Image, Speech and
Intelligent Systems
Research Group

Department of
Electronics
and Computer Science

ACOUSTIC ANALYSIS OF A SPEECH CORPUS OF EUROPEAN PORTUGUESE FRICATIVE CONSONANTS

Luis M. T. Jesus and Christine H. Shadle

**Department of Electronics and Computer Science,
University of Southampton, UK**

EuroSpeech'99
Budapest, Hungary
6 September 1999

Introduction

- Portuguese has:
  - vowel reduction, consonant clusters, and fricatization of plosives (Viana 1984).
- Few and limited studies of Portuguese fricatives (Lacerda 1982; Viana 1984; Andrade 1995; Martins et al. 1995):
  - duration, acoustic features, and perception.
- In this study:
  - Systematic study of European Portuguese fricatives in a range of contexts, with temporal and spectral analysis.
  - Use methodology of previous fricative studies, but shift focus to real words and phonology of Portuguese.

Recording Method

Subject: male adult Portuguese native speaker.

Corpus 1: 154 words, containing fricatives /f, v, s, z, ʃ, ʒ/ in combination with the non-nasal vowels /i, ɨ, e, ɛ, ɐ, a, ɔ, o, u/, e.g., "Diga seta, por favor.".

All fricatives elicited word-finally with a reduced /ɨ/ as final phoneme (Andrade 1994).

Corpus 2: 12 sentences (61 of the 154 words).

Corpus 3: nonsense words /pV_1FV_2/ (V_i = /ɨ, ɐ, u/) following Portuguese phonological rules.

Corpus 4: sustained fricatives at 3 effort levels.

Acoustic and laryngograph (Lx) signals recorded (16 bits, $f_s = 48$ kHz, DAT) in sound treated booth.

Temporal Analysis Method

- Identify VF, F and FV boundaries using the time waveforms (acoustic and Lx signals).
- Classify devoicing using a new criterion based on both the acoustic and Lx signals.
- Previous studies: Lx signal – (Smith 1997); acoustic signal – (Pirello et al. 1997).
- Empirical criteria for devoicing:

  devoiced – no periodic structure;

  partially devoiced – a few steady cycles;

  voiced – steady cycles, even if the amplitude is much lower than in the vowel.
- Quantitative criterion for devoicing: ...

Quantitative Criterion for Devoicing

Compute variance of laryngograph (Lx) signal

$\sigma_{\mathrm{t}}^2(\mathbf{x})$ – **in VF transition.**

$\sigma_{\mathrm{f}}^2(\mathbf{x})$ – **in fricative.**

Compute ratio of variances:

$$\mathbf{r}_\sigma(\mathbf{x}) = \frac{\sigma_{\mathrm{t}}^2(\mathbf{x})}{\sigma_{\mathrm{f}}^2(\mathbf{x})}$$

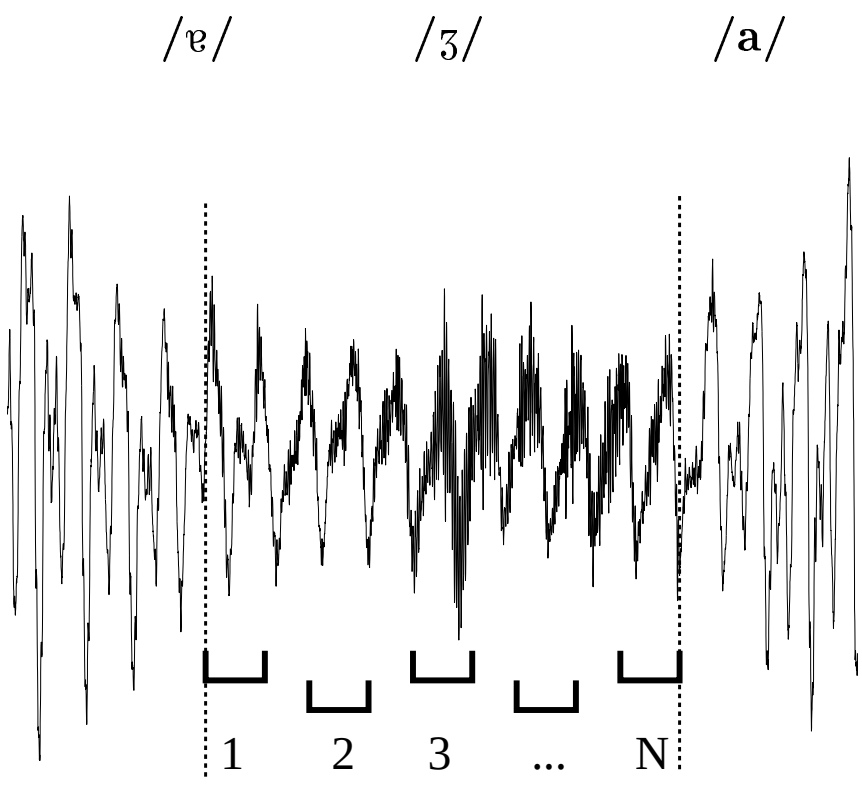
Apply heuristic thresholds:

If $\mathbf{r}_\sigma \geq 15$**, then** ***devoiced***

else ***voiced.***

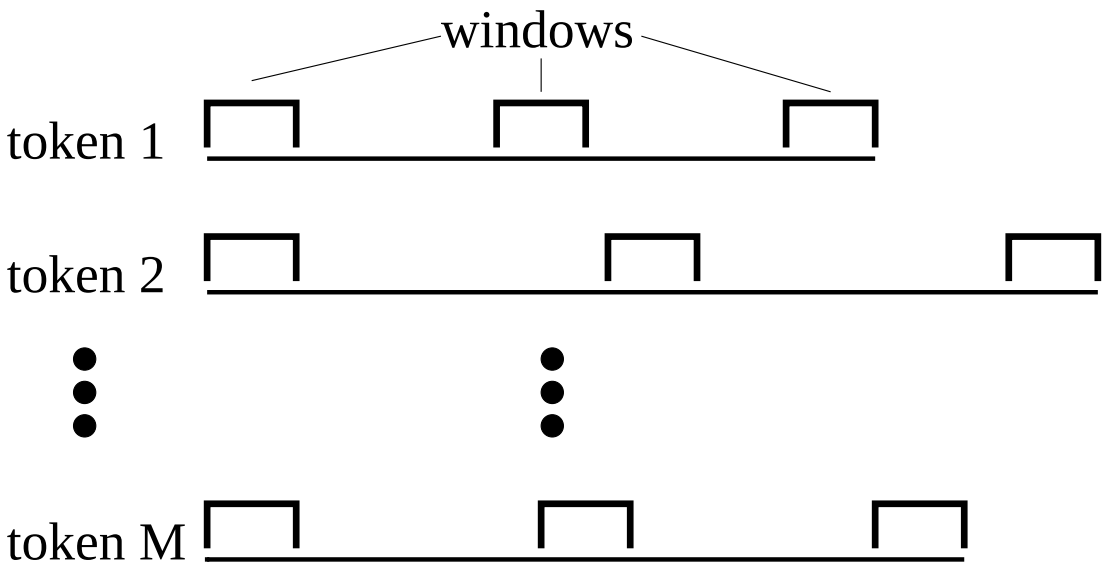
Time-averaged power spectrum (Corpus 1, 2 and 4)

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

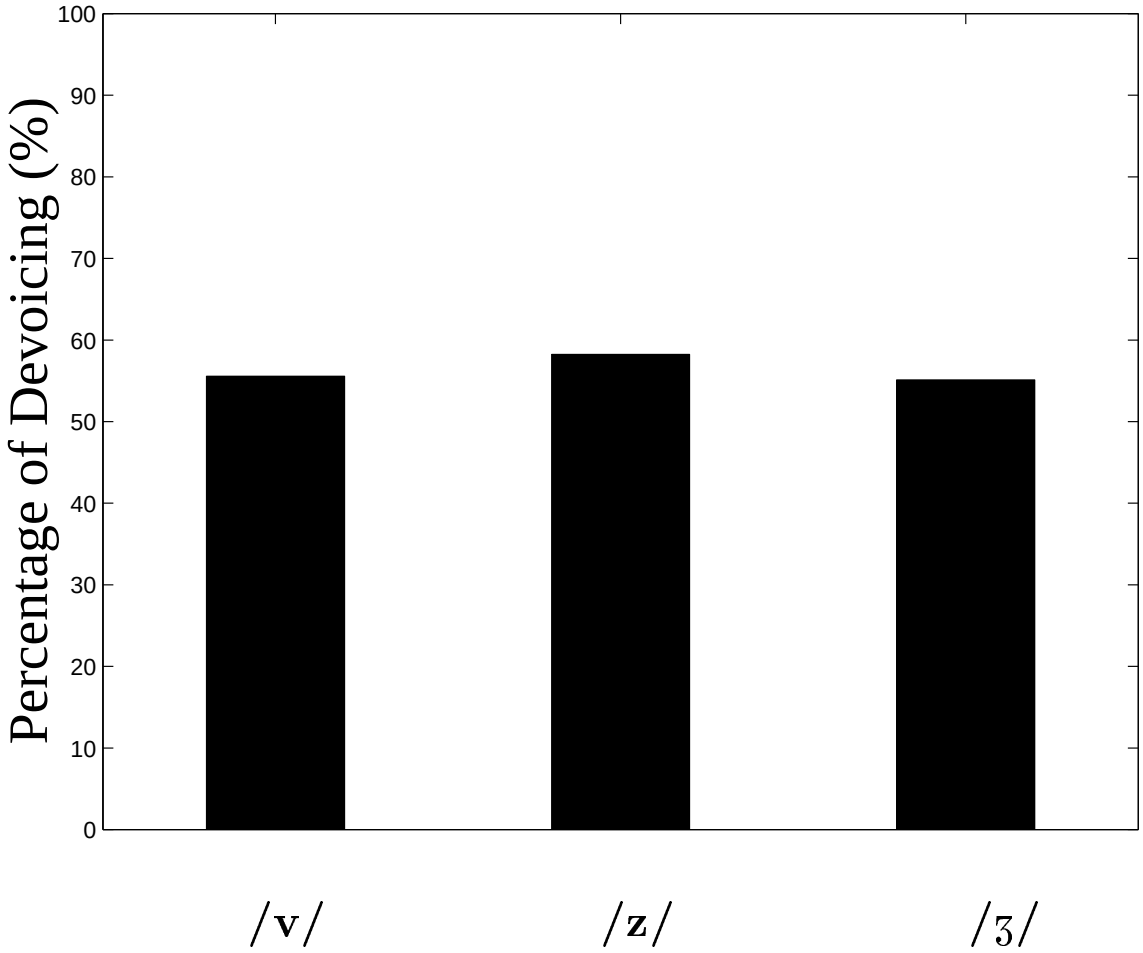


Ensemble-averaged power spectrum (Corpus 3)

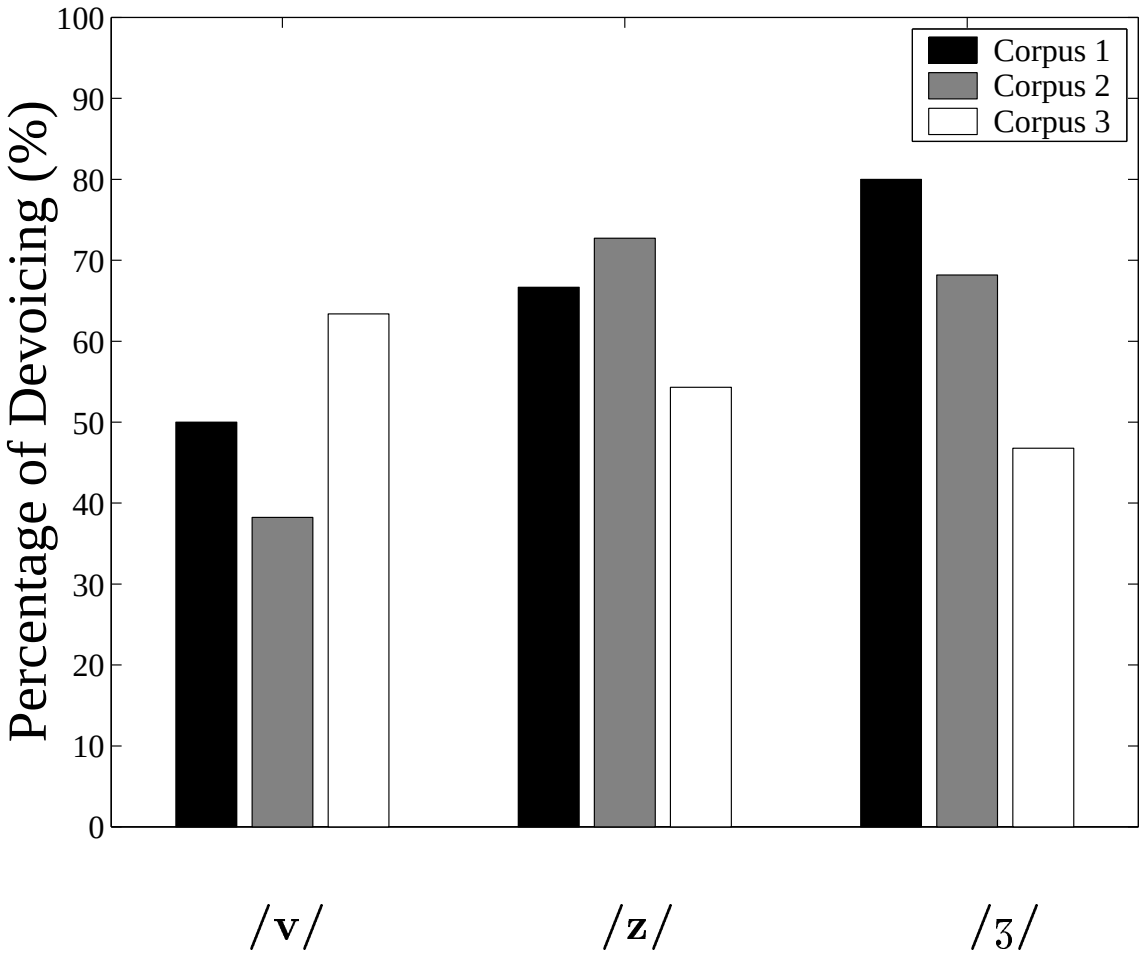
$$\mathbf{P_E(f) = \frac{1}{M} \sum_{k=1}^{M} |X_k(f)|^2}$$



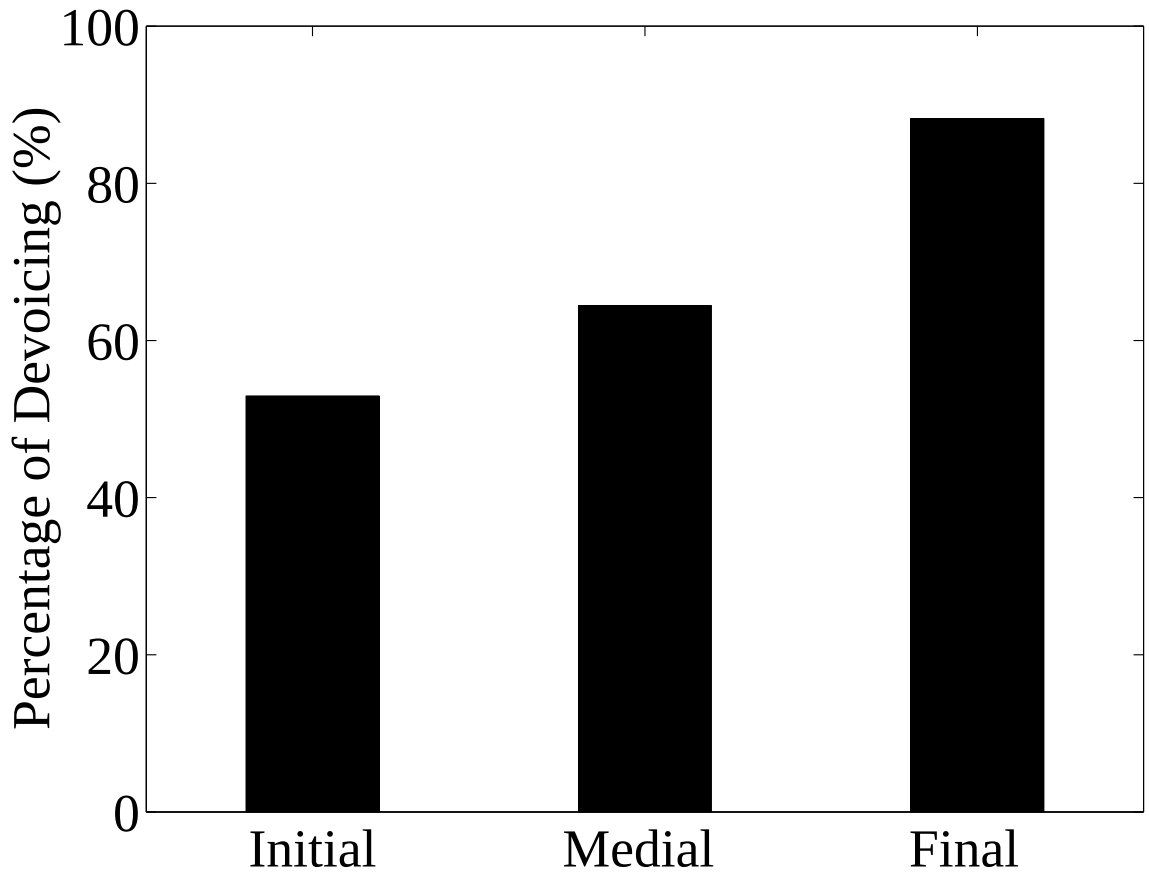
Temporal Analysis Results – Overall Devoicing (Corpus 1, 2 and 3)



Devoicing by Corpus

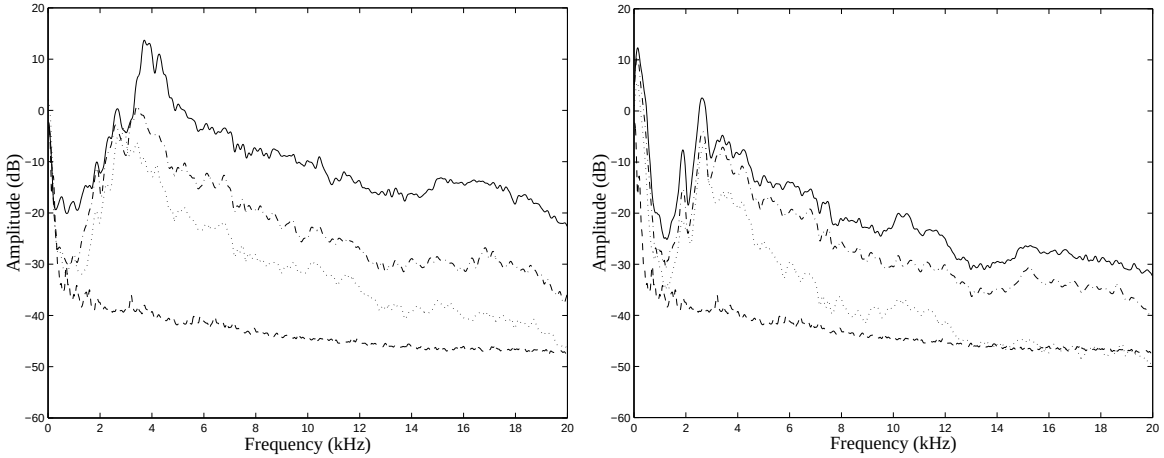


Devoicing by Position in Word (Corpus 1)



Spectral Analysis Results (Corpus 4)

Spectral shape differs for unvoiced – voiced pairs.



Frequency (kHz)

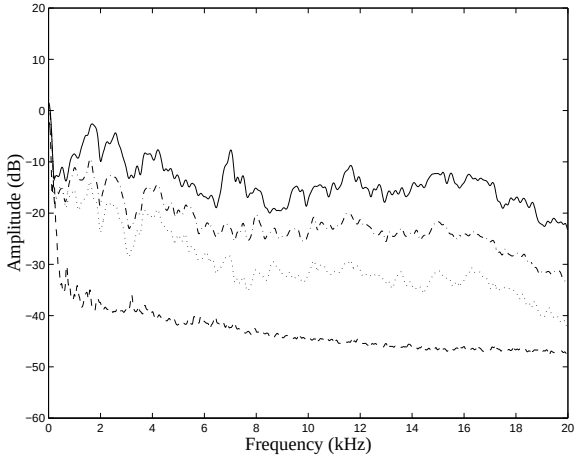
— high effort level

-.- medium

... low

Spectral Analysis Results (Corpus 4)

- **Amplitude increases with effort level: more at high frequencies.**

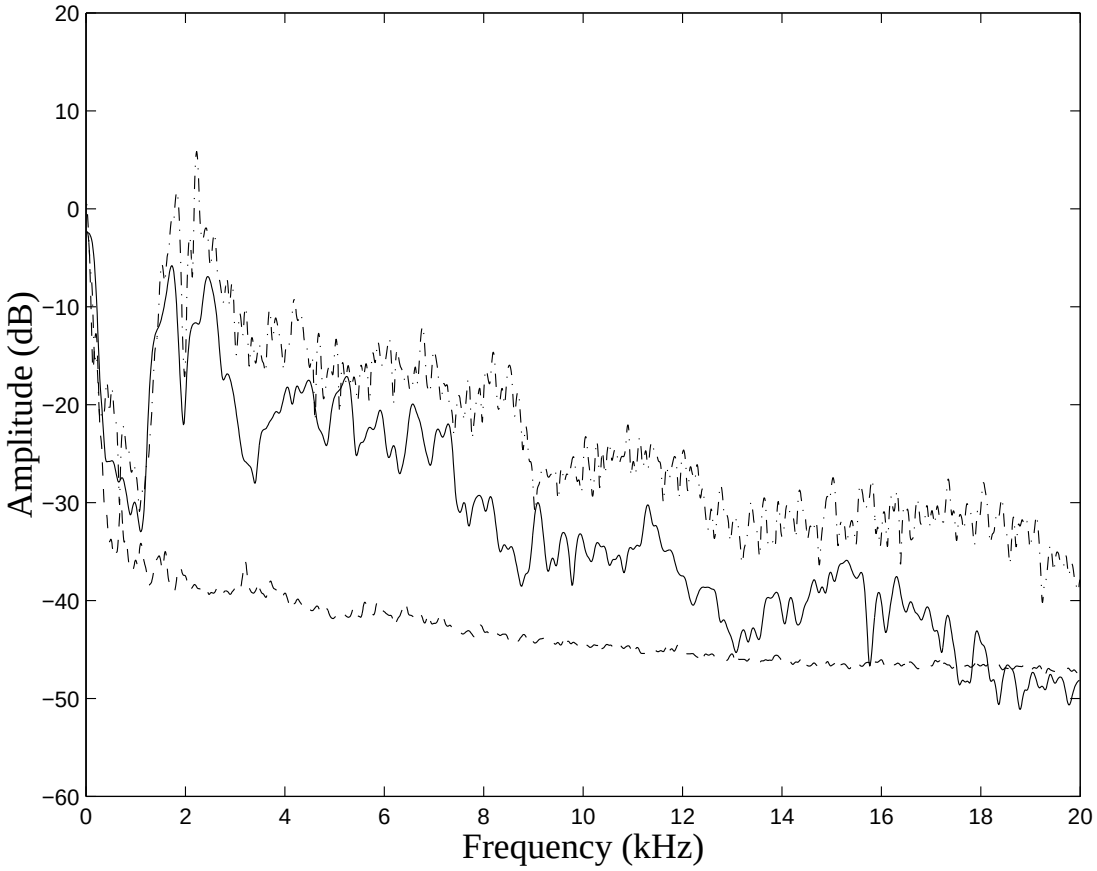


- **Amplitude difference at high frequencies: 10 dB for /f/, 5-10 dB for /v/, 15 dB for /ʃ/, and 10 dB for /ʒ/.**
- **These differences are associated with source type and strength, and are similar to results for American English and French subjects.**

Spectral Analysis Results (Corpus 1 and 3)

Does effort level of sustained fricatives correspond with stress of the syllable containing the fricative in Corpora 1 and 3, and possibly also position within the word in Corpus 1?

Spectral Analysis Results (Corpus 1 and 3)



— [ʃ] from [kɐˈpuʃ] – Corpus 1

-.- [ʃ] from [puʃu] – Corpus 3

Spectral Amplitude Comparisons

Same vowel context. Spectral amplitudes > 2 kHz.

Corpus 3 e.g. [puʃu]		Corpus 1 e.g. [kɐˈpuʃ]
	≈	3 cases, stressed
	>	4 cases, unstressed

Corpus 4 e.g. [uʃ... ʃu]		Corpus 1 and 3 e.g. [kɐˈpuʃ], [puʃu]
high effort		no equivalents in word or nonsense-word
medium	≈	all stressed
soft	≈	all destressed

The main peak for /uʃu/ context was at a significantly lower frequency than in, e.g., /ɨʃi/ context; this is as expected from previous work (Shadle and Scully 1995).

Conclusions

- Voiced fricatives devoice in over one half the cases in both nonsense and real words.
- Spectral analysis revealed a correspondence between the effect of effort level and of syllable stress, and showed some effect of vowel context.
- Corpus 3 is better controlled and easier to analyse than Corpus 1 or 2; validating its use would give an important advantage.
- Future/ongoing: analysis of other Portuguese speakers.