

Temporal and Devoicing Analysis of European Portuguese Fricatives

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

* Escola Superior de Saúde da Universidade de Aveiro, and
Instituto de Engenharia Electrónica e Telemática de Aveiro
Universidade de Aveiro, 3810-193 Aveiro, Portugal
e-mail: lmtj@essua.ua.pt

† Department of Electronics and Computer Science
University of Southampton, Southampton, SO17 1BJ, UK
e-mail: chs@ecs.soton.ac.uk

ABSTRACT

Duration and devoicing of Portuguese fricatives have been studied using a set of corpora that include nonsense words following Portuguese phonological rules, and real words; these were recorded by four subjects (2 male, 2 female). Results show that fricative duration varies most with voicing (voiceless are longer), and also significantly by speaker, place, and position within word. Devoicing occurs most often word-finally, and varies significantly by place; devoicing occurs more often than in English.

1 Introduction

There have been many studies in the past of duration and devoicing properties of fricatives in English, but none to date on Portuguese fricatives. Pirello et al. [1] showed that the American English alveolar fricatives /s, z/ tend to have longer durations than the labiodental fricatives /f, v/. Although vowel length played a role in the perception of voicing in word-final fricatives, it did not serve such a role in word-initial fricatives. Jongman et al. [2] showed that the mean duration of fricatives, produced by 20 American English speakers in nonsense words, differed significantly for sibilants versus nonsibilants.

The difference in duration between unvoiced and voiced fricatives has been reported in many studies. Baum and Blumstein's [3] acoustic and perceptual analysis of durational characteristics distinguishing American English fricatives in nonsense words showed that while the overall mean value of the voiceless fricatives was longer than the voiced fricatives, there was considerable overlap in the two distributions. This was later confirmed by Crystal and House [4]. The results of Docherty [5] showed that unvoiced British English

fricatives were longer than their voiced counterparts. Smith [6] reported that in American English /z/ was shorter than /s/, and that vowels were significantly longer before /z/ than before /s/. Mair and Shadle [7] studied voiced-voiceless distinction for a male French speaker by comparing the pairs /s, z/ and /ʃ, ʒ/ in nonsense words. The voiceless fricatives were longer than voiced, but for voiced fricatives, durations of the preceding and following vowels were longer. The total VCV duration was longer for voiced fricatives; voicing amplitude diminished earlier and resumed later in the voiceless compared with the voiced fricatives.

Many studies have also investigated the influence of stress on fricative durations, but results are inconclusive. Crystal and House [4] showed that the probability densities of the durations of American English fricatives occurring in stressed and unstressed syllables overlapped significantly.

The influence of vowel context on the duration of fricatives has also been studied. Behrens and Blumstein [8] showed minimal effects of vowel context on duration of American English fricatives. Mair and Shadle [7] reported that, for a male French speaker, the total VCV duration was affected by vowel context, and that duration was longer for /a-a/ than /i-i/ and longer for /i-i/ than /u-u/ in both the voiced and voiceless fricatives, but durations were inconsistent with regard to fricative place of articulation.

Devoicing of voiced fricatives has also been studied. Stevens et al. [9] reported a significant number of devoiced examples of American English /v/ and /z/, though /v/ remained fully voiced more often than /z/; fricatives in word-final position were virtually always devoiced. Smith [6] observed high percentages of devoicing for American English /z/, but also noted a large variability amongst speakers. /z/ devoiced more often word-finally than word-medially. There was no consistent effect of stress or context on devoicing across

speakers or tokens.

Pirello et al. [1] investigated whether or not systematic patterns of voicing could be identified as a function of phonetic context for American English fricatives. There was a greater preponderance of voicing throughout for the labiodental fricatives. Although contextual influences emerged, they did not necessarily occur in the majority of the utterances, and there was variation amongst different speakers.

Though Portuguese fricatives have been studied in the past, there have been no studies of the duration and devoicing, so we will be focusing on such analysis here.

2 Method

A speech corpus has been designed for European Portuguese, with the fricatives /f, v, s, z, ʃ, ʒ/ in the following contexts: sustained (Corpus 1), repeated nonsense words (Corpus 2), words containing fricatives in frame sentences (Corpus 3), and the same set of words in sentences (Corpus 4). Four subjects, two male (LMTJ and CFGA) and two female (ACC and ISSS), were recorded reading the corpora.

Recordings were made in a sound treated room using a Bruel & Kjaer 4165 1/2 inch microphone located 1 m in front of the subject's mouth, connected to a B & K 2690 preamplifier. The signal was amplified and filtered by a B & K 2636 measurement amplifier, with high-pass cut-on frequency of 22 Hz and low-pass cut-off frequency of 22 kHz. A laryngograph signal (Lx) was also collected using a laryngograph processor¹. The acoustic speech signal and Lx were recorded with a Sony TCD-D7 DAT system at 16 bits, with a sampling frequency of 48 kHz, and digitally transferred to a computer for post-processing.

The time waveforms of all the corpus words were manually analysed to detect the start of the vowel-fricative transition, the start and end of the fricative, and the start of the fricative-vowel transition. Both the acoustic and laryngograph signals were used to determine the VF and FV boundaries.

Voicing is often maintained only over part of the fricatives. A fricative was called *devoiced* when there was no periodic structure in the acoustic or laryngograph signals during the frication interval; *partially devoiced* when more than one-third but less than half of the frication interval contained steady acoustic and laryngograph signal cycles; *voiced* when there were steady acoustic and laryngograph signal cycles throughout the whole fricative [10].

¹Model LxProc, type PCLX produced by Laryngograph Ltd (UK).

3 Results

3.1 Temporal Analysis

The mean durations of the voiceless fricatives for all subjects in Corpus 3 are greater than all mean durations of the Corpus 3 voiced fricatives. There is no significant difference by place of articulation. For 3 of the 4 subjects (not CFGA) the mean duration of the fricative is greater than that of the VF and FV transitions, and the mean duration of the VF transition is greater than that of the FV transition.

For each voiceless fricative and subject, the mean durations in Corpus 4 are greater than the durations of their voiced counterparts, as shown in Figure 1, which agrees with results for English [11, 4, 9, 1]. The mean duration of the fricatives is greater than the duration of the VF and FV transitions; the VF and FV transitions do not differ significantly in duration for any of the speakers. Corpus 4 alveolar fricatives /s, z/ were on average longer than labiodentals /f, v/, as has been previously reported for English [1], but which was not observed in Corpus 3. The mean duration (within subject) of the fricatives in Corpus 3 was quite similar to that in Corpus 4.

3.2 Analysis of Variance of Duration

One-way analysis of variance (ANOVA) was used to study the effects of the independent variables (factors) speaker (LMTJ, CFGA, ACC and ISSS), place of articulation (labiodental, alveolar and postalveolar) and position in word (word-initial, -medial and -final) on the dependent variable duration of fricatives in Corpus 3 and 4. Two separate ANOVAs were run to study the correlation between place of articulation and duration, because from the analysis presented in previous section, it was clear that the duration of unvoiced fricatives was always significantly greater than their voiced counterparts. Since place was the independent variable and not voicing, the two subgroups /f, s, ʃ/ and /v, z, ʒ/ were analysed.

There was a significant effect of the factor **speaker** on the duration of fricatives /f, s, ʃ/ in both Corpus 3 and 4, as shown in Table 1. There was no significant effect for /v, z, ʒ/ in Corpus 3, but in Corpus 4 there was a significant effect.

There was a significant effect of the factor **place** on the duration of fricatives /f, s, ʃ/ and /v, z, ʒ/ in both Corpus 3 and 4, as shown in Table 1. There was a significant linear trend (Corpus 3 – $F(1, 338) = 21.787$, $p < 0.001$; Corpus 4 – $F(1, 410) = 26.091$, $p < 0.001$). In Corpus 3, as the place of articulation moved posteriorly, duration increased proportionately.

There was a significant effect of the factor **position in word** on the duration of fricatives /f, s, ʃ/ and /v, z, ʒ/ in both Corpus 3 and 4, as shown in Table 1. There was a significant linear trend (Cor-

pus 3 – $F(1, 338) = 91.513$, $p < 0.001$; Corpus 4 – $F(1, 410) = 10.415$, $p = 0.001$), indicating that as the position of the fricative moves from initial, through medial, to final word position, duration increased proportionately.

3.3 Devoicing Analysis

Figure 2 shows the percentage of total devoicing for each subject by fricative and corpus. We also investigated devoicing by position in word; ANOVA results are summarized in Table 1.

The Corpus 2 and 3 fricatives of speakers LMTJ and ACC were analyzed to see if vowel context interacted with devoicing. No consistent pattern was observed, so this comparison was not made for speakers CFGA and ISSS. This result, that there is no effect of vowel context on devoicing of Portuguese fricatives, agrees with two studies of English [1, 6].

Corpus 3 results for all four subjects showed that 55% (70 out of 127) of /v/ tokens, 74% (79 out of 107) of /z/ tokens, and 86% (92 out of 107) of /ʒ/ tokens were totally devoiced. Most word-final fricatives (93% – 55 out of 59) were totally devoiced, and the percentage of devoicing increased as the place of articulation moved posteriorly. Veatch’s [12] study of American English fricatives showed that word-final fricatives devoiced 25–100% of the time, depending on context. Although devoicing in our Corpus 3 ranges from 9 to 100%, word-final fricatives devoice 93% of the time.

In Corpus 4, 44% (77 out of 177) of /v/ tokens, 78% (86 out of 110) of /z/ tokens, and 71% (89 out of 126) of /ʒ/ tokens were totally devoiced. The Corpus 4 fricatives devoiced mostly word-finally, but less often than in Corpus 3. In Corpus 4, the words coming just after the word-final fricatives were not phonetically controlled, which might account for the lower word-final average percentage of devoicing in Corpus 4 when compared with Corpus 3. Indeed, some voiceless fricatives become voiced in Corpus 4, likely as a result of cross-word coarticulation.

In Corpus 4, /ʒ/ devoiced slightly less often than /z/, which contradicts the very clear results of Corpus 3 fricatives (in which the percentage of devoiced examples decreases as the place of articulation moves anteriorly). One possible explanation could be that /ʒ/ is produced in a more anterior place in continuous speech than in isolated word production. This hypothesis can only be confirmed with additional articulatory data, which is planned as future work.

3.4 Analysis of Variance of Devoicing

As in the analysis of duration (Section 3.2), one-way analysis of variance was used to study the effects of the independent variables (factors) speaker, place of articulation and position in word on the dependent variable amount of devoicing of fricatives in Corpus 3 and 4.

There was a significant effect of the factor **speaker** on the voicing of fricatives /v, z, ʒ/ in Corpus 3, but no significant effect in Corpus 4, as shown in Table 1.

There was a significant effect of the factor **place** on the voicing of fricatives in both Corpus 3 and 4, as shown in Table 1. There was a significant linear trend (Corpus 3 – $F(1, 338) = 21.182$, $p < 0.001$; Corpus 4 – $F(1, 410) = 42.167$, $p < 0.001$). In Corpus 3, as the place of articulation moved further back the amount of devoicing increased proportionately.

There was a significant effect of the factor **position in word** on the voicing of fricatives in Corpus 3, but no significant effect in Corpus 4, as shown in Table 1. There was a significant linear trend in Corpus 3 ($F(1, 338) = 13.285$, $p < 0.001$) indicating that as the position of the fricative moves from initial, through medial, to final word position, the amount of devoicing increased proportionately.

3.5 Duration and Devoicing Correlations

Two studies of American English tested whether duration and devoicing of fricatives were correlated. Smith found they were [6]; Crystal and House, that they weren’t [4].

In the present study, results of linear regression analysis of the amount of devoicing (dependent variable) vs. the duration of fricatives (independent variable), for all speakers, fricatives /v, z, ʒ/ and all positions in word, showed a significant linear trend both in Corpus 3 ($F(1, 339) = 49.153$, $p < 0.001$) and in Corpus 4 ($F(1, 411) = 27.455$, $p < 0.001$), i.e., as duration of fricatives increased the amount of devoicing increased proportionately. These four Portuguese speakers thus show the same effect exhibited by Smith’s four American speakers.

4 Conclusions

The mean duration of the unvoiced fricatives is greater than that of the voiced fricatives, and the mean duration of fricatives is greater than that of the VF and FV transitions. These characteristics are not particular of Portuguese, as similar results have been previously reported for English.

Devoicing rate is generally very high, especially when compared with studies of other languages. It is thought that this is an important characteristic of Portuguese, which would have to be incorporated in any production model to obtain more natural-sounding synthetic speech.

Analysis of variance showed that devoicing was significantly more likely for word-final fricatives and posterior place of articulation. Also, results of linear regression analysis showed that as duration of fricatives increased the amount of devoicing increased propor-

tionately.

ACKNOWLEDGMENTS

This work was partially supported by Fundação para a Ciência e a Tecnologia, Portugal.

REFERENCES

- [1] K. Pirello, S. E. Blumstein, and K. Kurowski, “The characteristics of voicing in syllable - initial fricatives in American English,” *JASA*, vol. 101, no. 6, pp. 3754–3765, 1997.
- [2] A. Jongman, R. Wayland, and S. Wong, “Acoustic characteristics of English fricatives,” *JASA*, vol. 108, no. 3, pp. 1252–1263, 2000.
- [3] S. R. Baum and S. E. Blumstein, “Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English,” *JASA*, vol. 82, no. 3, pp. 1073–1077, 1987.
- [4] T. H. Crystal and A. S. House, “A note on the durations of fricatives in American English,” *JASA*, vol. 84, no. 5, pp. 1932–1935, 1988.
- [5] G. J. Docherty, *The Timing of Voicing in British English Obstruents*, Foris Publications, Berlin, 1992.
- [6] C. L. Smith, “The devoicing of /z/ in American English: Effects of local and prosodic context,” *J. Phon.*, vol. 25, no. 4, pp. 471–500, 1997.
- [7] S. J. Mair and C. H. Shadle, “The voiced/voiceless distinction in fricatives: EPG, acoustic and aerodynamic data,” in *Proc. IoA*, 1996, vol. 18, pp. 163–169.
- [8] S. J. Behrens and S. E. Blumstein, “Acoustic characteristics of English voiceless fricatives: A descriptive analysis,” *J. Phon.*, vol. 16, pp. 295–298, 1988.
- [9] K. N. Stevens, S. E. Blumstein, L. B. Glicksman, M. Burton, and K. Kurowski, “Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters,” *JASA*, vol. 91, no. 5, pp. 2979–3000, 1992.
- [10] Luis M. T. Jesus and Christine H. Shadle, “A parametric study of the spectral characteristics of European Portuguese fricatives,” *Journal of Phonetics*, vol. 30, no. 3, pp. 437–464, 2002.
- [11] J. T. Hogan and A. J. Rozsypal, “Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant,” *JASA*, vol. 67, no. 5, pp. 1764–1771, 1980.
- [12] T. C. Veatch, “Word-final devoicing of fricatives in English,” in *Linguistic Society of America Winter Meeting Handbook*, Washington, USA, 1989.

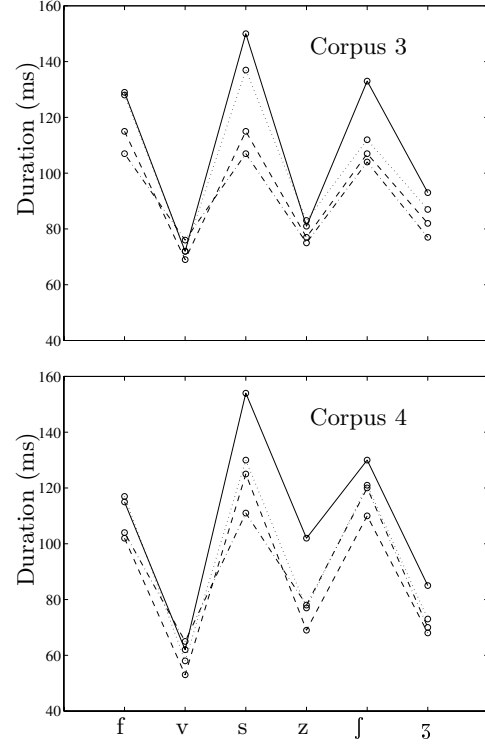


Figure 1: Mean duration of fricatives /f, v, s, z, ʃ, ʒ/. Speaker LMTJ – solid line; Speaker CFGA – dash-dotted line; Speaker ACC – dashed line; Speaker ISSS – dotted line.

Table 1: ANOVA of duration and devoicing. C. = Corpus. P. in W. = Position of fricative in word.

Factor	C.	Duration		Devoicing
		/f, s, ʃ/	/v, z, ʒ/	/v, z, ʒ/
Speaker	3	$p < .001$	N. S.	$p = .025$
	4	$p < .001$	$p < .001$	N. S.
Place	3	$p = .005$	$p < .001$	$p < .001$
	4	$p < .001$	$p < .001$	$p < .001$
P. in W.	3	$p < .001$	$p < .001$	$p < .001$
	4	$p < .001$	$p = .005$	N. S.

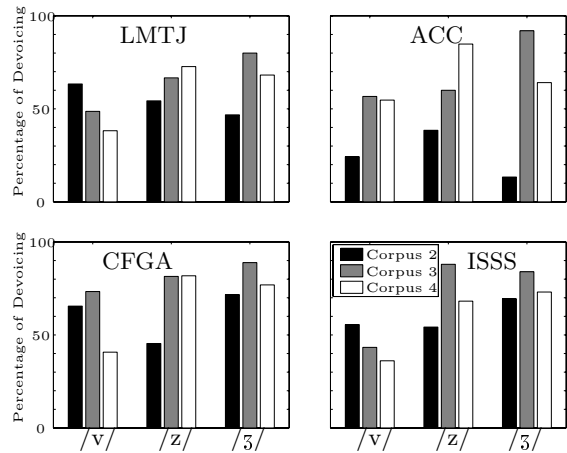


Figure 2: Percentage of total devoicing.