






ARTICLE

An acoustic exploration of sibilant contrasts and sibilant merger in Mixean Basque

Ander Egurtzegi¹ , Dorota Krajewska^{2,*} , Christopher Carignan³  and
Iñigo Urrestarazu-Porta⁴

¹CNRS – IKER UMR5478, ²University of the Basque Country UPV/EHU, ³University College London, and ⁴CNRS – IKER UMR5478, UPV/EHU, UPPA

*Corresponding author. Email: dorota.krajewska@ehu.eus

(Received 16 March 2023; revised 19 January 2024; accepted 6 February 2024; first published online 16 May 2024)

Abstract

This exploratory study investigates sibilants in Mixean Low Navarrese, an endangered variety of Basque. This variety has been described with ten different contrastive sibilants: /s̺, s̠, ʃ, ts̺, ts̠, tʃ, z̺, z̠, ʒ, dʒ/. The objective of the paper is to (a) provide a detailed description of the acoustics of Mixean sibilants, and (b) elucidate whether ten categories can be proposed based only on acoustical data, or whether fewer categories should be considered. The study is based on free-conversation data of ten subjects (three females, seven males) aged between 80 and 85 years. We analyze metrics reflecting the place of articulation (spectral moments, and especially the center of gravity (CoG)), including also the temporal dynamics of CoG (using the discrete cosine transform of CoG measurements of nine intervals of each phone). We also explore the acoustic correlates of the contrasts between (a) voiced and voiceless sounds and (b) fricative and affricate sounds. The results show that only seven categories can be proposed based on acoustic measurements. The lamino-alveolar series reliably contrasts with the rest, but the distinction does not hold between the apico-alveolar and the postalveolar series. We found minimal differences in the analysis of dynamic data, and none in the static analysis.

Keywords: Basque; sibilants; merger; acoustics

1 Introduction

Mixean Low Navarrese is an endangered variety of Basque that has been described with ten different contrastive sibilants, including oppositions based on three places of articulation, voicing, and fricatives vs. affricates. Nevertheless, a merger collapsing the three apico-alveolar sibilants with the three corresponding postalveolars has been (preliminarily) proposed in the literature (Egurtzegi & Carignan 2020b).

The current exploratory study investigates these segments, in order to (a) provide a detailed description of the acoustics of sibilants in this variety, and (b) elucidate whether ten categories can be proposed based only on acoustical data, or whether fewer categories should be considered. At the same time, this paper aims to present a thorough exploration of how to approach the acoustic description of a variety of a language including an undetermined number of sibilant contrasts. To that end, we take into account an array

© The Author(s), 2024. Published by Cambridge University Press on behalf of The International Phonetic Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

of acoustic measures. To begin with, we include metrics which reflect the place of articulation (spectral moments; especially, the spectral center of gravity, i.e. CoG), including also the temporal dynamics, i.e. the variation CoG values show throughout the sound segment. We also take into account the acoustic correlates of the contrasts between (a) voiced and voiceless sounds and (b) fricative and affricate sounds.

The paper is organized as follows: Section 2 provides the necessary background information on sibilants and sibilant systems in the world's languages, the Basque variety under study and its phonology, and sibilant contrasts in Basque and related acoustic studies. Section 3 lists our research questions. Section 4 introduces the data and the acoustic measures used to analyze it. Section 5 presents the results: a general exploration (5.1), the analysis of spectral moments (5.2), the temporal dynamics of CoG (5.3), the voicing contrast and analysis of duration (5.4), and distinction between fricative and affricate sounds (5.5). Section 6 discusses the results and Section 7 provides conclusions.

2 Background

2.1 Sibilants in the world's languages

Sibilants are usually defined as fricatives with high-frequency spectral energy, in which a turbulent, random-frequency sound is generated by the strike of a jet of air against the teeth after passing at a high velocity through a narrow constriction, which is produced by the closeness of the flexible part of the tongue and the passive articulator (see e.g. Ladefoged & Maddieson 1996: 138). In the case of affricate sibilants, this sound is the second component of the segment, namely the release of a preceding homorganic stop. Sibilants are coronals by definition, and can be produced from the dental to the palatal region. They are more common than other fricatives, potentially due to their higher acoustic energy (Ladefoged 2001: 167). Most of the world's languages have at least one fricative sound (93.4% in the UPSID database, Maddieson (1984: 42)), which is typically a sibilant. A notable exception is Australian languages, many of which lack fricatives altogether (Maddieson 1984: 42).

The most frequent sibilant segment in the world's languages is /s/, which typically denotes an alveolar fricative sibilant, although it might represent a dental or denti-alveolar in some linguistic descriptions. These sounds have often been collapsed in the literature, given that it is often difficult to discern them (Ladefoged & Maddieson 1996: 146), and they rarely contrast. In the UPSID database, where about 83% of the languages have some kind of s-sound, only four languages (1.3%) have both /s/ and /s̺/ (cf. Maddieson 1984: 44).

In the Phoible database (Moran & McCloy 2019), /s/ is represented in 2020 phonological inventories (out of 3020, i.e. 67%). Following /s/, we find /tʃ/ (1218; 40%), /ʃ/ (1104; 37%), /z/ (893; 30%), /dʒ/ (820; 27%), /ts/ (666; 22%), /ʒ/ (478; 16%), and /dz/ (312; 10%). Thus, the most frequent sibilants involve all possible voiceless/voiced and fricative/affricate combinations of the alveolar or postalveolar places of articulation (with no secondary articulation/release). Other sibilants are represented in less than 200 phonological inventories of the Phoible database (less than 7% each).

Sibilants are most usually voiceless, although voiced sibilants are not uncommon (see Žygis *et al.* 2012 for a phonetic account of this biased distribution in sibilant affricates). Maddieson (1984: 45) observed that the ratio of voiced to voiceless sibilant fricatives was 0.36 for the (dental-)alveolar and 0.34 for the postalveolar pair. Non-sibilant voiced fricatives are more frequent, and thus non-sibilant fricative pairs show higher voicing ratios. Although not necessarily so, the presence of a voiced fricative in a segmental inventory generally implies that of its voiceless counterpart (Maddieson 1984: 47).

While many languages, such as English or Spanish, have a reduced number of sibilants, other languages have larger sibilant inventories. Polish (Jassem 2003: 103) is a classic example of a large sibilant inventory, with 12 sibilants (fricative/affricate and voiced/voiceless

variants of dental/alveolar, (retroflex) postalveolar, and alveolo-palatal sibilants). Larger inventories imply more opportunities for mergers to occur, as in some dialects of Polish which reduce the three-way contrast to two places of articulation (Żygis 2003: 179). Mixean Basque, the language variety studied here, involves a lesser known but nonetheless comparable case. Although previously described with a ten-sibilant system (/s̺, š̺, ʃ, t̺s̺, t̺š̺, t̺ʃ, z̺, ž̺, ʒ, d̺z̺/), recent research has proposed a merger between the postalveolar and apico-alveolar categories (Egurtzegi & Carignan 2020b).

2.2 Contextualizing Mixean Basque

Within France, Mixe (*Amiküze* in Mixean Basque) is one of the regions forming the Pyrénées-Atlantiques department, in the South-West corner of France. Mixe is located in the North(-East) of the Basque Country, in the North of the historical province of Low Navarre. It shares borders with the Gascon Béarn to the North and the historical Basque province of Zuberoa to the East. The region is formed by 32 towns, the main one being Donapaleu (*Saint-Palais* in French). However, the total population of Mixe is relatively small (7856 people in 2015; L'institut national de la statistique et des études économiques (2019)) and the number of Basque speakers is even smaller. In fact, Mixean Basque (*amiküzera* in Mixean Basque), has been described as being on its way to disappearing (Camino 2016: 51): not even 10% of children are schooled in a Basque-speaking model (Zabalik 2016) and for those the education language is standard Basque, and not the local variety.

Mixean Basque has been understudied until recently, perhaps because it is underrepresented in the literature and most Basque dialectologists have considered it, in synchronic classifications, as (Eastern) Low Navarrese (Mitxelena 2011 [1977]) or as a transition variety between Low Navarrese and Zuberoan. The endangerment of this variety urges researchers to study Mixean Basque now, while it is still possible, before it definitively disappears. Nevertheless, the recent thorough dialectological study by Camino (2016) and the general acoustic description by Egurtzegi & Carignan (2020b) have greatly helped improve our knowledge of Mixean Basque.

2.3 Mixean phonology

The phonological inventory of Mixean Basque consists of 34 contrastive consonants, including 12 stops /p, t, c, k, p^h, t^h, c^h, k^h, b, d, ʃ, g/, ten sibilants /s̺, š̺, ʃ, t̺s̺, t̺š̺, t̺ʃ, z̺, ž̺, ʒ, d̺z̺/, a labiodental fricative /f/, nine sonorants /m, n, ɲ, l, ʎ, r, r̺, j, w/, two laryngeals /h, h̺/ and six contrastive vowels /a, e, o, i, u, u̺/. Besides, French loanwords include /R, v, ε, œ/ and nasalized vowels. However, vowel nasalization is not phonemic in Mixean Basque (Camino 2016: 200), but nasalized vowels are produced due to coarticulation with any neighboring nasal segment. Some of these phonemic contrasts have been acoustically established, including the three contrastive stop series – voiced, voiceless unaspirated and voiceless aspirated (Egurtzegi & Carignan 2020b: 2797–2798)–, a sixth vowel quality /u̺/, not present in other Basque dialects (Egurtzegi & Carignan 2020b: 2796–2797), or nasality as a contrastive feature distinguishing the two laryngeals /h, h̺/ (Egurtzegi & Carignan 2020a). However, other phonological oppositions have not been studied in enough detail. This is the case of the sibilant series, which is tentatively described with the high number of ten contrasting segments. Nevertheless, Egurtzegi & Carignan (2020b: 2797–2798) proposed a potential merger that would reduce their number to seven (see details below). This paper aims to fill these gaps with a more in-detail study of the recordings of Mixean Basque already examined by Egurtzegi & Carignan (2020b), which were collected by Camino (2016).

measurements (e.g., Yárnoz 2002a), while other varieties often show mergers of different kinds.¹

Nevertheless, many studies are reduced in scope (with regard to the number of speakers, tokens, and/or segments analyzed), and they typically analyze western and central varieties of the language. For example, Hualde (2010) reported the results of a single speaker; Gandarias et al. (2014) too reported the complete CoG results for a single speaker and they only analyzed three tokens for each sibilant; Iglesias et al. (2016) did not use lexical items but nonce words and analyzed just one speaker. Muxika-Loitzate (2017) and Beristain (2018) analyzed only fricatives, leaving affricates aside. Beristain (2021) reports results of a bigger corpus (including 18 female speakers, six per variety, and 80 tokens per speaker, i.e. 40 tokens per fricative), comparing the apico-alveolar and lamino-alveolar fricatives of three varieties of Basque (and Northern Spanish).

Concerning descriptions of sibilant sounds in our variety of interest, there are two previous works that deal with Mixean data. First, Urrutia et al. (1991) report the lower energy cut-off frequencies of the sibilants of three speakers of Eastern Low Navarrese, including one speaker of Mixean Low Navarrese. They report the mean lower energy cut-off frequency of three voiceless sibilant fricatives: apico-alveolar, dorso-alveolar, and palatal (Urrutia et al. 1991: 203–232) and their affricate counterparts (Urrutia et al. 1991: 233–272), with only four tokens of (partially) voiced sibilants across speakers, all of them of [ʒ] (Urrutia et al. 1991: 227). Comparing their data to other modern acoustical analyses of sibilants is not easy, since all recent studies report the spectral CoG as the main (or only) metric. Besides, Urrutia et al. (1991) only offer aggregated data of three speakers of Eastern Low Navarrese, which combines Donibane Garazi (Cicean Low Navarrese), Donapaleu (Mixean Low Navarrese), and Salazar Valley (Salazarese) varieties. However, Mixean is markedly different from those geographically adjacent varieties, so that the results reported in this work do not necessarily represent the Mixean variety.

More recently, Egurtzegi & Carignan (2020b) provide an extensive acoustic description of Mixean Basque, their work likely being the most complete acoustic description available on any variety of the Basque language. They used data originally recorded by Camino (2016) through fieldwork carried over the last 40 years, and analyzed the speech of ten speakers from ten towns within the Mixe region, for a total of 1494 sibilant tokens (ranging between 97 and 214 tokens per speaker). They found very similar CoG values for the three voiced fricative sibilants, /z/ being a little higher than /ʒ/ and /ʒ/. Lamino-alveolar sibilants /s/, /ʃ/ showed the highest CoG values, with no meaningful differences between apico-alveolar /s/, /ʃ/ and alveolo-palatal /j/, /tʃ/. The authors proposed that this result suggests a merger in the place of articulation of the apical and postalveolar categories; both for voiced and voiceless as well as fricative and affricate sibilants (Egurtzegi & Carignan 2020b: 2799). Regarding the resulting place of articulation, the authors tentatively suggested that alveolo-palatal sibilants may have merged to the apico-(post)alveolar series in the Mixean variety (Egurtzegi & Carignan 2020b: 2800). There were no significant differences between the CoG frequencies of fricative and affricate segments (in their fricative portion) sharing place of articulation. Nevertheless, the description of Mixean Basque sibilants by Egurtzegi & Carignan (2020b) has its limitations: for example, they do not report the differences between voiced and voiceless sibilants, and they do not investigate the acoustic cues that differentiate fricatives and affricates.

¹ Yárnoz (2002a) does not report aggregated CoG values, but, as an example, a 69-year-old female High Navarrese speaker from Etxarri-Aranatz from the general study of sibilant production in different varieties of Basque in Urrestarazu-Porta (in preparation) shows the following mean (and SD) CoG values for each sibilant phoneme when extracted with the same script we used for the current study: /s/ 6685 Hz (706), /ʃ/ 5316 Hz (490), /j/ 4527 Hz (286), /ʒ/ 6967 Hz (393), /ʒ/ 5287 Hz (341), /tʃ/ 4922 Hz (352).

In this paper, we go well beyond the limitations of all previous studies by examining variation in the acoustic properties of sibilant sounds, as well as acoustically differentiating fricative vs. affricate sibilants and voiced vs. voiceless sibilants. To this end, we report on metrics that go beyond CoG, including direct cosine transform (DCT) coefficients to capture temporal information, relative intensity to distinguish between fricatives and affricates, and auto-correlation (AC) coefficients to differentiate voiced from voiceless segments. Thus, we aim at providing a full description of Mixean Basque sibilants that could be used as a basis for the analysis of the sibilant systems of other Basque varieties, or other languages with similar oppositions.

3 Research questions

In this exploratory acoustic study, we aim to analyze the properties of Mixean Basque sibilants by answering the following questions in order to present a detailed description of the sibilant oppositions in this variety:

1. What are the spectral characteristics of sibilants in Mixean Basque? Which measure differentiates their place of articulation most accurately? Do we find evidence in line with the merger of the apico-alveolar and postalveolar sibilants proposed in Egurtzegi & Carignan (2020b)?
2. What are the spectro-temporal properties of sibilants, and especially how does their spectral center of gravity change over time?
3. Which acoustic metrics differentiate voiced and voiceless sibilants?
4. Which acoustic metrics differentiate fricative and affricate sibilants?
5. Taking all the acoustic measures into account, how many sibilants can we distinguish in Mixean Basque?

4 Methodology

4.1 Data

The present study takes a deeper look into the data examined by Egurtzegi & Carignan (2020b). It is the only data available that allows us to study the Mixean variety of Basque preceding the abrupt stop in transmission and major shift to French that started in the 1970s (Camino 2016). It is based on free-conversation data recorded in various villages of the Mixe region by Iñaki Camino (analyzed from a dialectological point of view in Camino (2016)). We selected recordings carried out between 2005 and 2015 with a SONY MZ-R30 minidisc recorder (at a sampling rate of 44.1 kHz), leaving out speakers over 85 years old as well as older recordings made with a DAT recorder, both because of the different recording method and the large timespan in the period of recording. Thus, our corpus comprises data from ten subjects (three female, seven male) aged between 80 and 85 years and from the following villages in the region of Mixe: Donapaleu, Uharteheri, Sorhapürü, Arrüeta, Martxüeta, Labetze, Amendüze, Gamue, Zohota, and Arberatze. For each speaker, an average of 5.5 minutes of audio was analyzed (with a 3.5–8.5 min. range). The recordings were initially transcribed by Iñaki Camino, and then expanded and, when necessary, corrected by the first author, guided by the agreed-upon etymological forms of the Eastern varieties. The transcriptions were force-aligned using the WebMAUS application (Kisler et al. 2017) with the Basque (FR) language setting, specifically designed for the north-eastern varieties of Basque by the first author of this paper. The automatically generated TextGrids were later carefully hand-corrected.

Table 1. Total number of sibilant tokens in the study

Phone	Tokens
/ʒ/	83
/ʒ̥/	45
/ʒ̥̥/	70
/dʒ̥̥/	4
/ʃ/	78
/tʃ/	88
/s̥/	336
/tʃ̥̥/	46
/s̥̥/	716
/tʃ̥̥̥/	446

In total, 1912 sibilants were gathered from the recordings. On average, speakers produced 191 sibilant tokens (ranging between 111 and 265, $SD = 46.5$). Between four and 716 tokens of each phone were extracted from the corpus, as shown in Table 1. This asymmetry is inherent to the frequency of the sounds in the language, where voiceless lamino-alveolars are the most frequent segments (being part of high-frequency grammatical words and affixes), while the voiced lamino-alveolar affricate /dʒ̥̥/ is limited to a handful of words (mostly recent loans). Due to the low number of examples, /dʒ̥̥/ was not used in regression analyses, but it is included whenever feasible in descriptive statistics.

The plots in the next section report speaker-normalized data. Normalization was performed to account for the physiological differences among speakers. We first computed by-speaker Lobanov normalization values (or z-scores) in R (R Core Team 2022), which we then converted back to their original scale by using the average standard deviation and the average of the means (*grand mean*) of the ten speakers. These normalized values retain the speaker-specific normalized structure of z-scores but in a more familiar scale. All statistical analyses were performed on non-normalized data. Instead of using speaker-normalized values, a random effect for ‘speaker’ was included in all models to account for this variation.

4.2 Acoustic metrics

4.2.1 Static values of spectral moments

The center of gravity of the fricative noise (CoG, in Hz) was measured in Praat (Boersma & Weenink 2023) for nine equal intervals of the phone between 5% and 95% of the duration of the phone, using a 300–19000 Hz pass Hann filter. For fricatives, the measurements used in the static analysis are those corresponding to the middle interval of the phone (i.e. 45–55% of the phone’s duration). For affricates, however, in order to make sure to draw data from the fricative phase, a window centered around 70% (65–75%) of the phone was used instead. CoG is the most commonly used measure to analyze spectral properties of sibilants (see, among others, Jongman et al. 2000; Gordon et al. 2002; or Muxika-Loitzate 2017; Beristain 2021 for Basque), but, for the sake of completeness, kurtosis, skewness, and spectral standard deviation (SSD) were also measured at the 45–55% window of the fricative sounds and 65–75% window of the affricate sounds in this study.

4.2.2 Temporal dynamics of CoG

Reidy (2016) has shown that dynamic spectral analysis of sibilants reveals language-specific and consonant-specific information not available when using only static measures. In this study, the nine measurements of CoG were compressed into four coefficients using Discrete Cosine Transformation (DCT) following methods used to track temporal changes in formants (Harrington *et al.* 2008; Harrington & Schiel 2017) or for the analysis of sibilants as in Stuart-Smith (2020). The *dct* function of the package *emuR* (Winkelmann *et al.* 2021) was used for this procedure. DCT decomposes a signal (in this case, a series of data points) into a series of cosinusoidal waves. The sum of these waves approximates the shape of the original data, so that the signal can be approximately reconstructed from them. Each coefficient (k_0 , k_1 , etc.) corresponds to the amplitude of its respective wave. The first three coefficients are typically used in acoustic studies: k_0 , k_1 and k_2 . They are proportional to the mean, linear slope, and curvature of the trajectory, respectively (Harrington 2010: 305). Additionally, we also report the fourth coefficient (k_3) to capture yet more fine-grained temporal change, which may potentially help distinguish affricates from fricatives. Figure 1 is an example showing the way DCT transforms the data. The four DCT coefficients for one affricate sound are plotted on the left. The plots on the right show the sum of the waves, comparing the effect of using three and four coefficients, and the original nine data points.

4.2.3 Voicing and duration

Voicing contrasts can be described acoustically in different forms. For this study, we measured the auto-correlation (AC) coefficients and segment duration.

Following the protocols in Blevins *et al.* (2020: 311), AC values were calculated with the program EMU (Harrington 2010). We used the ESPS method, with a frame spacing of 10 ms, a window length of 7.5 ms, and pitch ranges of 60–400 Hz for male speakers and 90–600 Hz for female speakers, for a total of 17,957 measurements from all sibilants in our corpus. To aggregate the results while avoiding coarticulatory effects, we calculated the median values of the AC coefficients of each sibilant. *Ceteris paribus*, voiced phones are expected to show higher AC coefficients and, conversely, voiceless phones should show lower AC coefficients.

Duration may be another cue that helps distinguish voiced and voiceless sibilants. Some studies point to speakers perceiving shorter fricatives as voiced (Cole & Cooper 1975; Widdison 1995) and acoustic studies have shown voiced consonants tend to be in fact shorter (e.g. Smith 1997; Żygis *et al.* 2012). On the other hand, we do not expect to find remarkable differences within the voiced and voiceless categories (Gordon *et al.* 2002; Kochetov 2017). We relied on raw duration data (log-transformed) because it would offer a straightforward interpretation while still showing voicing distinctions.

4.2.4 Fricative/affricate distinction

Fricatives and affricates can be distinguished by various acoustic measures, such as friction duration (Kluender & Walsh 1992; Mitani *et al.* 2006) or amplitude rise time (Howell & Rosen 1983; Mitani *et al.* 2006). However, these measures require annotating the fricative portion of the phone. For this study, we used relative intensity, i.e. the difference between the highest intensity in the following vowel and the lowest intensity in the sibilant. Hualde *et al.* (2015) showed that this measure differentiates fricatives and affricates in Catalan. To calculate the relative intensity, we used Praat to extract the maximum intensity in the following phone and the minimum intensity in the sibilant. We did this after discarding the initial and final 5% of each phone and filtering the signal in the same way described above for the CoG measurements (i.e. a 300–19000 Hz pass Hann filter).

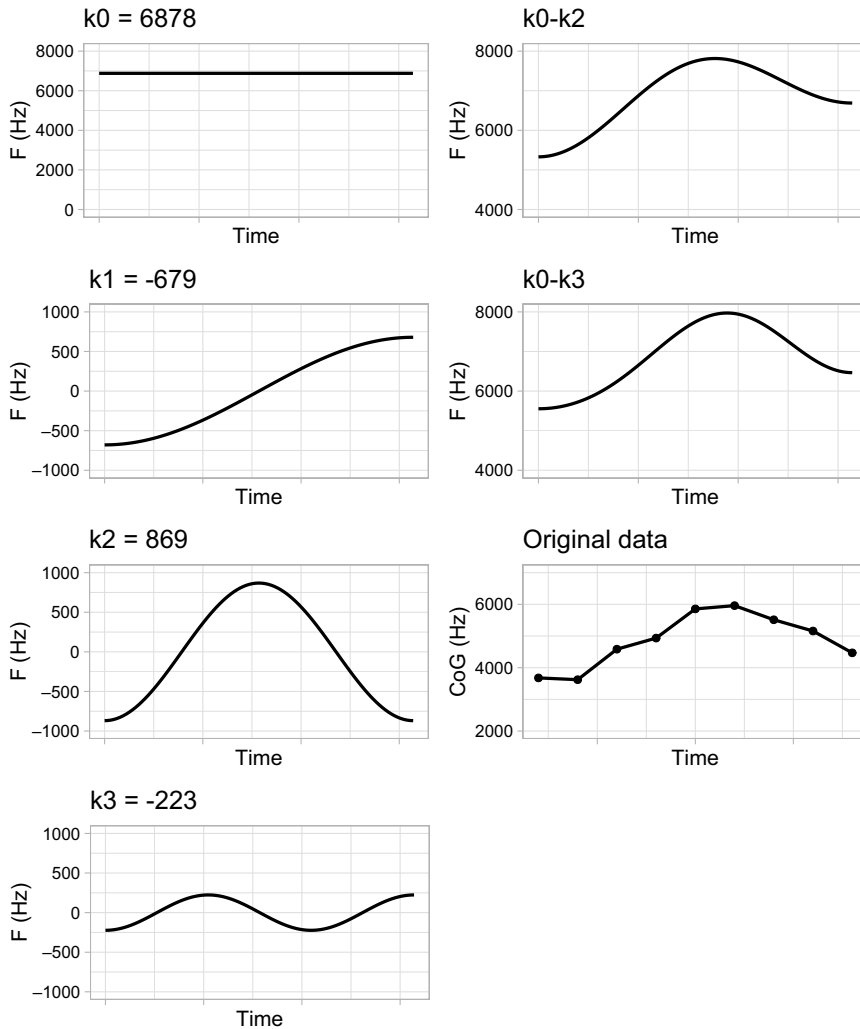


Figure 1. [ts] produced by Speaker 15 (item 1860). Left: plots of DCT coefficients (k0–k3). Right: the sum of k0 to k2 coefficients, the sum of k0 to k3 coefficients and the original nine values of CoG.

4.3 Statistical analyses

For the statistical analyses we used Bayesian mixed-effects models fitted with *brms* (Bürkner 2017). We used weakly informative priors, which means that the influence of extreme values on the posterior is minimized, but the prior does not have a strong influence on the posterior. For each of the relevant factors, we report mean estimates and their 95% credible intervals (CIs). When analyzing contrasts between phones we use the *emmeans* package (Lenth et al. 2023) to obtain median estimates of the differences between pairs of phones based on the posterior distribution, together with highest density intervals (HDI). Then, we calculate the percentage of 89% of HDI contained in the region of practical equivalence (ROPE) using the *bayestestR* package (Makowski et al. 2019). We take the ROPE to be a range of ± 0.1 SDs in the dependent variable. Our decision rule is that if less than 5% of HDI falls within the ROPE we take it as strong evidence for the difference.

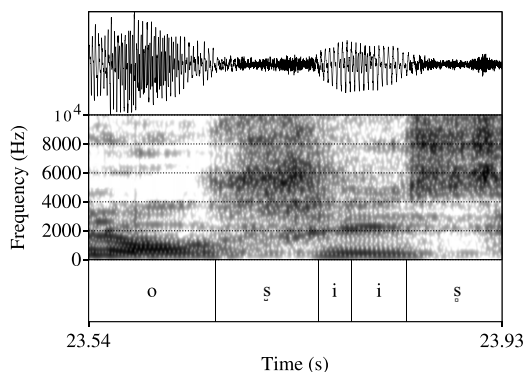


Figure 2. A spectrogram of the sequence [oʃiis̺] from *jelozi izaiteik* 'we didn't have to be jealous' (speaker 10).

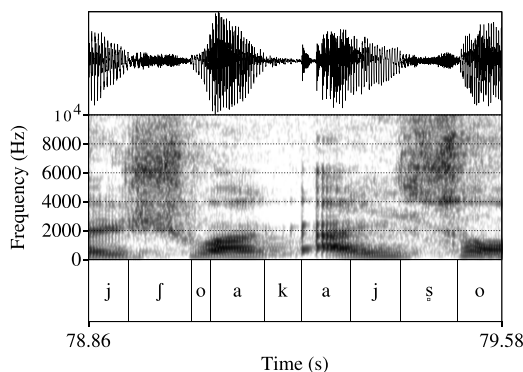


Figure 3. A spectrogram of the sequence [jʃoakaj̺so] from *gaixoak aizorat juiten* 'the poor going to the neighbor' (speaker 10).

Further details of the models, as well as their diagnostics, are included in the supplementary materials.

5 Acoustic properties of Mixean sibilants

5.1 Exploration

A first visual inspection of spectrograms of the Basque sibilant sounds as produced by Mixean speakers shows differences between apico-alveolar and lamino-alveolar voiceless fricatives (Figure 2) and lamino-alveolar and palato-alveolar voiceless fricatives (Figure 3), but no clear discernable difference between apico-alveolar and palato-alveolar voiceless fricatives (Figure 4) or their affricate counterparts (Figure 5). In the following subsection we will present quantitative data to support these observations.

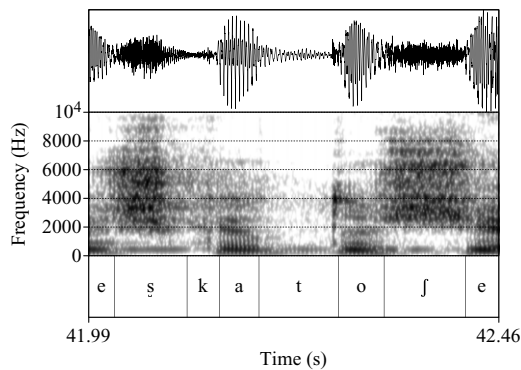
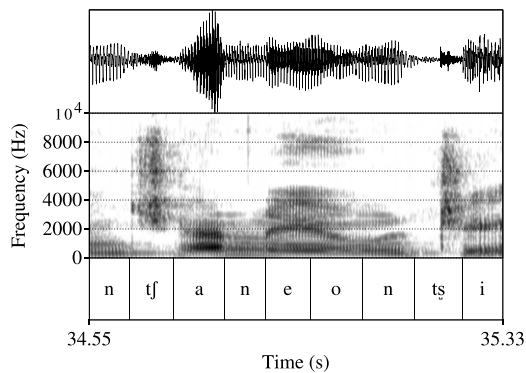
5.2 Spectral moments

Table 2 and Figures 6–9² show values of the spectral moments – CoG, SSD, kurtosis, and skewness – for each target sibilant. Lamino-alveolar voiceless sibilants have the highest

² In all boxplots in the paper the dot represents the mean and the horizontal line the median. The hinges correspond to Q1 and Q3 and whiskers to 1.5 IQR. The notches correspond to ~95% confidence interval for the median.

Table 2. Mean spectral moments (SD is shown between parenthesis)

Phone	N	CoG	SSD	Kurtosis	Skewness
/ɜ̜/	83	3289 (1030)	4733 (1013)	5.4 (4.4)	2.3 (0.7)
/ɜ̝/	45	3498 (898)	3686 (712)	6.7 (4)	2.2 (0.7)
/ɜ̞/	70	3998 (1055)	4010 (674)	4.6 (4.3)	1.8 (0.8)
/d̞ɜ̞/	4	4503 (366)	3864 (299)	3 (1.2)	1.4 (0.4)
/j/	78	4418 (779)	3080 (552)	5.1 (2.6)	1.7 (0.5)
/i̞j/	88	4536 (748)	3035 (497)	5.1 (3.2)	1.7 (0.6)
/s̞/	336	4418 (781)	3258 (495)	4.2 (2.4)	1.6 (0.5)
/t̞s̞/	46	4336 (833)	3167 (455)	4.4 (3.8)	1.6 (0.6)
/s̞̠/	716	5193 (1076)	3608 (549)	2.4 (2.4)	1.1 (0.6)
/t̞s̞̠/	446	5362 (791)	3382 (445)	2.4 (1.8)	1.1 (0.4)

**Figure 4.** A spectrogram of the sequence [ɛskatɔfɛ] from *neskatoxe* 'girl' (speaker 9).**Figure 5.** A spectrogram of the sequence [ntʃaneont̞si] from *laboantxan eontsi* 'be engaged in agriculture' (speaker 10).

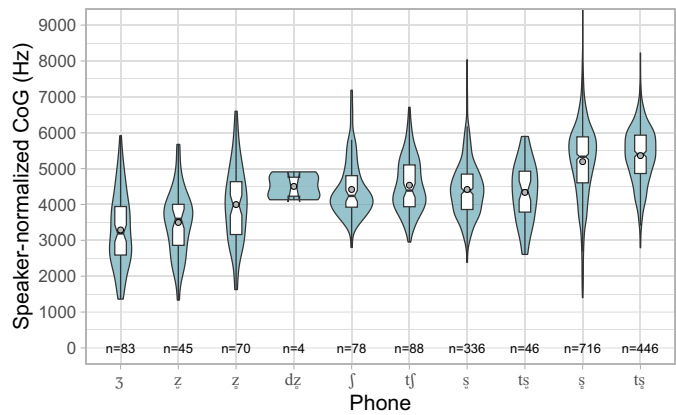


Figure 6. (Colour online) Violin plots and superimposed boxplots of CoG by phone.

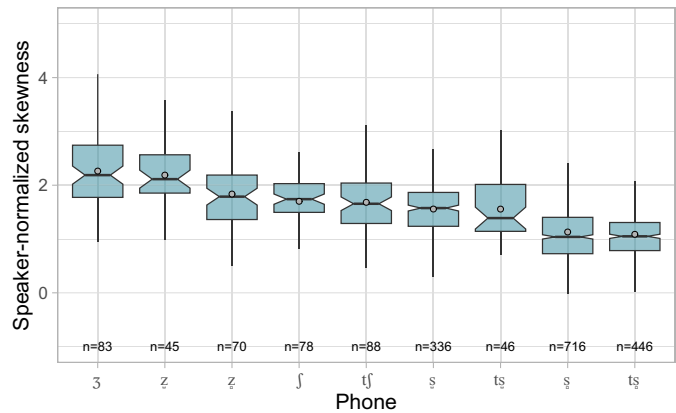


Figure 7. (Colour online) Boxplots of speaker-normalized skewness by phone.

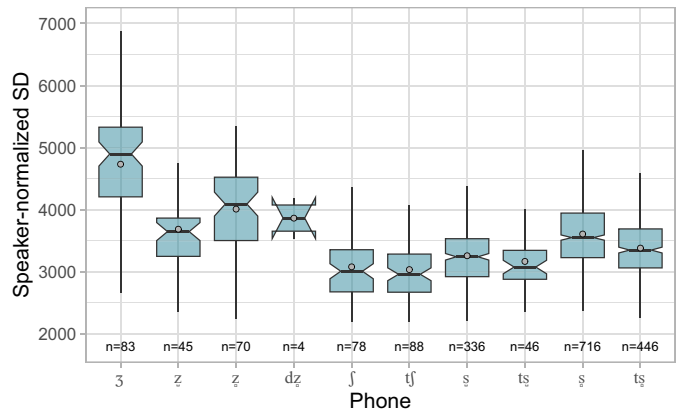


Figure 8. (Colour online) Boxplots of speaker-normalized Spectral Standard Deviation by phone.

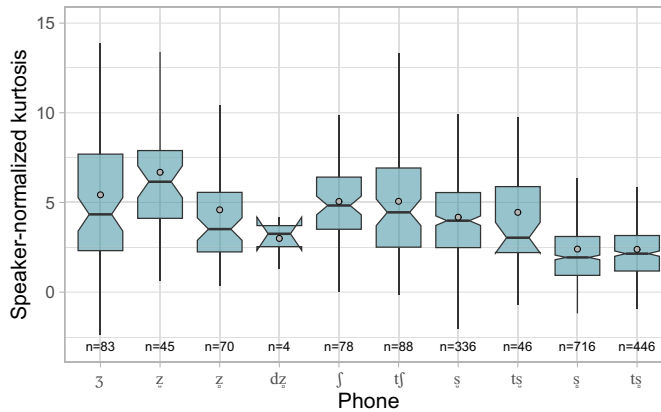


Figure 9. (Colour online) Boxplots of speaker-normalized kurtosis by phone.

mean CoG values (5362 Hz for the affricate and 5193 Hz for the fricative). Apico-alveolar and postalveolar voiceless sibilants have similar values (around 4400 Hz). The lowest CoG is that of voiced fricatives (ranging from 3289 Hz for /ʒ/ to 3998 Hz for /z/). SSD ranges from 3035 Hz for /tʃ/ to 4733 for /ʒ/. Mean kurtosis values are positive and range from 2.4 for laminal voiceless sibilants to 6.7 for /z/. Mean skewness values are also positive, suggesting a concentration of energy in the lower frequencies. The lowest values correspond to voiceless laminal sibilants and the highest to voiced apical and postalveolar sibilants.

A Bayesian linear mixed-effects model was fitted to the data with CoG value as the response variable and phone as the fixed effect. A random effect of word was included, as well as that of speaker with correlated varying slope for the variable phone: $\text{CoG} \sim \text{phone} + (1 | \text{word}) + (\text{phone} | \text{speaker})$. The following priors were specified for the CoG model:

Intercept $\sim \text{Normal}(4800, 1500)$
 $\beta, \tau, \sigma \sim \text{Normal}(0, 1500)$
 By-subject correlation $\sim \text{LKJcorr}(1)$

Results are presented in Table 3. Comparisons between the phones (Table 4) showed that there is no evidence for a difference between /s/ and /f/, on the one hand, and /ts/ and /tʃ/, on the other hand. The same holds for their voiced counterparts (/z/ and /dʒ/).

Models for the remaining spectral moments were also conducted (see supplementary materials for detailed results). None showed evidence of differences between apico-alveolar and postalveolar sibilants. Nevertheless, voiced apico-alveolar and postalveolar sibilants only differ in the second spectral moment, i.e. standard deviation, /ʒ/ having higher values than any other sibilant.

Finally, Figure 10 shows CoG values modeled for all speakers. The lack of difference between apical and postalveolar sibilants is generalized among the speakers.

5.3 Temporal dynamics of CoG

In this section we study the spectral temporal dynamics of sibilants, focusing on CoG values, which were measured in nine equal intervals of the phone. Figure 11 shows CoG trajectories for each phone (the mean for each speaker and the mean for all data). Figure 12 presents the mean CoG values for all data, separately for voiceless fricatives and affricates. It can be seen that:

Table 3. Model's estimates for CoG

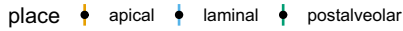
Predictors	Estimate [lower & upper 95% CI]
Intercept ($/s_{\text{g}}/$)	5220 [4733, 5699.1]
phone $/z/$	-1843.2 [-2245.2, -1399.5]
phone $/z_{\text{g}}/$	-1807.7 [-2334, -1284.8]
phone $/z_{\text{g}}/$	-1194.8 [-1619.4, -817.9]
phone $/j/$	-809.5 [-1151.5, -458.3]
phone $/tj/$	-689.4 [-1086.1, -270]
phone $/s_{\text{g}}/$	-815.1 [-1145.4, -477.4]
phone $/ts_{\text{g}}/$	-858.8 [-1219.6, -505.1]
phone $/ts_{\text{g}}/$	173.3 [-78.8, 415.1]

Table 4. Contrasts between the values of PHONE within each manner category

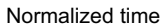
	Estimate	Lower HDI	Upper HDI	ROPE
$/s_{\text{g}}/-/j/$	809.9	466.2	1156.4	0.00
$/s_{\text{g}}/-/s_{\text{g}}/$	817.8	477.4	1145.3	0.00
$/j/-/s_{\text{g}}/$	7.0	-395.8	424.7	0.52
$/tj/-/ts_{\text{g}}/$	-864.5	-1322.0	-404.9	0.00
$/ts_{\text{g}}/-/ts_{\text{g}}/$	-1029.1	-1466.2	-620.9	0.00
$/tj/-/ts_{\text{g}}/$	165.9	-316.5	700.0	0.35
$/z_{\text{g}}/-/z_{\text{g}}/$	-618.4	-1227.4	-0.9	0.04
$/z/-/z_{\text{g}}/$	-34.6	-661.2	594.8	0.35
$/z/-/z_{\text{g}}/$	-653.8	-1198.0	-69.3	0.01

- The trajectories of voiced sounds are flatter than those of voiceless ones (though much less data is available for voiced sibilants).
- The time point of the highest mean CoG value is registered at 55–65% of the interval for both alveolar fricatives, and at 45–55% for the postalveolar fricative. For affricates, it is 65–75% of the whole segment for alveolar phones and 55–65 for the postalveolar.
- As regards the distinction between $/s_{\text{g}}/$ and $/j/$, CoG values are the same for both in the middle portion of the sound, but their overall trajectories appear to have different shapes, with lower values for $/s_{\text{g}}/$ than for $/j/$ in the onset. These higher values make $/j/$ more similar to $/s_{\text{g}}/$ at the onset.

The nine equidistant data-points were reduced to four coefficients using DCT. Recall that DCT coefficients reflect the different properties of the signal curve: k0 corresponds to its



and .95 CI).



represent the pooled mean.

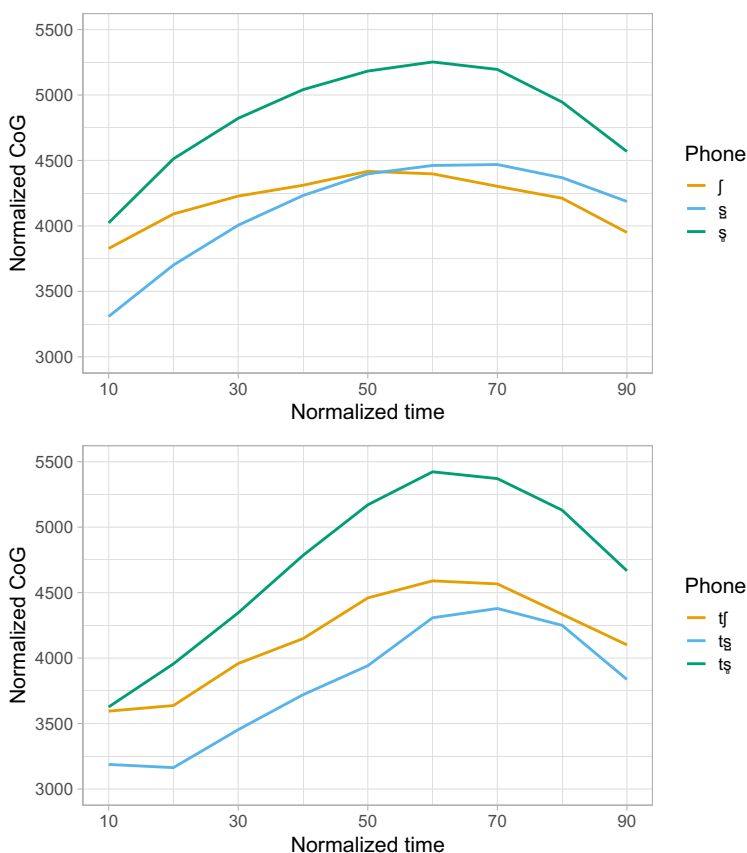


Figure 12. (Colour online) Average CoG values for voiceless sibilants (all data), fricatives plotted in the upper chart and affricates in the bottom one.

mean, k_1 reflects the signal's slope, and k_2 is proportional to its curvature (Harrington 2010: 305). Table 5 shows their mean values for each phone, followed by their standard deviation between parentheses.

Four models were fitted to the data, with k_0 , k_1 , k_2 , and k_3 as the response variable, respectively, all with phone as the predictor and speaker and word as random effects (e.g., $k_0 \sim \text{phone} + (\text{phone} \mid \text{speaker}) + (1 \mid \text{word})$). We provided different priors for each coefficient.

For k_0 :

Intercept $\sim \text{Normal}(6200, 1000)$
 $\beta, \tau, \sigma \sim \text{Normal}(0, 1000)$
 By-subject correlation $\sim \text{LKJcorr}(1)$

For k_1 :

Intercept $\sim \text{Normal}(-200, 500)$
 $\beta, \tau, \sigma \sim \text{Normal}(0, 500)$
 By-subject correlation $\sim \text{LKJcorr}(1)$

For k_2 :

Intercept $\sim \text{Normal}(-300, 300)$
 $\beta, \tau, \sigma \sim \text{Normal}(0, 300)$
 By-subject correlation $\sim \text{LKJcorr}(1)$

Table 5. Mean values of the first four DCT coefficients (SDs in parenthesis)

	k0	k1	k2	k3
/ɜ/	4555.3 (1204.1)	114.3 (593.7)	-46.4 (382.2)	16.8 (280.5)
/ʒ/	4769.6 (1020.3)	-53.4 (420.5)	-210.9 (237.3)	43.5 (221.8)
/z/	5314.2 (1226.3)	-101.9 (414.5)	-267.4 (313.2)	-20 (249.2)
/j/	5929.5 (1030.1)	-67.2 (459.9)	-246.5 (272.9)	7.5 (168.6)
/tʃ/	5875.6 (1073.3)	-365.3 (590.7)	-314.8 (352.6)	104.5 (232.9)
/ʃ/	5834.4 (965)	-404.3 (681.4)	-322.5 (317.5)	-35.7 (222.2)
/tʂ/	5380.3 (837.5)	-528.1 (570.3)	-254.3 (438.3)	166.4 (340)
/ʂ/	6842.1 (1322.4)	-272 (630.3)	-445.5 (366.4)	7.6 (240.7)
/tʂʰ/	6674.3 (1070.6)	-648.4 (845.5)	-520.1 (491.9)	119.9 (347.1)

Table 6. Models of DCT coefficients

	k0		k1		k2		k3	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept (/ɜ/)	6933.99	323.35	-280.97	70.14	-435.32	46.91	0.33	18.85
phone /ɜ/	-2356.57	152.1	434.53	91.25	377.9	49.85	11.11	33.66
phone /ʒ/	-2164.07	185.57	278.49	109.25	233.02	62	30.96	41.9
phone /z/	-1479.93	149.2	177.19	86.25	201.79	50.49	-28.11	34.11
phone /j/	-985.21	143.94	249.36	84.65	196.39	48.15	-0.11	32.53
phone /tʃ/	-1042.81	154.51	-42.15	92.97	114.76	50.37	100.48	34
phone /ʃ/	-1048.06	84.93	-102.76	50.72	140.79	27.97	-49.14	18.89
phone /tʂ/	-1512.3	187.11	-240.72	110.71	175.26	62.22	173.67	42.04
phone /tʂʰ/	-196.66	76.85	-368.78	45.65	-70.73	25.41	135.52	17.16

For k3:

Intercept, β , τ , $\sigma \sim \text{Normal}(0, 200)$

By-subject correlation $\sim \text{LKJcorr}(1)$

The estimates returned by the models are given in Table 6 and the most relevant contrasts (obtained with *emmeans*) are listed in Table 7 (see supplementary materials for all the data).

The main results are the following:

- k0. Results are similar to those obtained with CoG values.
- k1. Positive values of k1 correspond to a negative slope (i.e. the value is lower at the end). For the data analyzed here, estimates are highest for voiced sibilants and /j/, and, except for /ɜ/, all values are negative (i.e. we have an overall increase of CoG). As regards voiceless fricatives, the value is closest to zero for /f/, and it is lower for /ʃ/

Table 7. Most relevant contrasts between phones in the four models (contrasts with ROPE < .05 are in bold)

Contrast	k0		k1		k2		k3	
	Estimate	ROPE	Estimate	ROPE	Estimate	ROPE	Estimate	ROPE
/s̥/-/f/	877.8	0.00	-224.6	0.02	-189.0	0.00	-3.2	0.59
/s̥/-/s̥/	984.5	0.00	56.2	0.47	-113.1	0.00	44.6	0.18
/f/-/s̥/	106.7	0.46	284.3	0.04	75.7	0.25	47.9	0.29
/tʃ/-/tʃ̥/	-744.0	0.00	241.9	0.12	181.4	0.00	-16.3	0.43
/tʃ̥/-/tʃ̥/	-1272.7	0.00	92.1	0.29	257.6	0.00	33.5	0.31
/tʃ/-/tʃ̥/	527.4	0.10	148.9	0.22	-76.1	0.24	-49.9	0.25
/s̥/-/tʃ̥/	158.6	0.42	310.4	0.03	79.2	0.11	-110.7	0.00
/s̥̥/-/tʃ̥̥/	446.9	0.08	162.0	0.22	-67.6	0.27	-189.3	0.00
/f/-/tʃ̥/	25.6	0.36	291.2	0.04	85.8	0.23	-91.5	0.09
/ʒ/-/f/	-1221.4	0.00	177.3	0.18	148.3	0.12	-4.6	0.41
/ʒ/-/s̥/	-1045.8	0.00	278.4	0.06	102.8	0.17	81.0	0.12
/s̥/-/ʒ/	-611.8	0.03	48.2	0.38	48.3	0.33	57.4	0.24

than for /s̥/. In other words, the slope is steeper for /s̥/ than for other voiceless fricatives. For laminal sibilants, the value is more extreme for the affricate than for the fricative. The model shows some evidence of the difference between /f/ and /s̥/ (4% in ROPE), which was not obtained with static CoG measures, as well as that between the laminal and the postalveolar. However, k1 is the same for the laminal and the apical.

- k2. Positive values of this coefficient reflect a u-like trajectory, and negative values correspond to the inverse pattern. Values closer to zero represent flatter trajectories. All estimates are negative for our data. Laminal sibilants show the most arched trajectory. Values do not differ for fricative-affricate pairs. This model, however, shows few significant contrasts between phones.
- k3. This coefficient has not often been used in other phonetic studies, but it might be relevant for the distinction between fricatives and affricates, as it shows low values for fricatives, but higher positive values for all affricates. This appears to correspond to the position of the peak, with higher values having a later peak. The model fitted to our data suggests that k3 distinguished lamino-alveolar and apico-alveolar fricative vs. affricate pairs (but there is less evidence for the difference between the postalveolar fricative and affricate, with 9% in ROPE).

In general, looking at all pairs of sounds, k0 distinguishes more pairs than other coefficients. As regards the contrast between the apical and the postalveolar sounds, k1 is different for the fricatives, but none of the coefficients distinguishes the affricates.

5.4 Voicing

In this section we will present the results of the two cues we used for voicing, AC coefficient (Section 5.4.1) and duration (Section 5.4.2).

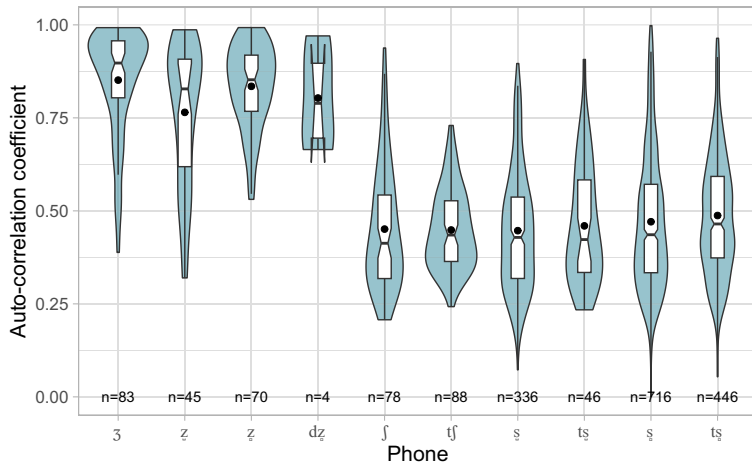


Figure 13. (Colour online) Distribution of the AC coefficients of each sibilant.

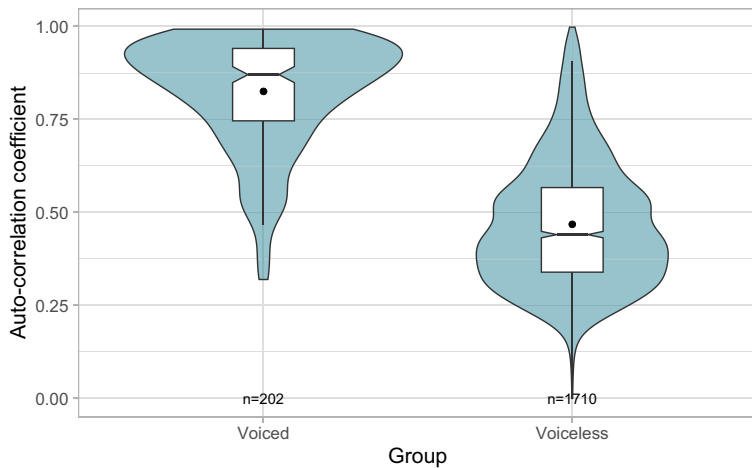


Figure 14. (Colour online) Aggregated distribution of the AC coefficients of voiced vs. voiceless sibilants.

5.4.1 Auto-correlation coefficients

An acoustic analysis of voicing probability based on auto-correlation (AC) values results in a clear distinction between voiced and voiceless sibilants. Note that the zero-to-one scale in the y axis in Figures 13–15 corresponds to fully voiceless (0, no correlation) to fully voiced and completely regular (1, perfect correlation) realizations, with extreme values (i.e. a perfect correlation between glottal cycles) being impossible in real data. Figure 13 illustrates the distribution of the realizations of each of the sibilants separately, while Figure 14 plots the aggregated AC coefficients of all voiced sibilants on the one hand, and all voiceless sibilants on the other.

Figure 15 shows the AC coefficient as a function of normalized time of a voiceless and a voiced lamino-alveolar sibilant. These were the first sibilant sounds in the corpus, produced by the first speaker analyzed for this study (S06, male). They are presented here as a means of illustration of the contrast in clear cases of thorough voicelessness/voicedness.

Table 8. Model's estimates for AC coefficients

Predictors	Estimate [lower & upper 95% CI]
Intercept ($/s/$)	0.48 [0.41, 0.56]
phone $/ʒ/$	0.38 [0.31, 0.45]
phone $/ʒ̥/$	0.26 [0.2, 0.32]
phone $/z/$	0.31 [0.25, 0.39]
phone $/j/$	−0.04 [−0.1, 0.02]
phone $/tj/$	−0.03 [−0.09, 0.02]
phone $/s̺/$	−0.02 [−0.06, 0.02]
phone $/ts̺/$	0 [−0.05, 0.05]
phone $/t̺s̺/$	0 [−0.03, 0.03]

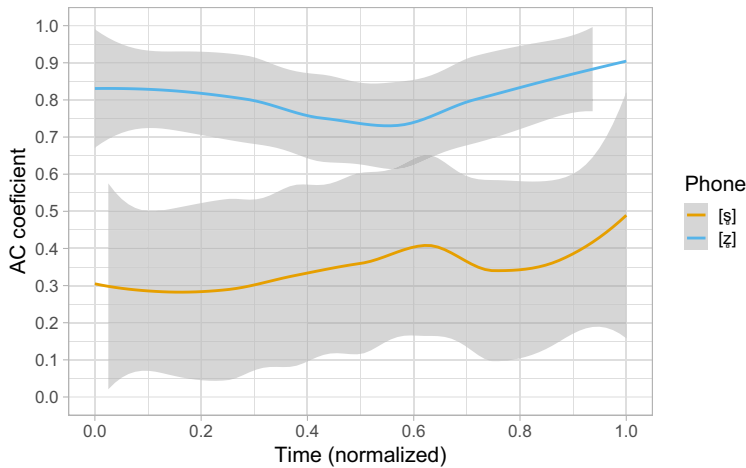


Figure 15. (Colour online) AC coefficient as a function of normalized time of a voiceless and a voiced lamino-alveolar sibilant (produced by the speaker S06).

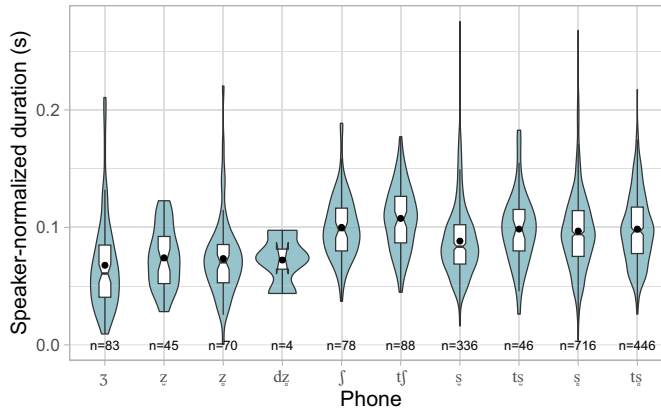
We fitted a Bayesian generalized mixed effects model to the data. The response variable was AC coefficient and phone was the fixed effect. We included word as an intercept-only random effect and speaker as a by-phone correlated varying intercept and slope effect: $AC \sim \text{phone} + (1 \mid \text{word}) + (\text{phone} \mid \text{speaker})$. We specified the following priors in the model:

$$\begin{aligned} \text{Intercept} &\sim \text{Normal}(0.5, 0.25) \\ \beta, \tau, \sigma &\sim (\text{Normal}(0, 0.25)) \\ \text{By-subject correlation} &\sim \text{LKJcorr}(1) \end{aligned}$$

The model estimates are summarized in Table 8. Table 9 shows the contrasts between sibilants with the same place of articulation. The posterior distributions of AC coefficients show a clear distinction between voiced and voiceless sibilants for every place of articulation.

Table 9. Contrasts between fricative values of PHONE within each place of articulation category

	Estimate	Lower HDI	Upper HDI	ROPE
/ʒ/-/ʃ/	0.42	0.34	0.52	0
/s/-/z/	-0.30	-0.38	-0.24	0
/ʒ/-/ʃ/	0.28	0.21	0.36	0

**Figure 16.** (Colour online) Boxplots of speaker-normalized duration by phone.

5.4.2 Duration

In this section we focus on the duration of each phone. Figure 16 shows speaker-normalized duration values by phone. Voiced sibilants appear to be shorter than voiceless sibilants. Within the voiceless sibilants, the apical fricative seems to have a slightly shorter duration.

A Bayesian mixed effects model was fitted to the data, with duration as the response variable, phone as the fixed effect, word as a random-intercept effect and speaker as a correlated varying intercept and slope random effect for phone: $\text{duration} \sim \text{phone} + (1|\text{word}) + (\text{phone}|\text{speaker})$. Unlike the rest of the models, we specified a lognormal family, to better account for the long tails produced by longer duration tokens. We specified the following priors:

Intercept $\sim \text{Normal}(-2.5, 2.5)$
 $\beta, \tau, \sigma \sim \text{Normal}(0, 0.5)$
 By-subject correlation $\sim \text{LKJcorr}(1)$

Table 10 summarizes the estimated duration of each phone (in seconds). Table 11 shows the contrasts between voiced and voiceless phones with the same place of articulation, between voiceless fricative and affricate phones with the same place of articulation, and the contrasts between voiceless postalveolars and apico-alveolars. On average, duration distinguishes between voiced and voiceless fricatives with the same place of articulation, but fails to distinguish fricatives and affricates. Apical fricatives appear to be shorter than other voiceless sibilant phones.

Table 10. Model's estimated mean duration for each phone

Predictors	Estimate [lower & upper 95% CI]
phone /ɜ/	0.06 [0.049, 0.073]
phone /ʒ/	0.073 [0.06, 0.089]
phone /z̥/	0.069 [0.056, 0.085]
phone /j/	0.098 [0.083, 0.114]
phone /tj/	0.101 [0.087, 0.118]
phone /s̥/	0.085 [0.075, 0.097]
phone /ts̥/	0.096 [0.081, 0.115]
phone /s̥/	0.093 [0.082, 0.105]
phone /ts̥/	0.095 [0.083, 0.109]

Table 11. Contrasts between the values of PHONE within each place of articulation category

Contrast	Estimate	Lower HDI	Upper HDI	ROPE
/ɜ/-/j/	-0.48	-0.69	-0.26	0.00
/ʒ/-/s̥/	-0.15	-0.32	0.01	0.01
/s̥/-/z̥/	0.29	0.12	0.47	0.00
/j/-/s̥/	0.14	0.03	0.27	0.00
/tj/-/ts̥/	0.05	-0.11	0.23	0.08
/j/-/tj/	-0.04	-0.19	0.11	0.10
/s̥/-/ts̥/	-0.13	-0.26	0.02	0.02
/s̥/-/ts̥/	-0.02	-0.10	0.06	0.18

5.5 Fricative/affricate distinction

In Section 5.3 we have shown that at least some pairs of fricative and its corresponding affricate sound differ in their temporal dynamics of CoG in that the value of the k3 DCT coefficient is higher for affricates.

In this section we focus on differences in relative intensity, i.e. the lowest intensity value in the sibilant as compared to the highest intensity value in the following vowel. As shown by Hualde et al. (2015), the difference in intensity is higher in affricates than in fricatives, due to the oral closure present in affricates. An example from our data is given in Figure 17, which compares the intensity curves and waveforms for an affricate and a fricative sound.

Figure 18 shows relative intensities for our data taking into account only prevocalic sibilants. As can be seen, affricates have higher values. A Bayesian mixed-effects model was fitted to the data, with relative intensity as the response variable and phone as the fixed effect and with a random intercept for word and a random intercept plus slope for subject:

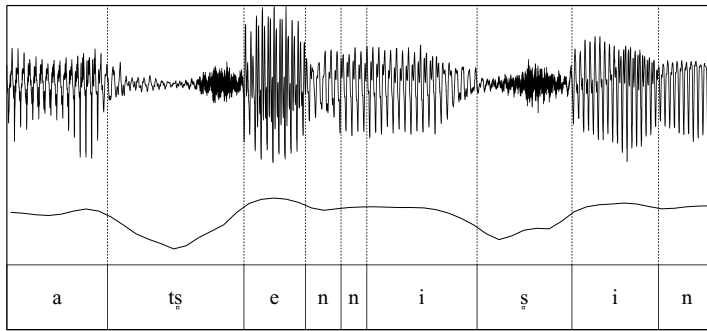


Figure 17. An example of the changes in intensity in an affricate and a fricative sound produced by Speaker 15.

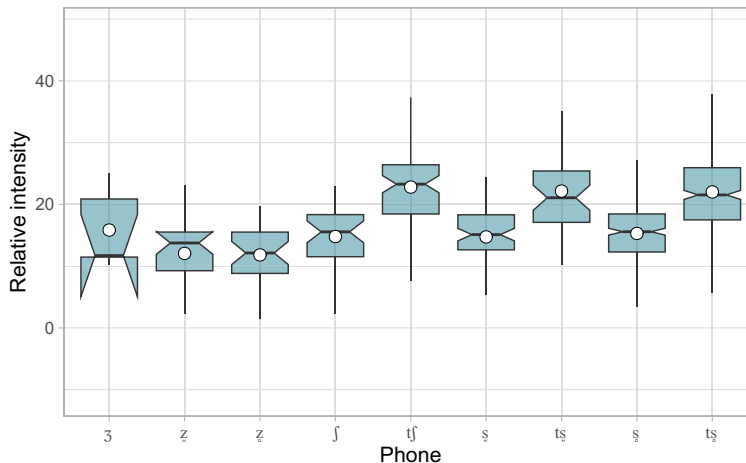


Figure 18. (Colour online) Boxplots of relative intensity by phone for prevocalic sibilants.

relative intensity \sim phone + (1 | word) + (phone | speaker). We specified the following priors:

Intercept \sim Normal(15, 5) β
 $\beta, \tau, \sigma \sim$ Normal(0, 5)
 By-subject correlation \sim LKJcorr(1)

Results are presented in Table 12. The model predicts higher relative intensity for affricates.

6 Discussion

As regards spectral moments, the CoG distinguishes lamino-alveolar sibilants from apico-alveolar and postalveolar sibilants. As in other acoustic studies of Basque sibilants, the CoG of lamino-alveolars is higher than that of other sibilants, which is expected due to the shorter front cavity found in the articulation of lamino-alveolars. We have also taken into account the remaining spectral moments, especially to find out whether they distinguish apico-alveolar and postalveolar sibilants. However, neither measure shows clear differences between those series, either for voiceless and voiced phones. This is in line with

Table 12. Model's estimates for relative intensity

Predictor	Estimate [lower & upper 95% CI]
Intercept (/3/)	16.78 [13.03, 20.51]
phone /z̥/	−4.73 [−8.52, −0.9]
phone /z̥/	−4.59 [−8.26, −0.67]
phone /j/	−2.6 [−6.24, 1.04]
phone /tʃ/	4.75 [1.28, 8.33]
phone /s̥/	−1.78 [−5.36, 1.97]
phone /ts̥/	4.02 [0.58, 7.46]
phone /s̥/	−2.15 [−5.27, 1.06]
phone /ts̥/	4.26 [1.13, 7.5]

Table 13. Comparison between the mean spectral CoG (SD between parenthesis) and model's estimates for the Mixean data and a speaker of High Navarrese

Mixean speakers (aggregated)				High Navarrese (fem. sp., 69 years)		
Phone	N	CoG	Estimate [95% CI]	N	CoG	Estimate [95% CI]
/j/	78	4418 (779)	4411 [3861, 4961]	67	4527 (286)	4536 [4418, 4644]
/tʃ/	88	4536 (748)	4532 [3984, 5059]	50	4922 (352)	4921 [4792, 5055]
/s̥/	336	4418 (781)	4401 [3881, 4994]	70	5316 (490)	5315 [5204, 5421]
/ts̥/	46	4336 (833)	4361 [3792, 4911]	46	5287 (341)	5284 [5146, 5417]
/s̥/	716	5193 (1076)	5223 [4747, 5707]	72	6685 (706)	6682 [6578, 6790]
/ts̥/	446	5362 (791)	5395 [4876, 5960]	45	6967 (393)	6961 [6828, 7102]

Egurtzegi & Carignan’s (2020b) study of the same data set, where, on the basis of CoG values only, a possible merger between apico-alveolar and postalveolar sibilants was proposed for Mixean Basque. These might be the first acoustic documentations of *xexeo* (Hualde 2010), i.e. a merger between apical and postalveolar sibilants in favor of the later due to contact with French. However, the lack of information on the directly preceding stage of the language does not help a suitable comparison of the resulting phone to the previously attested ones.

Nonetheless, differences in CoG between apico-alveolar and postalveolar sibilants have been documented in other varieties, mostly in High Navarrese, though they tend to be smaller than these between other sibilant pairs (Hualde 2010; Urrestarazu-Porta in prep.). As a means of comparison, Table 13 presents the mean and modeled CoG values of a 69-year-old female speaker of High Navarrese from Etxarri-Aranatz (from Urrestarazu-Porta in prep., plotted in Figure 19), alongside the values from Mixean Basque reported in Tables 2 and 3. Note that the High Navarrese measurements were extracted with the same script we used for the current study, and its model was constructed with the same population-level effect and the same priors as the Mixean one, but the experimental conditions and recording devices were different.

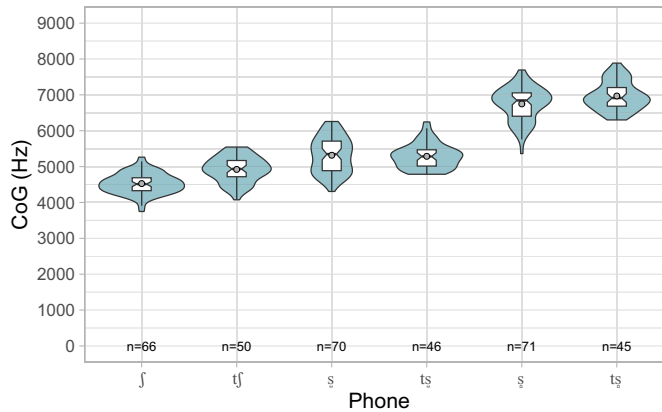


Figure 19. (Colour online) Violin plots and superimposed boxplots of CoG by phone (fem., High Navarrese).

The difference between the mean values of the CoG of / $\text{ʃ}/$ and / $\text{j}/$ is clearly smaller in Mixean (0 Hz) than in High Navarrese (789 Hz), and this is also true for their affricate counterparts (200 Hz in Mixean and 365 Hz in High Navarrese). In addition, we observe higher mean CoG values in / $\text{s}/$ and / $\text{tʂ}/$ in High Navarrese (6682 Hz and 6967 Hz, respectively) than in Mixean (5193 Hz and 5362 Hz, respectively), which is suggestive of a potential compression of the acoustic space of the fricatives and the displacement of lamino-alveolars after a potential merger between apico-alveolar and postalveolar segments.

Given that we are most interested in the difference in CoG between apico-alveolar and postalveolar fricative and affricate sibilants in Mixean (with a potentially lost opposition) and High Navarrese (where the opposition is still in effect), we have computed the contrast distribution between the posterior distributions of apico-alveolar and postalveolar sibilants from the two models' results. To this end, we subtracted the posterior distribution of the CoG of the postalveolar sibilant from that of the apico-alveolar sibilant. The differences between categories are plotted in Figure 20 for each case (fricatives and affricates) and variety (Mixean and High Navarrese). Note that, while the differences are reliably above 0 for the High Navarrese speaker (with a mean difference of 779 Hz for fricatives and 363 Hz for affricates), 0 is close to the center of the distribution in both cases in the Mixean data (with a mean difference of 174 Hz for fricatives and 13 Hz for affricates). While the former points to a clear difference between categories in High Navarrese, the latter suggests a merger between the two categories in Mixean Basque. The narrower distribution of the High Navarrese might be a consequence of the fact that it involves a single speaker as well as the different experimental setting.

In order to better understand the spectral differences between sibilants in Mixean Basque, we also studied the way CoG changes through time (using the DCT of CoG values measured in nine equal intervals of the phone). Recent studies have suggested that the temporal dynamics of frequency values might show divergent patterns between different languages (Reidy 2016 found that Japanese and English / $\text{s}/$ differed in the shape of their trajectories) or between genders or age groups (Stuart-Smith 2020). It is not clear, however, if differences in the way CoG varies throughout the phone can be a cue salient enough to distinguish between sibilants within a given variety.

In general, our results converge with results of other studies. As regards the overall characteristics of the trajectory, Iskarous et al. (2011) found that CoG increases in the first half of English / $\text{s}/$, and that the average increase is around 1500 Hz. Nine 30-ms intervals distributed evenly through each token were used in the study and, in the figures provided

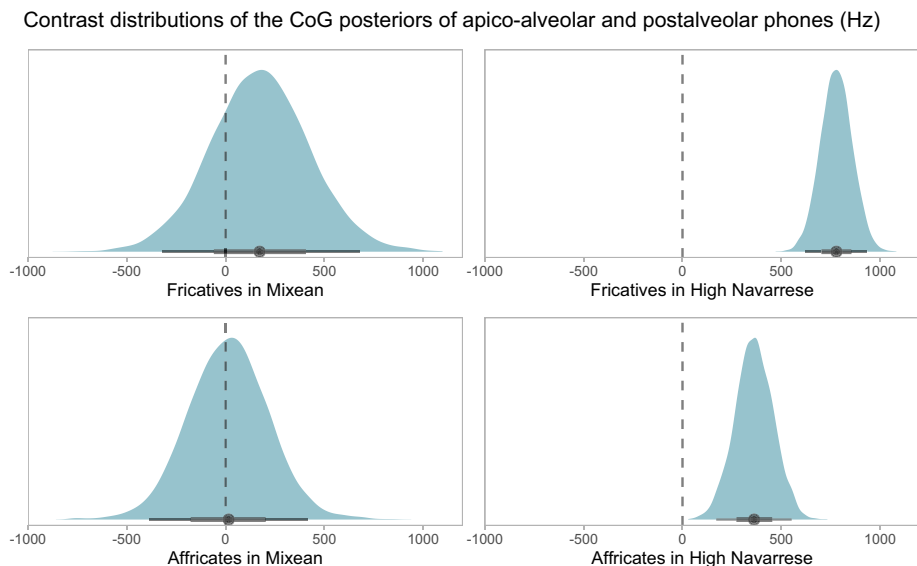


Figure 20. (Colour online) Distribution of the difference between posterior distributions of CoG by manner and variety.

in the paper, the first spectral moment appears to be highest around the sixth or seventh interval of the phone. This is in line with our results: for fricatives, the interval with the highest CoG was 55–65% for both alveolar voiceless fricatives, but 45–55% for the voiceless postalveolar fricative. As expected due to the initial closure, the CoG peak comes later in voiceless affricates; at 65–75% in both alveolar phones and 55–65% in the postalveolar.

As regards DCT coefficients, a study comparable to ours is Stuart-Smith (2020). Stuart-Smith reports that k_1 values are higher for English /*f*/ than for /*s*/. We have found that k_1 coefficients for all analyzed sounds except /*ʒ*/ were negative, which corresponds to a temporally increasing CoG. In our study, k_1 coefficients were also higher for /*f*/ than for /*ʃ*/ and /*ʒ*/, but we did not find any difference between /*ʃ*/ and /*ʒ*/. As for k_2 , in Stuart-Smith's (2020) study the curvature was similar for males, but females showed a more pronounced curvature for /*s*/ than for /*f*/. For the Basque data, k_2 is higher (thus pointing to a more curved pattern) for laminal phones than for the rest. Finally, the k_3 coefficient, which is related to the position of the peak CoG value, has proved relevant to distinguish between fricatives and affricates.

With regard to voicing, our study, based on auto-correlation coefficients, points to the distinction between voiced and voiceless sibilants being consistent in Mixean Basque. Besides AC coefficients, raw duration appears to be a reliable metric to distinguish Mixean voiced and voiceless sibilants with the same place of articulation. Voiced sibilants are around 30 ms shorter than voiceless sibilants.

In addition to the duration differences between voiced and voiceless sibilants, we found /*ʃ*/ to be shorter than /*f*/. This result points in the direction of duration potentially being another cue speakers might use to distinguish between these two segments alongside the difference in k_1 . Nonetheless, this result should be taken cautiously. Firstly, the difference is small. Some studies show speakers rely on duration for categorizing shorter voiceless fricatives as being voiced and longer as voiceless (Cole & Cooper 1975; Widdison 1995). However, the difference in duration between voiced and voiceless sibilants is around 30 ms, i.e. voiceless sibilants are almost 50% longer than their voiced counterparts or, conversely, voiced sibilants are around 40% shorter than voiceless sibilants. In our study, the median estimate

of the posterior distribution of /s̺/ is just around 14 ms shorter than the median estimate of the posterior distribution of /f/, and a difference between 30 and 3 ms is highly credible given our data. We are not aware of any study that concludes that speakers are able to only rely on such a small difference in duration to distinguish between a pair of voiceless sibilants.

Secondly, we used the raw duration of naturalistic speech. This means that we did not control for some factors that are known to affect duration, such as speech rate. We assumed each person's fluctuations of speech rate would be distributed normally and, thus, including the correlated varying intercept and slope effect for speaker would account for much of the variation produced by speech rate. However, the number of tokens for some phones was limited and there might be an effect related to speech rate that was unaccounted for.

Thirdly, /s̺/ and /f/ are not distributed evenly across the language. In our data, 47% of the occurrences of /f/ are between vowels (V.fV), 21% are word-initial (#fV), 17% are in syllable onsets preceded by a consonant (VC.fV) and 16% are in non-final syllable-codas (Vf.CV). In turn, 59% of /s̺/ in our data are in non-final syllable-codas (V̺s̺.CV), 25% are between vowels (V̺s̺V), 8% are in syllable-onsets preceded by a consonant (VC̺s̺V), 7% are word-initial (#̺s̺V) and around 1% are word-final (V̺s̺#). We speculate that the shorter estimation of /s̺/ may be due to it occurring mostly in coda position followed by a consonant.

Finally, the difference between prevocalic fricative and affricate sibilants seems to be appropriately accounted for through the analysis of relative intensity. In line with the results in Hualde et al. (2015), the difference in intensity was found to be higher in affricates than in fricatives, which is linked to the oral closure phase of the affricates.

In short, we can reliably distinguish between seven sibilants in Mixean Basque, but the acoustic difference between /s̺, t̺s̺, ʒ̺/ and /f, t̺f, ʒ̺/, respectively, observed in our data might not be enough to speak of a perceptible distinction. Our results call for further perceptual and articulatory analyses that would corroborate our findings. However, the linguistic landscape of Amiküze has changed dramatically since the 1970s; the transmission of Mixean Basque was almost entirely interrupted and French has replaced Basque almost completely for most social interactions (Camino 2016). Thus, it may be an impossible endeavor to find Mixean speakers with a variety of the language comparable to the one analyzed in our study. This highlights the urgent need for the documentation and study of other endangered varieties and languages.

7 Conclusion

This paper has presented a detailed study of the sibilant system of Mixean Basque, a variety which has been described with ten sibilants: /s̺, s̺, f, t̺s̺, t̺s̺, t̺f̺, ʒ̺, ʒ̺, ʒ̺, d̺ʒ̺/. Our results suggest that the phonological oppositions based on manner of articulation – i.e. voiced vs. voiceless and fricative vs. affricate – are still part of this variety, but two of the three historical places of articulation are not easily phonetically distinguishable in the modern language. While the lamino-alveolar series (/s̺, t̺s̺, ʒ̺/) reliably contrasts with the rest, only minimal differences between the apico-alveolar series (/s̺, t̺s̺, ʒ̺, d̺ʒ̺/) and the postalveolar series (/f̺, t̺f̺, ʒ̺/) have been found in the analysis of dynamic data and duration, and none in the static analysis. A difference in CoG in the onset of apico-alveolars vs. postalveolars and a small difference in duration might not be enough to reliably differentiate the two sets of segments today, but it can potentially be considered a historical remnant of what was a fully working opposition in earlier generations. Our analysis reminds us that studying synchronic acoustic data can lead to interesting historical observations, and help us better understand the gradual nature of sound change in general, and segmental mergers in particular.

Acknowledgments We are highly indebted to Iñaki Camino for letting us use his recordings, and two anonymous reviewers as well as the associate editor of JIPA for their many comments and suggestions, which have helped improve the final result of this paper. All remaining errors are ours. This research was partially funded by Modern approaches to diachronic phonology applied to Basque (MADPAB) (ANR-20-CE27-0007), Monumenta Linguae Vasconum VI (PID2020-118445GB-I00), and Diachronic Linguistics, Typology and the History of Basque (IT1534-22).

Author contribution Both AE and DK have contributed to this paper equally and should be considered first authors.

Replication material The measurements and code used for this research can be found at <https://osf.io/vctgw/>.

References

- Beristain, Ander. 2018. *Basque dialectal substrate in the realization of /s/ in L2 Spanish*. MA thesis, University of Illinois at Urbana-Champaign.
- Beristain, Ander. 2021. 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). 1–32. <https://doi.org/10.1017/S0025100320000274>.
- Blevins, Juliette, Ander Egurtzegi & Jan Ullrich. 2020. Final obstruent voicing in Lakota: Phonetic evidence and phonological implications. *Language* 96(2). 294–337. <https://doi.org/10.1353/lan.2020.0022>.
- Boersma, Paul & David Weenink. 2023. Praat: Doing phonetics by computer. <http://www.praat.org/>.
- Bürkner, Paul-Christian. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80. 1–28. <https://doi.org/10.18637/jss.v080.i01>.
- Camino, Iñaki. 2016. *Amiküze eskualdeko (h)eskuara*. Pamplona & Bilbao: Nafarroako Gobernua & Euskaltzaindia.
- Cole, Ronald A. & William E. Cooper. 1975. Perception of voicing in English affricates and fricatives. *The Journal of the Acoustical Society of America* 58(6). 1280–1287. <https://doi.org/10.1121/1.380810>.
- Dart, Sarah Northrop. 1991. *Articulatory and acoustic properties of apical and laminal articulations*. PhD thesis, University of California, Los Angeles. <https://www.proquest.com/docview/303942751/abstract/2D991346B41A4FBEPQ/1>.
- Egurtzegi, Ander. 2013. Phonetics and phonology. In Mikel Martínez-Areta (ed.), *Basque and proto-Basque. Language-internal and typological approaches to linguistic reconstruction*, 119–173. Frankfurt am Main: Peter Lang.
- Egurtzegi, Ander & Christopher Carignan. 2020a. A typological rarity: The /h/ versus /ħ/ contrast of Mixean Basque. In: The University of British Columbia, Vancouver (held online).
- Egurtzegi, Ander & Christopher Carignan. 2020b. An acoustic description of Mixean Basque. *The Journal of the Acoustical Society of America* 147(4). 2791–2802. <https://doi.org/10.1121/10.0000996>.
- Gandarias, Leire, Jone Plaza & Iñaki Gaminde. 2014. Lekeitioko txistukariez: frikariak eta afrikatuak. *Euskalingua* 24. 6–21.
- Gordon, Matthew, Paul Barthmaier & Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32(2). 141–174. <https://doi.org/10.1017/S0025100302001020>.
- Harrington, Jonathan. 2010. *Phonetic analysis of speech corpora*. Chichester: Wiley-Blackwell.
- Harrington, Jonathan, Felicitas Kleber & Ulrich Reubold. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America* 123(5). 2825–2835. <https://doi.org/10.1121/1.2897042>.
- Harrington, Jonathan & Florian Schiel. 2017. /u/-fronting and agent-based modeling: The relationship between the origin and spread of sound change. *Language* 93(2). 414–445. <https://doi.org/10.1353/lan.2017.0019>.
- Howell, Peter & Stuart Rosen. 1983. Production and perception of rise time in the voiceless affricate/fricative distinction. *The Journal of the Acoustical Society of America* 73(3). 976–984. <https://doi.org/10.1121/1.389023>.
- Hualde, José Ignacio. 2003. Segmental phonology. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 15–64. Berlin: De Gruyter Mouton.
- Hualde, José Ignacio. 2010. Neutralización de sibilantes vascas y seseo en castellano. *Oihenart* 25. 89–116.
- Hualde, José Ignacio. 2014. *Los sonidos del español*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511719943>.
- Hualde, José Ignacio, Christopher D. Eager & Marianna Nadeu. 2015. Catalan voiced prepalatals: Effects of nonphonetic factors on phonetic variation? *Journal of the International Phonetic Association* 45(3). 243–267. <https://doi.org/10.1017/S0025100315000031>.
- Hualde, José Ignacio, Oihana Lujanbio & Juan Joxe Zubiri. 2010. Goizueta Basque. *Journal of the International Phonetic Association* 40(1). 113–127. <https://doi.org/10.1017/S0025100309990260>.

- Iglesias, Aitor, Leire Gandarias & Lorea Unamuno. 2016. Euskararen txistukariak aztertze indize akustikoez. *Euskalingua* 28. 6–18.
- Iskarous, Khalil, Christine H. Shadle & Michael I. Proctor. 2011. Articulatory-acoustic kinematics: The production of American English /s/. *The Journal of the Acoustical Society of America* 129(2). 944–954. <https://doi.org/10.1121/1.3514537>.
- Jassem, Wiktor. 2003. Polish. *Journal of the International Phonetic Association* 33(1). 103–107. <https://doi.org/10.1017/S0025100303001191>.
- Jongman, Allard, Rاتree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America* 108(3). 1252–1263. <https://doi.org/10.1121/1.1288413>.
- Jurado, Mirari. 2011. Caracterización de sibilantes fricativas vascas y su percepción en el sistema fonético español. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 45(1). 81–137. <https://doi.org/10.1387/asju.9727>
- Kisler, Thomas, Uwe Reichel & Florian Schiel. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326–347. <https://doi.org/10.1016/j.csl.2017.01.005>.
- Kluender, Keith R. & Margaret A. Walsh. 1992. Amplitude rise time and the perception of the voiceless affricate/fricative distinction. *Perception & Psychophysics* 51(4). 328–333. <https://doi.org/10.3758/BF03211626>.
- Kochetov, Alexei. 2017. Acoustics of Russian voiceless sibilant fricatives. *Journal of the International Phonetic Association* 47(3). 321–348. <https://doi.org/10.1017/S0025100317000019>.
- Ladefoged, Peter. 2001. *Vowels and consonants: An introduction to the sounds of languages*. Malden, MA: Blackwell.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Lafon, René. 1999. Contribution à l'étude phonologique du parler basque de Larrau (Haute-Soule). In Jean Haritschelhar & Pierre Charritton (eds.), *Vasconiana*, 113–133. Bilbao: Euskaltzaindia.
- Larrasquet, Jean. 1932. Phonétique du basque du Larrajá (quartier du Barcus). *Revista Internacional de los Estudios Vascos* 23(1). 153–191.
- Larrasquet, Jean. 1934. *Le basque souletin nord-oriental*. Paris: Floch. <http://catalog.hathitrust.org/api/volumes/oclc/12866856.html>.
- Lenth, Russell V., Paul Buerkner, Iago Giné-Vázquez, Maxime Herve, Maarten Jung, Jonathon Love, Fernando Miguez, Hannes Riebl & Henrik Singmann. 2023. emmeans: Estimated Marginal Means, aka Least-Squares Means. <https://CRAN.R-project.org/package=emmeans>.
- L'Institut national de la statistique et des études économiques. 2019. Populations légales 2016. <https://www.insee.fr/fr/statistiques/zones/3681328>.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
- Makowski, Dominique, Mattan S. Ben-Shachar & Daniel Lüdtke. 2019. bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software* 4(40). 1541. <https://doi.org/10.21105/joss.01541>.
- Mitani, Shigeki, Toshihiro Kitama & Yu Sato. 2006. Voiceless affricate/fricative distinction by frication duration and amplitude rise slope. *The Journal of the Acoustical Society of America* 120(3). 1600–1607. <https://doi.org/10.1121/1.2221390>.
- Mitxelena, Koldo. 2011 [1977]. Fonética histórica vasca. In Joseba A. Lakarra & Iñigo Ruiz Arzalluz (eds.), *Obras completas*, vol. VI. Bilbao: UPV/EHU.
- Moran, Steven & Daniel McCloy (eds.). 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. <https://phoible.org/>.
- Muxika-Loitzate, Oihane. 2017. Sibilant merger in the variety of Basque spoken in Amorebieta-Etxano. *Languages* 2(4). 25. <https://doi.org/10.3390/languages2040025>.
- N'Diaye, Geneviève. 1970. *Structure du dialecte basque de Maya*. The Hague: Mouton. <https://doi.org/10.1515/9783111349725>.
- R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reidy, Patrick F. 2016. Spectral dynamics of sibilant fricatives are contrastive and language specific. *The Journal of the Acoustical Society of America* 140(4). 2518–2529. <https://doi.org/10.1121/1.4964510>.
- Rodrigues, Alexandra Soares. 2022. Phonotactic conditions and morphotactic transparency in Mirandese word formation. *Folia Linguistica* 56(1). 87–122. <https://doi.org/10.1515/flin-2021-2005>.
- Smith, Caroline L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25(4). 471–500. <https://doi.org/10.1006/jpho.1997.0053>.
- Stuart-Smith, Jane. 2020. Changing perspectives on /s/ and gender over time in Glasgow. *Linguistics Vanguard* 6(s1). <https://doi.org/10.1515/lingvan-2018-0064>.
- Txillardegi. 1982. Some acoustic data about the three Basque sibilants. In *Proceedings of the First International Basque Conference in North America*, 18–34. Fresno, CA: California State University.
- Urrestarazu-Porta, Iñigo. in preparation. The acoustic properties of Basque sibilants.

- Urrutia, Hernán, Maïtena Etxebarria, Itziar Túrrez & Juan Carlos Duque. 1991. *Fonética vasca 3: Las sibilantes en los dialectos orientales*. Bilbao: Universidad de Deusto.
- Widdison, Kirk A. 1995. The perception of voicing in Spanish sibilants. In *Proceedings of the 4th European Conference on Speech Communication and Technology (Eurospeech 1995)*, 2289–2292. <https://doi.org/10.21437/Eurospeech.1995-521>
- Winkelmann, Raphael, Klaus Jaensch, Steve Cassidy & Jonathan Harrington. 2021. *emuR: Main Package of the EMU Speech Database Management System*.
- Yárnnoz, Belén. 2002a. *Sibilants in the Basque dialect of Bortziri: An acoustic and perceptual study*. Pamplona: Gobierno de Navarra.
- Yárnnoz, Belén. 2002b. Descripción de las sibilantes vascas mediante el parámetro Tongue shape. *Euskalingua* 1. 25–31.
- Zabalik. 2016. Euskara gure eskoletan: dena eta ez dena. *Zabalik. Amikuzeko euskalgintza*. <http://zabalik-amikuze.eus/euskara-gure-eskoletan-dena-eta-ez-dena/>.
- Żygis, Marzena. 2003. Phonetic and phonological aspects of Slavic sibilant fricatives. *ZAS Papers in Linguistics* 32. 175–213. <https://doi.org/10.21248/zaspil.32.2003.191>.
- Żygis, Marzena, Susanne Fuchs & Laura L. Koenig. 2012. Phonetic explanations for the infrequency of voiced sibilant affricates across languages. *Laboratory Phonology* 3(2). 299–336. <https://doi.org/10.1515/lp-2012-0016>.