, *60*(2), 1–14.

doi Nearey, T. M., & Rochet, B. L. (1994). Effects of Place of Articulation and Vowel Context on VOT Production and Perception for French and English Stops. *Journal of the International Phonetic Association*, *24*(1), 1–18.

doi Oh, J. S., Jun, S. A., Knightly, L. M., & Au, T. K. F. (2003). Holding on to childhood language memory. *Cognition*, *86*(3).

doi Osborne, D. M., & Simonet, M. (2021). Foreign-language phonetic development leads to first-language phonetic drift: plosive consonants in native Portuguese speakers learning English as a foreign language in Brazil. *Languages*, *6*(3), 112.

doi Qi, R., Tsukada, K., & Hua, Z. (2012). An Acoustic Phonetic Analysis of Word-Initial Stop Production by Young Simultaneous Mandarin-English Bilingual Children. *Asia Pacific Journal of Speech, Language and Hearing*, *15*(1), 73–84.

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from https://www.R-project.org/

Robb, M., Gilbert, H., & Lerman, J. (2005). Influence of gender and environmental setting on voice onset time. *Folia Phoniatrica et Logopaedica*, *57*(3), 125–133.

Rosner, B. S., López-Bascuas, L. E., García-Albea, J. E., & Fahey, R. P. (2000). Voice-onset times for Castilian Spanish initial stops. *Journal of Phonetics*, *28*(2), 217–224.

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. Boston, MA. Retrieved from http://www.rstudio.com/

Ryalls, J., Zipprer, A., & Baldauff, P. (1997). A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research*, *40*(3), 642–645.

Schuhmann, K. S., & Huffman, M. K. (2019). Development of L2 Spanish VOT before and after a brief pronunciation training session. *Journal of Second Language Pronunciation*, *5*(3), 402–434.

Stewart, M. (1999). *The Spanish language today*. Routledge.

Stoehr, A., Benders, T., Van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch–German late bilinguals. *Second Language Research*, *33*(4), 483–518.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism*, *9*(1), 97–114.

Swartz, B. L. (1992). Gender Difference in Voice Onset Time. *Perceptual and Motor Skills*, *75*(3).

Torreira, F. (2012). Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association*, *42*(1), 49–63.

Troya Déniz, M. (2005). El VOT de las oclusivas sordas en la norma culta de Las Palmas de Gran Canaria. *Boletín de Lingüística de La Universidad Central de Venezuela*, *24*, 92–107.

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization: Electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological science*, *21*(10), 1532–1540.

Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *The Journal of the Acoustical Society of America*, *92*(2), 723–735.

Watson, I. (2007). Simultaneous bilinguals and Flege's Speech Learning Model. In *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrucken*, 6–10 August, 2007, 1533–1536.

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2).

Williams, L. (1977a). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics*, *21*(4), 289–297.

Williams, L. (1977b). The voicing contrast in Spanish. *Journal of Phonetics*, *5*(2), 169–184.

Yang, J. (2021). Comparison of VOTs in Mandarin–English bilingual children and corresponding monolingual children and adults. *Second Language Research*, *37*(1), 3–26.

Yeni-Komshian, G. H., Caramazza, A., & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics*, *5*(1), 35–48.

Appendix 1. Number of monosyllabic and disyllabic stimuli per plosive and language

	Basque		Spanish	
	Monosyllabic	Disyllabic	Monosyllabic	Disyllabic
/b/	6	4	0	10
/d/	1	9	1	9
/g/	2	8	0	10
/p/	1	9	0	10
/t/	1	9	0	10
/k/	2	8	0	10

Appendix 2. Items used in the PNT

	Basque	Spanish		Basque	Spanish
/b/	bat 'one'	barba 'beard'	/p/	pila 'battery'	pico 'beak'
	begi 'eye'	barca 'boat'		pintxo 'snack'	pina 'pinneapple'
	behi 'cow'	bata 'robe'		pipa 'pipe'	pino 'pine'
	beltz 'black'	bici 'bike'		pisu 'flat'	pinza 'clothespin'
	berde 'green'	boca 'mouth'		porru 'leak'	pipas 'sunflower seeds'
	beso 'arm'	bolsa 'bag'		pultsu 'pulse'	pollo 'chicken'
	bi 'two'	bolso 'purse'		puma 'puma'	pua 'guitar pick'
	bide 'road'	bosque 'forest'		punta 'point'	pulpo 'octopus'
	bihotz 'heart'	buho 'owl'		putre 'vulture'	puno 'fist'
	bost 'five'	burro 'donkey'		putz 'blow'	puzle 'puzzle'
/d/	dantza 'dance'	dado 'die'	/t/	talde 'group'	tanque 'tank'
	data 'date'	danza 'dance'		tanta 'drop	tarro 'jar'
	dei 'call'	dardo 'dart'		tanto 'point'	taza 'cup'
	denda 'shop'	dedo 'finger'		tarta 'cake'	techo 'ceiling'
	dike 'dike'	dia 'day'		te 'tea'	teja 'tile'
	diru 'money'	doble 'double'		teila 'tile'	timbre 'bell'
	done 'saint'	doce 'twelve'		tigre 'tiger'	tiza 'chalk'
	donuts 'donut'	dos 'two'		tinta 'paint'	toro 'bull'
	dorre 'tower'	dulces 'sweets'		tunel 'tunnel'	tubo 'tube'
	dutxa 'shower'	duro 'hard'		tutu 'pipe'	tumba 'grave'

Appendix 2. *(continued)*

	Basque	Spanish		Basque	Spanish
/g/	galtza 'pants'	gafas 'glasses'	/k/	kanpo 'out'	cabra 'goat'
	ganba 'shrimp'	gallo 'rooster'		karga 'weight'	calvo 'bald'
	garbi 'clean'	garra 'claw'		ke 'smoke'	cama 'bed'
	gatz 'salt'	gato 'cat'		keinu 'sign'	carcel 'prison'
	gazta 'cheese'	golf 'golf'		ken 'minus'	carne 'meat'
	gehi 'plus'	golpe 'hit'		koko 'coconut'	carro 'cart'
	gela 'room'	gordo 'fat'		kolpe 'hit'	carta 'letter'
	gezi 'arrow'	gorra 'hat'		kontu 'receipt'	casa 'house'
	giltza 'key'	gota 'drop'		kopa 'glass'	codo 'elbow'
	gorri 'red'	guino 'wink'		kutxa 'box'	culo 'bottom'

Appendix 3. Output of the LMEM for 'voiced' plosives

	VOT				
Predictor	**Estimate**	***SE***	**CI (95%)**	***t***	***p***
(Intercept)	−76.610	3.459	[−83.397, −69.822]	−22.146	**<.001**
Language [Spanish]	5.353	1.670	[2.075, 8.630]	3.204	**.001**
Gender [Woman]	4.465	2.662	[−0.757, 9.687]	1.678	.094
Place [Labial]	−3.869	1.972	[−7.738, 0.001]	−1.962	.050
Place [Dorsal]	6.916	1.815	[3.354, 10.478]	3.810	**<.001**
Number of Syllables [1]	−0.067	1.519	[−3.048, 2.913]	−0.044	.965
Height [High]	−0.896	1.646	[−4.126, 2.333]	−0.545	.586
Frontness [Front]	−3.488	1.569	[−6.567, −0.409]	−2.223	**.026**
English Score	2.863	2.662	[−2.360, 8.086]	1.076	.282
Marginal R^2 / Conditional R^2	.087 / .398				

Appendix 3. *(continued)*

Random Effects	
σ^2	566.24
τ00 Word	41.20
τ00 Participant	183.77
τ11 Word.scale(EnglishScore)	0.03
τ11 Participant.Language	22.92
τ11 Participant.Place.labial	25.80
τ11 Participant.Place.dorsal	14.95
τ11 Participant.Height	14.87
τ11 Participant.Frontness	7.55
τ11 Participant.NumSyllables	0.68
ρ01 Word	1.00
ρ01 Participant.Language	−0.52
ρ01 Participant.Place.sum.labial	0.32
ρ01 Participant.Place.dorsal	−0.43
ρ01 Participant.Height	−0.53
ρ01 Participant.Frontness	0.47
ρ01 Participant.NumSyllables	0.67

Appendix 4. Output of the LMEM for 'voiceless' plosives

	VOT				
Predictor	**Estimate**	***SE***	**CI (95%)**	***t***	***p***
(Intercept)	21.086	1.022	[19.081, 23.092]	20.630	**<.001**
Language [Spanish]	0.069	0.465	[−0.843, 0.981]	0.148	.882
Gender [Woman]	−0.205	0.544	[−1.272, 0.863]	−0.376	.707
Place [Labial]	−5.658	0.789	[−7.206, −4.111]	−7.175	**<.001**
Place [Dorsal]	9.010	0.868	[7.307, 10.712]	10.384	**<.001**
Number of Syllables [1]	1.331	0.896	[−0.426, 3.088]	1.486	.137
Height [High]	2.657	0.575	[1.528, 3.785]	4.620	**<.001**
Frontness [Front]	−1.784	0.505	[−2.775, −0.792]	−3.530	**<.001**
English Score	−0.078	0.563	[−1.183, 1.027]	−0.139	.889
Marginal R^2 / Conditional R^2	.400 / .670				

Appendix 4. *(continued)*

Random Effects	
σ^2	29.73
τ00 Word	8.12
τ00 Participant	7.05
τ11 Word.scale(EnglishScore)	0.39
τ11 Participant.Language	0.89
τ11 Participant.Place.labial	1.80
τ11 Participant.Place.dorsal	5.34
τ11 Participant.Height	0.73
τ11 Participant.Frontness	0.80
τ11 Participant.NumSyllables	0.93
ρ01 Word	−0.32
ρ01 Participant.Language	−0.05
ρ01 Participant.Place.sum.labial	−0.04
ρ01 Participant.Place.dorsal	0.46
ρ01 Participant.Height	0.18
ρ01 Participant.Frontness	−0.74
ρ01 Participant.NumSyllables	−0.31

Appendix 5. Bonferroni-corrected pairwise comparisons of the contrast Basque – Spanish syllabic rate (syllables per second), per participant

LMEM: Syllables per second ~ Participant * Language+(1|Word)					
Participant	**Basque-Spanish estimate**	**SE**	**df**	**t**	**p**
f1	0.0719	0.2906	140.0690	0.2473	.8050
f2	−0.3522	0.2911	141.5210	−1.2100	.2283
f3	−0.7292	0.2906	140.0690	−2.5096	**.0132**
f4	−0.2436	0.2906	140.0690	−0.8385	.4032
f5	0.7759	0.2906	140.0690	2.6704	**.0085**
f6	−0.1928	0.2911	141.5210	−0.6623	.5089
f7	−0.1490	0.2906	140.0690	−0.5128	.6089
f8	0.0145	0.2906	140.0690	0.0500	.9602
f9	−0.5515	0.2915	141.8772	−1.8918	.0606
f10	0.2957	0.2926	143.7967	1.0107	.3139

Appendix 5. *(continued)*

LMEM: Syllables per second ~ Participant * Language+(1|Word)					
Participant	**Basque-Spanish estimate**	**SE**	**df**	**t**	**p**
m1	−0.2791	0.2926	143.7967	−0.9538	.3418
m2	−0.6555	0.2906	140.0690	−2.2561	**.0256**
m3	0.6226	0.2906	140.0690	2.1428	**.0339**
m4	−1.1390	0.2906	140.0690	−3.9200	**.0001**
m5	−0.6935	0.2915	141.8772	−2.3788	**.0187**
m6	−0.5472	0.2915	141.8772	−1.8768	.0626
m7	0.0278	0.2906	140.0690	0.0957	.9239
m8	0.0309	0.2906	140.0690	0.1064	.9154
m9	−0.4089	0.2906	140.0690	−1.4073	.1616
m10	−0.5343	0.2915	141.8521	−1.8329	.0689

Address for correspondence

Christoforos Souganidis
Basque Center on Cognition, Brain and Language
Paseo Mikeletegi 69
Donostia-San Sebastián 20009
Spain

c.souganidis@bcbl.eu
https://orcid.org/0000-0002-5013-1442

Co-author information

Nicola Molinaro
Basque Center on Cognition, Brain and Language

n.molinaro@bcbl.eu
https://orcid.org/0000-0002-7549-6042

Antje Stoehr
Basque Center on Cognition, Brain and Language

a.stoehr@bcbl.eu
https://orcid.org/0000-0002-6065-5952

Publication history

Date received: 26 November 2021
Date accepted: 31 July 2022
Published online: 31 October 2022