

Analysis of stop consonant production in European Portuguese

Marisa Lobo Lousada¹ and Luis M. T. Jesus²

¹Escola Superior de Saúde da Universidade de Aveiro and Secção Autónoma de Ciências da Saúde, Universidade de Aveiro, Portugal

²Escola Superior de Saúde da Universidade de Aveiro and Instituto de Engenharia Electrónica e Telemática de Aveiro, Universidade de Aveiro, Portugal

<https://doi.org/10.36505/ExLing-2006/01/0038/000038>

Abstract

This paper investigates acoustic features correlated with voicing (VOT, stop duration, closure duration, release duration, voicing into closure duration, duration of preceding vowel and duration of following vowel) and with the place of articulation (spectral peaks) of stop consonants /p, b, t, d, k, g/. A corpus with these stops in initial, medial and final word position was recorded for six native speakers of European Portuguese.

Introduction

The present study examines the acoustic properties correlated with voicing and with place of articulation for European Portuguese stops.

Andrade (1980) compared VOT of homorganic stops, in initial position, before a vowel, in words produced by a speaker of European Portuguese. Results showed that some voiced stops had a period of prevoicing (120 ms to 130 ms) followed by a devoiced period (10 to 20 ms), and that VOT was larger for velars, than for labials and dentals, as in English (Klatt, 1975).

Stops are often devoiced (Alphen and Smits, 2004) and there multiple acoustic properties related with voicing distinction. Viana (1984) and Veloso (1995) observed that stop duration and duration of the preceding vowel, were acoustic properties that cued voicing in European. Fuchs (2005) also suggested closure duration, duration of following vowel, duration of preceding vowel, and voicing into closure duration, as voicing cues.

The shape of the spectrum of the stop release was analyzed by Blumstein and Stevens (1978). Labial stops had a diffuse-falling or diffuse-flat pattern, and alveolar stops, also had a diffuse spread of peaks of energy, but the amplitudes of these peaks were greater at high frequencies (diffuse-rising pattern). Velar stops had a mid-frequency spectral peak (compact pattern). Labial and alveolar stops shared the property of diffuseness and were distinguished by the shape of the spectral energy distribution.

Recording method

A corpus of fifty four Portuguese real words containing /p, b, t, d, k, g/ was recorded using a Philips SBC ME 400 unidirectional condenser microphone located 20 cm in front of the subject's mouth. A laryngograph signal (Lx) was also collected using a laryngograph processor (model EG-PC3 produced by Tiger DRS, Inc., USA). The acoustic and Lx signals were pre-amplified (Rane MS 1-b) and recorded with a Sony PCM-R300 DAT recorder, each with 16 bits and a sampling frequency of 48 kHz.

The corpus contained an equal number (eighteen) of words with stops in: initial position, followed by the vowels /a, i, u/; medial position, preceded by the vowels /a, i, u/ and followed by the vowel /ə/; final position, preceded by the vowels /ə, a/. The words were produced without any context and within the frame sentence "Diga,... por favor." by six native speakers of European Portuguese (three men and three women).

Analysis method

Temporal analysis

All corpus words were manually analyzed to detect the: beginning of the preceding vowel; end of preceding vowel and beginning of closure; voice offset; end of closure and beginning of release; beginning of prevoicing; end of the release and beginning of the following vowel; end of following vowel.

The following measurements were obtained: duration of preceding vowel, closure duration, voicing into closure duration, release duration, type of voicing (voiced, partially devoiced or voiceless), VOT, stop duration and duration of following vowel.

Spectral analysis

Multitaper spectra were calculated with 11ms windows left aligned to the release of the stop. We also calculated the frequency (F) at which the spectral amplitude was maximum, excluding the fundamental and its harmonics in voiced stops. It provided an endpoint for line fits used to determine the spectral slope. The average values for all Corpus 1 stops produced by speakers ML e LJ were: $F_{/p,b/} = 3,7$ kHz, $F_{/t,d/} = 3,9$ kHz and $F_{/k,g/} = 4,6$ kHz.

Results

Temporal analysis

Results of temporal analysis showed that when speakers ML and IM (female), LJ and HR (male) produced the words in a frame sentence, the stop duration, as shown in Figure 1, and the closure duration was longer for voiceless than for voiced stops in all word positions. The voicing into closure duration, the duration of preceding vowel and the duration of following vowel were generally shorter for voiceless than for voiced stops. VOT was generally shorter for bilabials than for dentals, and shorter for dentals than for velars except in final-word position.

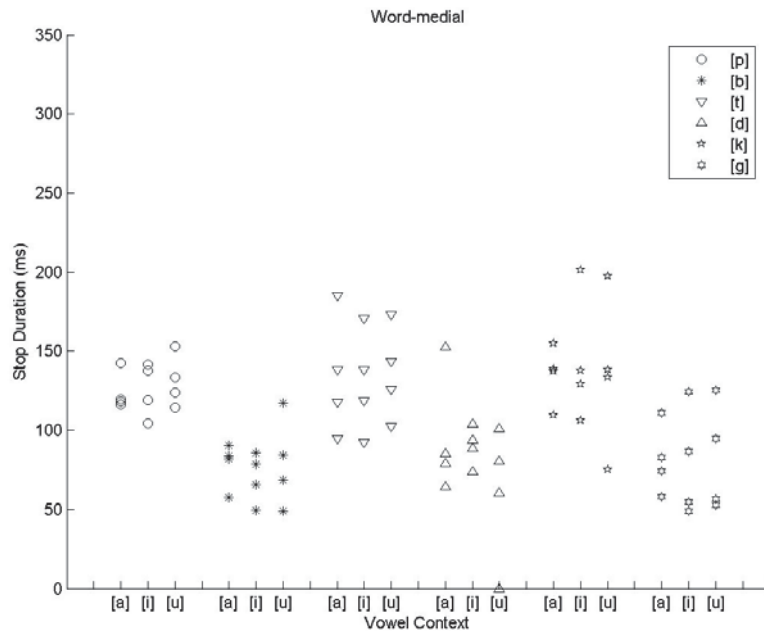


Figure 1. Stop duration for words in medial position produce by speakers ML, IM, LJ and HR.

Spectral analysis

[p] had spectral troughs at 0.8-4.6 kHz, and a broad peak at 1.4-5.6 kHz. [b] had spectral troughs at 0.7-5.0 kHz, and a broad peak at 1.5-5.6 kHz. [t] had spectral troughs at 1.7-7.0, a first peak at 0.3-3.7 kHz, a second peak at 0.2-4.6 kHz, and a broad peak at 6.0-10.4kHz. [d] had spectral troughs at 1.5-5.5 kHz, a first peak at 0.3-1.7kHz, a second peak at 2.4-5.6kHz, a first broad peak at 5.2-9.9 kHz, and a second broad peak at 11.0-12.8 kHz. [k] had

spectral troughs at 2.9-6.3 kHz, a first peak at 0.6-3.8 kHz, a second peak at 3.9-5.4 kHz, a first broad peak at 7.1-9.4 kHz, and a second broad peak at 10.0-13.2 kHz. [g] had a spectral troughs at 0.8-7.4 kHz, a first peak at 1.0-2.6 kHz, a second peak at 3.9-4.9 kHz, a first broad peak at 6.8-9.1, and a second broad peak at 12.0-13.6 kHz. Results of spectral analysis also showed that: [p, b] had a concentration of energy in the low frequencies (500 to 1500 Hz); [t, d] had flat spectrums or a concentration of energy in the high frequencies (above 4000 Hz); [k, g] had a concentration of energy in intermediate frequency regions (1500 to 4000 Hz).

Voiceless bilabial stops, in initial position, followed by vowel [a] had spectra with steeper negative slopes than dentals and velars. Spectra of velars followed by vowel [i] had positive slopes, bilabials were mostly flat and dentals had a negative slope. Dentals followed by vowel [u] had a positive, sometimes flat slope, and bilabials and velars had a negative slope. Velars in medial position had a less negative slope than bilabials and dentals. Word-final dentals had a less negative slope than bilabials and velars.

Conclusions

The results of stop duration agree with those presented by Viana (1984) and Veloso (1995). We observed the same correlation between place of articulation and VOT previously reported by Klatt (1975) and Andrade (1980). The results showed that different acoustic properties are important for voicing distinction in European Portuguese stops. We were not able to observe the spectral patterns reported by Blumstein and Stevens (1978).

References

- Alphen, P. and Smits, R. 2004. Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: the role of pre-voicing. *Journal of Phonetics* 32, 455-491.
- Andrade, A. 1980. Estudos experimentais aerodinâmicos, acústicos e palatográficos do vozeamento nas consoantes. CLUL, Lisboa, Portugal.
- Blumstein, S. and Stevens, K. 1979. Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, 66(4), 1001-1017.
- Fuchs, S. 2005. Articulatory correlates of the voicing contrast in alveolar obstruent production in German. *ZAS-Papers in Linguistics*, 41.
- Klatt, D. 1975. Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, 18(4), 686-706.
- Viana, M. 1984. *Étude de Deux Aspects du Consonantisme du Portugais: Fricatisation et Devoisement*, Doct., U. Sciences Humaines de Strasbourg, France.
- Veloso, J. 1995. Aspectos da percepção das “oclusivas fricativizadas” do Português: Contributo para a compreensão do processamento de contrastes alofônicos. Universidade do Porto, Portugal.