ASJU, 2023, 57 (1-2), 121-140
https://doi.org/10.1387/asju.25918

ISSN 0582-6152 – eISSN 2444-2992

A classroom study on the perception and production of Basque sibilant fricatives by native speakers of English[1]

Ander Beristain*
Saint Louis University

ABSTRACT: This study examines the production and perception of the three Basque sibilant fricatives by L1 English learners in a semester-long university course. Basque has three sibilant fricatives: lamino-alveolar (/s̻/), apico-alveolar (/s̺/), and pre-palatal (/ʃ/) sounds. Acoustically, the first two sounds share similarities with English /s/ and /ʃ/, respectively. L2 phonology theories posit that learners will map the production and perception of L2 sounds to (similar) sounds in their native language. Afterwards, they will create new categories based on the phonetic proximity of the new sounds and those in their native language. Ten students enrolled in a *Beginners' Basque* course and two native speakers of Basque serving as control group undertook a read-aloud task (Experiment 1) followed by an ABX discrimination task (Experiment 2) during Week 5 of classes. Center of Gravity and accuracy scores were extracted. Results show that learners tended to merge the two apical sibilants into an English /s/-like sound. Accuracy scores between the two alveolar sibilants in Basque were significantly lower than the other sound contrasts.

KEYWORDS: Basque sibilants; L1-English; production; perception; classroom acquisition.

[1] I want to thank José Ignacio Hualde for being an incredible mentor over the six years I spent at the University of Illinois Urbana-Champaign completing my MA and PhD. I only hope I can pay him back by becoming as a respected and respectful, enthusiastic, and encouraging mentor to my future students. I also want to thank Ane Icardo Isasa for her help building the perception experiment on PsychoPy, Aurore Mroz for her feedback on earlier versions of the manuscript, and my Basque students for agreeing to participate in this study. All errors remain my own.

*** *Corresponding author:*** Ander Beristain. Department of Languages, Literatures and Cultures - Morrissey Hall 1501 (Saint Louis University). 3700 Lindell Blvd (63108 St. Louis, MO, USA). – ander.beristain@slu.edu – https://orcid.org/0000-0001-9083-3871

How to cite: Beristain, Ander (2023). «A classroom study on the perception and production of Basque sibilant fricatives by native speakers of English», *ASJU*, 57 (1-2), 121-140. (https://doi.org/10.1387/asju.25918).

Received: 2022-07-22; Accepted: 2022-10-23.

ISSN 0582-6152 - eISSN 2444-2992 / © UPV/EHU Press

1. Literature review

1.1. Theoretical framework

This study explores how the first language (L1) of learners might influence the production of their foreign language (FL).[2] Most specifically, the two languages under study are (American) English and Basque, respectively. Whether a critical period for the acquisition of language exists or not remains a controversial topic in the field of linguistics, and according to MacSwan and Pray (2005), time is relevant in order to acquire language properly. Research has shown that each language aspect possesses a different critical period, with phonology having the shortest one (Pallier *et al.* 1997).

Models that attempt to explain the acquisition of L2 phonology are present in research. Flege's (1995, 2007) *Speech Learning Model* (SLM) explains that the sound inventory of the L1 will influence the production of the L2, and that phonetic proximity and distance will play a determining role in the production of L2 sounds. If a given L2 target sound is, phonetically speaking, close to that of the L1, overlapping will occur. In that case, L2 learners would employ the L1 closest sound.

As far as the perception of L2 sounds is concerned, Best's (1995) *Perceptual Assimilation Model* (PAM) considers three possible eventualities for the discrimination process. A given sound might be classified as a "CATEGORIZED EXEMPLAR" when the sound in the L1 and L2 are equally categorized, as an "UNCATEGORIZED EXEMPLAR" when the L2 sound to be categorized falls in between two L1 sounds, and as a "NON-ASSIMILABLE NON-SPEECH SOUND" where the L2 sound does not closely belong in the L1 and is therefore non-assimilable (Best *et al.* 2001: 777).

In this study's case, (i) regarding production, it is expected that in initial stages of the acquisition of FL Basque phonology, /s̺/ and /s̻/ would not be produced as such. Indeed, they would be produced as [ʃ] and [s], respectively; and (ii) regarding perception, there are two different discrimination possibilities: for the case of the lamino-alveolar sound /s̻/, its acoustic and articulatory similarity to English /s/ will make it an example of a "CATEGORIZED EXEMPLAR", that is, L1 speakers of English would find Basque /s̻/ perceptually assimilated to English /s/. With regards to the apico-alveolar sound /s̺/, since acoustically speaking this sound falls between English /s/ and /ʃ/, it would pose an example of an "UNCATEGORIZED EXEMPLAR". In other words, the discrimination of this sound is hypothesized to be harder than the rest of the sounds under study. Indeed, although previous research has claimed that perceptually /s̺/ is more similar to the pre-palatal /ʃ/, acoustically speaking, it is located between two phonetic categories of English-speakers' L1, which will complicate mapping this sound in their FL Basque.

[2] In this chapter, I will refer to Basque as a 'foreign language' (FL) and not 'second language' (L2). Klein (1986: 19) points out that "'foreign language' is used to denote a language acquired in a milieu where it is normally not in use […] A 'second language', on the other hand, is one that becomes another tool of communication alongside the first language; it is typically acquired in a social environment in which it is actually spoken." Considering the linguistic context under the current study happens, we will then refer to Basque as an FL.

1.2. Previous studies on Basque sibilant production and perception

Research has primarily focused on the production and analysis of Basque sibilants by native speakers of Basque and Spanish (Beristain 2018, 2022; Hualde 2010; Iribar & Isasi 2008; Iribar *et al.* 2005; Jurado 2011; Larraza 2015). Nevertheless, the production of Basque sibilant fricatives by speakers whose L1 is not Spanish has not been studied so exhaustively.

The production of sibilants has been observed by analyzing several variables, such as the Center of Gravity (henceforth, CoG). CoG is the mean value of all frequencies found in a segment, and it has been found to be correlated with place of articulation of fricative sounds, that is, the more fronted the place of articulation, the higher the CoG (Jongman *et al.* 2000). For example, in the study that Jongman *et al.* (2000) conducted, they found that English /s/ had a CoG of 6133 Hz while that of /ʃ/ was 4229 Hz. Beristain (2022) and Hualde (2010) also found such correlation with Goizueta Basque sibilant fricatives, the CoG for /s̺/ was, respectively, 5117 Hz and 4173 Hz, that for /s̻/ was 7914 Hz and 6645 Hz, and Hualde (2010) found that it was 3531 Hz for /ʃ/. It is worth mentioning that Goizueta Basque is a conservative Basque variety where no sibilant merger has occurred (see Beristain 2022 and Hualde 2010 for a review of Basque sibilant mergers).

As far as L2 Basque sibilants are concerned, Larraza (2015) compared the perception and production of the Basque sibilants /s̺/, /s̻/, /ts̺/ and /ts̻/ between early Spanish-Basque bilinguals and late bilinguals who learned Basque after puberty. According to her results, late bilinguals could not produce the voiceless lamino-alveolar affricate sound [ts̻]. Their realizations varied in a continuum of the respectively apico-alveolar and pre-palatal affricates [ts̺] and [tʃ], which is usually regarded as "Spanish-accented" Basque.

As far as the perception of sibilants in world's languages is concerned, Lisker (2001) explains that for L2 Polish the contrast between the alveopalatal /ɕ/ and retroflex /ʂ/ sibilant fricatives is difficult for L1-English speakers because they map those two sounds into their native category /ʃ/. That contrast is also found in Chinese, and has been described to be complicated for English listeners (McGuire 2007). Furthermore, it has been shown that the vocalic context in which sibilants appear plays a significant role in their discrimination, since listeners do not solely rely on the frication noise produced by the sibilants (Bladon, Clark, & Mickey 1987). In fact, Fujisaki & Kunisaki (1978), and previously Kunisaki & Fujisaki (1977), studied the perception of [s-ʃ] in Japanese in CV and VCV contexts. They also controlled the vocalic context and found that the feature [+round] in the vowel /u/ was a significant factor in the way the consonant was discriminated. They explained that anticipatory rounding would result in a lowering of the frequencies of the consonant, thus complicating its correct discrimination and deducing "that the influence of context in speech production is corrected in speech perception" (Kunisaki & Fujisaki 1977: 91). In other words, when an /s/ was in close contact to /u/, it could be perceived as /ʃ/. This finding had previously been pointed out by Carney and Moll (1971), who were studying the spectral characteristics of English fricatives. Other studies have opted for noise frequency manipulation between different sibilant sounds. Mann & Repp (1980) and Whalen (1981) also found a decrease in frequencies in /u/ contexts in phoneme

https://doi.org/10.1387/asju.25918

boundaries. Because of the lowering of frequencies in /u/ contexts, it could be expected that an apico-alveolar sound like /s̺/ could sound more like a post-alveolar or pre-palatal sound (i.e., /ʃ/).

Finally, Larraza *et al.* (2017) tested the perception of alveolar sibilants in Basque by different types of Basque speakers: (i) Standard Basque speakers who make the three-way sibilant distinction, versus (ii) Western Basque speakers (a *seseo* variety, i.e., merging in favor of [s̺]). Westerners showed lower accuracy scores and higher response times (Larraza *et al.* 2017: 96). This aligns with the theory since Westerners do not have the two sound categories in their phonetic inventories, but speakers of Standard Basque do. In this study, FL-Basque learners should show lower accuracy responses and higher response times than L1-Basque speakers who have the three sibilants in the L1.

The present study is divided into two main experiments: (i) the production experiment, and (ii) the perception experiment. Each section includes its methodology and results section. After that, a general discussion and conclusion is provided.

2. Methodology

2.1. Production experiment

2.1.1. Research questions and predictions

An important prediction that Flege's (1995) model makes is that when L2 speakers encounter new sound categories they will accommodate to the closest one(s) in the L1. Based on these findings, the present study intends to answer the following question:

— *Research question 1:* Can L1-English-speaking students acquire the production of the three sibilant fricatives of Standard Basque in a semester-long beginners' course?

Since both the apico-alveolar and lamino-alveolar sibilants of Basque are new sounds for learners, and more so, they are similar to English /ʃ/ and /s/, respectively, a "matched" production is not expected. Two different predictions are made in this regard:

- Prediction 1: Since the two alveolar sounds are new sound categories, participants might not be able to produce them and will map them onto their native language's closest alveolar sound, that is, [s].
- Prediction 2: If students are able to notice that the apico-alveolar sound is produced differently, another difficulty will be added: because of acoustic and articulatory similarities, it is expected that students should produce the apico-alveolar sound as the English pre-palatal sibilant [ʃ].

2.1.2. Methods

2.1.2.1. Sampling

A total of 12 participants took part in this study (N = 12). These participants were divided into two groups: L1-Basque speakers (n = 2: 2F; age mean: 26; age-range: 25-27), and FL-Basque learners (n = 10 students: 1M, 9F; age mean: 21.2; age-range: 19-26). Native speakers of a traditional variety of Basque with a three-way place contrast in sibilants served as control group. Initially, 14 students were recruited. However, four FL-learners were excluded: two for dropping the class early, and two for missing more than 10 classes (out of 30) in the semester.

FL-Basque learners belonged to a group of students who were registered in a Beginners' Basque course at a Midwest university. The course was an introductory-level class about Basque language and culture. Language classes were conducted entirely in Standard Basque, whereas cultural classes (5/30 = 17%) were conducted in English. The class consisted in bi-weekly 1h 20min sessions. The instructor of the course was a native speaker of Basque from Gipuzkoa (central Basque Country) who can produce the three-way place contrast.

Prior to the experiment, students filled out a linguistic background questionnaire where they explained what languages they speak and what kind of linguistic instruction they had had in the past. None of them spoke any other languages that included the alveolar sibilants of Basque. Taking part in the tasks that comprise the experiment was part of the final grade of students. They were presented as "oral activities" where students would have the chance to further practice the articulation of Basque sounds and develop the perception of sounds in Basque. Participants did not receive any monetary compensation from participating in this study. Participants signed up a consent form where they granted the researcher access to their data.[3] The decision was made to choose Week 5 to collect the data because this was a beginners' course and the first weeks of class were dedicated to the learning of basic vocabulary and grammar concepts, as well as the orthographic system of Basque and the letter-to-sound correspondence.

2.1.2.2. Data collection

The read-aloud task occurred before the perceptual task to avoid priming effects on production. Furthermore, it was considered that a read-aloud was the most optimal task for low level proficiency students. Data collection took place in a sound-attenuated booth at a Midwest university. An AKG C520 head-mounted microphone was used, which was positioned approximately 2 inches (5 cm) away from the participants' lips. A Marantz PMD570 solid state recorder with a sampling rate of 44.1 kHz and a sample size of 16 bits were used to record participants.

Each speaker produced a total number 54 filler items and 54 target items per session (3 phonemes (/s̺, s̻, ʃ/) × 3 vowels (/a, i, u/) × 3 words per category × 2 repetitions). Words were monosyllabic or disyllabic. Thus, within the control group, 54 target tokens × 2 participants = 108 tokens were collected. In the case of learners, 54 target tokens × 10 participants × 2 sessions = 540 tokens were collected. Words were produced within the carrier-phrase *Nik TARGET irakurri dut* 'I read TARGET' in Basque (see Appendix A for Basque word list.)

[3] IRB Protocol Number: 19445, University of Illinois Urbana-Champaign.

FL-Basque learners also conducted a similar type of task in English 10 weeks after the completion of the Basque task to serve as a control experiment for cross-linguistic comparison. Since English does not have the same corner vowels /a, i, u/ as Basque does, following Todd *et al.* (2011), English vowels producing similar coarticulatory effects were grouped together. Thus, words containing /ɔ/, /ɑ/, /ʌ/, and /æ/ were in the same category (resembling Basque /a/); /i/ and /ɪ/ were in the same category (resembling Basque /i/); and /u/ and /ʊ/ were in the same category (resembling Basque /u/). The list of words participants elicited was a modified list adapted from Todd *et al.* (2011). To keep the production of the sibilants as similar as possible between the two languages, monosyllabic or disyllabic words were targeted. In total, 3 SIBILANTS × 3 VOWELS × 3 WORDS per category × 2 REPETITIONS = 54 target tokens were produced per speaker. Since there were 10 participants, and one session took place in English, a total of 540 target tokens were collected for English (see APPENDIX B for English word list.)

2.1.2.3. Data analysis

The realization of the CoG in each phoneme was analyzed in *Praat* (Boersma & Weenink 2019). The first step was to annotate the target items. The first tier was used to segment the word. The second tier was used to segment the vowel. Last, the third tier was used to segment the sibilant. Sibilant segmentation was conducted by checking the F0 contour. Since sibilant sound frequencies are not periodic waves, they do not show any F0. However, for English sibilants, since /z/ is a voiced phoneme, and the vibration of the vocal folds produces a periodic energy wave added to the aperiodic fricative wave, the same methodology could not be employed. For these cases, the spectrogram and sound waves were checked. Then, using a script specifically designed for this project, CoG values were obtained.

Using the statistical programs *R* (R Core Team 2019) and *RStudio* (RStudio Team 2019), linear mixed-effect models were run under the *lme4* and *lmerTest* packages (Bates *et al.* 2015; Kuznetsova 2017 *et al.*, respectively) to attest of the statistical significance in CoG. *P*-values were obtained using Satterthwaite estimations. The decision was made to use linear mixed-effect models because several fixed factors as well as random factors were present while dealing with a continuous dependent variable.

For optimal model selection, *Akaike Information Criterion* (AIC) scores[4] were used in the *MuMIn* package (Burnham & Anderson 2002). The model with the lowest AIC score was chosen as the optimal one. Initially, for each dependent variable, models up to three fixed factors (SOUND (levels: <s>, <z>, <x>, /s/, /z/, /ʃ/), VOWEL (levels: /a, i, u/), and SESSION (levels: Session 1, English, OnlyBasque (control group)) and three random factors (SPEAKER, WORD, and REPETITION) were considered. Note that the fixed factor GROUP was not included as a predictor, but it was included within SESSION type.[5]

[4] See Akaike (1974).

[5] Statistical model formula: lmer(COG ~ Sound*Vowel*Session+ (1|Speaker)+(1|Word))

As can be observed, the model for the CoG includes an interaction between the fixed factors Sound × Vowel × Session and Speaker and Word as random factors. The statistical significance level was established at $\alpha = 0.05$. Notice that the L1-Basque data is included within the fixed factor Session, where learners' session will be compared to that of native speakers'.[6]

2.1.3. Results

For clarity and simplicity purposes, the orthographical representation of Basque fricatives will be used to refer to the sibilant fricatives in Basque throughout the description of the results: Basque <*s*> corresponds to the apico-alveolar phoneme /s̺/, <*z*> corresponds to the lamino-alveolar sound /s̻/, and <*x*> corresponds to the pre-palatal sound /ʃ/. However, when referencing English sibilant fricatives their phonetic representations will be used, that is, */s/, /z/,* and */ʃ/*.

RQ 1:

As far as the Center of Gravity (CoG) of sibilants is concerned, L1-Basque speakers produced three clearly distinguished sibilants (Table 1). As in previous literature, for native Basque speakers the lowest CoG belonged to the pre-palatal sound <x>, then that of the apico-alveolar <s>, and finally, the highest CoG was that of the lamino-alveolar sound <z>.

Table 1

Center of Gravity values (in Hz) of Basque sibilant fricatives by L1 Basque speakers

	Control (n = 2)	
	M	**SD**
<s>	6319	901
<x>	4184	322
<z>	11752	1406

Similarly, the control dataset of L1-English sibilants showed clearly distinguished sibilants' values (Table 2). English /s/ and /z/ have essentially identical values as these two English phonemes share the same place of articulation. The CoG values suggest that English alveolar fricatives /s/ are intermediate between Basque <z> (/s̻/) and Basque <s> (/s̺/) in their PoA (see Figure 1).

[6] The decision was made to only include Session as a predictor (as opposed to Group) because not every group took part in all sessions. While learners conducted both the Basque and English sessions, native speakers of Basque only participated in the Basque session.

https://doi.org/10.1387/asju.25918

Table 2

Center of Gravity values (in Hz) of English sibilants by L1 English speakers

	Learners (n = 10)	
	M	**SD**
/s/	8915	522
/ʃ/	4452	716
/z/	8675	648

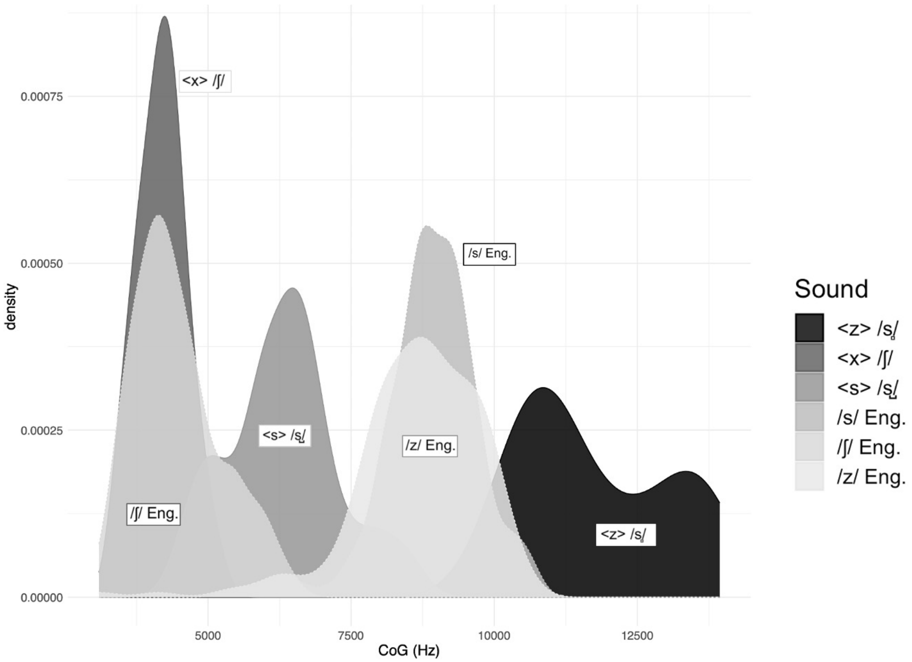

Figure 1

Production of sibilants in English and Basque by native speakers

In the case of FL-Basque learners' production, as can be seen in Table 3, <x> has a clearly distinguished realization, that is, below 5000 Hz. The other two sounds show higher CoG values, with <z> being the highest. However, <z> produced by learners is not as high as the <z> produced by native speakers of Basque. Figure 2 shows the production of learners' Basque sibilants (in yellow) and that of native speakers' (in grey).

Table 3

Center of Gravity values of Basque sibilants
by FL-learners (in Hz)

	Learners (n = 10)	
	M	**SD**
<s>	7771	1348
<x>	4946	965
<z>	8439	851

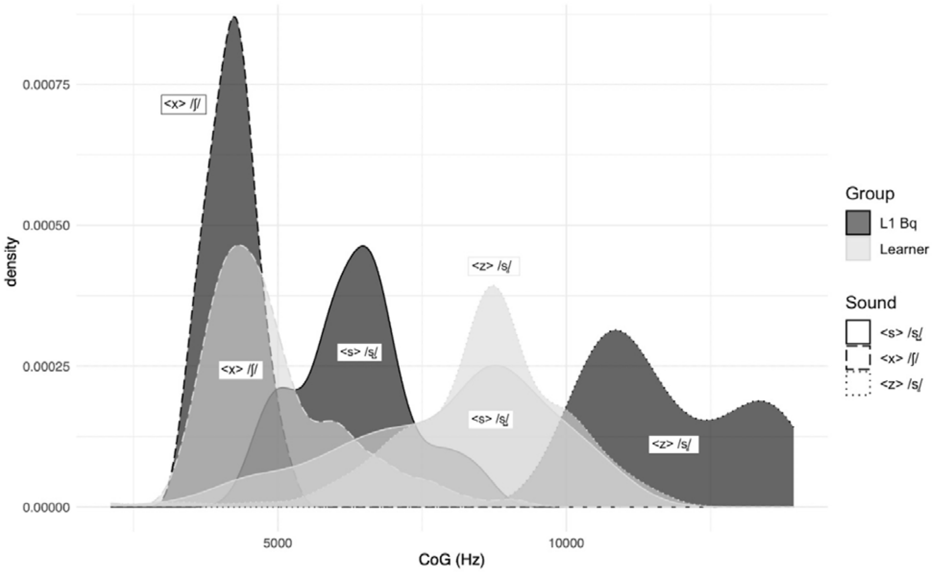

Figure 2

Center of Gravity values by Group

Regarding the effect of Session, the linear mixed-effects model revealed statistical significance. Table 4 shows the output of the effect of Session.

https://doi.org/10.1387/asju.25918

Table 4

Output of lmer model by Session (dependent variable: CoG)

	β	SE	df	t-value	p-value
English	1191.70	205.11	1732	5.810	< 0.001
OnlyBasque (NS)	–1769.67	562.06	26	–3.149	< 0.01

Results show that every session was significantly different from each other with regard to CoG. In the English session, higher CoG values both for /s/ and /z/ could explain its statistical significance, and in the OnlyBasque session the clearly distinguished CoG values (with considerably higher values in <z>) could explain the statistical significance obtained.

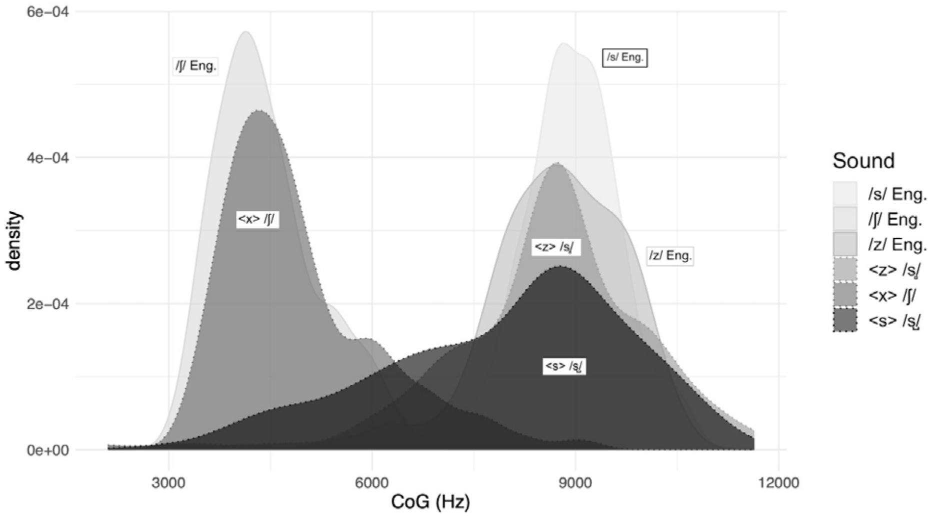

Figure 3

Center of Gravity values in Session 1 and English session

Furthermore, the control dataset of English sibilants revealed something remarkable. The mean CoG of [s] in English was higher than what has usually been encountered in previous literature. Jongman *et al.* (2000) placed the CoG of English [s] around 6000-7000 Hz. Here, it is almost 9000 Hz, and is similar to that of [z], showing that English [s] and [z] were somewhat more fronted for these participants. The pre-palatal sound [ʃ] shows lower frequency values than the rest, as expected, because of its more retracted point of articulation. Figure 3 provides a visualization of the sibilants produced by FL-Basque learners in Session 1 and English sibilants.

Regarding vocalic context, CoG decreases in the context of /u/ and increases in the context of /i/. The vowel /i/ showed main effects (*β*: 397.56, *SE*: 200.01, $t(1732) = 1.988$, $p < 0.05$), but /u/ did not (*β*: -167.67, *SE*: 198.52, $t(1732) = -0.845$, $p = 0.398$). Regardless, the decrease in frequencies in the context of /u/ confirms what previous literature showed about the influence of the feature [+round] on frequencies (Carney & Moll 1971; Fujisaki & Kunisaki 1978; Jesus & Shadle 2002; Nowak 2006).

2.2. Perception experiment

2.2.1. Research questions and predictions

Best's (1995) model predicts different outcomes with regard to new sound categorization: if a given FL sound is acoustically very similar to a sound in the L1, these will be a "CATEGORIZED EXEMPLAR", that is, an FL sound will be categorized as an L1 sound. If a given sound is between two L1 sounds, this is an "UNCATEGORIZED EXEMPLAR", and the perception of such sounds will vary, being perceived as one of the nearby sounds in the L1 (yet not one specifically). Studies such as Larraza *et al.* (2017) found that new sound categories are harder to process, and therefore take more time. Furthermore, according to Fujisaki and Kunisaki (1978), acoustic effects produced by the [+round] feature of /u/ in the context of sibilants can complicate the correct perception of these sounds. Based on those findings, this study intended to answer the following research questions:

— *Research question 1:* Are L1 English speakers able to perceptually discriminate between Basque sibilant fricatives?

- Prediction 1: based on previous research, it is expected that FL-Basque learners should find it difficult to discriminate between /ʃ/ and /s̺/ because of their acoustic proximity. More specifically, it is expected that lowest scores should be found in the context of /u/, as shown by Fujisaki and Kunisaki (1978).
- Prediction 2: since both /s̺/ and /s̻/ are new sound categories for L1-English learners of FL-Basque, lower accuracy scores in the contrast of these sounds are expected.

— *Research question 2:* Is there a correlation between production and perception?

- Prediction 1: It is expected that there will be a correlation between production and perception. Since it is usually stated that perception precedes production, the production results should correlate with those found in the perception experiment. Therefore, FL learners should show lower accuracy scores in the <s>-<z> alveolar contrast.

https://doi.org/10.1387/asju.25918

2.2.2. Methods

2.2.2.1. Sampling

The subjects were the same 12 speakers who participated in the production experiment.

2.2.2.2. Data collection

Data collection took place during Week 5 of classes and in a sound attenuated booth in the same laboratory as in the production experiment. Participants undertook an ABX discrimination task after taking the read-aloud task from the production experiment. The ABX discrimination task was administered to the participants using a 13-inch MacBook Pro Retina Display, 2015 model laptop. The *PsychoPy 2.0* software (Peirce & MacAskill 2018) was utilized to run the experiment. Participants heard 36 target pairs and 72 filler pairs divided into three blocks. The stimuli were produced by two female speakers who are native speakers of a conservative variety of Basque where the distinction between the three sounds is maintained.[7] Stimuli had been recorded with a sampling rate of 44.1 kHz and a sample size of 16 bits. SONY MDRZX110AP ZX headphones were used to listen to the stimuli. The volume was initially set at maximum level. Participants were informed they could adjust it at their desired level. Instructions were provided in English. Before the experiments took place, participants had a short training phase where they could get used to the program, procedure, and commands. Target pairs consisted in CV syllables with the structure of *sibilant + corner vowel*, like in the production experiment. Sibilants under study were the same ones as in the Basque production experiment, but the focus here is the opposition of <s-z> (/s̺/-/s̻/), <s-x> (/s̺/-/ʃ/), and <x-z> (/ʃ/-/s̻/).

Participants were first provided with two different stimuli (A & B) produced by the first female speaker. After this, a third stimulus (X), produced by a second female speaker, was played. Students had to decide whether this third stimulus (X) corresponded to the first stimulus (A) 'left arrow' or the second one (B) 'right arrow'. Each of the pairs in each condition appeared in all four possible stimuli combinations (ABA, ABB, BAA, BAB). As in Larraza *et al.* (2017), the intertrial interval was 1000 ms, and began after the subject's response. Participants did not have a maximum amount of time to respond. The interstimulus interval was 300 ms, between the first and second item, and between the second and third one. Participants were informed they had to respond as fast as they could but being certain about their answer.

2.2.2.3. Data analysis

PsychoPy 2.0. generated an excel sheet per speaker where accuracy scores were indicated for each stimuli contrast pair. The statistical significance of accuracy scores was obtained by means of binomial logistic regressions under the *glmmTMB* package

[7] These females were not the two speakers serving as control group.

(Brooks *et al.* 2017). The binomial logistic regression was used because the dependent variable was dichotomous, that is, "0" for *incorrect*, and "1" for *correct*.

The same procedure as in the production experiment was employed to for optimal model selection in the perception experiment, that is, AIC scores. Initially, for each dependent variable, models up to three fixed factors (Sounds (levels: s-z, s-x, x-z]), Vowel (levels: a, i, u), and Session (levels: Session 1; OnlyBasque) and one random factor (Speaker) were considered. The fixed factor Group was not included as such because, as will be seen in the results section, the control group obtained perfect accuracy (100/100), thus making it not possible to draw inferential results. The binomial logistic regression for accuracy response level included the interaction of the fixed factors Sounds (s-z, s-x, x-z) × Vowel (/a, i, u/) and Speaker as random factor. As previously, the statistical significance level was established at $\alpha = 0.05$.[8]

2.2.3. Results

As in the production experiment, the orthographical representation of Basque sibilant fricatives will be employed to refer to their phonetic realizations, i.e., <s> for the apico-alveolar /s̺/, <z> for the lamino-alveolar /s̻/, and <x> for the pre-palatal /ʃ/.

All groups scored above chance (50%) scores for all sound contrasts. However, while the control group reached perfect accuracy, that was not the case for FL-Basque learners. Comparing *filler* vs. *target* stimuli, it was more difficult to discern target words than filler words for FL-Basque learners. Table 5 shows the summary of the general results by word type:

Table 5

Accuracy scores and percentages by WordType

	Learners (n = 10)		**Control (n = 2)**	
	Session 1		**OnlyBasque**	
	N	**% (SD)**	**N**	**% (SD)**
filler	659/720	92 (6)	144/144	100 (0)
target	292/360	81 (11)	72/72	100 (0)

A two-way repeated measures ANOVA with the interaction of Group and WordType as predictors of accuracy was run to see whether the differences found between WordType by Group were statistically significant. Main effects were found for Group ($F(1, 1) = 29.228$, $p < 0.001$), and WordType ($F(1, 1) = 46.687$, $p < 0.001$). Moreover, the interaction of WordType × Group ($F(1, 1) = 4.669$, $p < 0.05$) yielded significant results because the effect of WordType was only perceivable in the FL-Basque learners group. This was confirmed by a post-hoc

[8] Statistical model formula: *glmer(Answer~Sounds*Vowel + (1|Speaker), data = targets, family = "binomial")*

https://doi.org/10.1387/asju.25918

TukeyHSD test since the significance between the two WordType categories was $p = 1$ for the control group but $p < 0.001$ for learners.

As one of the initial hypotheses predicted (RQ1, P2), for target sound contrasts, in both sessions the alveolar contrast <s-z> was the hardest one while the <x-z> contrast the easiest one at both sessions, which would be expected if we consider the articulations of each sound. Table 6 shows a summary of results and Figure 4 shows a visual representation of the results.

Table 6

Accuracy scores and percentages by Sound contrast

	Learners (n = 10)		Control (n = 2)	
	Session 1		OnlyBasque	
	N	% (SD)	N	% (SD)
<s-x>	97/120	81 (40)	24/24	100 (0)
<s-z>	87/120	72 (45)	24/24	100 (0)
<x-z>	108/120	90 (30)	24/24	100 (0)

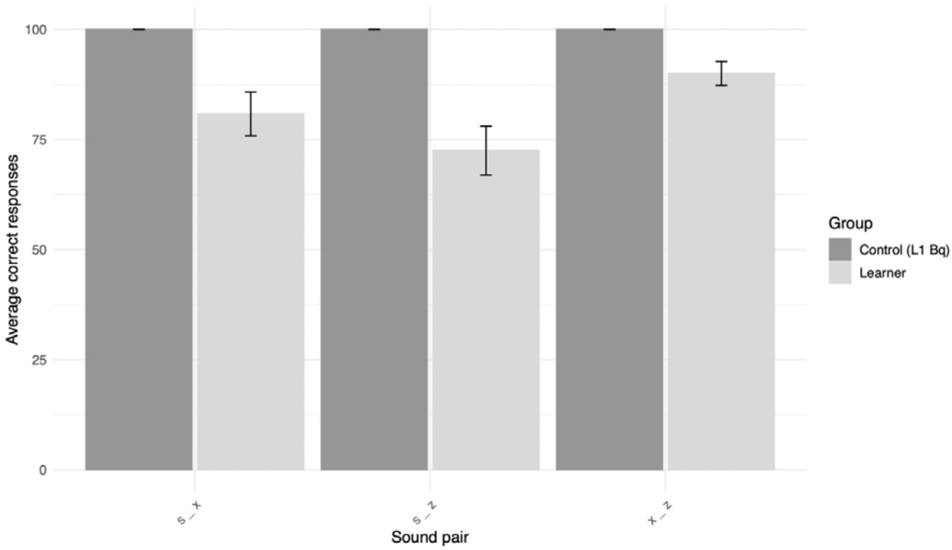

Figure 4

Accuracy percentages by target sound contrast

The binomial logistic regression model showed that the lower scores of <s-z> yielded statistical significance compared to the other pairs, β: –2.431, *SE*: 0.817, *Wald's z*: 4.08, $p < 0.01$.

Regarding the vocalic context in which the initial sibilant appeared, various tendencies are found by sound contrast. Table 7 shows the summary of the results and Figure 5 shows its visual representation.

Table 7

Accuracy scores and percentages by VOWEL

	Learners (n = 10)					
	<s-x>		**<s-z>**		**<x-z>**	
	N	**% (SD)**	**N**	**% (SD)**	**N**	**% (SD)**
a	38/40	95 (16)	27/40	68 (17)	37/40	93 (17)
i	33/40	83 (21)	33/40	83 (17)	34/40	85 (13)
u	26/40	65 (34)	27/40	68 (35)	37/40	93 (17)

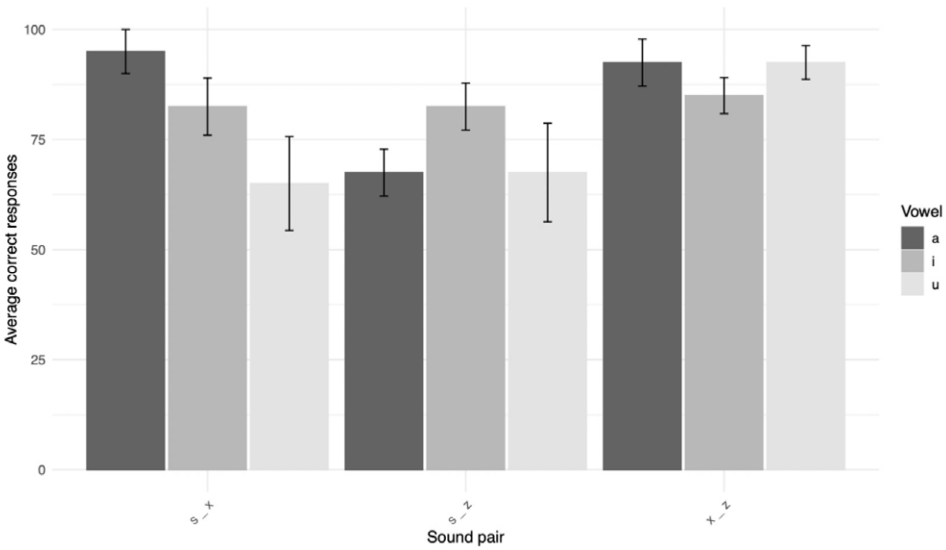

Figure 5

Accuracy scores and percentages by VOWEL

One of the initial predictions was that the sound contrast of <s-x> in the /u/ context would be the hardest to discern (RQ1, P1), because of the lowering of the frequencies due to the rounding of the vowel (Fujisaki & Kunisaki 1978;

https://doi.org/10.1387/asju.25918

Mann & Repp 1980), leading <s> to be perceived as <x>. The vowel /u/ showed main effects, *β*: –2.464, *SE*: 0.815, *Wald's z*: –3.02, $p < 0.01$. This can be understood because, overall (yet not always), accuracy scores were lower in the context of /u/.

As predicted, results reveal that the interactions between sound contrast of <s-x> and the /u/ context yielded the lowest level of accuracy (65%) and statistical significance of *β*: –2.431, *SE*: 0.817, *Wald's z*: 4.08, $p < 0.01$, and the sound contrast of <s-z> and the /i/ context also yielded statistical significance, *β*: 2.341, *SE*: 1.018, *Wald's z*: 2.299, $p < 0.05$, with higher accuracy scores in this context than in the other two contexts during Session 1 (83% vs. 68%). The /u/ context also showed statistically significant effects in the interaction with the sound contrasts <s-z> (*β*: 2.464, *SE*: 0.953, *Wald's z*: 2.58, $p < 0.01$) and <x-z> (*β*: 2.464, *SE*: 1.190, *Wald's z*: 2.07, $p < 0.05$). As can be observed in Figure 5, accuracy scores in these contexts were higher than for <s-x> by /u/.

3. Discussion and conclusion

Results showed that as far the production of the sibilants the two different hypotheses that were suggested are met. It was proposed that since the two alveolar sibilants of Basque are new categories for L1 English learners, they could either map both of them onto their English alveolar sibilants, or, in case of stepping ahead and noticing that the Basque apico-alveolar <s> has a different place of articulation (which is more retracted than that of their English /s/) they could easily map that new sound onto their English pre-palatal sound /ʃ/ due to acoustic similarity. Another possibility for those speakers that voiced <z> in Basque is that, because of spelling, they would apply the following mapping: Basque <s> = English /s/, Basque <x> = English /ʃ/, Basque <z> = English /s-z/.

Group results show that students merge the apical sounds into their English alveolar sibilants. It is worth mentioning that individual variation was found: 4/10 participants were able to produce the three sibilants in Basque, while 6/10 merged <s-z>. Due to length-related constraints, a section about how each subgroup behaved in production and perception experiments was not included but can be found in Beristain (2019), who exhibited that the sub-group that produced the three sibilants was significantly more successful at discriminating the three sibilants in Basque, too. Notably, participants also took part in a second session towards the end of the semester, and while general improvement was found for production, that was not the case for perception. It is hypothesized that the production task type (read-aloud task) could have had an effect in that regard. According to Flege (1995), those who have not created a new category and they have assimilated them to their L1 sound category. In other words, these participants map the two realizations of Basque alveolar sibilants onto their English alveolar /s/ (or realize both as alveolar, but Basque <s> as a voiceless alveolar and Basque <z> as a voiced alveolar). According to Best's (1995) PAM, that contrast is an example of a "CATEGORIZED EXEMPLAR", where the two sounds have been categorized as their English /s/ (or, as previously explained, /z/).

As far as difficulty of sound contrast perception is concerned, it was hypothesized that the sound contrasts <s-z> and <s-x> would the most difficult ones. The former because of it being a combination of two new sound categories, and the latter because it includes a new sound (Basque <s>) which is perceptually speaking similar to English /ʃ/. The <s-z> contrast was the hardest one for FL-Basque learners (and it was the sound that learners were merging as a group, showing the correlation between perception and production). As previously explained, because students are categorizing the two sounds as their English /s/. It could be hypothesized that since the two sounds are new sounds, students map both sounds into the same category in their L1, and as far as perception, that would be English /s/, since that is the only voiceless alveolar fricative (an example of a "CATEGORIZED EXEMPLAR", Best 1995). The case of <s-x> could be explained by the acoustic similarity that these sounds have, and how similar they would sound to a non-native speaker of Basque. According to Best (1995), this would be an "UNCATEGORIZED EXEMPLAR", because its perception falls in between English /s/ and /ʃ/. Lower accuracy scores could show that some FL-Basque learners perceive it as /s/ whereas others perceive it as /ʃ/. Following acoustic theories proposed by Fujisaki and Kunisaki (1978), as predicted, the lowest score was obtained when the apico-alveolar and pre-palatal sound were followed by /u/. Perceptually speaking, it could be that those two sounds sounded like /ʃ/ to learners, due to the lowering in frequencies in the sibilants.

To finalize, limitations of this study cannot be disregarded. The small sample size is a factor that prevented making broad generalizations. However, it should be noted that an experiment that included an L1-English FL-Basque population in the classroom had not yet been carried out. It is expected that this project will thus contribute to the field of L2/FL sound production and perception, classroom acquisition, and most specifically to the acquisition of L2/FL-Basque phonology by non L1-Spanish speakers. Furthermore, the use of a read-aloud task where participants could have been influenced by the spelling of the words could have had an influence in the overall results. Future research should consider examining task effects and observe whether different results are found in more naturalistic environments such as study abroad.

4. References

Akaike, Hirotogu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control 19*(6). 716-723.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67(1). 1-48.

Beristain, Ander. 2018. The acoustic realization of /s̺/ and /ts̺/ by L1 and L2 Basque speakers. In Lorea Unamuno, Asier Romero, Aintzane Etxebarria & Aitor Iglesias (eds.), *Linguistic Variation in the Basque Language and Education* – III, 70-82. Bilbao: UPV/EHU.

Beristain, Ander. 2019. The acquisition of Basque fricative sibilants by L1 English speakers. Poster presentation at *L2 Pronunciation Research Workshop: Bridging the Gap between Research and Practice.* Universitat de Barcelona.

https://doi.org/10.1387/asju.25918

Beristain, Ander. 2022. Spectral properties of anterior sibilant fricatives in Northern Peninsular Spanish and sibilant-merging and non-merging varieties of Basque. *Journal of the International Phonetic Association* 52(3), 421-452.

Best, Catherine. 1995. A direct realist perspective on cross-language speech perception. In Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech* Research, 167-200. Baltimore: York Press.

Best, Catherine, Gerald McRoberts & Elizabeth Goodell. 2001. Discrimination of non-native consonant contrast varying in perceptual assimilation to the listener's native phonological system. *Acoustical Society of America* 109(2). 775-794.

Bladon, Anthony, Christopher Clark & Katrina Mickey. 1987. Production and perception of sibilant fricatives: Shona data. *Journal of the International Phonetic Association* 17. 39-65.

Boersma, Paul & David Weenink. 2019. *Praat: doing phonetics by computer*, version 6.048. https://www.fon.hum.uva.nl/praat/.

Brooks, Mollie, Kasper Kristensen, Koen van Benthem, Arni Magnusson, Casper Berg, Anders Nielsen, Hans Skaug, Martin Maechler & Benjamin Bolker. 2017. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal* 9(2). 378-400.

Burnham, Kenneth & David Anderson. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer-Verlag.

Carney, Patrick & Kenneth Moll. 1971. A cinefluorographic investigation of fricative consonant-vowel coarticulation. *Phonetica* 23. 193-202.

Flege, James. 1995. Second language speech learning. Theory, findings, and problems. In Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233-277. Timonium, MD: York Press.

Flege, James. 2007. Language contact in bilingualism: Phonetic system interactions. In Jennifer Cole & Jose Ignacio Hualde (eds.), *Laboratory Phonology 9*, 353-381. New York: Mouton de Gruyter.

Fujisaki, Hiroya & Osamu Kunisaki. 1978. Analysis, recognition, and perception of voiceless fricative consonants in Japanese. *IEEE Transactions* 26. 21-27.

Hualde, José Ignacio. 2010. Neutralización de sibilantes vascas y seseo en castellano. *Oihenart* 25. 89-116.

Iribar, Alexander & Carmen Isasi. 2008. El seseo vasco: nuevos datos para una vieja cuestión. *Oihenart* 23. 241-258.

Iribar, Alexander, Carmen Isasi, Sara Gómez Seibane & Carmen Moral del Hoyo. 2005. Notas para la descripción acústica del seseo vizcaíno. In Manual González, Elisa Fernández & Begoña González (eds.), *III Congreso Internacional de Fonética Experimental*, 389-398. Santiago de Compostela: Xunta de Galicia.

Jesus, Luis M.T. & Christine H. Shadle. 2002. A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics* 30(3). 437-464.

Jongman, Allard, Retree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustic Society of America* 108(3). 1252-1263.

Jurado, Mirari. 2011. Caracterización de sibilantes fricativas vascas y su percepción en el sistema fonético español. *ASJU* 45(1). 81-137.

Klein, Wolfgang. 1986. *Second Language Acquisition.* Cambridge: Cambridge University Press.

Kunisaki, Osamu & Hiroya Fujisaki. 1977. On the influence of context upon perception of voiceless fricative consonants. *Annual Bulletin*, vol. II, 85-91. Tokyo: Research Institute for Logopedics and Phoniatrics, University of Tokyo.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune Haubo Bojesen Christensen. 2017. lmerTest: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13). https://doi.org/10.18637/jss.v082.i13 (Version: 2.0-25, Kuznetsova 2015).

Larraza, Saioa. 2015. Acquisition of Phonology and Spanish-Basque Bilinguals' Phonological Systems. In *Actes du colloque AcquisiLyon 09*, 1-4. Lyon: France. http://ohll.ish-lyon.cnrs.fr/Colloques/AcquisiLyon/pageweb/Fichier/Larraza.pdf (29/12/2023).

Larraza, Saioa, Arthur Samuel & Miren Oñederra. 2017. Where do dialectal effects on speech processing come from? Evidence from a cross-dialect investigation. *Quarterly Journal of Experimental Psychology* 70(1). 92-108.

Lisker, Leigh. 2001. Hearing the Polish sibilants [s š ś]: phonetic and auditory judgements. In Nina Grønnum & Jørgen Rischel (eds.), *Travaux du Cercle Linguistique de Copenhague XXXI*, 226-238. Copenhagen: C.A. Reitzel.

MacSwan, Jeff & Lisa Pray. 2005. Learning English bilingually: Age of onset of exposure and rate of acquisition of English among children in a bilingual education program. *Bilingual Research Journal* 29(3). 687-712.

Mann, Virginia & Bruno Repp. 1980. Influence of vocalic context on the perception of the [sh]-[s] distinction. *Perception & Psychophysics* 28. 213-228.

McGuire, Grant. 2007. English listeners' perception of Polish alveopalatal and retroflex voiceless sibilants: A pilot study. *UC Berkeley Phonology Lab Annual Report* 3(3). 391-417.

Nowak, Paweł. 2006. The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics* 34(2). 139-152.

Pallier, Cristophe, Laura Bosch & Núria Sebastian-Gallés. 1997. A Limit on Behavioral Plasticity in Speech Perception. *Cognition* 64. 9-17.

Peirce, Jonathan & Michael MacAskill. 2018. *Building Experiments in PsychoPy*. London: Sage.

R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org (29/12/2023).

RStudio Team. 2019. RStudio: *Integrated Development for R.* Boston, MA: RStudio, Inc. http://www.rstudio.com (29/12/2023).

Todd, Ann, Jan Edwards & Ruth Litovsky. 2011. Production of contrast between sibilant fricatives by children with cochlear implants. *The Journal of the Acoustical Society of America* 130. 3969-3979.

Whalen, Douglas. 1981. Effect of vocalic formant transitions and vowel quality on the English [s]-[sh] boundary. *Journal of Acoustical Society of America* 90(4). 275-282.

https://doi.org/10.1387/asju.25918

5. Appendices

Appendix A

List of target words in Basque

Carrier phrase: *Nik TARGET irakurri dut* '*I read$_{present}$ TARGET*'

	/a/	/i/	/u/
/s̺/ <s>	*sator* 'mole' *sasi* 'bramble' *sagu* 'mouse'	*sinatu* 'sign' *sifoi* 'siphon' *siku* 'dry'	*sumin* 'anger' *sukar* 'fever' *sute* 'fire'
/s̻/ <z>	*zaku* 'bag' *zati* 'piece' *zama* 'weight'	*zimel* 'withered' *zirin* 'bird excrement' *zilar* 'silver'	*zure* 'your (sg.)' *zubi* 'bridge' *zuku* 'juice'
/ʃ/ <x>	*xare* 'handball game' *xake* 'chess' *xafla* 'slice'	*xingar* 'bacon' *xira* 'ivy' *xima* 'sprout'	*xume* 'humble' *xuko* 'dry' *xuka* (informal addressing)

Appendix B

List of target words in English

Carrier phrase: *I read$_{present}$ TARGET*

	/ɔ/, /ɑ/, /ʌ/, /æ/	/i/, /ɪ/	/u/, /ʊ/
/s/	*soccer* *sauce* *sun*	*sea* *sister* *seal*	*super* *soup* *suit*
/z/	*zag* *zach* *zags*	*zeal* *zee* *zeals*	*zoo* *zoom* *zooms*
/ʃ/	*shark* *shop* *shovel*	*sheep* *shield* *ship*	*shoot* *shoe* *sugar*