A CASE STUDY OF PORTUGUESE AND ENGLISH BILINGUALITY

Luis M. T. Jesus †, 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

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

ABSTRACT

This study of the acoustic characteristics of European Portuguese and British English fricatives as produced by two bilingual subjects, consisted of time and frequency analysis of words in a carrier sentence. Time-averaged power spectra were calculated and parameterised in order to aid comparisons across speaker, across corpus, and across language, and to gain insight into the production mechanisms underlying the language-specific variations.

1. INTRODUCTION

Although some bilinguals seem to attain monolingual-like speech production in both languages, it is very likely that bilinguals choose different strategies from monolinguals, which “reduce the difficulties created by their need to use two systems, without thereby sounding in any way abnormal in either” [12] (p. 37). Therefore we must consider the particulars of bilingual speech when interpreting our cross-language results, and begin our study by establishing the type of bilingualism exhibited by our subjects.

We used one of the measures of bilinguality proposed by Hamers and Blanc [13] (p. 40), which involved the collection of language biographies, self-evaluation, and judgements of bilingual production by monolingual speakers of Portuguese and of English. There are several ways of classifying bilinguals in terms of their fluency and language dominance [14, 15], so we used a previously tested procedure. Subjects filled in a questionnaire [16] very similar to that originally designed by Hazan and Boulakia [11].

The subjects used in this study were two adult bilingual siblings, with no reported history of hearing or speech disorders. Subject PS was a 22-year-old male and Speaker RS was an 18-year-old female. The siblings’ mother is a European Portuguese speaker and the father a British English speaker, who reside in Cascais, Portugal. They have interacted with their parents since infancy in their mother tongues: in Portuguese with their mother and in English
with their father. The age and context of acquisition of both languages, their past and present use, and the degree of literacy were inferred from the questionnaires [16] (pp. 221 - 223).

The level of bilingual competence was evaluated informally by two speech researchers, where the naturalness of the recorded Portuguese and English sentence corpora was judged to be close to "native-like" for both speakers and for both languages. Considering all of this information it is most probable that our subjects have developed a balanced and compound bilinguality [13].

3. CORPORA DESIGN AND RECORDING

The Portuguese corpora had a very similar design to the one described by Jesus [16]. The English corpora were designed to provide valid data for cross-language comparisons with the Portuguese corpora. He also included sustained fricatives (Corpus 1a and 1b), and a set of nonsense words (Corpus 2), words (Corpus 3) and sentences (Corpus 4). Previously used English Corpora [2, 3] were augmented to match the Portuguese corpora.

Each speaker was recorded in two separate sessions (Portuguese and English sessions), where the subjects counted and talked in the language of the current session, and the order of corpora recording was one of decreasing naturalness: we started by recording the sentence corpus (Corpus 4), followed by the real word corpus (Corpus 3), nonsense word corpus (Corpus 2), and finally the sustained fricative corpus (Corpora 1a and 1b).

Technical aspects of the recording method were the same as described in [16]. 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 Bruel & Kjaer 2690 preamplifier. The signal was amplified and filtered by a Bruel & Kjaer 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 processor [1].

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.

4. METHOD

The segmentation techniques, temporal and spectral analysis methods, and parameterisation used in this Portuguese and English cross-language study, were the same as in the study of Portuguese by Jesus [16], and are described below.

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 end of the fricative-vowel transition [16]. 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 [16].

We used time-averaging in Corpus 3, with nine 10 ms Hamming windows and one hundred 10 ms windows to calculate the spectra of sustained fricatives in Corpus 1a and 1b [16]. The four parameters described by Jesus and Shadle [1] were also calculated: $F'$ – average of a manually calculated peak frequency of all sustained tokens which corresponds to the same cavity resonance for a particular place (labiodental, dental, alveolar and postalveolar); the dynamic amplitude $A_d$ – the difference between the maximum amplitude value of the averaged power spectrum occurring between 500 Hz and 20 kHz, and the minimum amplitude between 0 and 2 kHz; $S_p'$ – the slope of the line fit to all the spectral amplitude points from 500 Hz to $F'$; $S_p$ – the slope of the line fit to all the points from $F'$ to 20 kHz.

The value of $F'$ used for English dental fricatives was the same as used for Portuguese labiodental fricatives [1], that is, $F'/f, v, z'/ = 5$ kHz. This resulted from an analysis of Corpora 1a and 1b spectra, that consisted of a comparison of the spectra of fricatives $\theta, \partial$ with those of fricatives $f, v$. The overall amplitude and spectral peaks of $\theta, \partial$ did not differ substantially from $f, v$, so an $F' = 5$ kHz was considered adequate for both places of articulation.

5. RESULTS

5.1. Duration and Devoicing

The median duration of the unvoiced fricatives was always greater than the median duration of the voiced fricatives, which agrees with results for the English language [17, 18, 19, 20]. However, there is no significant difference by place of articulation or between Portuguese and English in the results presented in Figure 1.

For both Portuguese and English, most word-final fricative examples (48 out of 50) were totally devoiced. Overall results from the analysis of devoicing show that more than 50% of the fricatives devoice, except for $/\partial/ \ v$ produced by PS. When the percentage of devoicing was plotted by position in words, there was no significant difference between Portuguese and English.

5.2. Parameterisation of Spectra

Spectral parameters were examined only for Corpus 3 fricatives because these reflected language-specific characteristics. In plots of $A_d$ and $S_p$ by fricative, $/s, z, j, z/ \ have
higher \( A_d \) and lower \( S_p \) than /\( f, v, \theta, \delta/ \) for Speaker RS, as shown in Figure 2. The values of \( A_d \) and \( S_p \) for \( \theta, \delta \) produced by Speaker PS seem to fall in between the values for /\( f, v \)/ and /\( s, z, f, j, z/\( \). On \( A_d \) vs. \( S_p \) and \( S_p' \) vs. \( S_p \) plots, there are separate clusters of sibilants and /\( f, v, \theta, \delta/ \) for Speaker RS. However, the cluster for \( \theta, \delta \) produced by Speaker PS seems to fall in between the /\( f, v/ \) cluster and the /\( s, z, f, j, z/ \) cluster. Results from both subjects seem, for the most part, to be consistent, and the same for Portuguese and English fricatives.

### 6. CONCLUSIONS

In this study, we designed fricative corpora ranging from sustained fricatives to real Portuguese and English words, and recorded and analysed two bilingual speakers. Our goal was to test some conclusions from our study of four monolingual Portuguese subjects [1], and to compare our results to those of previous studies of English fricatives [2, 3].

Our principal findings are as follows. Devoicing occurs more often in word-final than word-initial position, both for Portuguese and English fricatives. The percentage of totally devoiced Portuguese examples produced by the four monolingual subjects was higher than for English examples produced by the two bilingual subjects, but Portuguese and English bilingual results were very similar.

The parameters spectral slope, frequency of maximum amplitude, and dynamic amplitude, were applied to the bilingual corpora. A combination of parameters \( A_d \) and \( S_p \) and of parameters \( S_p \) and \( S_p' \) was useful for separating the fricatives by sibilance. Results for Portuguese and English fricatives seem to be very similar. The parameters \( A_d \), \( S_p \) and \( S_p' \) are either capturing aspects of Portuguese that do not differ from English, or the subjects produce Portuguese and English fricatives the same way.

The parameters studied in this cross-language study might not be capturing subtle differences, and the time and frequency characteristics analysed for Portuguese and English fricatives appear to be quite invariant. It is possible that speakers PS and RS used different production strategies from monolinguals, without this being perceptible, but resulting in an attenuation of language acoustical contrasts. Therefore our British English corpus should be used to collect data on monolingual subjects in the future.

### 7. REFERENCES


**Fig. 1** Median duration of fricatives /f, v, θ, ð, s, z, ʃ, ʒ/ in Corpus 3. Portuguese – solid line; English – dashed line; × – Speaker PS; ○ – Speaker RS.

**Fig. 2** Corpus 3 (Speaker RS): Portuguese fricatives – top; English fricatives – bottom. ○ – /f/, × – /v/, ▽ – /θ/, △ – /ð/, ⋄ – /s/, ⋆ – /z/, × – /ʃ/ and □ – /ʒ/.