UNIVERSITAT DE BARCELONA

Estudios de Fonética Experimental
Journal of Experimental Phonetics
ISSN: 1575-5533 · ISSN-e: 2385-3573

EFE
Estudios de Fonética Experimental
Journal of Experimental Phonetics

Lenition of fricative sibilants in casual conversations in Basque

Dorota Krajewska [1] 0000-0002-6909-985X

[1] University of the Basque Country UPV/EHU (Spain)

DOI: 10.1344/efe-2024-33-91-120

Corresponding address: dorota.krajewska@ehu.eus

Received: 11/04/2024 Accepted: 10/05/2024 Published: 08/07/2024

Krajewska, D. (2024). Lenition of fricative sibilants in casual conversations in Basque. *Estudios de Fonética Experimental*, *33*, 91–120. https://doi.org/10.1344/efe-2024-33-91-120

Abstract

This paper analyses lenition of lamino-alveolar and apico-alveolar fricative sibilants in Basque in casual conversations between speakers of the variety of Beasain (Central Basque dialect). To describe the phonetic realisation of sibilants, the following measures were used: the proportion of voiced frames, the centre of gravity and relative intensity. Results show that 13% of sibilants can be classified as 'voiced' (i.e. they show uninterrupted voicing during the middle 50% of the duration) and another 25% has at least one voiced frame in the middle interval. Several factors were identified as important predictors of voicing: context, speech rate and the presence of the word boundary. It is also shown that voicing lowers the sibilants' centre of gravity and relative intensity. Finally, the paper discusses potential lexical effects in lenition phenomena.

Keywords

lenition; sibilants; voicing; Basque

Lenició de sibilants fricatives en converses informals en basc

RESUM

Aquest article analitza la lenició de les sibilants fricatives laminoalveolars i apicoalveolars en converses informals entre parlants de basc de Beasain (dialecte central). La realització fonètica de les sibilants es descriu a partir de la proporció d'intervals sonors, el centre de gravetat i la intensitat relativa. Els resultats mostren que el 13% de les sibilants es poden classificar com a "sonores" (i.e., sonoritat ininterrompuda al llarg del 50% central de la seva durada) i un altre 25% té com a mínim un *frame* sonor en l'interval central. Diversos factors constitueixen predictors importants de la sonoritat: el context, la velocitat de parla i la presència d'un límit de mot. També es mostra que la sonoritat redueix el centre de gravetat i la intensitat relativa de les sibilants. Finalment, l'article analitza els possibles efectes lèxics en els fenòmens de lenició.

MOTS CLAU

lenició; sibilants; sonoritat; basc

1. Introduction[1]

This paper analyses the phonetic realisation of fricative sibilants in casual speech in the Central Basque variety of Beasain. The lamino-alveolar /s̻/ and the apico-alveolar /s̺/ are phonologically voiceless, as are all Basque fricatives (in most varieties of the language). However, they can be realised as partially or totally voiced, especially in the vicinity of voiced sounds. This can be seen as a sign of lenition, a process in which sounds are realised in a more relaxed manner. In the case of obstruents, lenition typically involves shorter duration, wider constriction, lesser intensity and often also voicing.

Sibilants are among the topics which have received most attention in studies of Basque phonetics, but research has focused especially on the acoustic properties (e.g. the centre of gravity) of the different phonological categories and the acoustic consequences of various merger phenomena. Anticipatory voicing before a voiced consonant is often mentioned in the descriptions of the language, but has not been studied acoustically in detail. The existence of intervocalic voicing has been noted in Urrutia et al. (1988, 1989) and Hualde et al. (2019a, p. 97), but no study has focused on the topic.

The main goal of this exploratory paper is to contribute to the understanding of the acoustic properties of sibilants in Basque, and to the more general topic of lenition of Basque consonants, an issue which has recently received attention in Basque phonetics with reference to voiceless stops (Eguskiza et al., 2020; Hualde et al., 2019a, 2019b; Nadeu & Hualde, 2015).

The paper is organised as follows. Section 2 introduces the background on lenition and sibilants, both in general and with reference to Basque. Section 3 lists research questions and main predictions. Section 4 describes the methodological aspects of the study. Section 5 presents the results. The paper closes with discussion in Section 6 and conclusions in Section 7.

2. Preliminaries

This section introduces the phenomenon of lenition (Section 2.1). Then, Section 2.2 deals with fricatives taking into account their production and the conditions necessary to produce turbulent noise and phonation at the same time. Research on lenition of Spanish sibilants is summarised in Section 2.3. Basque sibilants and lenition phenomena described for the language are introduced in Section 2.4.

2.1. Lenition

Even though the label ‘lenition’ has been used to talk about a wide range of phenomena, most approaches to lenition of consonants agree that it can be understood as “reduction in constriction degree or duration” (Kirchner, 2004, p. 313).[2] Phenomena typically treated as consonant lenition include degemination, flapping, reduction to an approximant, debuccalisation, voicing and elision at its most extreme end (Kirchner 2004, p. 313).

Lenition of stops is the most common type of lenition (Gurevich, 2011, p. 3) and it has been extensively studied for various languages (e.g., Broś et

[1] 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/AEI/10.13039/501100011033) and Diachronic Linguistics, Typology and the History of Basque Research Group (IT1534-22). Data, code and supplementary materials are available at https://osf.io/9eaqz/.

[2] See Honeybone (2008) for a discussion of the history of the term and Bauer (2008) for a discussion of the problems with definition of lenition. Another frequently used term is ‘reduction’. According to DiCanio et al. (2022, p. 2), “The terms *speech reduction* and *lenition* are often used interchangeably in the speech production literature, the former being more common in phonetics and the latter being more common in describing discrete phonological patterns or processes of historical sound change”. Also, as noted by one of the reviewers, ‘lenition’ is most commonly used to describe the weakening of consonants, whereas for many researchers ‘reduction’ may also include the centralization and weakening of vowels.

al., 2021; Ennever et al., 2017; Hualde et al., 2010, 2011; Hualde & Zhang, 2022; Katz & Pitzanti, 2019; Kingston, 2008; Nadeu & Hualde, 2015; Torreira & Ernestus, 2011). Intervocalic voiced stops might, for example, be realised as approximants (as, e.g., in Spanish; see e.g. Hualde et al., 2011). Voiceless stops, in turn, may become (at least partially) voiced and can be realised, depending on the language, as voiced stops, fricatives or approximants (Gurevich, 2011; Hualde, 2014). The common characteristics of the weakened variants in both cases (voiceless and voiced stops) is that they have higher intensity and shorter duration compared to their non-weakened counterparts. Thus, they become more similar to the surrounding vowels.

Such phenomena affecting intervocalic obstruents can be, following Katz (2016), classified under the label of 'continuity lenition'. He distinguishes between 'continuity lenition' and 'loss lenition'. The former occurs in perceptually robust positions and does not normally trigger loss of phonological contrasts. The latter targets segments in positions where they might be perceptually difficult to distinguish and might lead to neutralisations or mergers. For example, it has been observed that stop lenition in Spanish and Italian does not cause the two series /p t k/ and /b d g/ to merge, despite some phonetic overlap (Hualde et al., 2011; Hualde & Nadeu, 2012).

Various linguistic factors have been shown to favour lenition: the most often mentioned are speech style, duration and rate of speech, prosodic factors and lexical factors.

With regards to the rate of speech and style, lenition is more likely to occur at faster speech rates and in more spontaneous productions (as opposed to careful pronunciation) (DiCanio, 2012; Hualde & Zhang, 2022; Lewis, 2001; Warner & Tucker, 2011). More generally, shorter duration of weakened consonants has been observed in most studies dealing with lenition (e.g. Broś et al., 2021; DiCanio et al., 2022; Katz, 2016; Torreira & Ernestus, 2011). The most frequently given explanation is that, as the duration of the consonant decreases, the speaker might not reach the articulatory target, and this might cause, for instance, an incomplete closure instead of a full one. Some studies conclude that duration is the most important factor which allows to predict lenition (e.g. Katz & Pitzanti, 2019). Cohen Priva and Gleason (2020) identified changes in duration as the most important reason of lenition.

As for prosodic factors, a commonly observed regularity is that consonants in stressed syllables are less likely to lenite (e.g. Torreira & Ernestus (2012) for fricatives, and Lewis (2001) and Torreira & Ernestus (2011) for stops in Spanish). Moreover, word-initial or phrase-initial consonants are often found to be less prone to show signs of lenition. This might be to some extent explained by longer duration of word-initial consonants, which was observed for various languages (e.g. Keating et al., 2004). Language-specific factors might also be important: for example, consonants might lenite more in word-final position if the relevant phonological contrast is neutralised in this position (Hualde et al., 2019b). DiCanio et al. (2022) also emphasize that language-specific stress patterns might be responsible for some lenition patterns (in their study of Yoloxóchitl Mixtec, word-initial consonants were not less likely to weaken). Ennever et al. (2017), concluded that, after controlling for the effect of duration, medial consonants were not more likely to show lenition that initial consonants in Gurindji. The effects of boundaries are therefore complex and language-dependant.

Lexical effects have also been found in lenition phenomena, with more frequent words or phrases and grammatical words typically showing stronger weakening (Bybee et al., 2016; Warner & Tucker, 2011).

Extralinguistic factors can also play a role in lenition phenomena. For example, gender has been found to be relevant in many studies. Some authors argue that the fact that men often lenite more might be related to physiological factors (see Nadeu & Hualde, 2015 for an anysis of the lenition of voiceless stops), but in some studies it has been proposed that there might be a sociolinguistic explanation, for some varieties at least (Chappell et al., 2023; Chappell & García, 2017).

2.2. Articulation and phonation of fricatives

Lenition of fricatives has attracted less attention than that of stops: it is not as common, in part due to the nature of fricative sounds, and, in particular, to the conditions necessary to produce turbulence. At least for some languages, though, lenition of fricatives can be seen as part of a larger phenomenon, that of lenition of voiceless consonants (for example; for Spanish, Hualde & Prieto 2015, p. 110).

As explained by Ohala and Solé (2010), in order to produce the turbulent airflow present in fricative sounds, multiple conditions have to be met, the most important being the speed of air passing through the oral constriction. It also has to be taken into account that the volume of the air which passes through a constriction depends on the size of the aperture and the difference in pressure on both sides (Ohala & Solé 2010). For the frication noise to have high amplitude, there pressure before the oral constriction must be higher than in the atmosphere. However, for voiced fricatives the constriction at the glottis must also be considered, and the same applies: the pressure must be higher in the subglottal cavity than in the supraglottal cavity. Voiceless obstruents have a larger glottal aperture and the flow of air is continuous. In the production of voiced obstruents, however, due to the vibration of the vocal folds, the rate of flow diminishes and the pressure in the oral cavity is lower. This, in turn, causes a lower intensity of fricative noise in voiced fricatives. Because of all these considerations, Ohala and Solé (2010, p. 53) conclude that "Voiced fricatives are hard to make; if voicing is strong, there is a tendency to defricate; if friction is achieved, there is a tendency to devoice". This incompatibility between voicing and turbulence can be solved in different ways: if voicing is to be maintained, friction might be reduced (resulting in a lower intensity of noise) or lost altogether and the fricative might become a continuant (glide, rhotic or approximant) (Ohala & Solé, 2010, p. 54).

In articulatory studies (Gráczi et al., 2023; Liker & Gibbon, 2013; Narayanan et al., 1995; Proctor et al., 2010; Tabain, 2019, p. 268) it has been observed that the configuration of the tongue changes in voiced fricatives: the tongue root tends to be more advanced in their production as compared to voiceless counterparts. Another difference is that voiced sibilants show increased anterior contact and a smaller groove through which the air passes (Dixit & Hoffman, 2004; Fuchs et al., 2007; Liker & Gibbon, 2013). These modifications help to maintain the necessary differences in pressure, and thus enhance voicing and frication. Thus, articulatory studies show that the articulatory distinction between voiced and voiceless pairs of sounds might go beyond the action of the vocal folds. As a result, the acoustic properties of voiced and voiceless fricatives also show various differences apart from the periodicity added by the glottal source in the voiced sound.

Many studies have tackled the acoustics of voiced and voiceless fricatives in languages where the difference is phonological (as for the English /s/ and /z/). Voiced fricatives are systematically shorter than voiceless fricatives (among others, Jongman et al., 2000; Nirgianaki, 2014; Silbert & de Jong, 2008; Smith, 1997). Results regarding spectral properties are more complex. The problem with most commonly used spectral measurements is that if the whole frequency range is used, the periodicity affects the measurements. Spectral peak was found to be the same for voiced and voiceless fricatives in Jongman et al. (2000), but the centre of gravity was lower for voiced sibilants (it appears that they used the whole range of frequencies). Studies which used filtered sound to exclude lower frequencies resulting from the vibration of vocal folds showed mixed results. Nirgianaki (2014) found that Greek voiced fricatives show lower spectral peak (though this difference was not found for alveolar sibilants). Silbert and de Jong (2008) did not found any difference in the spectral mean for English fricatives. Also Chodroff (2017) found that the spectral peak in the middle range of frequencies (4000-7000 Hz) was the same for /s/ and /z/ in English. Another commonly studied property of fricatives is the amplitude of the whole phone or of the fricative noise (often measured in comparison to surrounding vowels). In Jongman et al. (2000) and Nirgianaki (2014), for example, voiced fricatives had smaller relative amplitude.

2.3. Lenition of sibilants in Spanish

Several studies have dealt with lenition of intervocalic fricatives in Spanish, especially /s/. Their results are relevant for the present study because of the topic itself, but also because the speakers analysed here are Spanish-Basque bilinguals.

File-Muriel and Brown (2011) analysed the realisation of /s/ in interviews with speakers of Caleño (Colombia) Spanish. They measured duration, centre of gravity (above 750 Hz) and voicelessness (percentage of voiceless frames). They looked at an array of dependent variables and concluded that especially important were speaking rate, word position, the following phonological context and stress. For faster speaking rates, duration and centre of gravity decreased and the percentage of voiced frames increased. The same effect was observed for word-final sibilants and those followed by a non-high vowels. In tonic syllables or when a pause followed /s/, the opposite happened: duration and centre of gravity increased, and voicing was less likely. Moreover, there were important differences between speakers.

Torreira and Ernestus (2012) used casual conversations in Peninsular Spanish to study several aspects of sibilants. First, they measured the low-frequency band (0-1.5 kHz) intensity dip duration: when the constriction is formed, the intensity in the lower frequencies falls, and the duration of this dip can be used to analyse the temporal aspect of weakening (it is short in weakened /s/ and longer in voiceless unreduced ones). Second, they analysed voicing, and more specifically, whether uninterrupted voicing was present during the dip. Finally, they measured the high-frequency band (4-8 kHz) intensity difference between the highest intensity in the sibilant and the lowest in the vowel. Thus, they only used the high frequency region of the spectrum, precisely where the noise produced in sibilants shows up: the bigger the difference, the less weakened the consonant. Results show that 34% of the analysed intervocalic sibilants had uninterrupted voicing (during the intensity dip). Speech rate, low-frequency band intensity dip duration and word position (final sibilants showed more voicing) were found to be statistically significant predictors of voicing. For the intensity, the most influential variables were speech rate, voicing and low-frequency band dip duration. Generally speaking, they concluded that voicing was the most common sign of weakening.

Hualde and Prieto (2015) compared lenition of intervocalic fricatives in (Madrid) Spanish and Catalan, using map task data. Their prediction was that Catalan would exhibit less lenition of /s/ than Spanish because of the existing contrast between /s/ and /z/. They also wanted to explore the role of word boundaries, predicting more lenition in VC#V context. Results showed that 8.3% of Spanish intervocalic sibilants were voiced in their data (10.6% in word-initial position, 5.9% in medial position and 12.5% in final position). For Catalan, as predicted, the occurrence of voiced /s/ was lower at 4.1%.

Chappell and García (2017) studied the voicing of intervocalic /s/ in Costa Rican Spanish, and found that it is very common, with 44.6% of fully voiced sibilants in data from interviews (and 25.5% in read data). Similarly to other studies, they found that [z] is more likely to occur in faster speech, in word-final position and before an unstressed vowels. They also concluded that men tend to lenite more than women. The authors consider two possible explanations: a physiological one (men tend to have bigger larynges and longer vocal tracts) proposed in several studies (Nadeu & Hualde, 2015) and a sociolinguistic one ([z] may have a social meaning), and conclude that, even though both are plausible and disentangling the physiological and social factors is difficult, the sociolinguistic motivation might be more important for this variety of Spanish.

Spanish /s/ was also extensively studied in coda position due to the large variation in its realisation among the varieties of the language (see Núñez-

Méndez, 2022 for an overview). The most widely analysed phenomena concerns the weakening of /s/ to [h] or Ø. Another topic of research on Spanish coda /s/, especially relevant for this paper, concerns voicing assimilation: voiceless coda fricatives can be realised as voiced if followed by a voiced consonant (Navarro Tomás, 1957, p. 108; Quilis, 1993, p. 251). This process is not fully regular and categorical, and phonetic analyses show that there is a great deal of variation. For instance, Sedó et al. (2020) have found that around half of the /s/ were produced as [z] when followed by a voiced consonant. Several factors were identified as influencing voicing in Spanish /s/ in different studies (Sedó et al. 2020 and references therein), such as the manner of articulation of the following consonant (there was more voicing before a stop or an approximant than before a lateral or nasal), the preceding vowel (some dialects show more voicing after /e o/), the stress (fricatives in unstressed syllables are more likely to weaken), the position within the word (word-final fricatives are more likely to undergo lenition than those in the middle of the word), speaking style and speech rate (more informal and faster speech favour lenition). Thus, in general, the factors favouring voicing in the intervocalic and preconsonantal positions are similar for Spanish.

2.4. Basque sibilants and lenition phenomena

Conservative varieties of Basque distinguish six sibilants: the dorso- or lamino-alveolar /s̻/, apico-alveolar /s̺/, and palato-alveolar /ʃ/ fricatives, and their affricate counterparts /t͡s̻ t͡s̺ t͡ʃ/ (Hualde, 2003; Mitxelena, 1977/2011, etc.). However, there is significant dialectal and sociolinguistic variation related to sibilants: (a) various merger phenomena have reduced the number of distinctions in many varieties (Beristain, 2022; Hualde, 2010), and (b) some varieties have added voiced sibilants to their phonemic inventory, due to the historical evolution of the palatal glide or due to contact with Romance languages in the Northern Basque Country (Hualde, 2003, p. 26).

This paper studies the variety of Beasain, a town located in the province of Gipuzkoa. It belongs to the Central Basque dialect (Zuazo, 2014). Speakers from Beasain do not have phonemically voiced sibilants. In the fricative series the distinctions are well preserved, but it is not entirely clear if the postalveolars are different from the apicals in the affricate series. Lack of distinction between apical and postalveolar affricates is in fact common in parts of Gipuzkoa (Hualde, 2010). The status of the palato-alveolar sibilants, and more generally of all palatal sounds in Basque (Oñederra, 1990), is different from the rest: their frequency is low and they often appear in expressive or diminutive words (*sagu* ‘mouse’ vs. *xagu* ‘little mouse’, Hualde, 2003, p. 22). Due to their lower frequency, it is difficult to analyse them in a corpus study like the one presented here, and because of that I will focus on apico-alveolar and lamino-alveolar fricatives.

The acoustics of Basque sibilants have received increased attention in recent years (among others, Beristain, 2022; Egurtzegi et al., 2024; Hualde, 2010; Jurado, 2011; Muxika-Loitzate, 2017), with most studies focusing on the ways the contrasts between the categories manifest themselves in the acoustic signal, and on the results of mergers. Centre of gravity (CoG) has been the most widely used measure to capture those contrasts: it has been confirmed that, as expected from the different articulatory configurations, apical sibilants show lower CoG values than laminal ones. Nevertheless, it has to be taken into account that the relation between measures such as CoG and articulatory configurations is complex: recent articulatory studies on Basque sibilants (Iribar et al., 2020, 2022) show that even though the predominant articulatory model is dento-alveolar laminal for /s̻/ and apico-alveolar for /s̺/, other realisations are also common. According to Iribar et al. (2020, 2022) other articulatory differences between the two sounds include bigger external cavity for /s̺/, longer and narrower articulatory canal and contact with lower for incisors for /s̻/.

The voicing of Basque (phonologically voiceless) sibilants has not been studied acoustically in detail.[3] The only phenomenon often found in descriptions of many varieties is anticipatory voicing assimilation. Thus, coda sibilants might be realised as voiced when followed by a voiced consonant, for example in *esne* 'milk' [ez̺ne] (Hualde & Bilbao, 1992, p. 4). The same might happen at the word boundary: *eskuz garbitu* 'clean by hand' can be pronounced with [z̺ɣ][4] or [s̺ɣ] (Hualde, 2003, p. 42). Another known phenomenon concerns the final sibilant of the negative particle *ez*, which has historically caused the devoicing of the consonant in the finite verb following it: *ez da* 'it is not' [es̺ta], *ez gara* 'we are not' [es̺kara], etc. However, it appears that the devoicing is not systematic any more (see also Oñederra (2004, p. 29):

> Devoicing after *ez* is nowadays an optional process and pronunciations where, instead, the sibilant voices appear to be increasingly common: *e*[z̺ð]*oa* 'she is not going', *e*[z̺ɣ]*ara* 'we are not', *e*[z̺β]*ada* 'if it is not'. In some western areas (e.g. Ondarroa and Lekeitio), there is (optional) devoicing of /d-/ and /b-/ in this context, but /g-/ never devoices. (Hualde, 2003, p. 42)

Additionally, Hualde (2019, pp. 349–350) mentions a few instances of allophony in fricatives in examples such as *eztaki* 'he/she does not know' (pronounced [eztai]) and *asko* 'much' ([azko]) with the sibilants realised as voiced. In both, Hualde explains, it appears that the sibilant assimilated to the preceding vowel (the following consonant being voiceless). In the case of *eztaki*, moreover, it appears that the historical assimilation pattern (devoicing) and the newer one (voicing of the sibilant) overlap, producing a voiced sibilant followed by a voiced stop. Hualde adds that the fact that *eztaki* shows a weakened realisation might be related to the high frequency of the verb *jakin* 'to know' and some of its finite forms: such common words and expressions are more likely to undergo various reduction processes. Hualde mentions the following reduced variants of *eztakit*: [estaɣit͡], [estajt͡] and [estajʔ].

As regards the acoustics, a wide range of ways sibilants can be realised in Basque is presented in Urrutia et al. (1988, 1989). For both Western and Central varieties of Basque, the authors report several tokens of voiced sibilants, both before voiced consonants and intervocalically. The frequencies of voiced sibilants were low in this study (less than 5%), but the type of data should be taken into account (interviews where speakers were asked to translate words). More recently, also Hualde et al. (2019a, p. 97) mention in passing the existence intervocalic voicing, adding a footnote to an example containing the transcription [βezela] 'like, as' to say that [z] is not a typo and that sibilants can be voiced between vowels in Basque.

Even though the lenition of sibilants was not analysed for Basque, a few studies dealt with weakening phenomena in voiceless intervocalic stops /p t k/. Similarly to what happens in Spanish, they can have voiced realisations (usually as approximants) and studies on this phenomenon show that voicing is fairly frequent. Iribar and Túrrez (2008) analysed data from one speaker and concluded that voiced realisations were common, especially for /p/ and /k/ (over a third of examples). Nadeu and Hualde (2015) analysed /p t k/ in Goizueta Basque and concluded that around a third was voiced and that there were significant differences between speakers (males tended to show more weakening). Hualde et al. (2019b) analysed intervocalic /t k/ in Azpeitia Basque to compare lenition in consonants which appear at the word boundary (word-initial or word-final) and those in the middle of the word. They concluded that word-final consonants have a lesser degree of constriction and are more often voiced than those in other contexts. The reason for

[3] Voicing was taken into account, however, in the acoustic analysis of Mixean Basque (Egurtzegi & Carignan, 2020), which is one of the varieties with phonological contrast between voiced and voiceless sibilants.

[4] In the Basque transcriptions, [β ð ɣ] are used to represent approximants.

that, as Hualde et al. (2019b) explain, might be that in Basque there is no phonological opposition between /t/ and /d/ or /k/ and /g/ in the word-final position (voiced stops do not appear in this context). Hualde et al. (2019a) analysed consonant clusters at word boundaries when one word ends with /t/ or /k/ and the following starts with a consonant. The main conclusion was that, on the basis of intensity measures, such clusters tend to be simplified, with the first consonant often deleted or realised as an approximant. Finally, Eguskiza et al. (2020) used data from Arratia (Biscay) to analyse the lenition of /p t k/. 14% of the tokens in their study were voiced. In that study, voiced allophones had shorter duration, higher mean intensity and, compared to adjacent vowels, showed lower relative intensity.

3. Research questions

The main question treated here is whether Basque fricative sibilants (apico-alveolars and lamino-alveolars) undergo lenition.

First of all, I want to find out how common is voicing in sibilants and what factors favour the appearance of voicing.

Rather than focusing on only intervocalic or preconsonantal sibilants, I include fricatives in all contexts, especially because there is no previous research on the topic for Basque. On the basis of the previous literature, we can expect vocal fold vibration to spread to sibilants preceding voiced stops, though whether it is a regular phenomenon is unknown. The extent to which sibilants are voiced in other contexts is a question not previously posed for Basque.

Moreover, I consider a range of other factors: speech rate, the presence of a word boundary and the type of word (grammatical vs lexical). Based on previous research on other languages and works concerned with lenition of stop consonants in Basque, we can expect those factors to influence voicing.

Secondly, I analyse acoustic consequences of lenition other that those directly related to the presence of periodicity: changes in spectral properties and in the intensity of sibilants. Previous research on other languages suggests that the influence of the vibration of the vocal folds is not limited to the addition of a periodic component to the sound wave. It also causes other adjustments in the production of fricatives, which might result in a lowering of the CoG and the intensity of the fricative noise.

4. Methods

4.1. Speakers and procedure

Ten Basque-Spanish bilingual speakers (19-22 years old, 4 males and 6 females) from Beasain participated in the study, though eventually 9 were used in the analysis due to technical problems with one of the recordings (one of the males). Participants were recorded in 48 kHz and 24 bit rate in pairs in a quiet environment with head-mounted Shure WH20XLR microphones and a ZOOM H5 recorder (each speaker was recorded on a separate channel). According to the Bilingual Language Profile questionnaires (Birdsong et al., 2012) the participants filled, speakers F1 and F2 can be classified as balanced bilinguals, and the remaining ones as Basque-dominant.

Speakers were instructed to engage in conversation with their partner (who they knew well). They were told that they could talk about anything, and that they could start from a general question like "How is your week going?". They were also given a list of topics they could use (e.g. holidays, education, free time, etc.). After setting up the recording, the researcher left the room. 40-50 minutes were recorded for each pair. The goal of this procedure was to obtain a conversation as informal as possible in experimental settings. This kind of material is especially suitable to study lenitions, which are more common in casual speech (Tucker & Ernestus, 2016).

This study was approved by the ethics committee of the University of the Basque Country (UPV/EHU), and the participants provided their informed consent to participate.

4.2. Segmentation and acoustic analysis

Recordings were segmented manually in *Praat*. I decided to annotate the first 600 fricatives (laminal and apical taken together) for each speaker (with the exception of one speaker, F5, which participated less in the conversation and only had 500 examples). Fragments with unclear, highly glottalised or whispered speech, disfluencies, laughter and overlaps between the speakers were omitted. The sibilants were segmented manually by inspecting the waveform and the spectrogram. The main cue was the onset and offset of high-frequency aperiodic noise. Preceding and following phones were also segmented and annotated. A separate tier contains the limits of the word and its orthographic transcription.

Praat scripts were used to extract data from recordings and TextGrids.

First, a script by de Jong et al. (2021) was used to automatically determine the speech rate. The script looks for silences and establishes phrases, and then identifies syllable nuclei with them. I decided to use an automatic analysis of speech rate rather than, e.g. duration of phones,[5] in order to avoid relying too much on the manually placed boundaries. The speech rate was calculated as number of syllables per second.

Another script was used to extract the label of the sibilant, context in which it appears (i.e. preceding and following phones), word in which it appears, durations of the phone, word and the automatically identified phrase, the number of syllables in the phrase, whether the phone is word-initial or -final, spectral measures and intensity measures. For the measurement of the centre of gravity, the sound was filtered (Hahn pass band 750-16 000 Hz) in order to focus on the noise component and reduce the effect of voicing (File-Muriel & Brown, 2011; following Silbert & de Jong, 2008). Measurements were taken in the middle 10% interval of the phone. A 4000-10000 Hahn pass band filter was used for the measurement of intensity of frication. In order to be able to compute relative intensity, maximum, minimum and mean intensity were measured for the sibilant and for the previous and the following phone (initial and final 20% of phones were discarded). In this paper relative intensity is understood as in Torreira and Ernestus (2012), i.e. as the difference between the maximum intensity in the sibilant and the minimum intensity of the previous or following vowel. Figure 1[6] is helpful to understand this approach. In the spectrogram two lines are superposed: the blue one is the high-frequency band (4-10 kHz) intensity contour and the red one is the low-frequency band (0-1000 Hz) intensity contour. For the voiceless fricative, the noise is especially intense in the higher parts of the spectrum, and this is why intensity rises there. For the voiced sibilants intensity also decreases in the lower band (but less than in the voiceless one) and in the higher band, intensity does not increase as in the voiceless sibilant, suggesting that the frication noise is weaker for that sound. Thus, comparing the highest intensity in the fricative with the lowest intensity in the vowel can give us a measure of lenition.

The spreadsheet produced by this script was further enriched manually to add lemma for each word, and whether it is grammatical or lexical. The following elements were classified as grammatical: the auxiliary verbs and all forms of *izan* 'be', pronouns, conjunctions, particles, postpositions and case markers, suffixes and discourse markers. The information whether a sibilant is word-initial or word-final was extracted from the TextGrids automatically, and corrected manually when necessary.

Finally, another script measures pitch throughout the sibilant following the method of measuring voicing in *Praat* described by Eager (2015). The

[5] See supplementary materials for some details concerning duration. As expected, voiceless sibilants are shorter. Additionally, preconsonantal sibilants are shorter than prevocalic ones, both when the following consonant is voiced or voiceless.

[6] Note that in the segmentation shown below the spectrograms all sibilants are transcribed as voiceless.

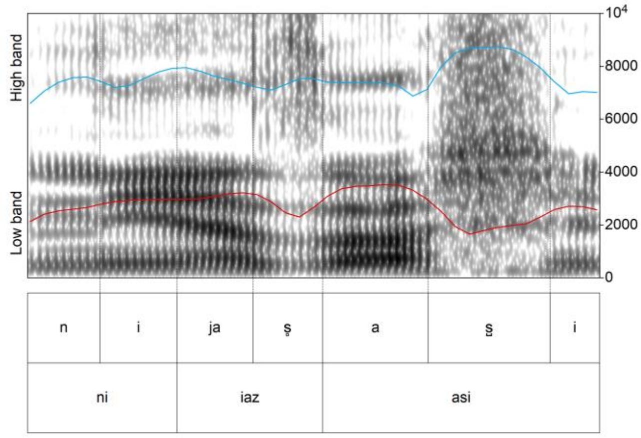

Figure 1. Spectrogram, high-band intensity contour (blue) and low band intensity contour (red).

floor for pitch measurement was set to 100 Hz for females and 70 Hz for men, and the ceiling to 300 Hz and 250 Hz, respectively. Time step was set to 0.001 (1 measurement per frame). Additionally, the same settings were used to extract 20 measurements at equal intervals of the phone. All frames were used for modelling and the 20-point dataset was used in plots.

4.3. Statistical analyses

Further analyses were conducted in R (R Core Team, 2024), using especially *ggplot2* (Wickham, 2016) for visualisation and *brms* (Bürkner et al., 2022) for statistical analyses.

For plotting of the centre of gravity, data was normalised in order eliminate the effect of anatomical differences between speakers. Normalisation was performed using the Lobanov method (following Egurtzegi & Carignan, 2020). In the statistical analyses non-normalised data were used. In order to account for variation between speakers, a group-level effect of 'speaker' (varying intercept) was included in all models.

As for the statistical analyses, I use Bayesian mixed-effects models fitted with the *brms* package. For each model, I have chosen weakly informative priors, which means that the prior does not have a strong influence on the posterior. The models' formulas are given in Section 5. When describing the results, I report mean estimates and their 95% credible intervals (CrIs) for each of the relevant factors. Additional details of the models (including priors and models diagnostics) are included in the supplementary materials.

5. Results

5.1. Overview

5300 tokens of apical and laminal fricatives were extracted from the recordings (600 examples for each speaker, except for one –F5– which has 500), of which 2391 are /s̺/ and 2909 are /s̻/.

I will first present the main ways in which sibilants are realised in the data. Figure 2 illustrates fairly typical intervocalic voiceless sibilants, with a short voiced interval at the onset. Figure 3 shows the realisation of sibilants before a voiceless stop. The higher frequency of the laminal as compared to the apical can be observed in both spectrograms.

Figure 4 shows a fully voiced laminal and an apical with the voicing weakening mid-phone. It can also be seen that the amplitude of noise in the (more) voiceless interval is higher.

Figures 5 and 6 show examples of negated verbs which start with *d*-. Both are variations of *ez dakit* 'I don't know', realized as [es̻taɣit] and [ez̻ᶞaɣit] The first illustrates the devoicing of the first consonant of the verb. In the second the voicing does not stop. Moreover, it is often impossible to delimit the sibilant and the first consonant of the verb, [ð]. It is unclear whether there are two consonantal gestures or rather just one voiced sibilant. In such examples only the sibilant was annotated in the TextGrid, though I decided to treat them as preceding a voiced consonant for further analyses (see Section 5.5 for more discussion of the phonetic aspects of negation).

Finally, even though other fricatives are not analysed in this paper, it is important to mention that voicing is also found in those (e.g. in alveolo-palatal sibilants as in Figure 7).

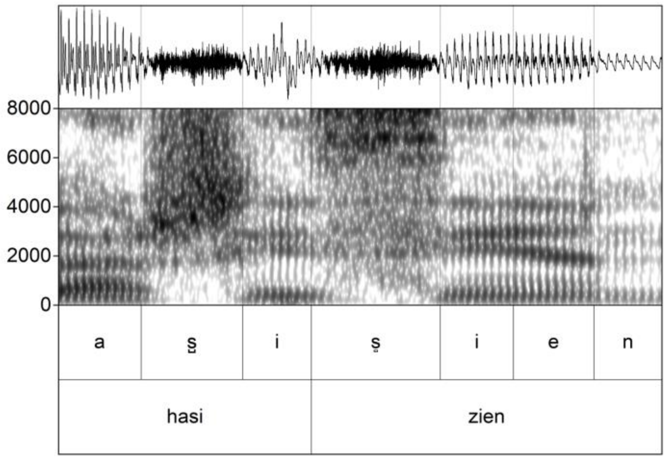

Figure 2. Intervocalic voiceless sibilants.

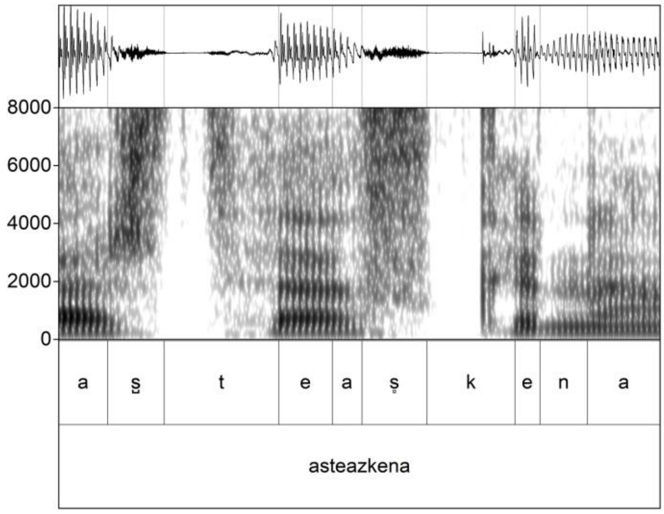

Figure 3. Sibilants before voiceless consonants.

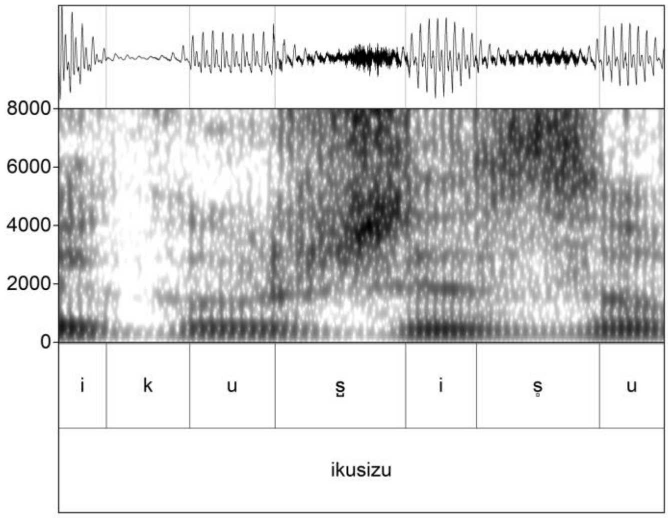

Figure 4. Intervocalic sibilants.

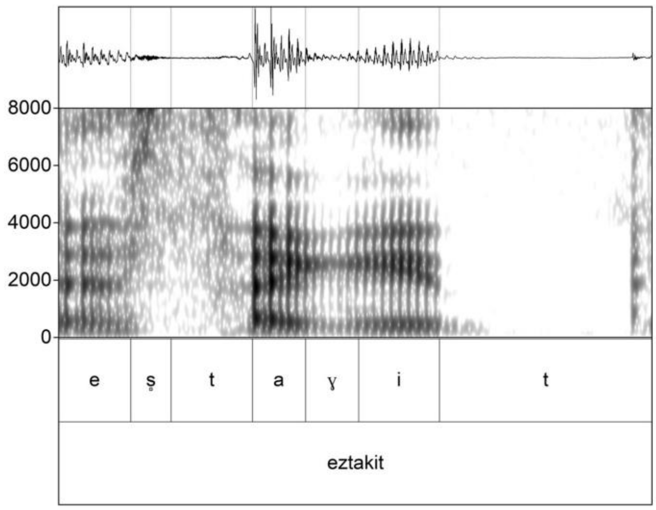

Figure 5. Negation with a voiceless sibilant.

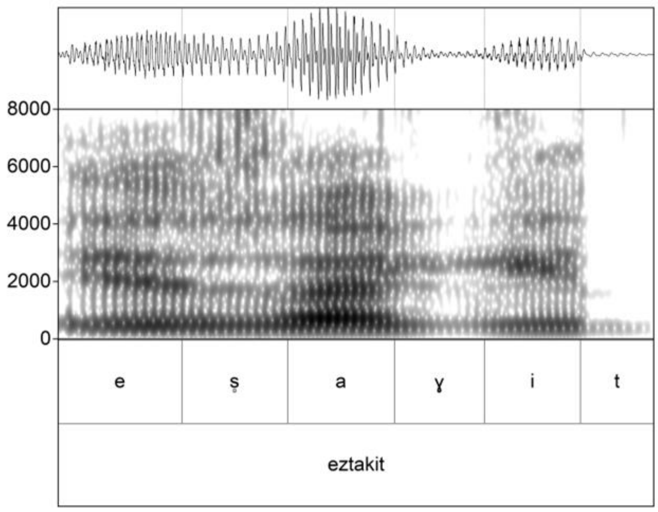

Figure 6. Negation with a voiced sibilant.

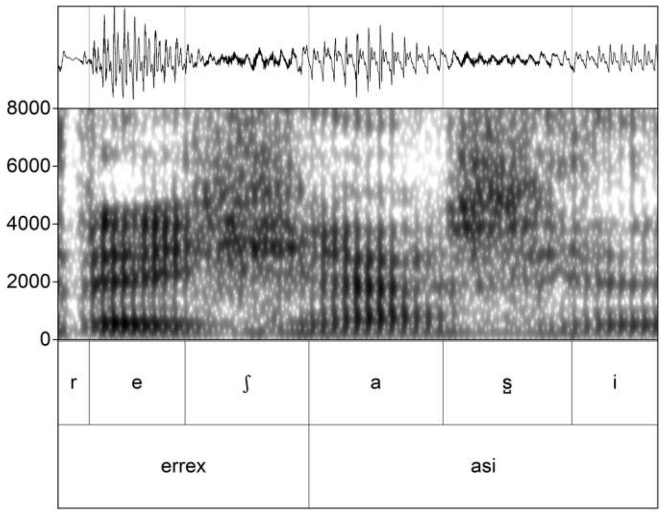

Figure 7. A voiced prepalatal sibilant.

5.2. Voicing

5.2.1. Exploration

In order to find out how common voicing is in Basque sibilants, we can start with the proportion of voiced frames. Table 1 shows the distribution of sibilants with different amount of voiced frames, separately for the middle half of the phone (50% of the phone centred on the mid-point) and for the whole duration (see also histograms in Figures 8 and 9 for more details). Thus, for the whole phone, 83% of examples show at least one voiced frame (71% partial voicing and 12% full voicing). If we only consider the middle 50% of the duration of phone, most examples show no voicing (62%), some examples are fully voiced (13%) and some show partial voicing (25%). Taking into account the difference in proportions of completely voiceless sibilants in the middle interval and in the whole phone, it is clear that an important proportion of voicing happens at the margins, in the onset and in the offset.

Interval	N	Voiceless	P. voiced	Voiced
Middle 50%	5300	0.62	0.25	0.13
Whole phone	5300	0.17	0.71	0.12

Table 1. Proportions of fully voiceless, partially voiced and fully voiced sibilants.

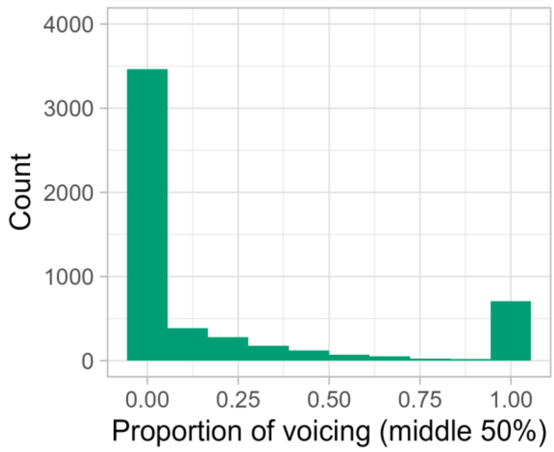

Figure 8. Histogram of proportion of voicing for the middle 50% of the phone.

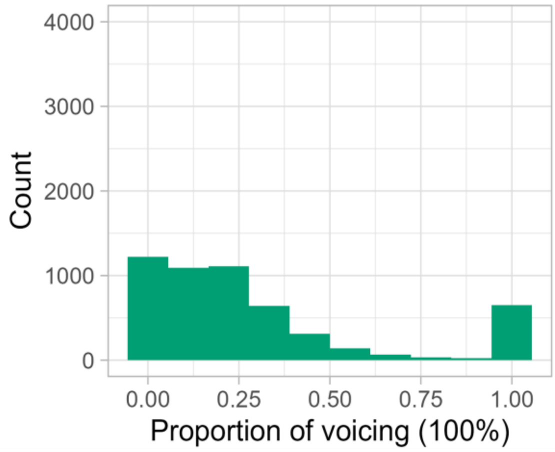

Figure 9. Histogram of proportion of voicing for the whole duration of the phone.

There are important differences between speakers (Table 2): full voicing of the middle interval ranges from 5% (F4 and F5) to 26% (M1).

Speaker	Voiceless	P. voiced	Voiced
F1	0.57	0.25	0.18
F2	0.58	0.32	0.11
F3	0.53	0.30	0.16
F4	0.64	0.31	0.05
F5	0.72	0.23	0.05
F6	0.66	0.26	0.07
M1	0.59	0.15	0.26
M3	0.73	0.13	0.14
M4	0.58	0.28	0.14

Table 2. Proportions of fully voiceless, partially voiced and fully voiced sibilants for each speaker (for the middle half of the phone).

We can also briefly consider the temporal aspect of voicing. Figure 10 shows the proportion of voiced frames throughout the phone averaged over all tokens for each speaker (the shaded area corresponds to the middle 50% of the phone). Voicing appears stronger in the onset than in the offset, which is partially related to the phonotactics: fricative sibilants very rarely appear after a voiceless consonant, but they can be followed by such a sound.

Figure 11 focuses precisely on the context, as it shows separately tokens followed by voiceless consonants (or pause),[7] voiced consonants and vowels. The phonetic environment is important for the presence of periodicity. Voicing is uncommon in the offsets of sibilants which precede voiceless segments, though the initial and middle portions of the phone might show some voiced frames. Prevocalic sibilants show an important proportion of voiced frames also in the offset. Sibilants followed by a voiced consonant are less common than others in the data, but it is nevertheless quite clear that sibilants are more often voiced in this context than in others.

Table 3 provides more data on the role of context. 13% of prevocalic sibilants are 100% voiced in the middle interval, 4% of those occurring before a voiceless segment and as much as 60% of those followed by voiced consonant. At the same time only 56% of sibilants before a voiceless consonant have no voiced frames, which shows that partial voicing is common even in this context.

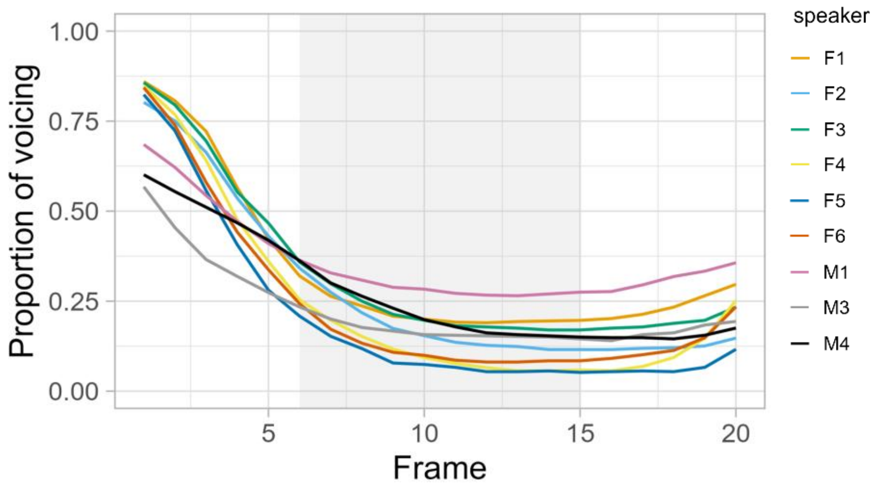

Figure 10. Mean proportion of voiced frames for each speaker for 20 points along the sibilants.

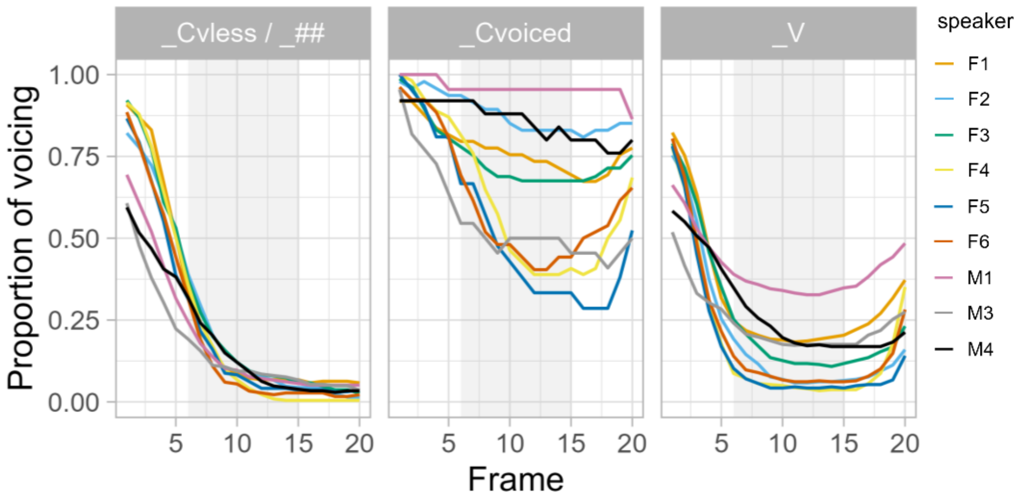

Figure 11. Mean proportion of voiced frames for each speaker for 20 points along the sibilants: for sibilants preceding a voiceless consonant or a pause (left), prevocalic ones (middle), and those preceding a voiced consonant (right).

[7] The utterance-final (i.e. before a pause) sibilants are analysed together with voiceless consonants. It would be preferable to treat them separately, but sibilants preceding a pause are not very common in the corpus (only 215 tokens, 4%). The prevalence of full voicing for those examples is 4%, the same as for those which precede a voiceless consonant.

Context	N	Voiceless	P. voiced	Voiced
Vowel	3025	0.70	0.17	0.13
Voiceless c.	1691	0.56	0.40	0.04
Voiced c.	369	0.15	0.25	0.60
Pause	215	0.76	0.20	0.04

Table 3. Proportions of fully voiceless, partially voiced and fully voiced sibilants for each following context (for the middle half of the phone).

An interesting aspect is the apparent difference in voicing in the onset and offset of prevocalic sibilants which can be seen in Figure 11: earlier I have mentioned that since most sibilants follow a vowel, more voicing is expected at the beginning of the phone. In a prevocalic context, however, we would not expect voicing to be stronger in the onset, unless carry-over voicing is stronger than anticipatory voicing.

For the statistical analyses I only take into account the influence of the phone following the sibilant, because the majority of sibilants are postvocalic (86% of the examples in the corpus). However, the preceding context also plays a role, as illustrated in Table 4: sibilants which do not follow vowels are much less commonly voiced. This effect is nevertheless indirectly taken into account in the statistical analyses: as already mentioned, postconsonantal sibilants practically only occur at word boundaries (when a word ends with a consonant and the following one starts with a vowel), and the statistical modelling includes the variable boundary (word-initial / word-internal / word-final).

Prec. context	N	Voiceless	P. voiced	Voiced
Vowel	4635	0.58	0.28	0.15
Voiceless c.	113	0.92	0.04	0.04
Voiced c.	211	0.80	0.14	0.06
Pause	341	0.96	0.02	0.01

Table 4. Proportions of fully voiceless, partially voiced and fully voiced sibilants for each preceding context (for the middle half of the phone).

5.2.2. Modelling

This section presents a statistical analysis of factors which influence voicing. In particular, the following predictors are taken into account:

a) PHONE: *apical* or *laminal*
b) CONTEXT_RIGHT: *_Cvless* (voiceless consonant or pause), *_Cvoiced* (voiced consonant) or *_V* (vowel)
c) BOUNDARY: *#_* (word-initial), *mid* (word-internal) or *_#* (word-final)
d) SPEECH_RATE: speech rate in syllables per second
e) LEX_TYPE: *grammatical* or *lexical*

The outcome variable is VOICED_FRAMES_MID, the proportion of voiced frames in the middle 50% of the phone. As can be seen in Figure 8, the distribution of voicing is highly skewed towards 0 (fully voiceless) and 1 (fully voiced), with some tokens in between. In order to take into account both the binary and the continuous aspects, the data was modelled using zero-one inflated beta regression, which is mixture of logistic and beta regression.[8] The model is described by four parameters. It estimates the probability that an observation is either 0 or 1 (the parameter α), the probability that it is 1 (γ), the mean μ and the precision parameter ϕ of the continuous beta distribution for observations between 0 and 1.

The model's formula for the mean μ is the following:

```
voiced_frames_mid ~ phone + speech_rate + context_right
+ boundary + lex_type + (1 | lexeme) + (1 | speaker)
```

The same set of predictors was used to model the parameters α, γ and ϕ. Thus, the models include the variables listed above and also group-level effects. The motivation for including a varying intercept of LEXEME is that different lexemes might show different patterns of voicing. Speakers are also likely to behave differently (use more or less voiced sibilants) and this has to be taken into account. The

[8] Such solution was used to model voicing by DiCanio et al (2022).

model does not include any interaction term, as I have not found theoretical motivation for any.

All of the model's estimates are listed in Table i in the appendix. To analyse the role of the predictors, I will first present results for the models for *α*, *γ* and *μ*, focusing on the predictors which are significant.

The sub-model for *α* estimates the probability of the extreme values (fully voiced or fully voiceless). In the model, an increase in speech rate (by one syllable per second) decreases the probability of the proportion of voiced frames being 0 or 1 (β = -0.09,[9] 95% CrI = [-0.14, -0.05]). Sibilants followed by a voiceless segment have lower probability of extreme values compared to those followed by a vowel (β = -1.12, 95% CrI = [-1.32, -0.92]). The same holds for sibilants before a voiced consonant, though the CrI is close to 0 (β = -0.40, 95% CrI = [-0.73, -0.06]). As compared to word-internal sibilants, the probability is higher for word-initial sibilants (β = 0.51, 95% CrI = [0.27, 0.75]) and word-final sibilants (β = 0.35, 95% CrI = [0.03, 0.68], though note that the credible interval approaches 0 here). Another result worth emphasising is that, averaging over the posterior distribution, the median probability of extreme values is 0.79 (95% CrI = [0.52, 0.92]). Thus, there is evidence to say that full voicing or lack of voicing are more likely than partial voicing.

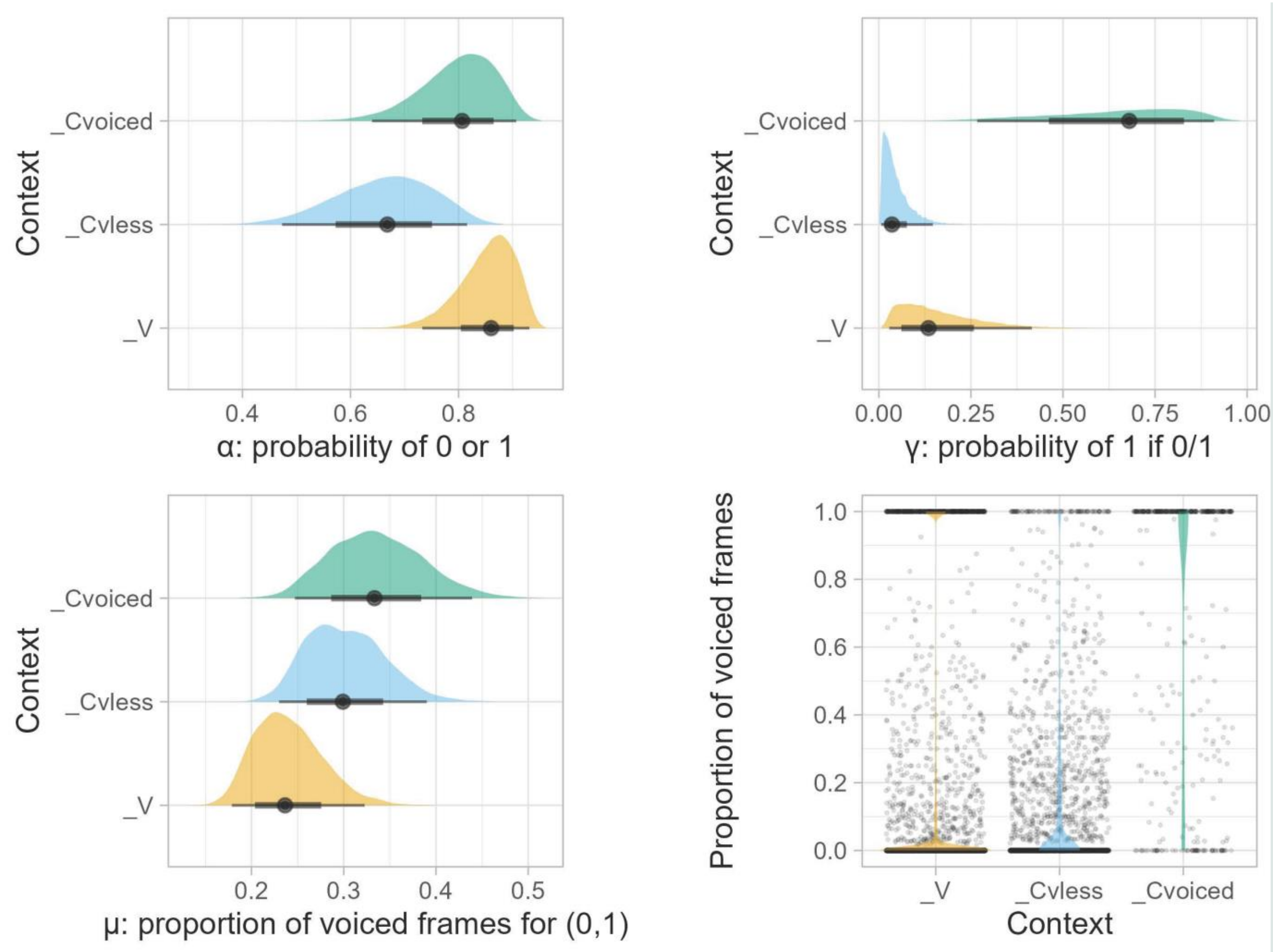

Figure 12. Posterior distributions for the predictor CONTEXT_RIGHT and distribution of the original data.

9 The results of logistic regressions are given in log-odds.

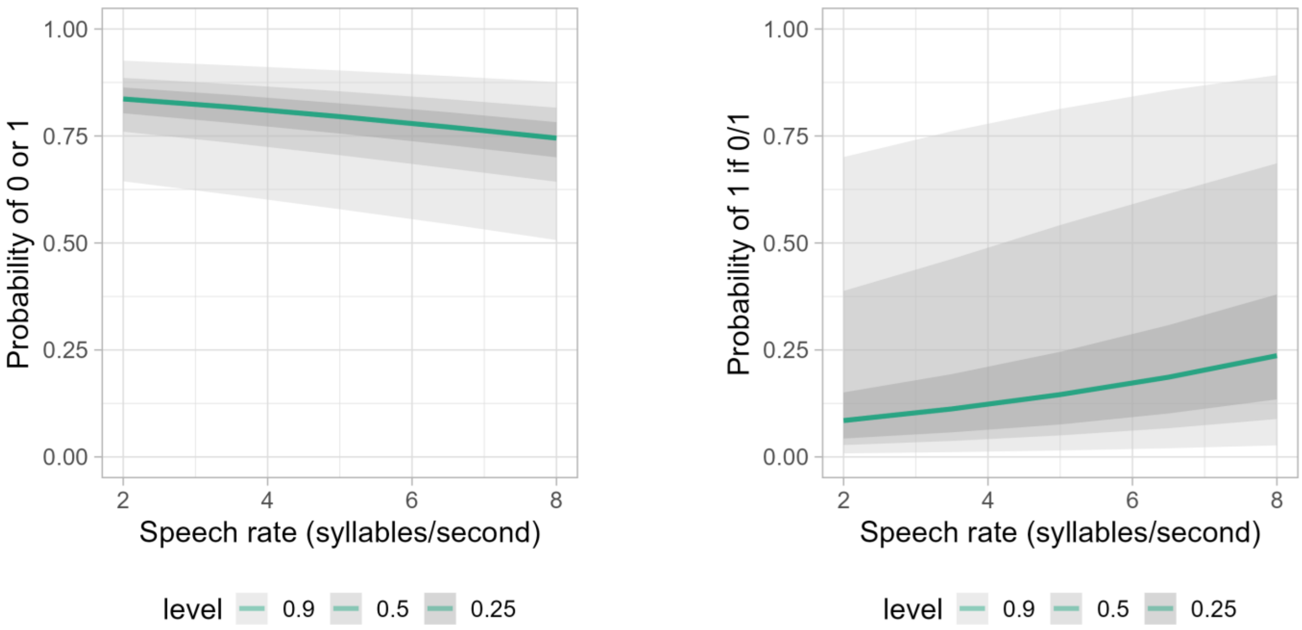

Figure 13. The effect of speech rate.

The sub-model for γ deals with the probability of full voicing. Here increasing the speech rate brings higher probability of the outcome being 1 (β = 0.21, 95% CrI = [0.14, 0.27]). There is also an effect of context: as compared to sibilants followed by a vowel, the presence of a voiceless segment to the right decreases the probability of the outcome being fully voiced (β = -1.46, 95% CrI = [-1.80, -1.13]) and the opposite happens if there is a voiced consonant (β = 2.63, 95% CrI = [2.19, 3.08]). Moreover, if the sibilant is word-initial, the probability of 1 is lower than when it is medial (β = -1.02, 95% CrI = [-1.34, -0.71]).

Finally, the sub-model for μ predicts the proportion of voiced frames if the outcome is neither 0 nor 1. Also here we have an effect of context: when followed by both voiceless and voiced consonants, the proportion of voicing is higher than for prevocalic sibilants, with the effect being stronger for voiced consonant (β = 0.31, 95% CrI = [0.17, 0.45] for voiceless, β = 0.48, 95% CrI = [0.20, 0.76] for voiced). The presence of a boundary to the left of the sibilant also increases the proportion of voiced frames (β = 0.3, 95% CrI = [0.07, 0.53]).

The remaining predictors, phone (apical vs laminal) and type of lexeme, have not proven important in the models. This means that apicals and laminals exhibit similar patterns with respect to voicing. Also there is no difference between grammatical and lexical words.

Thus, context is the most important predictor in all models. Figure 12 compares posterior distributions for the levels of this variable for the three sub-models (additionally it shows the distribution of the data). We can see that sibilants before a voiceless segment pattern differently from those followed by a vowel or a voiced consonant: they are less likely to be fully voiced or fully voiceless. In other words, they are more likely to be partially voiced, and in this case they are expected to have around a third of the frames voiced (more than prevocalic sibilants which are not fully voiced or voiceless). If they happen to fall into the 0 or 1 group, then the probability of them being voiced is very low. Sibilants which precede a voiced consonant or a vowel are very likely to be fully voiced or voiceless, and if this is the case, the probability of being voiced is higher for those before a voiced consonant (being voiceless is also possible, though). For those before a vowel, being voiced is possible, as can be seen from the wide interval, but less likely than being voiceless. Interestingly, for partially voiced sibilants preceding a vowel the proportion of voiced frames is expected to be lower than that of sibilants before any consonant.

Speech rate is also important in the model (Figure 13). First of all, the probability of extreme outcomes decreases with an increase in speech rate. Put differently, in faster speech partially voiced sibilants are more likely to occur. Secondly, the probability of a sibilants being fully voiced (as opposed to fully voiceless) increases as the speech rate increases (in this case, though, the intervals are rather wide).

Finally, the presence of a word boundary also affects the results. Word-initial sibilants are more likely to be fully voiced or voiceless than medial or word-final ones, and the probability of them being voiceless is high. Word-final sibilants show higher proportion of voiced frames than word-initial or medial ones.

As regards differences between speakers, the posterior distribution for the three models are shown in Figure 14. In general, speakers show similar tendencies in the proportion of voiced frames. Looking at the standard deviations of the model's parameters for the effect of SPEAKER, the highest value is that for γ, suggesting that speakers differ most in their probability of using fully voiced sibilants. Even though too few speakers were analyses in this study to be able to draw firm conclusions on the role of gender, males appear more likely to use fully voiced sibilants.

5.3. Spectral measures

Mean speaker-normalised CoG (measured using 750-16000 Hz Hahn pass filter, as explained) for the laminal is 7720 Hz (sd = 1068) and the for apical it is 5471 Hz (sd = 723). These results are in line with previous acoustic studies of Basque sibilants, which are usually interpreted as reflecting different articulatory configurations: in the production of laminal sibilants, the cavity in front of the constriction is generally smaller, and this increases the frequency of the fricative noise (though see Iribar et al. 2020, 2022 for an articulatory study of Basque sibilants). In this section I want to focus on how voicing affects the CoG: the prediction is that, due to the weaker frication expected when vocal folds vibrate, the frequency should be lower in voiced sibilants.

Figure 15 presents violin plots for voiced and voiceless sibilants ('voiced' are those with all frames voiced in the middle 50% of the phone). Generally speaking, an increased proportion of voiced frames lowers the value of the CoG. Additionally, the figure suggests that there is greater variability in the

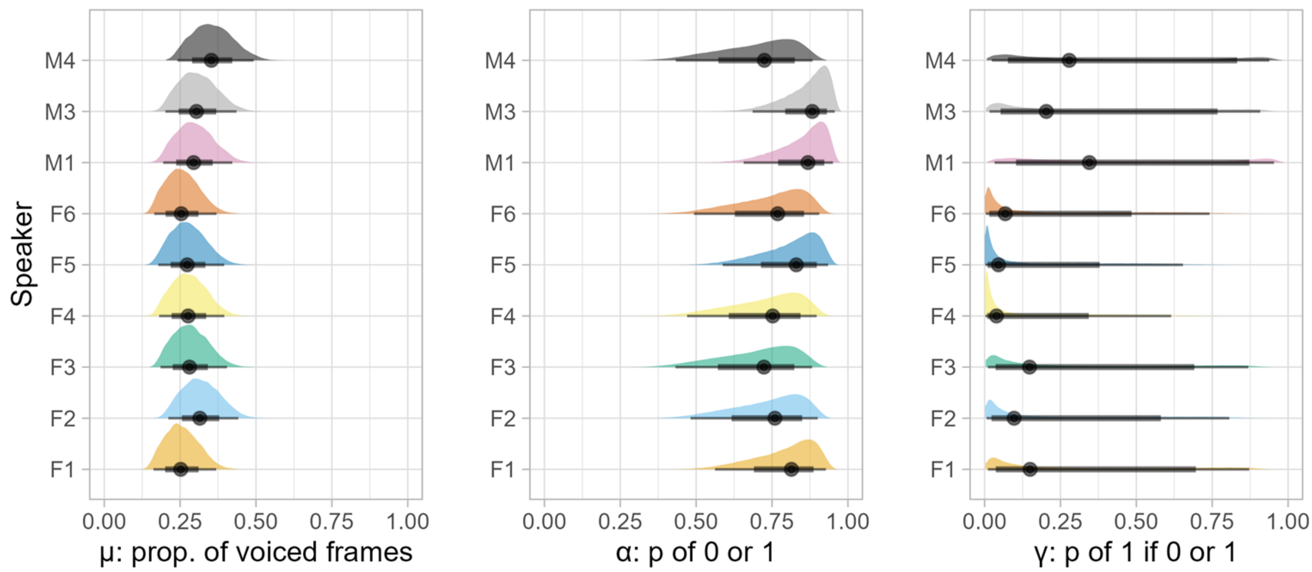

Figure 14. The effect of speaker.

realisation of [s̻] than in that of [s̺]. Mean values of CoG for voiced and voiceless tokens are given in Table 4.

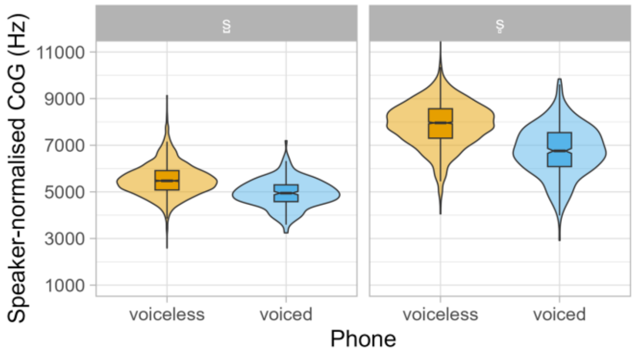

Figure 15. Violin and box plots of CoG.

Phone	Voice	CoG Mean	CoG SD
s̻	voiceless	7900	965
s̻	voiced	6762	1080
s̺	voiceless	5529	711
s̺	voiced	4938	600

Table 4. Mean values of speaker-normalised CoG for voiced and voiceless sibilants (voiced = 100 % voiced frames in the middle 50% of the phone).

To further establish the relationship between CoG and voicing, a model was fitted to the data with the following formula: `CoG ~ phone * voiced_frames_mid + (1|lexeme) + (1|speaker)`. Thus, the main question is how the proportion of voiced frames modulates CoG.

Table ii in the appendix lists the model's results for population-level effects. Compared to laminal sibilants, the CoG of apical sibilant is lower by 2426 Hz (95% CrI = [-2514, -2335]). The proportion of voiced frames impacts the CoG. For laminal fricatives, a phone with 100% of voiced frames has CoG lower by 1295 Hz (95% CrI = [-1383, -1208]). Additionally, there is compelling evidence for an interaction between phone and the proportion of voiced frames. For apical sibilants, 100% of voicing brings a lowering in the CoG smaller by 712 Hz (95% CrI = [569, 855]) as compared to laminals, i.e. of 583 Hz. This interaction is illustrated in Figure 16, which shows the regression lines for laminal and apical sibilants: for laminal sibilants the CoG is predicted to decrease more with the increase of the proportion of voiced frames.

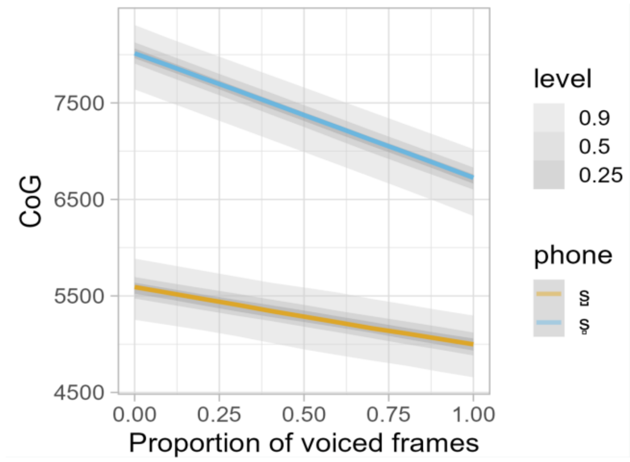

Figure 16. CoG and proportion of voiced frames for apical and laminal sibilants.

5.4. Relative intensity

As explained in Section 4.2, following Torreira and Ernestus (2012), I will compare the maximum intensity in the sibilant with the minimum in the adjacent vowel. I will focus on postvocalic sibilants, because in the corpus there are more examples which follow a vowel than examples preceding a vowel (4635 postvocalic vs 3025 prevocalic tokens). Figure 17 shows plots for relative intensity for postvocalic sibilants. It can be seen that laminals have lower intensity than apicals, and that the intensity is also lower for voiced sibilants.

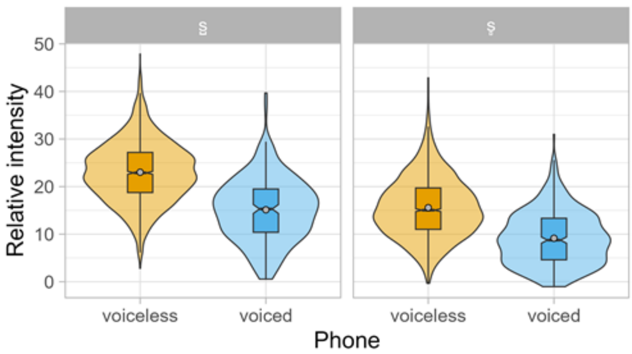

Figure 17. Relative intensity (dB) for postvocalic voiced and voiceless sibilants (voiced = 100 % voiced frames in the middle 50% of the phone).

Using only sibilants which follow a vowel (4635 examples), the following model was fitted: `rel_intensity ~ phone * voiced_frames_mid + (1|lexeme) + (1|speaker)`. Thus, the model is analogous to the model used in the previous section for CoG, and the question is whether relative intensity differs for the two sibilants and how voicing influences it.

The results (Table iii in the appendix) show that there is an effect of phone: apicals have relative intensity higher by 6.01 dB (95% CrI = [5.22, 6.80]). If 100% of the frames are voiced, the relative intensity decreases by 6.90 dB (95% CrI = [-7.52, -6.28]). There is also some evidence for an interaction between phone and the number of voiced frames: fully voiced apical sibilants lower the relative intensity by additional 1.51 dB (95% CrI = [-2.54, -0.52])

5.5. Lexical effects

This section explores possible lexical effects, or the question whether the lenition of sibilants might be more common in some words or classes of words. A bigger corpus would be necessary to perform a statistical analysis of such effects. Nevertheless, in a more informal manner, we can try to identify common traits in the lexemes which undergo lenition more than others.

Lexemes with more than 50 instances in the corpus are listed in Table 5 together with the proportion of occurrences with a voiced sibilant (i.e. full voicing in the middle interval). Particularly frequently voiced (33%) is the sibilant in the negative particle *ez*. The instrumental suffix -*z* comes second. A few common verbs –**edun* 'have' (used as a transitive auxiliary), *esan* 'say', *izan* 'be' (and an intransitive auxiliary)– also frequently contain voiced sibilants. The verb *uste* 'think' is quite interesting. *Uste*, when carefully pronounced is not very likely to show voicing because of the voiceless stop. However, as illustrated in Figure 18, in casual speech the very common expression *uste det* 'I think' is often shortened to [us̺tet] and probably through [uz̺ðet] to [uz̺et]. As mentioned earlier, when *ez* 'no' is followed by a verb starting with a voiced consonant, especially /d/, the resulting consonant cluster also appears to be simplified to a single voiced fricative. There are, however, a few common words which rather rarely contain a voiced allophone.

Lexeme	Phone	N	% voic.
ez 'no'	s̻	599	0.33
instrumental -*z*	s̻	87	0.21
zeozer 'something'	s̻	62	0.19
edun 'have'	s̻	301	0.18
esan 'say'	s̺	149	0.17
uste 'think'	s̺	80	0.17
izan 'be'	s̻	474	0.15
eduki 'have'	s̻	109	0.11
kriston 'very'	s̺	65	0.11
o sea discourse particle (Sp.)	s̺	232	0.08
zer 'what'	s̻	209	0.06
beste 'other'	s̺	87	0.05
hasi 'start'	s̺	56	0.05
zu 'you'	s̻	77	0.05
es ke discourse particle (Sp.)	s̺	143	0.03
ikasi 'learn'	s̺	65	0.03
azken 'last'	s̻	76	0

Table 5. Proportion of voiced sibilants for the most common lexemes in the corpus.

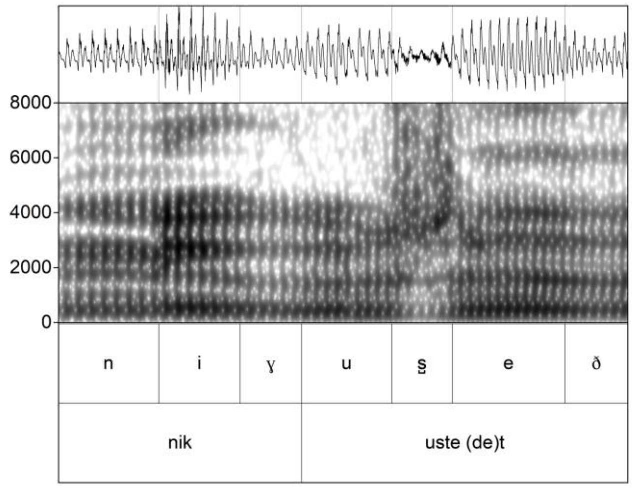

Figure 18. Example of *uste det* 'I think' contracted to [uz̺eð].

Finally, it is interesting to focus on the negative particle *ez*, the lexeme which most frequently contains a voiced sibilant among the most common words in the corpus. Because of the different items that might go after the particle, the sibilant in *ez* might be followed by a vowel, a voiceless or a voiced consonant. Table 6 shows the percent of examples of *ez* with a fully voiced sibilant (in the middle 50% interval) according to the context. The proportion of voiced examples is important in all environments, but it is clearly highest in the context of a voiced consonant.

CONTEXT_RIGHT	N	Fully voiced (%)
_V	81	17
_Cvless	269	10
_Cvoiced	249	62

Table 6. The percent of voiced sibilants in the word *ez* 'no' in different contexts.

As already mentioned, when *ez* is followed by a verb starting with /b d g/, two main outcomes are possible: (a) devoicing of the initial consonant of the verb and (b) maintaining voicing in that consonant. Voicing in the particle *ez* is more common if the voicing in the verb is maintained (but there are also some examples where the first consonant of the verb is voiceless and the sibilant in *ez* is voiced). Furthermore, in the data analysed, when the sibilant is voiced the following /d/ is usually highly reduced (and often impossible to segment in the spectrogram, see Figure 6). Forms like [ez̪ðu] are uncommon compared to those which rather look like [ez̪u], but a few cases are found with a segment which appears to correspond to [ð]. Figure 19 shows a more fricative [ð] and Figure 20 a more approximant [ð]. In finite verbs dental consonants outnumber those with other places of articulation, but for labials there are examples with two consonants clearly differentiated (Figure 21). The explanation is probably that an alveolar consonant and a dental one can be easily assimilated into one consonant, but it is not possible for an alveolar and a labial or velar to assimilate.

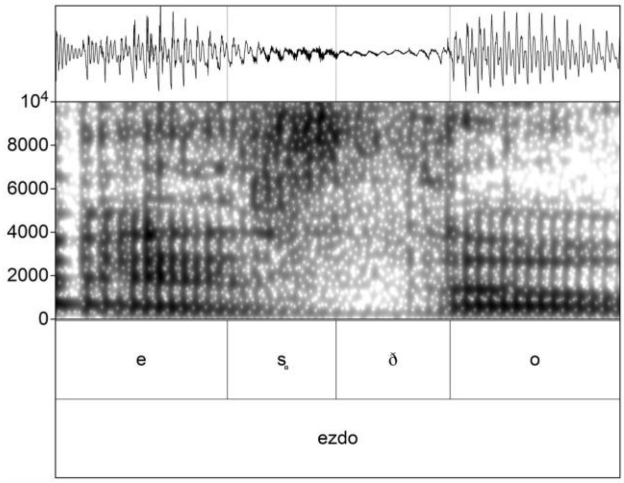

Figure 19. Laminal sibilant followed by a dental approximant (or fricative).

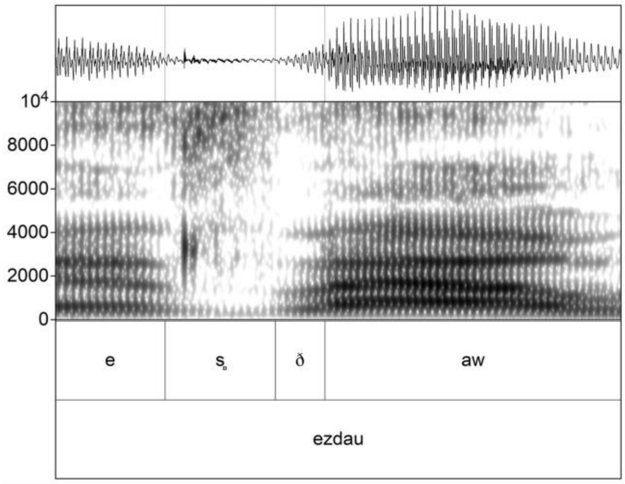

Figure 20. Laminal sibilant followed by a dental approximant.

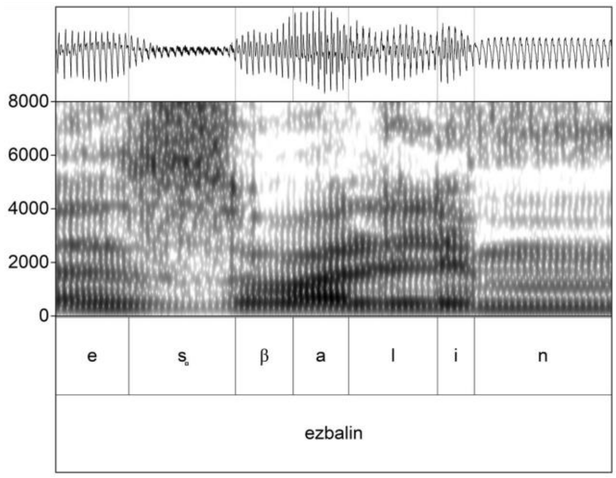

Figure 21. Laminal sibilant followed by a bilabial approximant.

Additionally, data show important inter-speaker and intra-speaker variation in the realisation of negation and the following verb. On the one hand, 60% of tokens with negation + a verb starting with /b d g/ are realised with the initial consonant voiced, but the speakers vary from 14% to 92% (speakers uttered on average 38 examples). On the other hand, 44% of those tokens contain a voiced sibilant, and the speakers vary from 22% to 72%. Figure 22 plots the four possibilities (voiced/voiceless sibilant + voiced/voiceless first consonant of the verb) for each speaker. It is clear that there are important differences between speakers, but most fall into two groups: (1) those who prefer voiceless sibilant and /p t k/ in the verb and (2) those who most often voice both the sibilant and the first consonant of the verb. Interestingly, males appear to prefer option (1), while females rather opt for (2).

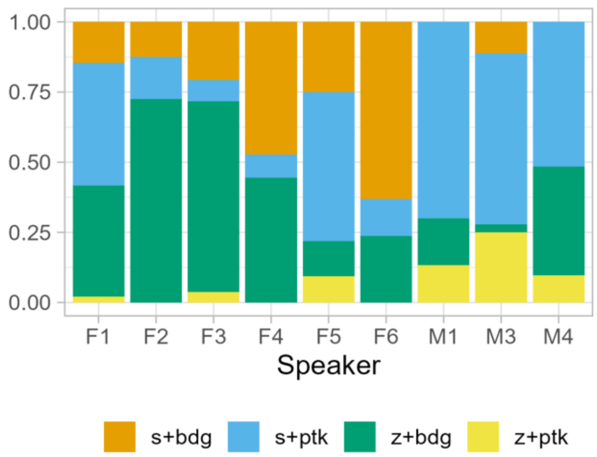

Figure 22. The realisation of negation *ez* + verbs starting with /b d g/.

Given this variability, a possible explanation is that the devoicing of the first consonant of the verb operates in this variety. It might then feed into lenition of /p t k/, which might in turn feed into anticipatory voicing and, finally, when the consonant is dental, into assimilation of that consonant. Schematically, the most common realisations of, e.g., *ez da* 'it is not' can be derived in the following way:

/es̺da/	
[es̺ta]	devoicing of /b d g/ / es̺ __
[es̺ða]	lenition of /p t k/
[ez̺ða]	voicing of the sibilant
[ez̺^{ð}a] / [ez̺a]	assimilation of [ð]

6. Discussion

One of the main results of this study is that (young) speakers of the analysed variety of Basque (Beasain, a variety of the Central Basque) realise some voiceless sibilants as voiced or partially voiced. Of all analysed sibilants, 13% can be classified as 'voiced' (i.e. they show uninterrupted voicing during the middle 50% of the duration of the phone) and another 25% have at least one voiced frame in the middle interval. The context in which the sibilants appears is important. 13% of the prevocalic sibilants are fully voiced, 60% before a voiced consonant and 4% before a voiceless consonant and pause. The statistical model used in this paper suggests that speakers are more likely to use fully voiced or fully voiceless sibilants rather than partially voiced ones (again taking into account only the middle 50% interval). Interspeaker differences were also registered, as the proportion of fully voiced sibilants ranges from 5% to 26%. Many studies of lenition found that males tend to weaken more than females (e.g. Nadeu & Hualde, 2015). More research would be necessary to establish if it is also the case with Basque sibilants, but the data presented in this paper appear to be in agreement with the previous research in that the lower proportions of voiced sibilants were found among females. Nevertheless, there might be a different pattern in the way the negation *ez* interacts with the verb which follows it: generally speaking, females tend to voice the sibilants and the first consonant of the verb, while males rather stick to the more conservative pattern of voiceless sibilant and voiceless consonant.

No other study has analysed this topic in Basque, but in Spanish the proportions of voiced intervocalic sibilants range from around a third in Torreira and Ernestus (2012) to about 8% in Hualde and Prieto (2015). It is not straightforward to compare results from different studies because of the different characteristics of the situations recorded (more or less controlled) and varying ways in which 'voicedness' is actually defined. However, we can conclude that the voicing of Basque sibilants occurs at rates roughly similar to those registered for Spanish in most studies. However, it is not as frequent as in Torreira and Ernestus (2012), and the conversations used here resemble the kind of data analysed in that paper.

Basque and Spanish differ in the inventory of sibilants, and one of the questions posed in this paper was whether the Basque apico-alveolar and lamino-alveolar sibilants show voicing to the same extent. No difference was found between the two sibilants in their probability of voicing. Additionally, although not included in the analysis, examples of other voiced fricatives (/f/ or /ʃ/) were also found in the corpus. This suggests that lenition might be a general phenomenon concerning all Basque fricatives in a similar way. More generally, this phenomenon can be seen in the wider perspective of lenition of voiceless obstruents. Hualde and Prieto (2015) made a similar observation for Spanish: lenition of /s/ and that of /p t k/ can be seen as part of a broader phenomenon (though they also emphasise that the difference in the articulatory characteristics of the two classes of sounds must not be ignored). Spanish and Basque are similar in another feature as well: neither has a phonological contrast between voiced and voiceless fricatives, which might facilitate lenition (as compared to languages such as Catalan, where there is such a contrast; see Hualde and Prieto, 2015).

So far no study has compared the voicing of voiceless stops and fricatives in Basque, but the available results for stops suggest (roughly) similar rates of voicing (about a third of the stops were voiced in Hualde and Prieto 2015 and about 14% in Eguskiza et al., 2020). A preliminary analysis of part of the data from two speakers (F1 and F2) recorded for this study suggests that the amount of voicing might be lower in fricatives than in /p t k/ (see details in the supplementary materials): out of 442 intervocalic tokens of /p t k/ uttered by these two speakers 22% are fully voiced. Speaker F1 voiced 23% of the instances of /p t k/ and 18% of the sibilants, and for F2 it was 21% and 11%, respectively. More research is necessary to establish the relation between lenition in different classes of obstruents.

Some studies of sibilant lenition have typically only analysed the intervocalic environment. The position before a voiced consonant was studied separately to establish how common voice assimilation is. In this paper all sibilants were included in the analysis for the sake of completeness and to determine how different the three main environments are: prevocalic, before a voiced and a voiceless segment (typically a consonant, but the rather unfrequent cases of sibilants before pause were analysed together with voiceless consonants). Results show that in Basque voiced sibilants are found in all of those contexts, suggesting that voicing is rather generalised, but there are also important differences.

Statistical modelling suggests that sibilants followed by a voiced consonant and those followed by a vowel are very likely to be fully voiced or voiceless (rather than partially voiced). Moreover, a following voiced consonant greatly increases the probability of the sibilant being voiced. Prevocalic sibilants are more often voiceless than voiced, but fully voiced examples are common. Sibilants which appear before a voiceless segment are much more likely to be fully voiceless than to be fully voiced. Additionally, they are more likely to be partially voiced than those in the remaining contexts. In such situation, they typically have a third of frames voiced. Temporal analysis shows that voicing is stronger in the onset of the sibilant.

As regards the interpretation of these facts, voicing before a voiced consonant might be seen as anticipatory voicing assimilation. This assimilation has often been mentioned in the literature on Basque phonology, and, even though not completely regular, it is clearly usual in the analysed variety. More interesting is the voicing attested in sibilants preceding a voiceless segment: full voicing is not common in this position, but partial voicing occurs in around a third of examples. This might be interpreted as carry-over assimilation. There are two arguments pointing to voicing being caused by the previous sound: (1) voicing is common in the onset of sibilants which appear before a voiceless consonant and it is stopped mid-phone and (2) in the prevocalic context voicing seems stronger in the onset than in the offset of the sibilant. On the whole, this results are in line with the analysis of voicing of coda sibilants for Spanish in Sedó et al. (2020), where the authors propose that the reason for the voicing of /s/ before a voiced consonant might actually be the preceding vowel, and thus the process should rather be seen as progressive assimilation. In that study, similarly to what has been found here, most sibilants before a voiceless consonant showed some partial voicing, and, the authors argue, this can only be caused by the preceding vowel. For Basque this topic has not been analysed elsewhere, but Hualde (2019) does mention examples of voiced sibilants in the context of a voiceless consonant. In general, it seems that in Basque both carry-over and anticipatory processes might be at work: voicing is not stopped immediately after the vowel and, if a voiced sounds follows, voicing tend to start early, and, in some cases, does not stop at all and continues throughout the sibilant.

Speech rate was measured automatically in this study. Doing so provides an interesting alternative to using manual segmentation of duration, which might not be totally reliable. Speech rate turned out to be important in the model of voicing. First of all, in faster speech the probability of having partially voiced sibilants increases: tokens which would be voiceless at slower rates can become partially voiced when the speech rate goes up. Secondly, the probability of full voicing increases with speech rate.

Apart from the effects of context and speech rate, the position within the word is also important for voicing, as word-initial sibilants are less likely to be fully voiced. There is also some evidence for word-final sibilants showing a slightly higher proportion of voiced frames, but, on the whole, word-final position does not differ much from word-internal position. Spanish word-final sibilants were observed to be weaker in Hualde & Prieto (2015), Torreira and Ernestus (2012) and File-Muriel and Brown (2011). In Hualde & Prieto (2015), however, there was a significant difference between word-final and medial position, but not for other comparisons (e.g. between word-initial and -final position). Torreira and Ernestus (2012) report a difference between word-final and other positions (initial and medial). More research is needed to find out whether there might be a difference between Spanish and Basque in the ways word-initial and -final sibilants are realised. Moreover, it might be interesting to compare the patterns characteristic of different classes of obstruents: in the analysis of Basque /t k/ in Hualde et al. (2019b) word-final consonants were weaker than medial and word-initial (smaller relative intensity and more voicing), and the word-initial position showed least weakening.

Voicing is known to affect other acoustic properties of a sound. Two such acoustic properties were also analysed in this paper: the centre of gravity and intensity (focusing especially on high frequencies, where fricative noise is generated). Voicing and frication are argued to be difficult to maintain at the same time (Ohala & Solé, 2010); because of that, if voicing is present in a fricative, the frication might be less intense and the centre of gravity lower (even if the parts of the spectrum which directly reflect the vibration of the vocal folds are filtered out). Voiced sibilants in this study had a lower centre of gravity (taking into account frequencies above 750 Hz) and lower relative intensity (comparing the minimum intensity of the preceding vowel to the maximum intensity of the sibilant in the 4-10 kHz band).

As regards lexical effects, it is not an easy task to find clear evidence for them. For the Basque data analysed here, the results are not straightforward, but interesting. The variable of word type (lexical vs grammatical) was not found influential in the model of voicing, but a more detailed analysis suggests that lexical factors might be important at the level of lexeme, as some words frequently contain a voiced sibilant. It is however difficult to tease apart possible reasons for that (lexeme, context, speech rate…). Among the lexemes with a particular tendency for voicing we find the negative particle *ez*. It is involved in two rules. One is the traditional devoicing of the initial consonant of the finite verb which follows the negation. The other, argued to be more recent, is the realisation of the initial consonant as voiced. In this paper I have shown that when the first consonant is dental, it is frequently indistinguishable in the spectrogram, which suggests its assimilation to the preceding sibilant. Whether this assimilation is complete or there is some remaining articulatory gesture of that dental consonant is difficult to determine in an acoustic study and would require further, preferably articulatory, analysis. The result of an assimilation of a lamino-alveolar sibilant and a dental approximant might be a sibilant with a more fronted place of articulation.

In a broader perspective, the behaviour of certain frequent expressions can be seen as examples of "special reduction" (Bybee et al., 2016). This refers to changes that happen in frequent words and expressions: "discrete variants that exist only for particular words or phrases in particular contexts" (Bybee et al., 2016, p. 423), for instance in discourse markers such as English *I don't know* and Spanish *o sea*. For instance, Bybee et al. (2016) found that in *o sea* the properties of the sibilant differed from that of the verb *sea* used as subjunctive: it tends to be shorter, have lower CoG and show more voicing. Bybee et al. (2016, p. 423) argue that "special reduction is determined by phonetic trends already present in the language which are accelerated in high-frequency phrases, and that special reduction is phonetically gradual". Data analysed in this paper suggest that the lenition attested in certain Basque expressions such as *ez dakit* 'I don't know' or *uste det* 'I think' which involve the assimilation of the consonant following the sibilant might be understood as such special reductions.

7. Conclusions

Analysing informal conversations provides an opportunity to dig into various lenition processes and better understand the ways phonemes can be realised. This paper has explored Basque fricative sibilants /s̺/ and /s̻/ in the variety of Beasain (Central Basque). It is the first acoustic analysis of Basque sibilants which focuses on voicing. The analysis has revealed that /s̺/ and /s̻/ are frequently partially or completely voiced. Factors which favour voicing were also identified: the most important were speech rate, the following phone (vowel vs consonant, voiced vs voiceless) and the presence of a boundary (less lenition in word-initial position). Additional spectral characteristics were also analysed, and the general outcome is that voiced sibilants differ from the voiceless ones in the centre of gravity and relative intensity. Thus, the presence of periodicity, lower centre of gravity and lower relative intensity are signs of lenition of the analysed sibilants in Basque.

Finally, it must be emphasised that more research is necessary to better understand lenition phenomena in Basque. For example, it is important to include more varieties and speakers of different age groups, but also to compare lenition phenomena in different classes of obstruents.

References

Bauer, L. (2008). Lenition revisited. *Journal of Linguistics, 44*(3), 605–624. https://doi.org/10.1017/S0022226708005331

Beristain, A. (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. https://doi.org/10.1017/S0025100320000274

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/

Broś, K., Żygis, M., Sikorski, A., & Wołłejko, J. (2021). Phonological contrasts and gradient effects in ongoing lenition in the Spanish of Gran Canaria. *Phonology*, *38*(1), 1–40. https://doi.org/10.1017/S0952675721000038

Bürkner, P.-C., Gabry, J., Weber, S., Johnson, A., Modrak, M., Badr, H. S., Weber, F., Ben-Shachar, M. S., & Rabel, H. (2022). *brms: Bayesian Regression Models using 'Stan'* (2.17.0) [Computer software]. https://CRAN.R-project.org/package=brms

Bybee, J., File-Muriel, R. J., & Souza, R. N. D. (2016). Special reduction: A usage-based approach. *Language and Cognition*, *8*(3), 421–446. https://doi.org/10.1017/langcog.2016.19

Chappell, W., & García, C. (2017). Variable production and indexical social meaning: On the potential physiological origin of intervocalic /s/ voicing in Costa Rican Spanish. *Studies in Hispanic and Lusophone Linguistics*, *10*(1), 1–37. https://doi.org/10.1515/shll-2017-0001

Chappell, W., García, C., & Davidson, J. (2023). Sociophonetics and fricatives. In C. Strelluf (Ed.), *The Routledge handbook of sociophonetics*. Routledge. https://doi.org/10.4324/9781003034636-9

Chodroff, E. R. (2017). *Structured variation in obstruent production and perception* [Doctoral Thesis, Johns Hopkins University]. JScholarship. https://jscholarship.library.jhu.edu/handle/1774.2/44696

Cohen Priva, U., & Gleason, E. (2020). The causal structure of lenition: A case for the causal precedence of durational shortening. *Language*, *96*(2), 413–448. https://doi.org/10.1353/lan.2020.0025

de Jong, N. H., Pacilly, J., & Heeren, W. (2021). PRAAT scripts to measure speed fluency and breakdown fluency in speech automatically. *Assessment in Education: Principles, Policy & Practice*, *28*(4), 456–476. https://doi.org/10.1080/0969594X.2021.1951162

DiCanio, C., Chen, W.-R., Benn, J., Amith, J. D., & Castillo García, R. (2022). Extreme stop allophony in Mixtec spontaneous speech: Data, word prosody, and modelling. *Journal of Phonetics*, *92*, 101147. https://doi.org/10.1016/j.wocn.2022.101147

DiCanio, C. (2012). The phonetics of fortis and lenis consonants in Itunyoso Trique. *International Journal of American Linguistics*, *78*(2), 239–272. https://doi.org/10.1086/664481

Dixit, R. P., & Hoffman, P. R. (2004). Articulatory characteristics of fricatives and affricates in Hindi: An electropalatographic study. *Journal of the International Phonetic Association*, *34*(2), 141–159. https://doi.org/10.1017/S0025100304001707

Eager, C. (2015). Automated voicing analysis in Praat: Statistically equivalent to manual segmentation. In M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. M. Scobbie (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0083.pdf

Egurtzegi, A., & Carignan, C. (2020). 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

Egurtzegi, A., Krajewska, D., Carignan, C., & Urrestarazu-Porta, I. (2024, in press). An acoustic exploration of sibilant contrasts and sibilant merger in Mixean Basque. *Journal of the International Phonetic Association*. https://doi.org/10.1017/S0025100324000045

Eguskiza, N., Gaminde, I., & de Pablo, I. (2020). Bokalen arteko herskari ahoskabeen ahostuntzea Arratian (Bizkaia). *Uztaro*, *113*, 51–64. https://doi.org/10.26876/uztaro.113.2020.3

Ennever, T., Meakins, F., & Round, E. R. (2017). A replicable acoustic measure of lenition and the nature of variability in Gurindji stops. *Laboratory Phonology*, *8*(1), 20. https://doi.org/10.5334/labphon.18

File-Muriel, R. J., & Brown, E. K. (2011). The gradient nature of s-lenition in Caleño Spanish. *Language Variation and Change*, *23*(2), 223–243. https://doi.org/10.1017/S0954394511000056

Fuchs, S., Brunner, J., & Busler, A. (2007). Temporal and spatial aspects concerning the realizations of the voicing contrast in German alveolar and postalveolar fricatives. *Advances in Speech Language Pathology*, *9*(1), 90–100. https://doi.org/10.1080/14417040601094315

Gráczi, T. E., Juhász, K., Csapó, T. G., Deme, A., & Markó, A. (2023). Dynamic articulatory and acoustic features of Hungarian sibilants as a function of phonological voicing. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 1077–1081). Guarant International.

Gurevich, N. (2011). Lenition. In M. van Oostendorp, C. J. Ewen, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 1–17). John Wiley & Sons. https://doi.org/10.1002/9781444335262.wbctp0066

Honeybone, P. (2008). Lenition, weakening and consonantal strength: Tracing concepts through the history of phonology. In J. Brandão de Carvalho, T. Scheer, & P. Ségéral (Eds.), *Lenition and fortition* (pp. 9–92). De Gruyter Mouton. https://doi.org/10.1515/9783110211443.1.9

Hualde, J. I. (2003). Segmental phonology. In J. I. Hualde & J. Ortiz de Urbina (Eds.), *A grammar of Basque* (pp. 15–64). De Gruyter Mouton. https://doi.org/10.1515/9783110895285.15

Hualde, J. I. (2010). Neutralización de sibilantes vascas y seseo en castellano. *Oihenart*, *25*, 89–116.

Hualde, J. I. (2014). Lenición de obstruyentes sordas intervocálicas en español: Estado de la cuestión. In Y.

Congosto, M. L. Montero, & A. Salvador (Eds.), *Fonética experimental, educación superior e investigación, Vol. 1* (pp. 113–136). Arco Libros.

Hualde, J. I. (2019). Fonologo baten galderak espektrogramaren aurrean. In I. Epelde & O. Jauregi (Eds.), *Bihotz ahots. M. L. Oñederra irakaslearen omenez* (pp. 337–351). UPV/EHU.

Hualde, J. I., Beristain, A., Icardo Isasa, A., & Zhang, J. (2019a). Hitz-bukaerako herskariak: Kontsonante talkak. *ASJU*, *53*(1/2), 83–106. https://doi.org/10.1387/asju.22412

Hualde, J. I., Beristain, A., Icardo Isasa, A., & Zhang, J. (2019b). Lenition of word-final plosives in Basque. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of International Congress of Phonetic Sciences 2019* (pp. 642–646). Australasian Speech Science and Technology Association Inc.

Hualde, J. I., & Bilbao, X. (1992). A phonological study of the Basque dialect of Getxo. *ASJU*, *26*(1), 1–117. https://doi.org/10.1387/asju.8283

Hualde, J. I., & Nadeu, M. (2012). Lenition and phonemic overlap in Rome Italian. *Phonetica*, *68*(4), 215–242. https://doi.org/10.1159/000334303

Hualde, J. I., Nadeu, M., & Simonet, M. (2010). Lenition and phonemic contrast in Majorcan Catalan. In S. Colina, A. Olarrea, & A. M. Carvalho (Eds.), *Romance linguistics 2009: Selected papers from the 39th Linguistic Symposium on Romance Languages* (pp. 63–79). John Benjamins. https://doi.org/10.1075/cilt.315.04hua

Hualde, J. I., & Prieto, P. (2015). Lenition of intervocalic alveolar fricatives in Catalan and Spanish. *Phonetica*, *71*(2), 109–127. https://doi.org/10.1159/000368197

Hualde, J. I., Simonet, M., & Nadeu, M. (2011). Consonant lenition and phonological recategorization. *Laboratory Phonology*, *2*(2), 301–329. https://doi.org/10.1515/labphon.2011.011

Hualde, J. I., & Zhang, J. (2022). Intervocalic lenition, contrastiveness and neutralization in Catalan. *Isogloss. Open Journal of Romance Linguistics*, *8*(4), 1–20. https://doi.org/10.5565/rev/isogloss.181

Iribar, A., Pagola, R. M., & Túrrez, I. (2020). Caracterización articulatoria de las fricativas sibilantes alveolares del euskera: Primera parte. *Loquens*, *7*(2), e073. https://doi.org/10.3989/loquens.2020.073

Iribar, A., Pagola, R. M., & Túrrez, I. (2022). Caracterización articulatoria de las fricativas sibilantes alveolares del euskera: Segunda parte. *Loquens*, *9*(1–2), e088. https://doi.org/10.3989/loquens.2022.088

Iribar, A., & Túrrez, I. (2008). Sobre las oclusivas sordas en el ámbito vasco-románico. *Oihenart*, *23*, 259–276.

Jongman, A., Wayland, R., & Wong, S. (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, M. (2011). Caracterización de sibilantes fricativas vascas y su percepción en el sistema fonético español. *ASJU*, *45*(1), 81–137.

Katz, J. (2016). Lenition, perception and neutralisation. *Phonology*, *33*(1), 43–85. https://doi.org/10.1017/S0952675716000038

Katz, J., & Pitzanti, G. (2019). The phonetics and phonology of lenition: A Campidanese Sardinian case study. *Laboratory Phonology*, *10*(1). https://doi.org/10.5334/labphon.184

Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2004). Domain-initial articulatory strengthening in four languages. *Phonetic Interpretation: Papers in Laboratory Phonology VI*, 143–161.

Kingston, J. (2008). Lenition. In L. Colantoni & J. Steele (Eds.), *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 1–31). Cascadilla Proceedings Project. https://www.lingref.com/cpp/lasp/3/abstract1711.html

Kirchner, R. (2004). Consonant lenition. In B. Hayes, D. Steriade, & R. Kirchner (Eds.), *Phonetically based phonology* (pp. 313–345). Cambridge University Press. https://doi.org/10.1017/CBO9780511486401.010

Lewis, A. M. (2001). *Weakening of intervocalic /p, t, k/ in two Spanish dialects: Toward the quantification of lenition processes* [PhD Thesis, University of Illinois at Urbana-Champaign]. ProQuest. https://www.proquest.com/docview/304695161/abstract/32F367C501C142CBPQ/1

Liker, M., & Gibbon, F. E. (2013). Differences in EPG contact dynamics between voiced and voiceless lingual fricatives. *Journal of the International Phonetic Association*, *43*(1), 49–64. https://doi.org/10.1017/S0025100312000436

Mitxelena, K. (2011). Fonética histórica vasca. In J. A. Lakarra & I. Ruiz Arzalluz (Eds.), *Obras completas* (Vol. 4). UPV/EHU. (Original work published 1977)

Muxika-Loitzate, O. (2017). Sibilant merger in the variety of Basque spoken in Amorebieta-Etxano. *Languages*, *2*(4), 25. https://doi.org/10.3390/languages2040025

Nadeu, M., & Hualde, J. I. (2015). Biomechanically conditioned variation at the origin of diachronic intervocalic voicing. *Language and Speech*, *58*(3), 351–370. https://doi.org/10.1177/0023830914554727

Narayanan, S. S., Alwan, A. A., & Haker, K. (1995). An articulatory study of fricative consonants using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *98*(3), 1325–1347. https://doi.org/10.1121/1.413469

Navarro Tomás, T. (1957). *Manual de pronunciación española*. Hafner. http://archive.org/details/manualdepronunci00nava

Nirgianaki, E. (2014). Acoustic characteristics of Greek fricatives. *The Journal of the Acoustical Society of America*, *135*(5), 2964–2976. https://doi.org/10.1121/1.4870487

Núñez-Méndez, E. (2022). Variation in Spanish /s/: Overview and new perspectives. *Languages*, *7*(2), 77. https://doi.org/10.3390/languages7020077

Ohala, J. J., & Solé, M.-J. (2010). Turbulence and phonology. In *Turbulence and phonology* (pp. 37–102). De Gruyter Mouton. https://doi.org/10.1515/9783110226584.37

Oñederra, M. L. (1990). *Euskal fonologia: Palatalizazioa. Asimlazioa eta hots sinbolismoa*. UPV/EHU.

Oñederra, M. L. (2004). *Fonetika fonologia hitzez hitz*. UPV/EHU.

Proctor, M. I., Shadle, C. H., & Iskarous, K. (2010). Pharyngeal articulation in the production of voiced and voiceless fricativesa). *The Journal of the Acoustical Society of America*, *127*(3), 1507–1518. https://doi.org/10.1121/1.3299199

Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Gredos.

R Core Team. (2024). *R: A language and environment for statistical computing* (Version 4.3.3) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Sedó, B., Schmidt, L. B., & Willis, E. W. (2020). Rethinking the phonological process of /s/ voicing assimilation in Spanish: An acoustic comparison of three regional varieties. *Studies in Hispanic and Lusophone Linguistics*, *13*(1), 167–217. https://doi.org/10.1515/shll-2020-2027

Silbert, N., & de Jong, K. (2008). Focus, prosodic context, and phonological feature specification: Patterns of variation in fricative production. *The Journal of the Acoustical Society of America*, *123*(5), 2769–2779. https://doi.org/10.1121/1.2890736

Smith, C. 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

Tabain, M. (2019). The phonetic properties of consonants. In W. F. Katz & P. F. Assmann (Eds.), *The Routledge handbook of phonetics* (pp. 65–105). Routledge. https://doi.org/10.1017/9781108644198.004

Torreira, F., & Ernestus, M. (2011). Realization of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology*, *2*(2), 331–353. https://doi.org/10.1515/labphon.2011.012

Torreira, F., & Ernestus, M. (2012). Weakening of Intervocalic /s/ in the Nijmegen Corpus of Casual Spanish. *Phonetica*, *69*(3), 124–148. https://doi.org/10.1159/000343635

Tucker, B. V., & Ernestus, M. (2016). Why we need to investigate casual speech to truly understand language production, processing and the mental lexicon. *The Mental Lexicon*, *11*(3), 375–400. https://doi.org/10.1075/ml.11.3.03tuc

Urrutia, S. H., Etxebarria, M., Túrrez, I., & Duque, J. C. (1988). *Fonética vasca I: Las sibilantes en el vizcaíno*. Universidad de Deusto.

Urrutia, S. H., Etxebarria, M., Túrrez, I., & Duque, J. C. (1989). *Fonética vasca II: Las sibilantes en el guipuzcoano*. Universidad de Deusto.

Warner, N., & Tucker, B. V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *The Journal of the Acoustical Society of America*, *130*(3), 1606–1617. https://doi.org/10.1121/1.3621306

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Zuazo, K. (2014). *Euskalkiak*. Elkar.

Appendix

Table i. Population-level effects in the model of voicing.

Parameter	Estimate	Estimate error	95% CrI	
			Lower	Higher
μ				
Intercept	-1.39	0.13	-1.66	-1.13
phone [s̺]	-0.11	0.08	-0.26	0.05
speech_rate	0.02	0.02	-0.02	0.05
context_right [_Cvless]	0.31	0.07	0.17	0.45
context_right [_Cvoiced]	0.48	0.14	0.2	0.76
boundary [#_]	0.15	0.1	-0.05	0.35
boundary [_#]	0.3	0.12	0.07	0.53
lex_type [gram]	0.09	0.08	-0.06	0.24
ϕ				
Intercept	1.88	0.17	1.56	2.21
phone [s̺]	0.22	0.11	0.02	0.43
speech_rate	-0.05	0.03	-0.1	0
context_right [_Cvless]	-0.27	0.1	-0.46	-0.07
context_right [_Cvoiced]	-0.38	0.18	-0.74	-0.03
boundary [#_]	-0.27	0.14	-0.54	0.01
boundary [_#]	-0.31	0.15	-0.6	-0.03
lex_type [gram]	-0.04	0.1	-0.23	0.16
α				
Intercept	1.82	0.23	1.36	2.27
phone [s̺]	0.35	0.13	0.1	0.59
speech_rate	-0.09	0.02	-0.14	-0.05
context_right [_Cvless]	-1.12	0.1	-1.32	-0.92
context_right [_Cvoiced]	-0.4	0.17	-0.73	-0.06
boundary [#_]	0.51	0.12	0.27	0.75
boundary [_#]	0.35	0.17	0.03	0.68
lex_type [gram]	0.06	0.14	-0.2	0.34
γ				
Intercept	-2.84	0.45	-3.71	-1.91
phone [s̺]	0.01	0.2	-0.39	0.41
speech_rate	0.21	0.03	0.14	0.27
context_right [_Cvless]	-1.46	0.17	-1.8	-1.13
context_right [_Cvoiced]	2.63	0.23	2.19	3.08
boundary [#_]	-1.02	0.16	-1.34	-0.71
boundary [_#]	0.31	0.24	-0.15	0.78
lex_type [gram]	0.28	0.22	-0.14	0.71

Table ii. Population-level effects in the model of CoG.

Predictor	Estimate	Estimate error	95% CrI	
			Lower	Higher
Intercept	8010	215	7579	8431
phone [s̪]	-2426	46	-2514	-2335
voiced_frames_mid	-1295	44	-1383	-1208
phone [s̪] × voiced_frames_mid	712	73	569	855

Table iii. Population-level effects in the model of intensity.

Predictor	Estimate	Estimate error	95% CrI	
			Lower	Higher
Intercept	17.2	0.69	15.84	18.58
phone [s̪]	6.01	0.4	5.22	6.8
voiced_frames_mid	-6.9	0.32	-7.52	-6.28
phone [s̪] × voiced_frames_mid	-1.51	0.52	-2.54	-0.52