Devoicing of phonologically voiced obstruents: Is European Portuguese different from other Romance languages?

Daniel Pape[1], *Luis M.T. Jesus*[12]

[1]IEETA (Universidade de Aveiro, Portugal), [2]ESSUA (Universidade de Aveiro, Portugal)
danielpape@ua.pt, lmtj@ua.pt

ABSTRACT

This paper presents results for voicing maintenance during European Portuguese (EP) stop and fricative production. Results showed that EP presented a high amount of devoicing for all phonologically voiced stops and fricatives. This is in contrast to classical literature reporting high voicing maintenance during stop closure for Romance languages, but confirms our preliminary results from previous work. Further, for the first time results are presented for the analyses of (time-dependent) voicing profiles for EP. These profiles show differences in manner or place of articulation. However, a strong speaker-dependency on both devoicing percentage and voicing profiles is observed. Analysis of the effect of height of the following vowel on devoicing showed no difference for three of four speakers. However, for one speaker higher devoicing for all consonants when followed by a low vowel was observed.

Keywords: phonetic obstruent voicing, devoicing, voicing profiles, European Portuguese.

1. INTRODUCTION

It has long been known that, across languages, different perceptual cues are responsible for the voicing distinction. For stops, it is generally agreed that voice onset time (VOT) is the most dominant (perceptual) cue for voicing distinction for English [7], Korean [2] and French [3]. Closure duration and contextual vowel duration are also seen as dominant cues in a number of languages [2, 8]. However, when the VOT cue is weak or missing, it's likely the perceptual system relies on other available cues and thus applies different weighting techniques. The interaction and weighting of different available cues could be highly language-dependent: While some languages merely rely on cues like VOT or vowel duration, other language could rather rely on voicing maintenance or closure duration in acoustically challenging conditions, e.g., noisy environments or multi-talker conditions.

Our current research project aims to examine the importance of the perceptual cue *voicing maintenance* for European Portuguese (EP) in comparison to other languages, and to compare these results to speech production data.

In this paper, we focus on the speech production part, i.e. the different phonetic realisations of consonant voicing maintenance for EP in different conditions. The available literature (e.g. [6, 11, 12, 13, 14]) shows inconclusive results on cross-linguistic differences for voicing maintenance. When examining an extensive cross-linguistic database, Shih et al. [12] showed that for all voiced stops the percentage of devoicing was considerably higher for German than for Italian or Spanish. For these two Romance languages, no devoicing occurred throughout the complete closure of the voiced stops. This finding for Romance languages is in line with the phonological view that for Romance languages voicing during consonant closure is more important than for Germanic languages. However, for EP as another member of the Romance languages, it has been shown [5, 6] that the devoicing rate for fricatives is surprisingly high. In fact, when comparing the data from [5, 6] with the devoicing results previously published for German stops [11], it can be seen that both languages show a similar percentage of phonetically devoiced items, despite the different language families and different manner of articulation.

Voiced fricatives tend to devoice easily [9], so for a valid cross-linguistic comparison of EP and German corpora, both manner and place of articulation have to be matched carefully. The aim of the current study is to shed light to the question whether the discrepancy of obstruent voicing for EP (on one hand the assumed high voicing status anticipated as a member of the Romance languages, on the other hand the high percentage of devoicing) can be attributed to the methodological and experimental differences or are in fact an important feature of the (presently not well-examined) EP.

Classically, the decision whether a given obstruent is considered voiced or devoiced is based on the absence of a detectable voicing bar for at least one period of the closure/obstruction. However, this binary decision of *voiced vs. devoiced* lacks the important time-varying characteristics during the process of devoicing. When focusing on cross-linguistic differences and the underlying processes it is quite different whether an item shows devoicing right at the beginning or rather in the last milliseconds of the closure (the latter resulting in a nearly complete voicing during closure). Although the underlying process causing voicing offset might be completely different, with the binary decision both items are equally categorised as devoiced. Therefore we are interested in a reliable measure for the voicing status throughout the complete obstruent that uses a time-varying strategy in contrast to the classical binary voicing decision.

2. METHOD

In order to be able to cross-linguistically compare our results, several key factors have to be taken into account, given that the amount of devoicing varies with place of articulation and vowel context [10, 11, 12] and phoneme position [14]. Thus, we carefully controlled for these issues when constructing our corpus.

We recorded CVCV items in the frame context *Diga CVCV outra vez*. Each consonant and vowel in the CVCV sequence was identical (thus resulting in an initial and a medial consonant position). Our recordings consisted of all EP stops and fricatives /p b t k b d g f v s z ʃ ʒ/ with identical vowel contexts /i e o a/, for the initial and medial position. Each item within the frame sentence was repeated 9 times by 4 different speakers (three females, one male). In sum, a total of 3456 items (4 speakers, 12 consonants, 4 vowel conditions, two positions, 9 repetitions) were produced. Recordings were made in a sound-proof room using a Cirrus Research MK224 microphone located 1m in front of the speaker's mouth. The acoustic signal was preamplified (Cirrus Research MV 181 A) and then recorded using a Marantz PMD671 Solid State Recorder with a sampling frequency of 48kHz. All speakers were from Aveiro (central Portugal) and did not spend extended periods in other regions of Portugal.

Each of the 3456 items was labelled at the following landmarks: beginning of preceding vowel, beginning of closure/obstruction of the first obstruent, beginning of obstruent noise/burst, beginning of first target vowel, end of first target vowel, beginning of closure/obstruction of the second obstruent, beginning of second target vowel, and end of second target vowel. All vowel onsets/offsets were regarded as the onsets/offsets of the corresponding obstruent and were labelled by the appearance/disappearance of a clear higher formant structure (F2 and F3).

In analogy to the work of Shih et al. [12], we computed the voicing status for each item, sampled at 10 equal-distant landmarks throughout the complete consonant duration. The first landmark corresponds to the beginning of the consonant obstruction, whereas the 10th landmark corresponds to the obstruction offset and onset of the following vowel. We used the PRAAT (v.5.2) [1] *AC pitch extraction* algorithm with the settings *voiceless decision = 0.55* and *Silence threshold =0.1* (resulting in the most accurate and reliable F0 tracking and corresponding voiced decision for our purpose [4]). Throughout the whole analysis process, we constantly checked manually for errors of the algorithm and corrected if necessary.

3. RESULTS AND DISCUSSION

In the following first three subsections, we only report results for the medial consonant position.

3.1. Temporal measures

The mean duration for the voiceless consonants was 180ms (*std* 30ms) and 134ms (*std* 29ms) for the voiced consonants. The mean duration of the preceding vowel was 118ms (*std* 21ms) for the voiceless consonants and 78ms (*std* 25ms) for the voiced consonants. There were no significant differences in duration across the speakers, when comparing stops with fricatives, or between different places of articulation.

3.2. Devoicing decisions

For compatibility with the classical literature, we first present the classical binary voicing decision, i.e., a stop/fricative is regarded as devoiced when at least one pitch period can be found without noticeable periodic signal in the acoustic waveform during the closure/obstruction. In table 1 we present the percentage count of devoiced items for all phonologically voiced consonants, collapsed over the four vowel contexts /i e o a/.

As can be seen for all speakers in table1, the percentage of devoiced items is very high for both stops and fricatives. When comparing stops with fricatives, it can be seen that stops have a higher count of devoiced items than fricatives. Interestingly, in contrast to results found in the literature [10, 11, 12] the high percentages of devoicing for EP are independent of the place of articulation, i.e., velars are not more devoiced than dentals or bilabials. When comparing the difference between speakers, speaker3 shows (for all consonants) less devoicing than all other speakers.

Table 1: Percentages of devoiced items in reference to the complete database count (n), split by consonant and speaker (sp). Highest percentages are bold printed.

consonant	n	sp.1 [%]	sp.2 [%]	sp.3 [%]	sp.4 [%]
/b/	36	**59**	**89**	39	**92**
/d/	36	**61**	**92**	**78**	**98**
/g/	36	**58**	80	**53**	**98**
/v/	36	33	39	25	**85**
/z/	36	53	58	40	**85**
/ʒ/	36	14	**83**	47	68

3.3. Devoicing patterns (voicing profiles)

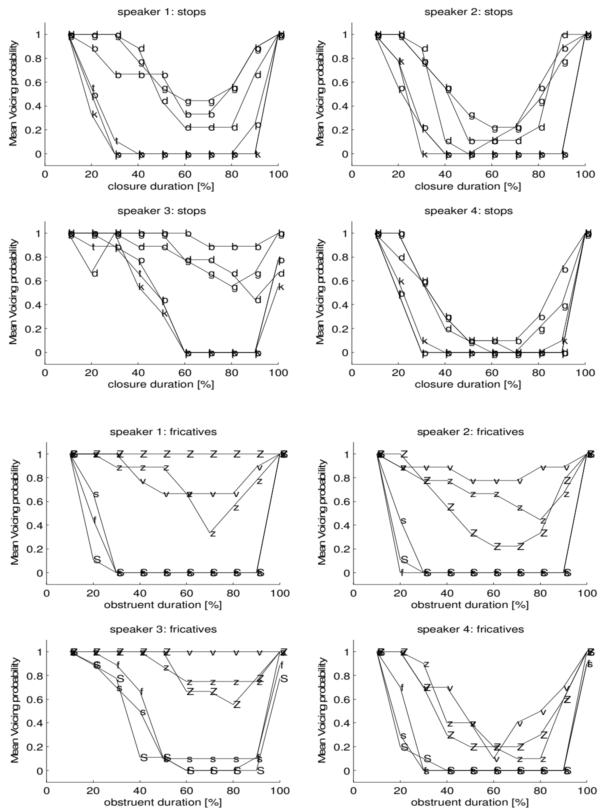
In figure 1 we present results for the voicing profiles sampled at the 10 equidistant points throughout the consonant closure/obstruction. The data for the voiceless obstruents is given as a reference for the difference of the voiced/voiceless voicing maintenance.

When examining the difference between voiceless and voiced consonants of all voicing profiles, figure 1 shows that voiceless consonants generally cease voicing very rapidly starting at the consonant onset and remain unvoiced until the offset.

Most phonologically voiced stops and fricatives show a high percentage of devoicing starting around 30% (after consonant onset). The only exceptions are the completely voiced /ʒ/ productions of speaker1 and /v/ productions of speaker3. Generally, stops showed more devoicing with a steeper slope than fricatives.

With respect to speaker-dependency, there is no clear pattern when comparing place of articulation. Only two speakers (speaker2 and speaker3) showed more and steeper devoicing for posterior place of articulation. However, for all consonants speaker3 shows less devoicing and less steep slopes than all other speakers. Further, for this speaker there is a difference in the transitions for his voiceless consonants: As can be seen in figure 1, the transition into the devoiced status decreases slowly towards the acoustic mid of the obstruction, whereas for all other speakers the devoiced status is reached very fast at the consonant onset.

Figure 1: Devoicing patterns for EP. Each data point is the mean of 36 items (9 repetitions x 4 vowels).

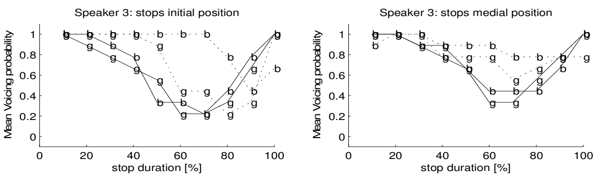


3.4. Dependence on context and position

To find out what are the reasons for the different devoicing patterns of speaker3, we analysed the effect of vowel height (see figure 2) and consonant position in all voicing profiles. We found that for all speakers the dental voiced stop had a substantially higher voicing probability when followed by high vowels, consistently occurring for both initial and medial consonant position.

Apart from this effect, there was no consistent influence of vowel height on the voicing profiles for three speakers. However, we found that the voicing profiles of speaker3 are different from all other speakers: For a high vowel, a very high voicing maintenance for all places of articulation can be observed (see figure 2). In contrast, when followed by a low vowel, there was consistently stronger stop devoicing for both initial and medial position. In sum, the weight of the strong voicing maintenance for high vowels explains the difference in voicing profiles for speaker3 when compared to the other speakers.

Figure 2: Stop devoicing patterns for speaker3 (effect of following vowel height): High vowel context /i/ is printed as dotted, low vowel context /a/ as a solid line.



4. CONCLUSIONS

With respect to EP stops, we found surprisingly high devoicing occurrences: For both manners and all places of articulation, we observed a consistent and strong devoicing. With respect to EP fricatives, we observed a consistent high devoicing percentage, based on a larger database than previously reported [5, 6].

Our analysis of the voicing profiles showed that EP clearly differs from other Romance languages as reported in [12]. In fact, when comparing our data to the voicing profiles given in [12], it is apparent that (with respect to devoicing) EP is more similar to a Germanic language than to a Romance language: For the two Romance languages Spanish and Italian reported in [12], all voiced consonants were highly voiced throughout the complete consonant duration, whereas in EP our data shows very strong devoicing patterns, independent of place or manner of articulation.

It is observed that the EP stops are more prone to devoicing than fricatives. Given that the simultaneous maintenance of voicing and frication is rather challenging [9], our high devoicing patterns for voiced fricatives in EP is not that surprising. However, for the stops one would not expect the consistent and high devoicing patterns reported here, thus marking an important difference to other Romance languages. The reason for this observed difference is not clear, so additional cross-linguistic comparisons and perceptual tests are necessary to find the underlying mechanisms.

We did not observe a consistent increase in devoicing with more posterior place of articulation, as would be expected [9, 10]. This is rather surprising, since the aerodynamic requirements and properties are similar across languages. Thus, for EP it could be the case that the high amount of devoicing is an important feature of this language, and thus overrides the expected higher voicing probabilities for bilabials and dentals. It has to be noted that there is still a solid and consistent difference between both voicing patterns and slopes when comparing voiceless consonants with its voiced counterparts. Thus, in EP the expected difference between voiceless and voiced consonant is maintained for all speakers and consonants. It is shown that the higher devoicing percentages and patterns for speaker3 can be explained by the influence of following high vowels. In contrast, for all other speakers there was no influence of context or position on the devoicing patterns.

5. ACKNOLEDGEMENTS

This work was partially supported by the Portuguese FCT (grant SFRH/BPD/ 48002/2008).

6. REFERENCES

[1] Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5(9/10)*, 341-345.

[2] Brunner J., Fuchs S. & Perrier P. (in press). Supralaryngeal control in Korean velar stops. *J Phon.*

[3] Caramazza, A. & Grace, Y. K. 1974. Voice onset time in two French dialects, *J Phon* 2, 239–245.

[4] Jesus, L. & Jackson, P. 2008. Frication and voicing classification, In: Teixeira et al. (Eds): *PROPOR 2008*, Springer: Heidelberg.

[5] Jesus, L. & Shadle, C. 2002. A parametric study of the spectral characteristics of fricatives, *J Phon* 30, 437-464.

[6] Jesus, L. & Shadle. C. 2003. Devoicing Measures of EP Fricatives, In: Mamede, Baptista, Trancoso, Nunes (eds.): *Comp. Proc. of the Portuguese Lang.*, Springer: Berlin.

[7] Lisker L. & Abramson A. (1964). A Cross-language study of voicing in initial stops. *Word*, 20, 384-422.

[8] Luce, P. & Charles-Luce, J.1985. Contextual effects on vowel duration, closure duration, and the consonant vowel ratio in speech production. *JASA* 78, 1949-1957.

[9] Ohala, J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.) *The production of speech*, 189–216. Springer: New York.

[10] Ohala, J. & Riordan, C. 1980. Passive vocal tract enlargement during voiced stops. *Rep. PhonLab Berkeley* 78-87

[11] Pape, D., Mooshammer, T., Hoole, P. & Fuchs, S. 2006. Devoicing of word-initial stops: A consequence of the following vowel? In Harrington, J. & Tabain, M. (eds). *Speech production: models, phonetic processes, and techniques*, Psychology Press: New York.

[12] Shih, C., Möbius, B. & Narasimhan, B. 1999. Contextual effects on consonantal voicing profiles: A cross-linguistic study. *Proc. 14th ICPhS*, 2, 989-992.

[13] Viana, M. 1984. Etude de deux aspects de consonantisme du portugais: fricatisation et devoisement. Ph.D. dissertation, LSHA Strasbourg, France.

[14] Westbury, J. & Keating, P. 1985. On the naturalness of stop consonant voicing. *J Ling* 22, 145-166.