Devoicing Measures of European Portuguese Fricatives

Luís 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

Introduction

Studies of Portuguese fricatives have been limited in comparison with other languages, yet are important for applications such as speech synthesis.

In this presentation we focus on devoicing of fricatives using recordings by four subjects of a set of corpora.

Method

A speech corpus has been designed for European Portuguese, with the fricatives /f, v, s, z, ʃ, ʒ/ in the following contexts:

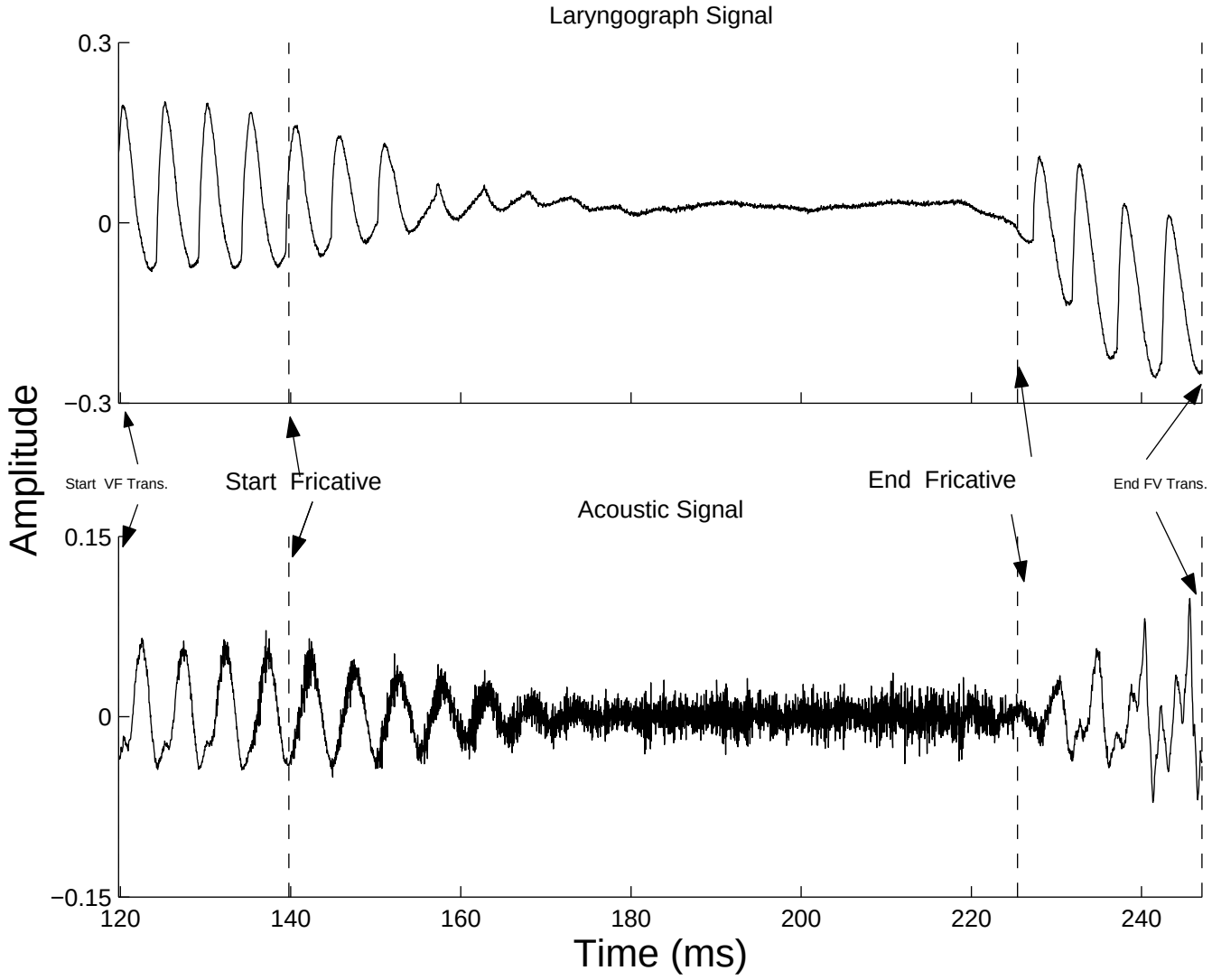
- sustained (Corpus 1),
- repeated nonsense words (Corpus 2),
- words containing fricatives in frame sentences (Corpus 3), and
- the same set of words in sentences (Corpus 4).

Method

Four subjects, two male (LMTJ and CFGA) and two female (ACC and ISSS), were recorded reading the corpora.

Recordings of acoustic and laryngograph (Lx) signals were made in a sound treated room, and digitally transferred to a computer for post-processing.

Laryngograph signal and acoustic signal of fricative /z/ in *azar* /ɐ'zaɾ/. Corpus 3 (Speaker ISSS).



Manual Criterion for Devoicing

Voicing is often maintained over only part of the fricative.

When acoustic or Lx signal has periodic structure for:

- $< 1/3$ **of frication interval** $\longrightarrow$ **devoiced**
- **between** $1/3$ **and** $1/2$ $\longrightarrow$ **partially devoiced**
- $> 1/2$ $\longrightarrow$ **voiced**

Automatic Criterion for Devoicing

Usually, techniques for automatically detecting whether a portion of a signal is voiced or not, are unsuitable for fricatives because in this class of speech sounds voicing has low energy (Docherty 1992, p. 102).

Therefore a new criterion, based on the Lx signal, was tested for the corpora used in the present study.

Automatic Criterion for Devoicing

The sample mean

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{1}$$

and sample variance

$$\sigma^2(x) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{x})^2 \tag{2}$$

of the Lx signal samples x_i, $i = 1, ...N$ were calculated during the VF transition and during the fricative.

Automatic Criterion for Devoicing

The ratio of variances of the two intervals,

$$r_{\sigma^2(x)} = \frac{\sigma_t^2(x)}{\sigma_f^2(x)}, \tag{3}$$

was calculated.

$\sigma_t^2(x)$ = variance of the signal during the VF transition

$\sigma_f^2(x)$ = variance of the signal during the fricative

Automatic Criterion for Devoicing

A heuristic threshold of 15 was used:

– for $r_{\sigma^2(x)} \geq 15$, the fricative is labelled devoiced;

– if $r_{\sigma^2(x)} < 15$, voiced.

partially voiced → devoiced category

The ratio of variances was used as the criterion for devoicing for Corpus 3 and 4 fricatives of Speaker LMTJ.

Automatic Criterion for Devoicing

For some examples the Lx signal slowly increases or decreases in amplitude over the frication interval.

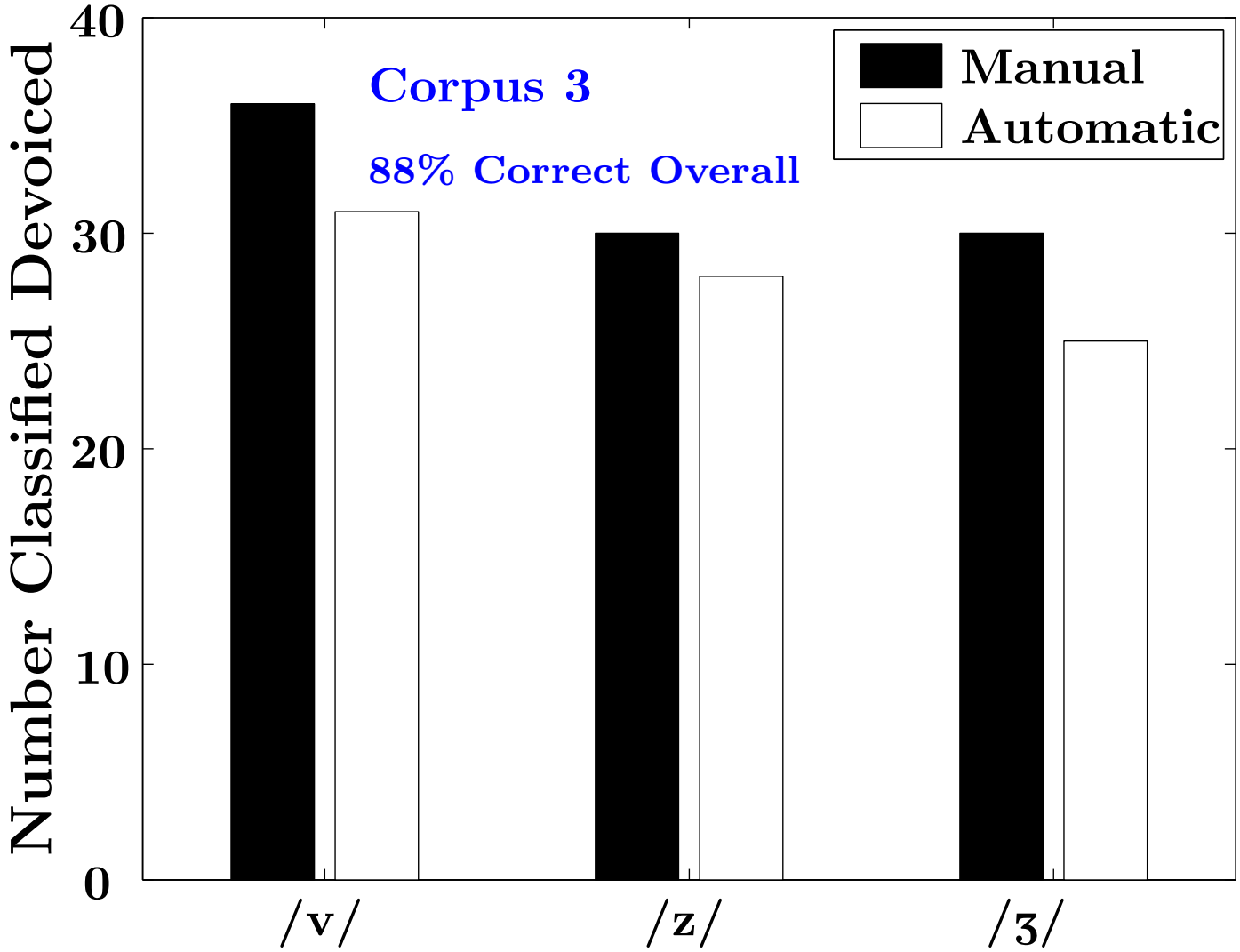
This results in a large variance, and therefore a misclassification as voiced.

Automatic Criterion for Devoicing

This problem has been solved using a new ratio of variances computed from the average frication interval variance calculated for three consecutive equal length sections of the frication interval.

Using a larger number of sections over which we calculate the averaged $r_{\sigma^2(x)}$ did not significantly improve the efficiency of this measure of devoicing.

Evaluation of the Automatic Devoicing Criterion (Corpus 3)

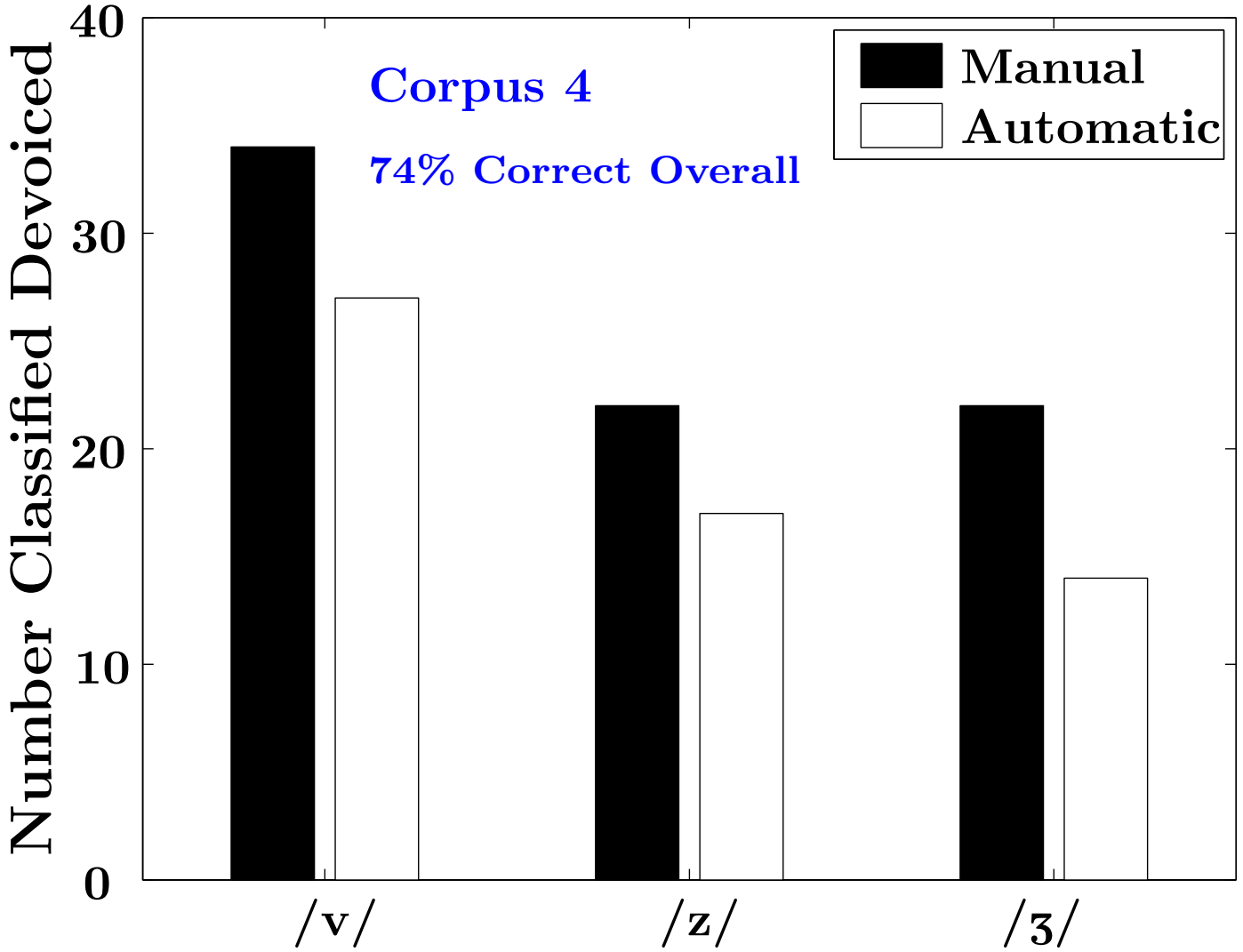


Evaluation of the Automatic Devoicing Criterion (Corpus 3)

Discrepancies in Corpus 3 classification caused by:

1. **Cases on devoiced / partially devoiced borderline (17%)**
2. **A few large peaks in Lx increase σ^2 – segmentation criteria become crucial (33%)**
3. **Low - amplitude Lx in completely voiced fricatives (50%)**

Evaluation of the Automatic Devoicing Criterion (Corpus 4)



Evaluation of the Automatic Devoicing Criterion (Corpus 4)

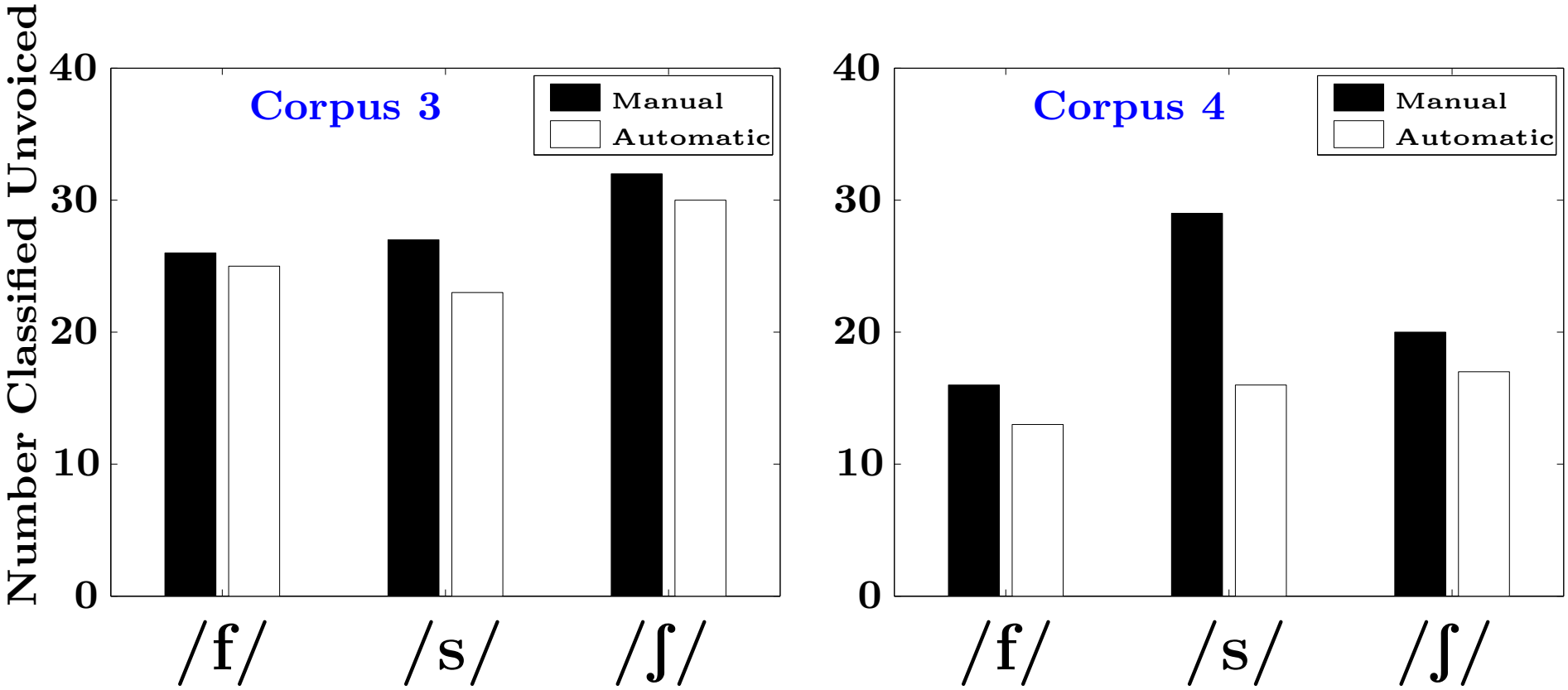
Discrepancies in Corpus 4 classification caused by:

1. **Devoicing during the VF transition (51%)**
2. **A few Lx cycles during fricative (41%)**
3. **DC drift in Lx (8%)**

Evaluation of the Automatic Devoicing Criterion

The r_{σ^2} metric was also successful when used for the unvoiced fricatives /f, s, ʃ/ of Corpus 3 and 4, with an overall 83% correctly classified.

Classification of Unvoiced Fricatives



Conclusions

A preliminary evaluation of the automatic criterion for devoicing showed great potential for the use of this technique in future work.

The percentage of voiced fricative tokens from Corpus 3 and 4 which were classified in the same category using the two methods (manual and automatic) is quite high (overall 83%; range 64 - 93%).

Acknowledgements

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

References

Docherty, G. J. (1992). *The Timing of Voicing in British English Obstruents*. Berlin: Foris Publications.