

A parametric study of the spectral characteristics of European Portuguese fricatives

Luis M. T. Jesus* and Christine H. Shadle

*Department of Electronics and Computer Science, University of Southampton,
Highfield, Southampton SO17 1BJ, UK*

Received 9th November 2001, and accepted 3rd January 2002

Studies of Portuguese phonetics and phonology indicate that fricatives are central to some interesting features of the language, yet studies of Portuguese fricatives have been few and limited. In this study, Portuguese fricatives were analyzed in ways designed to enhance our description of the language and to increase our understanding of the production of fricatives. Corpora of Portuguese words containing /f, v, s, z, ʃ, ʒ/, nonsense words of the pattern /V₁FV₂/ that follow Portuguese phonological rules, and sustained fricatives were recorded by four native speakers of European Portuguese (two men, two women). Results of analysis show that more than half of the voiced fricatives devoice; devoicing occurs more often in word-final fricatives. Averaged power spectra were computed for all fricatives and parameterized in order to aid comparisons across speaker and across corpus, and to gain insight into the production mechanisms underlying the language-specific variations. Substantial differences were found between spectra of voiced and unvoiced, same-place fricatives. The parameters spectral slope, frequency of maximum amplitude, and dynamic amplitude, derived from previous studies, behaved as predicted for changes in effort level, voicing, and location within the fricative. Changes in syllable stress, however, did not affect the fricatives in a manner consistent with effort level variation. Some combinations were also useful for separating the fricatives by place or by sibilance.

© 2002 Elsevier Science Ltd. All rights reserved.

1. Introduction

There have been many studies of the phonetics and phonology of Portuguese, which have shown some interesting features of the language; it is unusually rich in instances of vowel reduction (Andrade, 1995), consonant clusters (Martins, Harmegnies & Poch, 1995), and plosives that are realized as fricatives (Viana,

*Present address: Escola Superior de Saúde da Universidade de Aveiro, and Instituto de Engenharia Electrónica e Telemática, de Aveiro, Universidade de Aveiro, 3810-193 Aveiro, Portugal.

1984). The study of Portuguese fricative consonants constitutes a challenging and complex research area, which is as yet incompletely explored (Lacerda, 1982). This is in part due to the lack of specific speech corpora that reflect the variety of phonetic contexts in which these speech sounds occur, and the large number of variations in fluent Portuguese speech.

In this paper, we present a study of the acoustic properties of Portuguese fricatives, based on a large corpus that includes nonsense and real words, and sustained fricatives. The multiple comparisons possible in such a study, across speakers, corpora, place, vowel context, syllable stress, time within fricative, etc., demanded a systematic approach, because our interest is primarily in the production mechanisms of the fricatives and the language-specific variation of these mechanisms. We parameterized the fricative spectra using parameters developed in previous studies: spectral slope, frequency of maximum amplitude, and dynamic amplitude. These are related to the underlying aeroacoustic mechanisms and allow for more effective comparisons of the far-field acoustic spectra.

While our long-range goal is improved synthesis of Portuguese, the analysis methods described in this paper are more closely related to other goals that have long proven difficult for fricatives. Distinguishing fricatives in spite of changes caused by context effects is the fundamental task of speech recognition. Identifying differences in production when the intended fricative is known is needed for studies of disordered speech, and changes in speech over time, e.g., after a cochlear implant. Both of these applications require some knowledge of the distinguishing acoustic characteristics of fricatives that hold across speaker, context, and speaking style. We consider previous studies of distinguishing characteristics of fricatives below, grouped by type of parameter. Locus equations have been used by many authors. The method consists of finding the slope and the intercept of a line between the same acoustic feature in a vowel and an adjoining consonant. For instance, Sussman (1994) used $F_2(\text{onset}) = kF_2(\text{vowel}) + c$ on both stops and fricatives. The corpus consisted of a small number of nonsense words produced by four speakers of American English. k and c could be used to distinguish stops, but did not distinguish fricatives well. Jongman, Wayland & Wong (2000) also studied locus equations as cues to place of articulation. The corpus consisted of nonsense words produced by 20 speakers of American English. The slope of locus equations could be used to differentiate labiodentals from the other three places of articulation.

Forrest, Weismer, Milenkovic & Dougall (1988) used spectral moments to characterize normal speech with the intent of using them on disordered speech. In this method, the spectral envelope was treated as a probability density function and the first four moments of that function (the mean, variance, skewness and kurtosis) were found and used in a cluster analysis. Results showed that spectral moments worked well to classify stops but could not distinguish all fricatives. However, the authors used a very limited corpus: only five words contained fricatives. They were produced by 10 speakers of American English.

Shadle & Mair (1996) used spectral moments, as in Forrest *et al.* (1988), on a large fricative corpus recorded by one American English and one French native speaker. These researchers examined multiple tokens, varying effort levels, different vowel contexts, and three different locations within a fricative. They showed that the moments that Forrest *et al.* (1988) found to be the most useful for distinguishing

fricatives (especially skewness) were not so useful with their data; the differences across context within place were typically much greater than the differences across place. Jongman *et al.* (2000) seemingly contradicted this finding in their study using 20 American speakers, and six vowel contexts for each fricative. ANOVAs showed that spectral skewness differed significantly by place for all four places, and the other moments were significantly different for two or three place-groups. For these studies, all subjects, voicing, vowel contexts, and location within fricative were lumped together. The implication then is that the distributions for each place overlapped substantially, and classification based on these measures would likely have a high failure rate. Very little information is given about the effect of voicing, vowel context, etc., on the values of each moment.

Wilde (1995) studied acoustic cues (place and voiced/voiceless categorization) in fricative–vowel boundaries and assessed perceptual importance of various time- and frequency-domain parameters via synthesis. She used nonsense words produced by four speakers of American English. She concluded that temporal and spectral characteristics of voiceless fricatives are more dependent on vowel context than those of voiced fricatives, and what she referred to as “formant onset time of fricatives” provides important place information. She also showed that the amplitudes of fricative noise in restricted frequency regions can distinguish sibilants¹ from nonsibilants.

Funatsu (1995), in a cross-language study of Japanese and Russian fricatives in nonsense words, measured two spectral parameters: the frequency of the most prominent fricative peak (F_F) and the onset frequency of the second formant transition of the following vowel (F_V), both measured over a 0–10 kHz range. Within each language /s, ʃ/ could be distinguished using F_F (for Russian) or F_F and F_V (for Japanese). However, the parameter values varied with vowel context and the nature of this variation was language-dependent.

Shadle & Mair (1996) defined two parameters, dynamic amplitude and spectral slope, which will be discussed in more detail later. These did not distinguish the fricatives completely but did vary with source location and effort level as predicted.

Evers, Reetz & Lahiri (1998) tried to distinguish and characterize the fricatives /s, ʃ/ produced by two speakers each of English, Bengali and Dutch (12 real words). They used power spectra computed from a single 40 ms window placed mid-fricative, and calculated the slopes of linear regression lines fit to spectra from 0 to 2.5 kHz (S_a) and from 2.5 to 8 kHz (S_b). Their results showed that it was possible to separate /s/ from /ʃ/ by using the difference in slope below and above 2.5 kHz, i.e., $(S_a - S_b)_f > (S_a - S_b)_s$. The slope difference was successful in categorizing the two sibilants within a range of 7–15 dB/kHz across the three languages. Results also showed that there is no vowel influence in the discrimination, and that there is a variation between speakers.

Choo & Huckvale (1997; Choo, 1999) studied the correlations between perceptual and physical spaces of voiceless English fricatives using a multidimensional scaling technique. Results from perceptual tests suggested that a two-dimensional solution was the most appropriate to model the data. Dimension 1, shown to be related to

¹Sibilant — a fricative sound made by producing a narrow, groove-like stricture between the blade of the tongue and the back part of the alveolar ridge (Crystal, 1997); /s, z/ and /ʃ, ʒ/ are examples. This is a phonetic classification based on the manner of articulation.

the “peakiness” of spectra (the difference between maximum amplitude and mean amplitude), clearly separated the sibilants from nonsibilants. Dimension 2, related to the center of gravity of the spectra, separated fricatives according to their place of articulation. These representations were constructed from both perceptual similarity judgments and a Euclidean spectral distance metric obtained from 1/3-octave bandpass filtering.

Jongman *et al.* (2000) studied noise duration, noise amplitude, and the frequency of the highest amplitude spectral peak, as cues to place of articulation. Spectral peak location decreased in frequency as place of articulation moved posteriorly. The amplitude of the highest amplitude spectral peak differed significantly for all four places of articulation, for values computed across speakers, voicing and vowel context. Voiced fricatives had smaller amplitude relative to the vowel preceding them than unvoiced fricatives, with a larger difference between voiced and unvoiced for nonsibilants than for sibilants. Noise duration differed significantly for sibilants and nonsibilants.

Parameters similar to those used by Shadle & Mair (1996), and a parameter similar to S_a used by Evers *et al.* (1998), were used in this study in order to compare fricatives across-speaker, relate the more controlled productions (sustained and nonsense words) to those of real words, and gain insight into the production mechanisms underlying the variations specific to Portuguese. In the process, we have not only enhanced our understanding of Portuguese fricatives, but have contributed to the methodology for studying fricatives of any language.

2. Corpus

2.1. Design

A rich variety of phonetic contexts using both real Portuguese words and nonsense words was selected to study the most relevant phoneme variants, and fully describe the spectral and articulatory characteristics of Portuguese fricative consonants. The corpora also included sustained fricatives, which are better controlled and easier to analyze than those occurring in words.

To produce examples that would be possible words in Portuguese, the nonsense words all followed these generally accepted (for European Portuguese) language-specific phonological rules (Mateus & Andrade, 2000, p. 11):

1. vowels /ɐ/ and /u/ can occur in the tonic syllable;²
2. vowels /i/, /ɐ/ and /u/ can occur before and after the tonic syllable;
3. the fricatives /f, v, s, z, ʃ, ʒ/ can all occur in initial and medial positions;
4. /ʃ/ is the only fricative that can occur in word-final position.

Some of the nonsense words had unambiguous stress patterns due to the two vowels chosen (/pɐFi, puFi/ = syllable 1 stressed; /piFɐ, piFu/ = syllable 2 stressed). In ambiguous cases (where $V_1 = V_2 = /i/$, and where neither V_1 nor $V_2 = /i/$), the stress was marked for subjects 2–4 according to that used by subject 1 (LMTJ).

²Tonic syllable-syllable which carries maximal prominence, usually owing to a major pitch change (Crystal, 1997).

In addition to these constraints, the corpora were designed to be compatible where possible with the fricative corpora recorded for English, American, French and German subjects (Shadle, 1992; Shadle & Carter, 1993) to facilitate comparisons.

2.2. Corpus 1: Sustained fricatives

Corpus 1a consisted of a set of VCV sequences, where V belongs to the reduced set of Portuguese vowels /i, e, u/, and C is one of the Portuguese fricative consonants /f, v, s, z, ʃ, ʒ/ sustained for 5 s. As shown by Shadle, Mair & Carter (1996), the vowel context, even for sustained examples, influences the articulatory and spectral characteristics of fricatives. The vocalic contexts of Corpus 1a overlap with those of Corpus 3 (set of Portuguese words), which facilitated comparison.

A separate set of Portuguese fricative consonants, sustained for 3 s, at medium, soft and loud effort levels, was also recorded (and is called Corpus 1b). Ideally, we would like the articulation to be held constant, and only the flow velocity through the constriction to be varied. We attempted to elicit this by asking for a variation in effort level.

2.3. Corpus 2: Nonsense words

Corpus 2 consisted of /pV₁CV₂/ nonsense words, where V₁, V₂ were one of the vowels /i, e, u/. The set comprised all possible vowel and fricative permutations, each repeated about 12 times in one breath. The phoneme /p/ is an easily identifiable marker for segmentation and spectral analysis, and has been used in previous fricative recordings by Shadle *et al.* (Shadle, 1992; Shadle & Carter, 1993).

The stress was placed according to language-specific phonological rules, and subjects were instructed to keep it the same through all the repetitions. The subjects were not always able either to produce the indicated stress pattern or to produce a different pattern consistently, so there were some instances with equal stress in both syllables, and with deleted vowels.

2.4. Corpus 3: Real words

Corpus 3 consisted of 154 words, each said within the frame sentence “Diga ..., por favor.”, /diɡə ... pur ‘fəvor/, which was used to record the words in the corpus in a balanced phonetic context and with a neutral prosody. The words, listed in Appendix A, were presented in randomized order.

The 154 words consist of eight words forming nearly minimal pairs with the pattern /FV₁FV₂/; 54 words with the pattern /FV_/_/ (fricative in initial position); 69 words with the pattern /_V₁FV₂_/_/ (fricative in medial position); and 23 words with the pattern /_VF/_/ (fricative in final position).

The vowels in words with sequences /FV₁FV₂/, /FV/_/_/, /V₁FV₂_/_/ and /VF/_/_/ have been divided into three groups according to their location in the vowel triangle: /i, e/ — group 1; /ɛ, ɐ, a/ — group 2; /ɔ, o, u/ — group 3. Appendix A includes examples with nearly all Portuguese nonnasal vowels preceding each of the fricatives, followed by one vowel from each of the vowel groups. The vowel /i/ is generally deleted in final position, as shown by Andrade (1994), so the resulting allophone is not expected to influence the preceding fricative. Therefore, words such

as “chefe” /'ʃɛfi/, “ave” /avi/ and “asse” /'asi/, were used to “simulate” final position contexts. As mentioned by Mateus & Andrade (2000), phonologically, only /f/ can occur in final position, but phonetically any fricative can be found in word final position as a consequence of deletion of unstressed vowels.

2.5. *Corpus 4: Real words in connected speech*

Corpus 4 consisted of a set of sentences including 60 words from Corpus 3 (see Appendix B). Ten of the sentences are meaningful; two include word boundaries within some of the phonetic sequences in Corpus 3, but are semantically nonsense.

3. Recording and preliminary analysis method

3.1. *Recording*

The subjects used in this study were two male (LMTJ and CFGA) and two female (ACC and ISSS) adult Portuguese native speakers. Recordings were made in a sound-treated room using a Bruel & Kjaer 4165 1/2 in microphone located 1 m in front of the subject's mouth. The signal was preamplified (Bruel & Kjaer 2639), then amplified and filtered by a Bruel & Kjaer 2636 measurement amplifier, with high-pass cuton frequency of 22 Hz and lowpass cutoff frequency of 22 kHz. A laryngograph signal (Lx) was also collected using a laryngograph processor (model LxProc, type PCLX produced by Laryngograph Ltd, U.K.). The acoustic speech signal and Lx were recorded with a Sony TCD-D7 DAT recorder, each with 16 bits and a sampling frequency of 48 kHz.

3.2. *Segmentation and annotation*

The time waveforms of all the corpus words were manually analyzed to detect the start of the vowel–fricative transition, the start and end of the fricative, and the end of the fricative–vowel transition. During the vowel–fricative transition, there is a decrease in amplitude, voicing ceases (for unvoiced fricatives) and frication noise starts. During the fricative–vowel transition, there is an increase in amplitude, voicing starts (for unvoiced fricatives) and frication noise ceases. These events do not occur simultaneously or always in the same order, making the segmentation a somewhat subjective process. However, it is important to segment consistently, because the results of the analysis methods depend on where the boundaries are placed (Docherty, 1992, pp. 103–110). The amplitude and voicing changes appear in both acoustic and Lx signals, which aids the segmentation process.

3.3. *Devoicing*

While segmenting, we noticed a large number of devoiced examples. To study this phenomenon further we devised a new criterion for devoicing based on both the acoustic and laryngograph signals.

Smith (1997, p. 478) used a criterion for devoicing in American English based on the amplitude of the electroglottograph (EGG) cycles:

The fricative was considered to be voiced during the portion of its duration that the amplitude of the EGG cycles exceeded one-tenth of the EGG cycle amplitude at the time of maximum energy in the preceding vowel.

A study by Pirello, Blumstein & Kurowski (1997, p. 3756) used an alternative measure of voicing based on the acoustic signal:

An amplitude difference greater than 10 dB between the amplitude of the vowel and frication noise was classified as voiceless. A difference of less than or equal to 10 dB sustained over 30 ms was classified as voiced.

We used both the acoustic signal and the laryngograph signal to determine if a fricative was devoiced. A fricative was called *devoiced* when less than one-third of the frication interval showed periodic structure in the acoustic or laryngograph signals. The term *partially devoiced* was used when more than one-third but less than half of the frication interval contained steady acoustic and laryngograph signal cycles. A fricative was called *voiced* when more than half of the frication interval showed steady acoustic and laryngograph signal cycles, even if the amplitude was much lower than in the vowel. If the laryngograph signal was clearly periodic, the fricative was classified as voiced; if the laryngograph signal was zero or distorted, the signal was classified as voiced only if the acoustic signal was unambiguously periodic.

3.4. Spectral analysis

Stochastic signals require some form of averaging for their spectra to be both consistent and low-error estimates of the underlying distribution (Bendat & Piersol, 2000, p. 423–442). For a stationary signal, time-averaging can be used; for nonstationary signals where an ensemble exists, ensemble-averaging can be used; or a single spectrum can be smoothed, and the averaging achieved at the expense of frequency resolution. We varied the method according to corpus. In Corpus 2, where multiple tokens of the nonsense words existed, ensemble-averaging was used; time-averaging was used elsewhere (Corpus 1a, 1b, 3 and 4), as shown in Fig. 1. The duration of the windows (10 ms both for time and ensemble-averaging) was chosen so that a reasonable number of windows could be used to cover adequately the wide range of fricative durations in the corpora (from 40 to 200 ms). For the shorter fricatives the windows overlapped. We used nine windows to calculate the averaged spectra, because observation of Corpus 2 indicated that there were always at least nine valid repetitions of the nonsense words. This allowed us to compare the spectra of fricatives calculated from real Portuguese words (Corpus 3 and 4) with the ensemble-averaged spectra of nonsense words.

We used time-averaging with nine 10 ms Hamming windows, one left-aligned to the start of the fricative, one right-aligned to the end of the fricative, one centered at the middle of the fricative, and the rest evenly distributed in between. This meant that the amount of overlap varied according to fricative duration. The longest fricatives (~ 200 ms) had zero overlap; the shortest (~ 40 ms) had approximately 60% overlap, between successive windows. The time-averaged power spectrum for

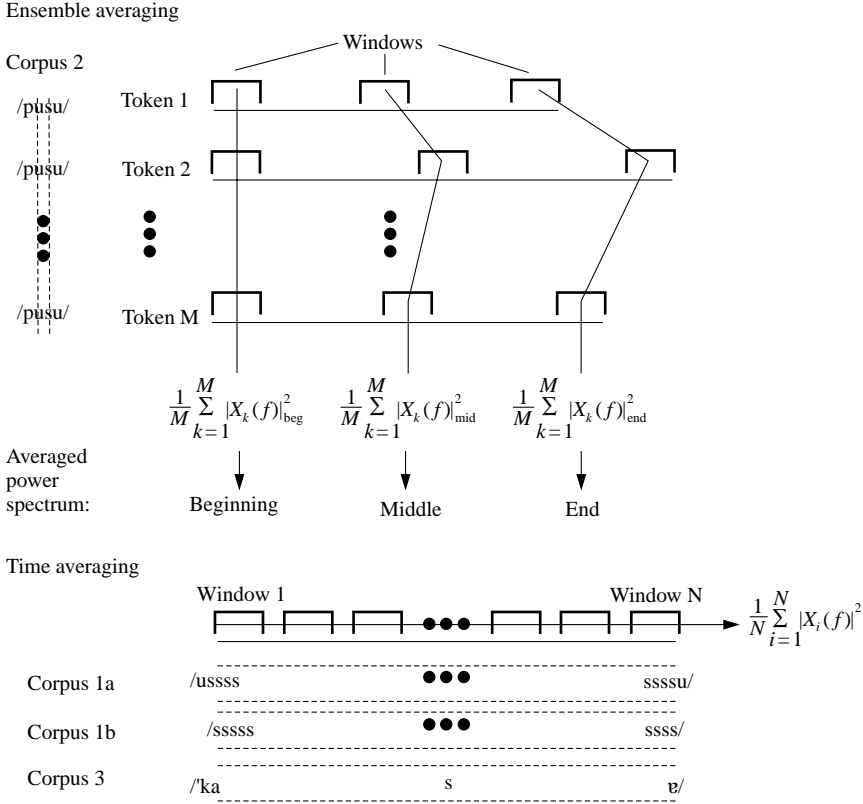


Figure 1. Diagram showing how windows are placed within fricative tokens, and how the corresponding power spectra $X_k(f)$ are combined to compute ensemble and time-averaged spectra.

each fricative is given by

$$P_T(f) = \frac{1}{N} \sum_{i=1}^N |X_i(f)|^2 \quad (1)$$

where X_i is the DFT of a portion of the fricative signal, x_i , corresponding to the i th windowed segment.

We used the same time-averaging technique to calculate the spectra of sustained fricatives in Corpus 1a and 1b, but since the sustained fricatives were so long, we used $N = 100$ windows (each 10 ms long).

Ensemble averaging, based on one DFT computed at the same event in each of nine tokens, was used for Corpus 2. The four speakers produced, on average, more than 12 repetitions of each nonsense word. However, the first, last, and any atypical tokens were eliminated, thus resulting in ensembles of nine tokens each. The 10 ms windows were located relative to events within one fricative: one left-aligned to the start of the fricative, one centered mid-fricative, and one right-aligned to the end of the fricative (see Fig. 1). The ensemble-averaged power spectrum of each fricative is

given by

$$P_E(f) = \frac{1}{9} \sum_{k=1}^9 |X_k(f)|^2 \quad (2)$$

where X_k is the DFT of a portion of the fricative signal, x_k , corresponding to the windowed segment (at the beginning, middle or end of the fricative) of the k th token.

4. Parameterization of fricative spectra

The mechanical model results of Shadle (1985) were used to define parameters that characterize the fricatives in the present study. These parameters have already been developed as a potential tool for classifying fricatives using real speech (Shadle & Mair, 1996). They consist of measures of spectral slope and of the dynamic range of the spectrum, and are applied to the spectrum of the far-field acoustic signal.

The far-field acoustic signal is the result of the excitation of the vocal tract transfer function by the source (for unvoiced) or sources (for voiced fricatives). The transfer function consists of poles, which are the resonances of the entire vocal tract, and zeros, which are antiresonances, related to the position of the source with respect to the tract.

It can be shown (Shadle, 1985; Stevens, 1998) that a source located in an intermediate position (i.e., not at the glottal end of the tract) always produces a zero at low frequencies. In a typical fricative configuration, articulators form a small constriction that separates the upstream back cavity from the downstream front cavity. Noise is generated somewhere downstream of that constriction. A set of antiresonances will be generated that nearly cancel back-cavity resonances. Front-cavity resonances are not cancelled, however, and a set of antiresonances are generated at frequencies related to the distance between noise source and constriction. These antiresonances result in sharply-defined troughs in the spectrum if the noise source is localized; if the source is distributed, the troughs will be correspondingly smeared.

The spectral prominence of the uncanceled resonances will depend on a number of factors: the particular ordering of resonances and antiresonances in the transfer function as a whole, the losses (particularly radiation losses) and the noise source strength. Above approximately 5 kHz nonplanar modes begin to propagate; the cuton frequency is inversely proportional to the longest cross-dimension of the tract. The radiation impedance decreases, and losses due to radiation thus decrease, above that frequency.

The noise source spectrum depends on the shape of the constriction, the tract downstream of it, and the flow velocity through it. This source spectrum has been described as having a broad peak, with maximum amplitude at a frequency proportional to the mean velocity through the constriction (Stevens, 1971). This is the spectrum of turbulence noise generated by a free jet, and the parameters that allow spectra for any jet diameter or velocity to be collapsed into a single curve have been amply described in the literature (e.g., Goldstein, 1976). However, with most fricatives the jet emerging from the constriction cannot freely expand, but

impinges on the tract walls downstream, generating additional noise; at speech dimensions and velocities, these sources of noise are generally of much higher amplitude than the self-noise of the jet.

Nelson & Morfey (1981) investigated the noise generated by a spoiler in a duct. These spectra likewise can be collapsed onto a single curve, but with an important difference: the spectrum does not show a broad peak, but instead shows a progressive decrease in amplitude as frequency increases. The source spectrum shape was similar to that derived by Shadle (1990) for obstacle-source mechanical models, which were shown to be good models for /s, ʃ/ (Shadle, 1991).

Fig. 2 shows an idealized fricative noise source spectrum. If the tract geometry, including constriction area, remains the same and flow velocity is increased, the spectral envelope of the noise increases in amplitude at all frequencies, *but more so at higher frequencies* (Shadle & Mair, 1996; Krane, 1999). A greater pressure drop across the constriction (Δp_c) results in a higher velocity through it, and thus an increased source amplitude. For unvoiced fricatives, greater subglottal pressure results in a greater Δp_c ; for voiced fricatives, a pressure drop is required to drive the vocal-fold oscillation, and thus the noise source in voiced fricatives is generally weaker on average than in their unvoiced counterparts.

In a sustained fricative, the essentially static tract configuration produces static source and filter characteristics. Increasing the effort level should closely approximate altering the source while keeping the filter constant. For fricatives in vowel context, however, the articulators move at least at the start and end of the fricative, setting up filter and source characteristics simultaneously, and may move during the fricative; for instance, the tongue constriction may be held while the lips round in preparation for the following vowel. While we hope to capture the source differences that occur during a fricative by using ensemble averaging at its

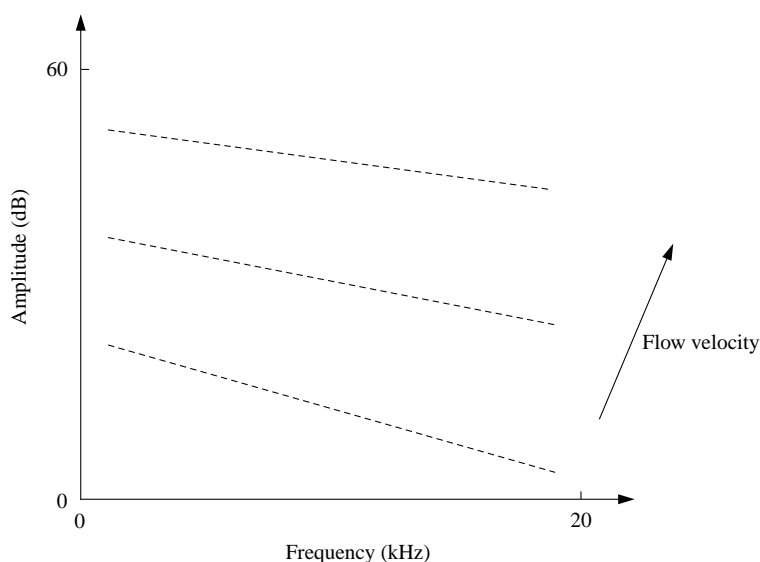


Figure 2. Schematic illustration of the effect of increasing airflow velocity on the noise source spectral envelope.

beginning, middle and end, we must remember that the filter is changing during this time as well. When we are forced to time-average through a fricative all of these effects are averaged together.

In this study, we are primarily interested in describing the acoustic variation caused by the context or the way in which a particular fricative is spoken, rather than identifying the fricative regardless of its context. We are thus interested primarily in changes in the source spectrum, since it offers clues to the source variations across subject and place. We have devised parameters accordingly. However, by restricting ourselves to the far-field acoustic spectrum, it is not always possible to differentiate source and filter characteristics unambiguously.

The parameter F was defined as the frequency at which the spectral amplitude was maximum, excluding the fundamental and its harmonics in voiced fricatives. The expectation was that F would correspond to the frequency of the first front-cavity resonance; it would thus shift with place significantly, and with rounded vowel context somewhat. It would also provide an endpoint for the lines fit to determine spectral slope. However, particularly for /f, v/, the values of F for the subjects ranged widely, from 4.2 to 7.8 kHz. These differences were not interesting, since the spectra were relatively flat. We therefore computed parameter \bar{F} , the average (rounded to the nearest kHz) of the values of F of all sustained tokens (Corpora 1a and 1b) for each place for all four speakers $\bar{F}_{/f,v/} = 5$ kHz, $\bar{F}_{/s,z/} = 6$ kHz and $\bar{F}_{/ʃ,ʒ/} = 4$ kHz. In Fig. 3, F differs from \bar{F} , but the spectral

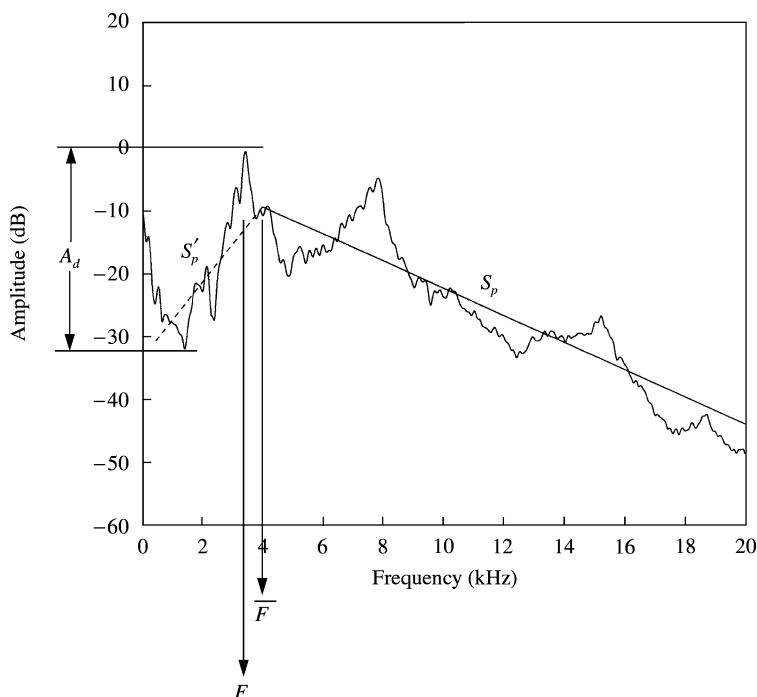


Figure 3. Dynamic amplitude A_d , and regression lines used to calculate low-frequency (500 Hz to \bar{F} kHz) slope S'_p (dashed line) and high-frequency (\bar{F} kHz to 20 kHz) slope S_p (solid line). Sustained fricative /ʃ/ (Corpus 1a) produced by Speaker ISSS.

slopes resulting appear to characterize the spectral shape fairly, and are computed using the same frequency range (and number of points) for all subjects.

By using \bar{F} we are ignoring changes in the spectral shape with vowel context. In Corpus 1a, in a majority of cases the rounded-vowel context had the lowest F , as expected, but this was not consistent either within subject for a given fricative, or across fricatives.

The dynamic amplitude, A_d , is the difference between the maximum amplitude value of the averaged power spectrum occurring between 500 Hz (lower limit set so that neither room noise, nor the peaks corresponding to the fundamental frequency and its first few harmonics, are used in the calculation of A_d) and 20 kHz, and the minimum amplitude between 0 and 2 kHz. Two linear regression lines are fit to the spectrum; S'_p is the slope of the line fit to all the spectral amplitude points from 500 Hz to \bar{F} (shown in Fig. 3 as a dashed line), and S_p is the slope of the line fit to all the points from \bar{F} to 20 kHz (solid line). This frequency range allowed us to capture relevant variations in the slope of the spectrum which was not possible in previous studies such as the one by Badin, Shadle, Ngoc, Carter, Chiu, Scully, & Stromberg (1994), in which spectral tilts were measured only up to 5 kHz.

Given these definitions, and what is known about parameter and articulatory behavior, we can make the following predictions. The parameter A_d should be maximized for a localized source, and for higher relative noise source strength, as in sibilants and unvoiced fricatives. A_d may not be very large for the fricatives /f, v/. The parameter S_p should be related to the source strength. Although the resonance peaks will affect the line fit, they should affect the fit in the same way for within-fricative comparisons. Thus, for a given fricative where the transfer function is assumed to vary only slightly from token to token, S_p should increase, i.e., become less negative, as flow velocity through the constriction increases.

Effort level and syllable stress should be correlated with increased flow velocity; the velocity should also be at a maximum mid-fricative, when constriction area is the smallest and pressure across the constriction the highest (Shadle & Scully, 1995; Scully, Georges & Castelli, 1992). The behavior of parameter S'_p should be similar to that of A_d , but unlike A_d , S'_p should interact with F . For a fricative with a localized source and posterior place, S'_p will be the largest. Within a fricative, increased S'_p should be correlated with either a more posterior place or greater source strength, because a more posterior place (or rounding) lowers F and a greater source strength increases the amplitude of the peak at F . Table I summarizes the predicted effects on parameters.

5. Results

5.1. Devoicing

Overall results from the analysis of devoicing in Corpora 2, 3 and 4 show that more than 50% of the voiced fricatives devoice (see Fig. 4) for all speakers, except ACC, who has a very low percentage of devoiced tokens in Corpus 2. There is more final devoicing in Corpus 3 than in Corpus 4 (see Fig. 5), but the real words of Corpus 4 do exhibit more than 50% total devoicing for all subjects. Devoicing rate differs

TABLE I. Predicted effects on parameters, considering the effects of one phonetic class at a time. A higher source strength is produced by higher volume velocity for the same constriction area A_c , or a constant volume velocity for a smaller A_c . The net result is a higher particle velocity in the constriction. For rounded vowel contexts the lips form a second constriction and so the first constriction (that intrinsic to the fricative) generates a noise source with lower strength (Shadle & Scully, 1995)

Phonetic class	Aeroacoustic effects	Predictions
Sibilants (relative to nonsibilants)	Localized source; Higher source strength	A_d and S'_p higher; S_p lower
Posterior place (relative to more anterior place)	Longer front cavity	F lower; S'_p higher; A_d and S_p same
Unvoiced	Higher source strength	A_d , S_p and S'_p higher
Voiced	Lower source strength	A_d , S_p and S'_p lower
Loud effort level	Higher source strength	A_d , S_p and S'_p higher
(relative to medium and soft)		
Middle of fricative	Higher source strength	A_d , S_p and S'_p higher
(relative to beginning and end)		
Stressed syllable	Higher source strength	A_d , S_p and S'_p higher
(relative to unstressed syllable)		
Medial word position	Higher source strength	A_d , S_p and S'_p higher
(relative to initial and final)		
Rounded vowel context	Longer front cavity; possibly lower source strength	F lower; A_d same; S_p lower; S'_p higher
(relative to unrounded)		
Male subjects	Longer front cavity	F lower; A_d same; S_p lower; S'_p higher
Female subjects	Shorter front cavity	F higher; A_d same; S_p higher; S'_p lower

among the three fricatives, and among Corpora 2, 3 and 4, but no consistent pattern could be detected.

Fig. 5 shows the percentage of devoicing as a function of position within a word, and relates it to syllable stress. There is a significant increase in devoiced examples from word-initial, through word-medial to word-final positions, for Corpus 3 fricatives. Although we might expect that voicing contrast weakens in unstressed syllables and, thus, unstressed fricatives would devoice disproportionately, the opposite appears to be the case. Our totals for all fricatives of Corpora 3 and 4, all subjects, are: /v/—48.4% (147 out of 304) devoiced and 12.2% (37 out of 304) partially devoiced; /z/—76.5% (166 out of 217) devoiced and 12% (26 out of 217) partially devoiced; /ʒ/—77.7% (181 out of 233) devoiced and 12.9% (30 out of 233) partially devoiced.

These results can be compared to those of Smith (1997) and Pirello *et al.* (1997) for American English. Smith (1997) studied only /z/ in a range of contexts and measured 47% devoiced and 36% partially devoiced. Pirello *et al.* (1997) studied /v, z/ in nonsense words (fricatives in initial and stressed position) and measured, respectively, 5%, 20% devoiced, and 35%, 40% partially devoiced. The comparable figures for initial stressed /v, z/ in Corpora 3 and 4 are 32%, 55% devoiced and 16%, 23% partially devoiced. It thus appears that there is more devoicing in Portuguese than in American English.

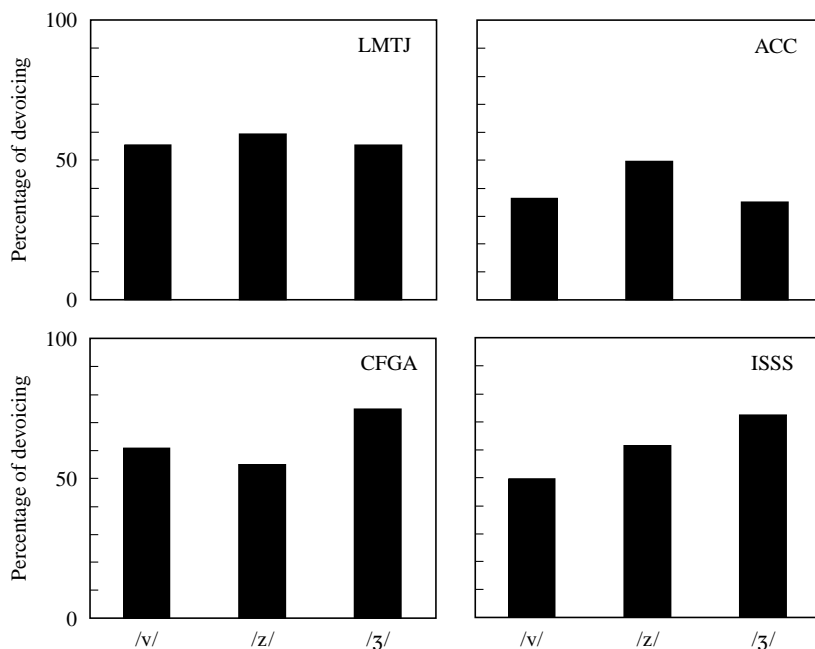


Figure 4. Percentage of total devoicing by fricative for each subject; Corpora 2–4 combined.

Other factors that might be correlated with devoicing were investigated using Corpora 3 and 4 for two of the subjects, LMTJ and ACC. First, there is no consistent pattern between duration of the fricatives /v, z, ʒ/ and percentage of devoicing. Second, there seems to be no particular vowel context that is primarily associated with devoicing. We note that the speakers all produced a large number of repeated tokens of nonsense words in one breath (more than 12 tokens). This high rate of speech (compared with previous recordings of similar corpora by French, American English and German speakers) could be one of the reasons why there are so many devoiced examples.

5.2. Parameterization of spectra

Fig. 6 shows average regression line fits (from \bar{F} to 20 kHz) to the spectra of the sustained fricatives in Corpus 1b for Subject ISSS. Each graph corresponds to a single place, and shows lines for three effort levels, voiced and unvoiced. Clearly, each place has a different “family” of nearly-parallel lines; higher effort level increases amplitude significantly and slope slightly, as predicted. For all subjects, the families of lines for the voiced and unvoiced fricatives always overlap, with the voiced cases mostly lower in amplitude and occupying a smaller range of amplitudes than the unvoiced cases. The results were similar for the other subjects.

In a plot of A_d by fricative, as shown in Fig. 7 for one subject, /s, z, ʃ, ʒ/ have higher A_d than /f, v/ as predicted. This holds for Corpora 1a and 1b for all subjects. A_d also tends to be lower for voiced fricatives than for their unvoiced counterparts, but this is less consistent across subject.

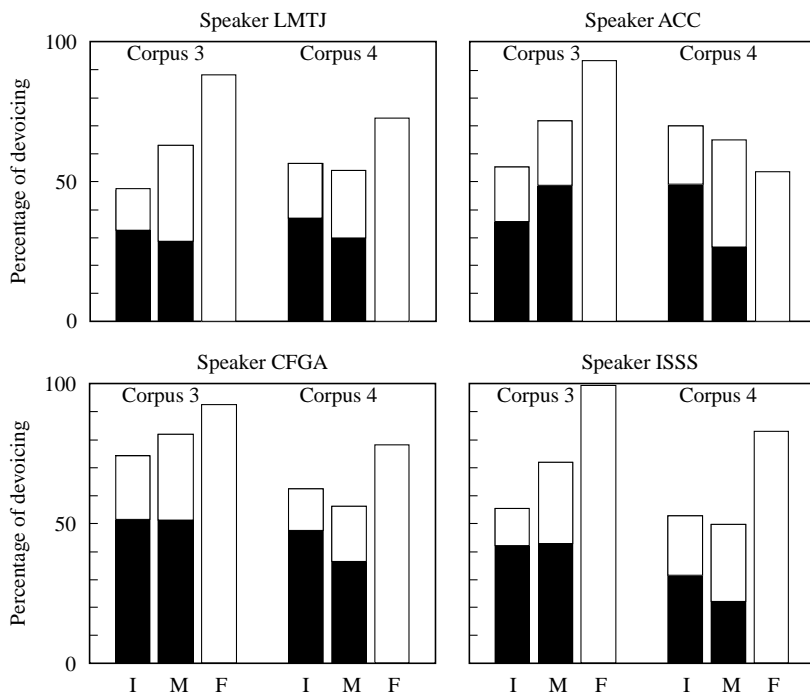


Figure 5. Percentage of total devoicing by position in word, word-initial (I), word-medial (M) and word-final (F), fricatives /v, z, ʒ/ combined. The black portion of each bar in the graph corresponds to fricatives in a stressed syllable and the white portion to fricatives in an unstressed syllable. There are no Portuguese fricatives in final stressed position.

Slope (S_p) generally increases with increased effort level, though this pattern is much more consistent for unvoiced fricatives. This is consistent with results in Shadle & Mair (1996) and the predictions in Table I.

The predictions made for A_d and S_p in Section 4 are shown in the schematic diagram of Fig. 8. This expected clustering of sibilants separately from /f, v/ is borne out by the results for all subjects, as shown in Fig. 9. Within the clusters, the expected relation of voiced and voiceless sibilants occurred, but was not as predicted for /f, v/.

As shown in Fig. 10, we predict that on an S'_p vs. S_p plot each place will cluster separately, with voiced tokens having lower S'_p but similar S_p relative to their unvoiced counterparts. Fig. 11 shows S'_p vs. S_p values plotted for Corpus 1a. For all speakers except ACC, both predictions were borne out. For ACC the voicing relationship was maintained, but /s, z/ tokens fell in between the /ʃ/ and /ʒ/ tokens.

In Fig. 12, A_d and S_p are plotted vs. location of the analysis window within the fricative (i.e., beginning, middle, or end) for Corpus 2. For /f, v/ there is no consistent pattern; results in Shadle, Tiede, Masaki, Shimada & Fujimoto (1996) indicate that for these the vowel context may play more of a role. As for the sustained fricatives, A_d separates sibilants from /f, v/. A_d is higher on average at the middle of the fricative than at the beginning and end for /s, z, ʃ, ʒ/, as predicted. S_p

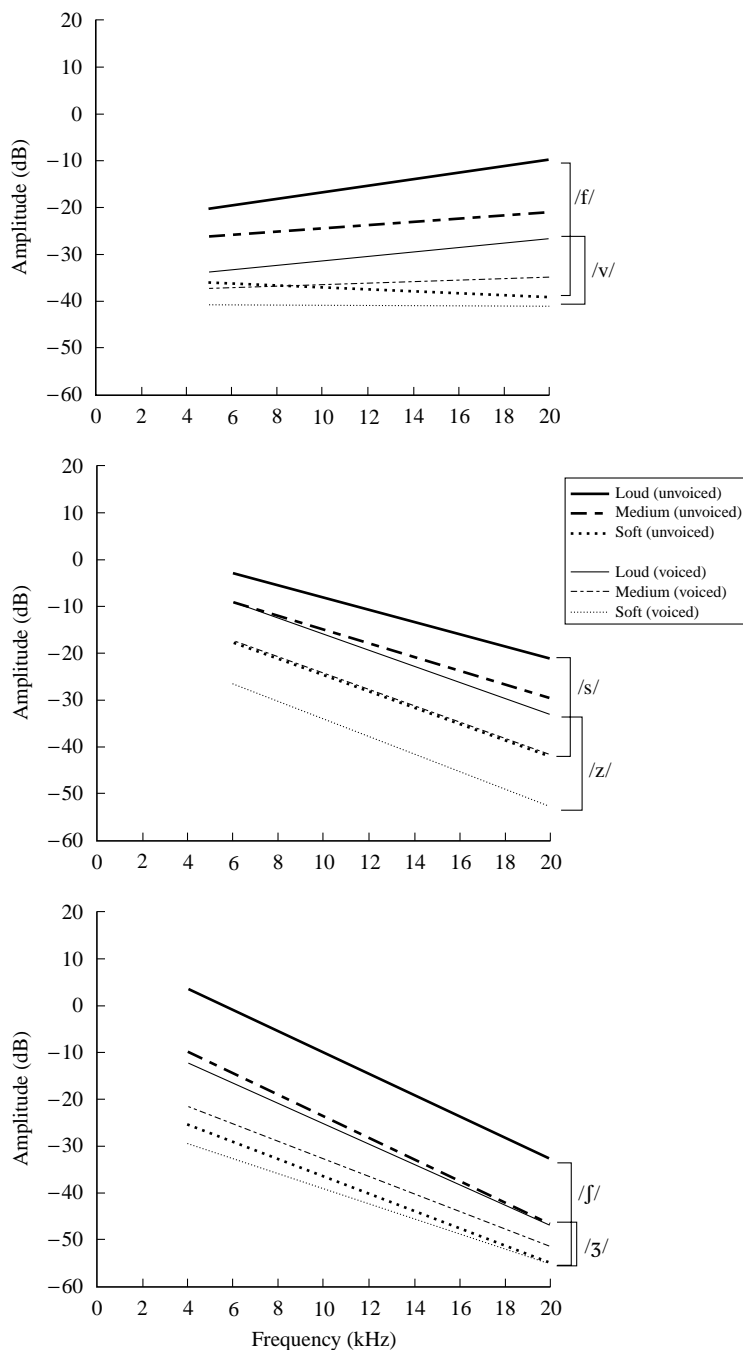


Figure 6. Average regression line fits (from \bar{F} to 20 kHz) of sustained labiodental (top), alveolar (middle) and postalveolar (bottom) fricatives from Corpus 1b at loud, medium and soft effort levels. Speaker ISSS.

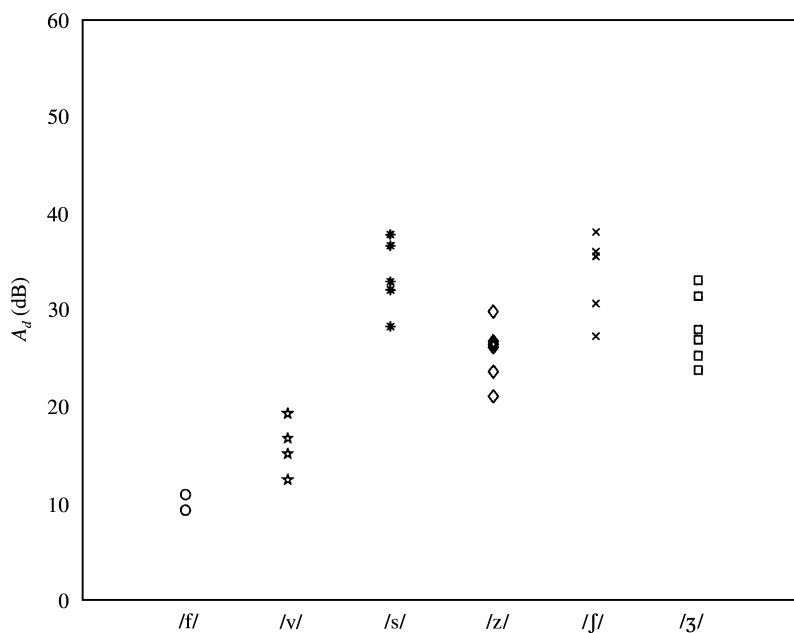


Figure 7. Dynamic amplitude of fricatives from Corpus 1a. Speaker LMTJ. Number of tokens per fricative: /f/ – 3; /v/ – 4; /s/ – 5; /z/ – 6; /ʃ/ – 5; /ʒ/ – 6.

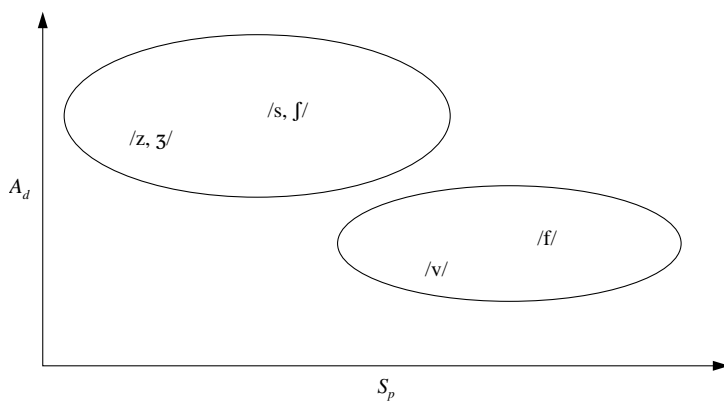


Figure 8. Predicted A_d vs. S_p relations for the fricatives.

tends to be lower for the sibilants than for /f, v/, but has no consistent trend with regard to location of the analysis window within the fricative.

Comparisons of stressed and unstressed fricatives indicate little or no change in A_d and S_p , as shown in Fig. 12 for Corpus 2 fricatives, which is not as predicted. We expected stress to act as increased effort level does, as in, for example, Fig. 6; that is, we expected both A_d and S_p to be higher in stressed than in unstressed syllables. However, it is clear that syllable stress does not affect fricatives in context the way effort level affects sustained fricatives.

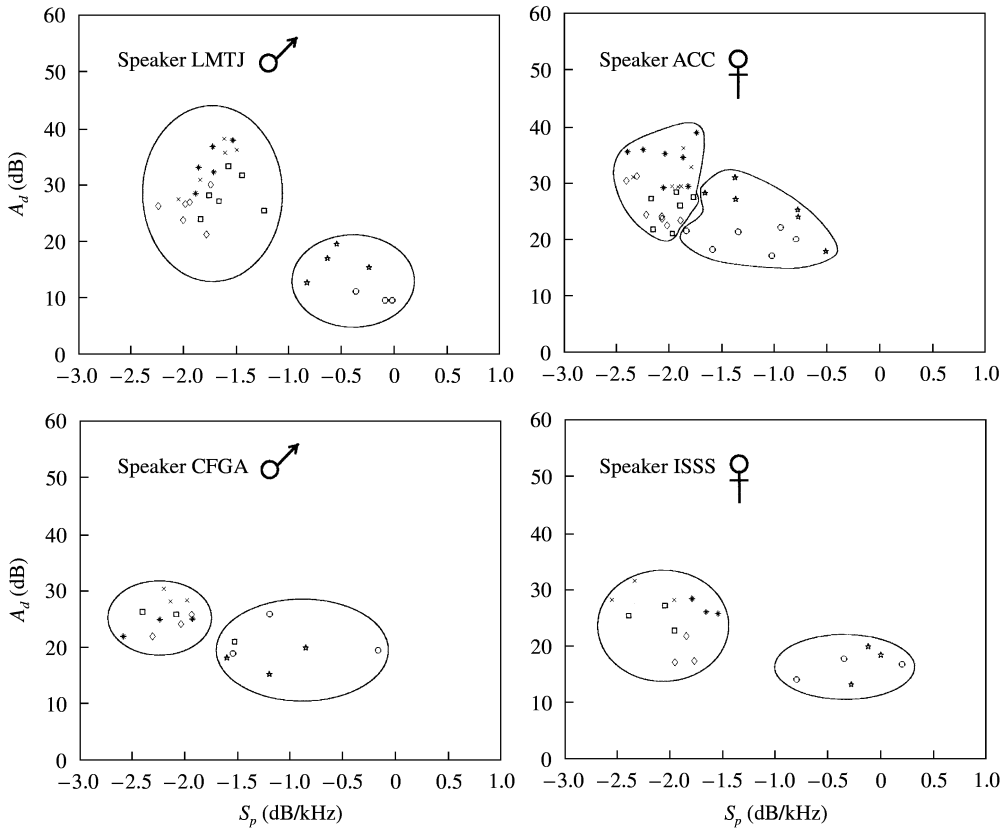


Figure 9. A_d vs. S_p for Corpus 1a (sustained fricatives), by subject. \circ – /f/, \star – /v/, \ast – /s/, \diamond – /z/, \times – /f/ and \square – /ʒ/. Three different vowel contexts were used with each fricative.

Analysis of Corpus 2 fricatives, which occur in rounded and unrounded vowel context, showed no consistent effect of rounding on the values of A_d and S_p . It is possible that F , if measured for this corpus, would show an effect, but the results from Corpus 1a indicate it unlikely to be a strong or significant one.

We were also interested in the following questions:

- Does stress affect A_d or S_p noticeably?
- Does the overall amplitude of the spectrum of a fricative vary noticeably with stress?

To answer these questions we superimposed the ensemble-averaged spectra computed at the middle of all examples of each fricative in Corpus 2 (Speaker LMTJ). The overall amplitude of stressed and unstressed fricatives is the same. The only significant difference seems to be the amplitude of the fundamental frequency component of voiced fricatives, which is 10–15 dB higher for stressed than for unstressed examples. It thus appears that stress results in a stronger voicing source, but does not affect the fricative noise source.

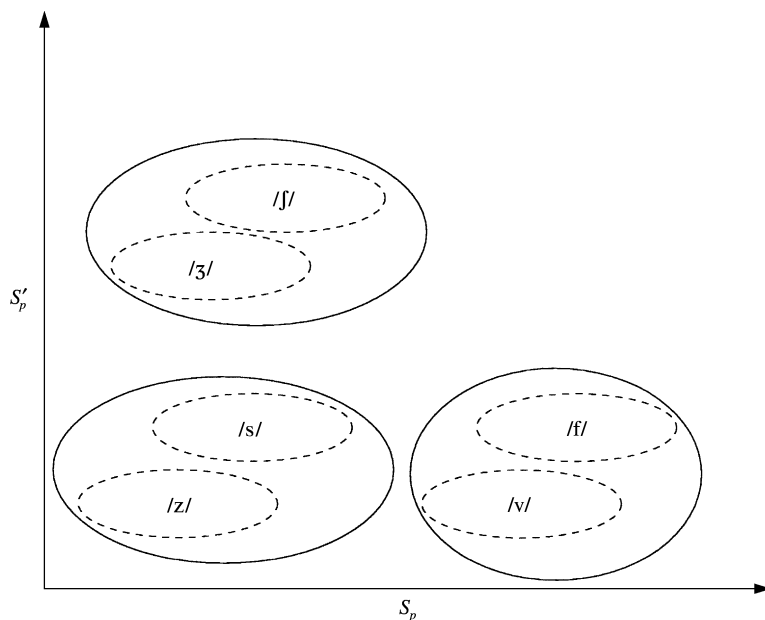


Figure 10. Predicted S'_p vs. S_p relations for the fricatives.

On \bar{F} vs. S'_p vs. S_p graphs the fricatives in real words (Corpus 4), produced by all four speakers, cluster by place (see Fig. 13). Similar results hold for Corpus 3, or when using A_d instead of S'_p . Of course, \bar{F} is not an independent parameter, but given the large number of tokens, using \bar{F} as in Fig. 13 allows us to check for other relationships more easily. Unlike in Corpus 1a, there is no clear relationship between voiced and unvoiced versions of the same place.

There does not seem to be any correlation between fricative duration, amount of devoicing, and the values of A_d and S_p . We also studied the relationships between the values of A_d , S_p , and duration, and various other contextual factors (stressed/unstressed syllables; word-initial/medial/final fricatives; voiced/partially-devoiced/devoiced fricatives), but without identifying any significant trends.

6. Conclusions and future work

In this study, we have designed fricative corpora with contents ranging from sustained fricatives to nonsense words following Portuguese phonological rules to real Portuguese words, and recorded and analyzed four native speakers of European Portuguese. Our principal findings are as follows.

Devoicing occurs more often in word-final than word-initial position, but is unrelated to a particular vowel context. Devoicing rate by fricative differs between Corpora 2 and 3, but it is generally very high, especially when compared with studies of other languages. It appears that this is an important characteristic of European Portuguese, which would have to be incorporated in any production model to obtain more natural-sounding synthetic speech.

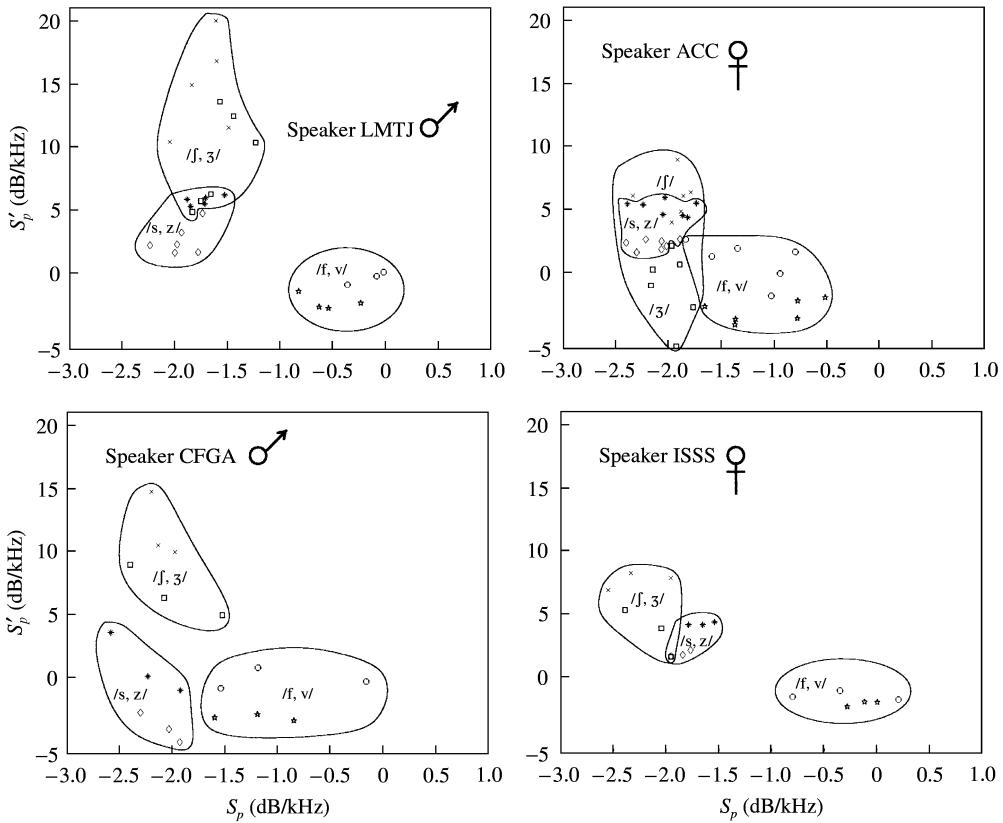


Figure 11. S'_p vs. S_p for Corpus 1a (sustained fricatives), by subject. \circ – /f/, \star – /v/, \ast – /s/, \diamond – /z/, \times – /ʃ/ and \square – /ʒ/.

The parameters spectral slope, frequency of maximum amplitude, and dynamic amplitude were developed to characterize fricative spectra, and applied to the recorded corpora. The parameters behaved as predicted for changes in effort level, voicing and location within the fricative. Some combinations were also useful for separating the fricatives by place or by sibilance.

The parameters capture source-related changes for the most part *as predicted*; for the sustained fricatives, they also separate fricatives by place. However, for the nonsense words of Corpus 2, comparisons of stressed and unstressed fricatives indicate little or no change in A_d and S_p . Since this pattern occurs also in real words of Corpus 3, this may be a characteristic of Portuguese; the effect on a fricative of syllable stress is not the same as that of increased effort level.

A combination of parameters A_d and S_p was also useful for separating the fricatives by sibilance, and a combination of parameters S_p and S'_p separated the fricatives both by place and sibilance.

Results from the four subjects seem, for the most part, to be consistent. Therefore, it is possible that these parameters are capturing aspects of Portuguese that differ from English or French fricatives. The quantified spectral characteristics

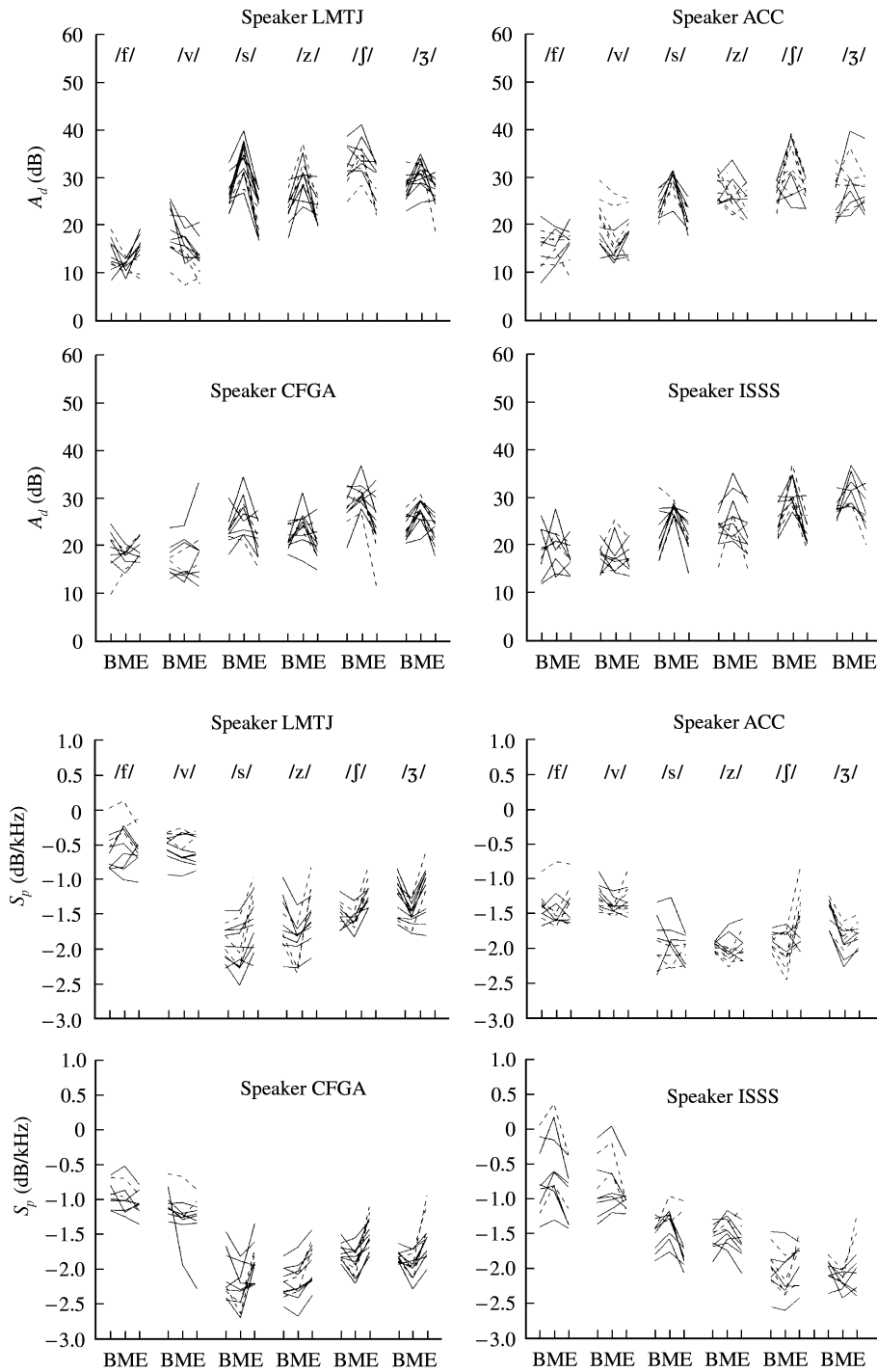


Figure 12. Dynamic amplitude and spectral slope of fricatives from Corpus 2, at the Beginning (B), Middle (M) and End (E) of the fricative. Fricatives in stressed syllables (solid line) and unstressed syllables (dashed line).

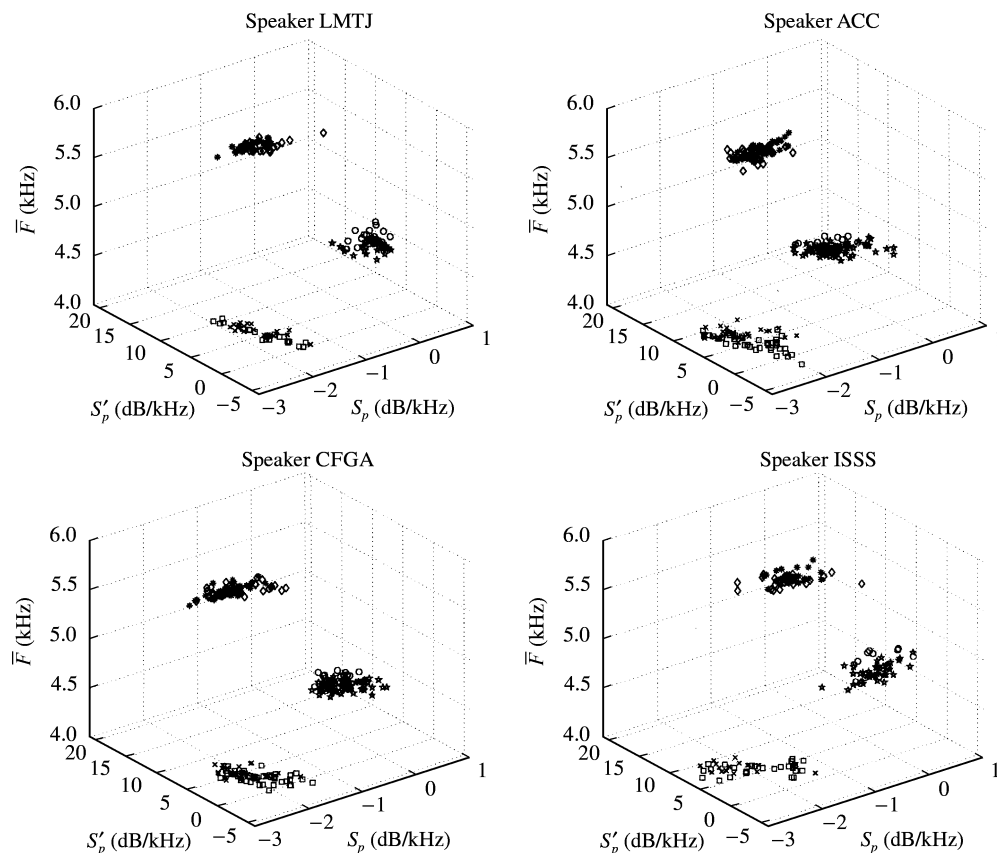


Figure 13. \bar{F} vs. S_p' vs. S_p , Corpus 4, by subject. \circ – /f/, \star – /v/, \ast – /s/, \diamond – /z/, \times – /ʒ/ and \square – /ʒ/.

of Portuguese fricatives can be related to specific properties of the transfer function and source spectrum during the production of these sounds. Implications for modelling Portuguese fricatives, and fricative production mechanisms, will be considered in the future.

The authors would like to thank Dr Lurdes de Castro Moutinho, at the Universidade de Aveiro in Portugal, for help in designing the corpus. The authors would also like to thank Dr Richard McGowan, Dr Michael Krane and Dr Christine Mooshammer, for comments on an earlier draft of this paper. This work was partially supported by Fundação para a Ciência e a Tecnologia, Portugal.

References

- Andrade, A. (1994) Reflexões sobre o 'e mudo' em português europeu. In *Actas do Congresso Internacional Sobre o Português*, Vol. 2, Lisboa, Portugal, pp. 303–344.
- Andrade, A. (1995) Percepção de C ou CC oclusivas por ouvintes nativos de português europeu. In *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística*, Vol. 3, Lisboa, Portugal, pp. 153–186.

- Badin, P., Shadle, C. H., Ngoc, Y. P. T., Carter, J. N., Chiu, W. S. C., Scully, C. & Stromberg, K. (1994) Frication and aspiration noise sources: contribution of experimental data to articulatory synthesis. In *Proceedings of the International Conference on Spoken Language Processing (ICSLP 94)*, Vol. 1, Yokohama, Japan, pp. 163–166.
- Bendat, J. S. & Piersol, A. G. (2000) *Random data: analysis and measurement procedures*, 3rd edition. New York: John Wiley.
- Choo, W. (1999) The relationship between perceptual and physical space of fricatives. In *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS 99)*, San Francisco, U.S.A., pp. 163–166.
- Choo, W. & Huckvale, M. A. (1997) Spatial relationships in fricative perception. *Speech Hearing and Language: Work in Progress* 10, University College London, Department of Phonetics and Linguistics, London, U.K. <http://www.phon.ucl.ac.uk/home/shl10/won/space.htm>.
- Crystal, D. (1997) *A dictionary of linguistics and phonetics*, 4th edition. Oxford: Blackwell.
- Docherty, G. J. (1992) *The timing of voicing in British English obstruents*. Berlin: Foris Publications.
- Evers, V., Reetz, H. & Lahiri, A. (1998) Crosslinguistic acoustic categorization of sibilants independent of phonological status, *Journal of Phonetics*, **26**, 345–370.
- Forrest, K., Weismer, G., Milenkovic, P. & Dougall, R. N. (1988) Statistical analysis of word-initial voiceless obstruents: preliminary data, *Journal of the Acoustical Society of America*, **84**(1), 115–123.
- Funatsu, S. (1995) Cross language study of perception of dental fricatives in Japanese and Russian. In *Proceedings of the 13th International Congress of Phonetic Sciences (ICPhS' 95)*, Vol. 4, Stockholm, Sweden, pp.124–127.
- Goldstein, M. E. (1976) *Aeroacoustics*. New York: McGraw-Hill.
- Jongman, A., Wayland, R. & Wong, S. (2000) Acoustic characteristics of English fricatives, *Journal of the Acoustical Society of America*, **108**(3), 1252–1263.
- Krane, M. H. (1999) Fluid dynamic effects in speech, *Journal of the Acoustical Society of America*, **105**(2, Part 2), 1159.
- Lacerda, F. P. (1982) Acoustic perceptual study of the Portuguese voiceless fricatives, *Journal of Phonetics*, **10**, 11–22.
- Martins, M. R. D., Harmegnies, B. & Poch, D. (1995) Changement phonétique en cours du portugais européen. In *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística*, Vol. 3, Lisboa, Portugal, pp. 249–259.
- Mateus, M. H. M. & Andrade, E. (2000) *The Phonology of Portuguese*. Oxford: Oxford University Press.
- Nelson, P. A. & Morfey, C. L. (1981) Aerodynamic sound production in low speed flow ducts, *Journal of Sound and Vibration*, **79**(2), 263–289.
- Pirello, K., Blumstein, S. E. & Kurowski, K. (1997) The characteristics of voicing in syllable-initial fricatives in American English, *Journal of the Acoustical Society of America*, **101**(6), 3754–3765.
- Scully, C., Georges, E. G. & Castelli, E. (1992) Articulatory paths for some fricatives in connected speech, *Speech Communication*, **11**(4–5), 411–416.
- Shadle, C. H. (1985) *The acoustics of fricative consonants*. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, U.S.A. Released as Research Laboratory of Electronics Technical Report 506.
- Shadle, C. H. (1990) Articulatory-acoustic relationships in fricative consonants. In *Speech production and speech modelling* (W. J. Hardcastle & A. Marchal, editor), pp. 187–209. Dordrecht: Kluwer Academic.
- Shadle, C. H. (1991) The effect of geometry on source mechanisms of fricative consonants, *Journal of Phonetics* **19**(3–4), 409–424.
- Shadle, C. H. (1992) Progress reports 1990–92. In *Mesure, caractérisation et modélisation des Sons Fricatifs* (B. Guérin, editor). EC SCIENCE Project SCI*0147-C(EDB).
- Shadle, C. H. & Carter, J. N. (1993) WPI: from speech signal to acoustic sources. In *Speech MAPS Year 1 Report* (P. Badin, C. Abry & C. Scully, editors). ESPRIT project 6975, v.2.
- Shadle, C. H. & Mair, S. J. (1996) Quantifying spectral characteristics of fricatives. In *Proceedings of the International Conference on Spoken Language Processing (ICSLP 96)*, Philadelphia, U.S.A., pp. 1517–1520.
- Shadle, C. H., Mair, S. J. & Carter, J. N. (1996) Acoustic characteristics of the front fricatives [f, v, θ, ð], In *Proceedings of the 1st ESCA Tutorial and Research Workshop (ETRW) on Speech Production Modeling—4th Speech Production Seminar*, Autrans, France, pp. 193–196.
- Shadle, C. H. & Scully, C. (1995) An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences, *Journal of Phonetics*, **23**(1,2), pp. 53–66.
- Shadle, C. H., Tiede, M., Masaki, S., Shimada, Y. & Fujimoto, I. (1996) An MRI study of the effects of vowel context on fricatives. In *Proceedings of the Institute of Acoustics Autumn Conference (Speech and Hearing 96)*, book 1, Vol. 18, Part 9, Windermere, U.K., pp. 187–194.
- Smith, C. L. (1997) The devoicing of /z/ in American English: Effects of local and prosodic context, *Journal of Phonetics*, **25**(4), 471–500.

- Stevens, K. N. (1971) Airflow and turbulence noise for fricative and stop consonants: static considerations, *Journal of the Acoustical Society of America*, **50**(4), 1180–1192.
- Stevens, K. N. (1998) *Acoustic phonetics*. Cambridge: MIT Press.
- Sussman, H. M. (1994) The phonological reality of locus equations across manner class distinctions: preliminary observations, *Phonetica*, **51**, 119–131.
- Viana, M. C. (1984) *Etude de deux aspects du consonantisme du portugais: fricatisation et dévoisement*, PhD. thesis (Doct. 3ème Cycle), Université des Sciences Humaines de Strasbourg, Strasbourg, France.
- Wilde, L. F. (1995) *Analysis and synthesis of fricative consonants*. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, U.S.A.

Appendix A

Corpus 3 uses the real words listed here. They are presented here according to position of the fricative within the word.

Portuguese words with fricatives /f, v, s, z, ʃ, ʒ/ in initial and medial position (nearly minimal pairs):

fofa—/'fofə/—English definition: soft.
viver—/vi'ver/—to live.
viva—/'vivə/—hurrah!, live (imperative), bless you.
cessa—/'sɛsə/—ceases (verb form).
Zézé—/'zɛzɛ/—(diminutive of Joseph = Joe).
chocha—/'ʃofə/—spineless, insipid, empty.
bochecha—/bu'ʃɛʃə/—cheek.
Gigi—/ʒi'ʒi/—(diminutive of a woman's first name).

Portuguese words with fricatives in initial position:

figo—/'figu/—fig.
ferir—/fɛ'rir/—to hurt, to injure.
febra—/'febrə/—a joint of pork.
ferro—/'fɛRu/—iron.
falir—/fə'lir/—to go bankrupt.
fala—/'falə/—speech.
foco—/'fɔku/—torch.
fogo—/'fogu/—fire.
furar—/fu'rar/—to drill.
vila—/'vilə/—small town.
vermelho—/vɛr'mɛʎu/—red.
ver—/vɛr/—to see.
vêu—/'vɛw/—veil.
veia—/'vɛjə/—vein.
vaca—/'vakə/—cow.
volta—/'vɔltə/—go/come back! (imperative), turn.
voo—/vow/—flight.
vogar—/'vugar/—to row, to float.
sítio—/'sitju/—place.
secar—/'i'kar/—to dry.
sede—/'sɛdi/—thirst.
seta—/'sɛtə/—arrow.
saber—/sə'ber/—knowledge, to know.
sala—/'salə/—room, suite.
só—/sɔ/—alone, lonely, only.
sopa—/'sopə/—soup.
subir—/su'bir/—to climb, to mount, to rise.

Zita—/'zite/—(woman's name, diminutive of a woman's name).
zelar—/'zilar/—to watch over, to pay great attention to.
zelo—/'zelu/—zeal.
Zé—/ze/—(diminutive of José = Joseph).
zarpar—/zər'par/—to escape, to run away, to lift anchor.
Zaire—/'zajri/—(proper noun).
Zópiro—/'zɔpiru/—(proper noun).
zona—/'zone/—zone.
zurrar—/zu'Rar/—to bray.
chicote—/ʃi'koti/—whip.
chegar—/ʃi'gar/—to arrive.
cheta—/'fete/—“não ter cheta” → “to be penniless”.
cheque—/'ʃeki/—cheque.
chamar—/ʃə'mar/—to call.
chá—/ʃa/—tea.
choca—/'ʃokə/—brooding (feminine), shocks (verb form).
choco—/'ʃoku/—brooding (masculine), cuttle fish.
chorar—/ʃu'rar/—to cry.
girar—/ʒi'rar/—to spin.
gelado—/ʒi'ladu/—ice cream.
gelo—/'ʒelu/—ice.
germe—/'ʒɛrmi/—germ.
jaqueta—/ʒə'ketə/—short jacket.
jacto—/'ʒatu/—jet.
jóia—/'ʒɔje/—jewel.
jogo—/'ʒogu/—game.
judeu—/ʒu'dew/—Jew.

Portuguese words with fricatives in medial position:

efetuar—/i'fetwar/—to accomplish.
benefício—/bini'fɨʃju/—benefit.
trefo—/'trefu/—cunning, astute.
benéfico—/bi'nɛfiku/—beneficial.
afiar—/ɐfi'ar/—to sharpen.
café—/kə'fɛ/—coffee.
garrafa—/gə'Rafɐ/—bottle.
bafo—/'bafu/—breath.
galhofa—/gə'ʎɔfɐ/—amusement, frolic.
mofo—/'mofu/—mould.
bufa—/'bufɐ/—to blow (verb form).
altivo—/al'tivu/—haughty, arrogant.
dever—/di'ver/—duty, to owe.
levar—/li'var/—to take.
relevo—/Ri'levu/—relief.
leva—/'lɛvɐ/—takes (verb form).
avó—/ɐ'vɔ/—grandmother.
cava—/'kavɐ/—digs (verb form).
bravo—/'bravu/—brave, wild.
nova—/'nɔvɐ/—new.
ovelha—/o'veʎɐ/—sheep.
mover—/mu'ver/—to move.
uva—/'uvɐ/—grape.
iça—/'isɐ/—lifts (verb form).
ressaca—/Rɨ'sakɐ/—hangover.
condessa—/kõ'desɐ/—countess.
pêssego—/'pesigu/—peach.
aquecer—/ɐkɛ'ser/—to heat.

passar—/pə'sjar/—to walk, to go for a walk.
assar—/ə'sar/—to roast.
caça—/'kasə/—hunting (verb form), game.
possa—/'pɔsə/—can (verb form).
moça—/'mosə/—girl.
possível—/pu'ʃivəl/—possible.
exato—/i'zatu/—exact.
mesinha—/mɛ'ziɲə/—small table.
beleza—/bi'lezə/—beauty.
peso—/'pezu/—weight.
mezinha—/mɛ'ziɲə/—traditional medicine.
Brasil—/brɛ'zil/—(proper noun).
azar—/turna'zar/—bad luck.
azul—/ə'zul/—blue.
mazinha—/ma'ziɲə/—pest, bad girl/woman.
asa—/'azə/—wing.
rosa—/'rɔzə/—rose.
amoroso—/əmu'rozu/—amorous, sweet.
acusar—/əku'zar/—to accuse.
bicha—/'bifə/—queue.
bexiga—/bi'figə/—bladder.
este—/'ɛfti/—this one.
meche—/'mɛʃi/—touches (verb form).
achar—/ə'far/—to find, to think.
bolacha—/bu'lafə/—biscuit.
tacho—/'tafu/—pot, pan.
tocha—/'tɔfə/—torch.
mocho—/'moʃu/—owl.
capucho—/kə'puʃu/—hood.
originar—/oriʒi'nar/—to originate, to generate.
tijolo—/ti'ʒolu/—brick.
arejar—/əri'ʒar/—to ventilate.
pejo—/'peʒu/—modesty.
Beja—/'bɛʒə/—(proper noun).
agir—/ə'ʒir/—to act.
cajado—/kə'ʒadu/—crook.
ajudar—/ə'ʒudar/—to help.
haja—/'aʒə/—there is (verb form).
aloja—/ə'lɔʒə/—lodges (verb form).
tojo—/'toʒu/—gorse.
tugir—/tu'ʒir/—to speak low.

Portuguese words with fricative /ʃ/ in final position:

diz—/diʃ/—says (verb form), tell me (imperative).
mares—/'marɪʃ/—seas.
mês—/'mɛʃ/—month.
pés—/pɛʃ/—feet.
perdas—/'pɛrdɛʃ/—losses.
capaz—/kə'paʃ/—capable.
pós—/pɔʃ/—powders.
pôs—/poʃ/—put (verb form).
dos—/'duʃ/—of (the).

Portuguese words with fricatives /f, v, s, z, ʒ/ in “simulated” final position:

chefe—/'ʃɛfi/—chief.
Fafe—/'fafɪ/—(proper noun).
teve—/'tevi/—had (verb form).

leve—/'levi/—light.
ave—/'avi/—bird.
move—/'mɔvi/—moves (verb form).
partisse—/par'tisi/—left (verb form).
batesse—/bɐ'tesi/—hit (verb form).
asse—/'asi/—roast (verb form).
posse—/'posi/—possession.
doce—/'dosi/—sweet.
doze—/'dozi/—twelve.
age—/'aʒi/—acts (verb form).
hoje—/'oʒi/—today.

Appendix B

Corpus 4 sentences, listed in this appendix, were constructed using about half of the words from Corpus 3 (see Appendix A), with sentences 1–10 making some sense in Portuguese. Sentences 11 and 12 were devised to reproduce some of the vocalic contexts used for Corpus 3 across word boundaries (this is signaled in the phonetic transcription by boxes), but they make no sense.

1. “A Gigi é uma chocha e age em benefício da avó doce.”
 /ɐ ʒi'ʒi ɛ 'umɐ 'ʃoʃɐ i 'aʒi ẽʒ bi'ni'fisju dɐ ɐ'vɔ 'dosi/
 (Gigi is spineless and acts for the benefit of the sweet grandmother.)
2. “A vaca fuge do gelo na zona.”
 /ɐ 'vakɐ 'fɔʒi du 'ʒɛlu nɐ 'zonɐ/
 (The cow runs from the ice in the zone.)
3. “A ave, no voo a subir, move a asa para zarpar da seta.”
 /ɐ 'avi nu vɔw ɐ su'bir 'mɔvi ɐ 'azɐ 'pɐrɐ zɐr'par dɐ 'sɛtɐ/
 (The bird, in a rising flight, moves his wing to escape from the arrow.)
4. “O chefe altivo fala à rosa de sede de beleza.”
 /u 'ʃɛʃi al'tivu 'falɐ a 'rɔzɐ di 'sɛdi di bi'lezɐ/
 (The haughty chief speaks to the rose about thirst and beauty.)
5. “Quero chorar hoje, arejar o sítio e vogar.”
 /'keru ʃu'rar 'oʒi ɐri'ʒar u 'sitju i 'vugar/
 (I want to cry today, air the place out and float.)
6. “O bravo do Zé quer ajudar o judeu só.”
 /u 'bravu du zɛ kɐr ɐ'ʒudar u ʒu'dɛw sɔ/
 (Brave Zé wants to help the lonely Jew.)
7. “O café cura a ressaca ao chegar da caça.”
 /u kɐ'fɛ kurɐ ɐ Ri'sakɐ aw ʃi'gar dɐ 'kasɐ/
 (Coffee cures a hangover when you arrive from hunting.)
8. “Ver a mesinha nova de volta à sala, é benéfico para o modo de viver dos doze.”
 /vɐr ɐ mi'ziɲɐ 'nɔvɐ di 'vɔltɐ a 'salɐ ɛ bi'nɛfikɐ 'pɐrɐ u 'mɔdu di vi'vɐr duf
 'dozi/
 (It is beneficial for the way of living of the twelve, to see the new table back in the room.)
9. “Furar uma jóia choca a condessa Zita.”
 /fu'rar umɐ 'ʒɔʲɐ 'ʃɔkɐ ɐ kɔ'dɛsɐ 'zitɐ/
 (To drill a jewel shocks Countess Zita.)

10. “O vermelho do fogo e o azul dos mares do Brasil.”
 /u vɪr'mɐxu du 'fogu i u ɐ'zul duʃ 'marɨʃ du brɐ'zil/
 (The red of fire and the blue of the sea of Brazil.)
11. “A chá no sítio é possível achar.”
 /ɐ ʃa n u'si tʃu ɛ p u'si vɐl ɐ'ʃa r/
 (Tea in the place is possible to find.)
12. “Cava sala meche ver o dever de assar.”
 /'kav ɐ 'sa lɐ 'mɛʃ i ver u d i'ver dɨ ɐ'sa r/
 (Dig room touches see the duty to roast.)