Mapping variation in European Portuguese: intonation, phrasing and rhythm

Pedro Oliveira[1], Pedro Palma[2], Nádia Barros[1], Bruno Neto[1], Marisa Cruz[1] & Sónia Frota[1]
[1]Centro de Linguística, University of Lisbon
[2]Centro de Estudos Geográfios, University of Lisbon

According to Goebl (2006, 2007) and Maguire & McMahon (2011), the choice of methods and models to use, how to quantify relations between dialects, or how many features are needed to define a dialectal area are key questions in the mapping of linguistic variation.

The present study uses data from the ongoing project *Interactive Atlas of the Prosody of Portuguese* (Frota (Coord.) 2012-2015) to address these questions, including different prosodic dimensions: intonation (Crespo-Sendra *et al.* 2014, Frota *et al.* 2015), phrasing (Elordieta *et al.* 2005, Cruz & Frota 2013, Barros in progress), and rhythm (Cruz & Frota 2014, Oliveira *et al.* 2014). The analysis covers 8 urban regions distributed along the territory of Portugal. ArcGIS for Desktop was used and proved to be an innovative system for mapping linguistic variation, since the software of GIS allows the computation of linguistic data combined with geographical and statistical information, providing an important contribution to the knowledge of areas of variation. Two procedures were used to map prosodic variation in European Portuguese (EP): (i) Huff model (for intonation and phrasing), testing the representation of nuclear contours (dominant and alternative) and the main phrasing pattern (Fig. 1); and (ii) Inverse Distance Weight method (IDW) for rhythm, using spatial interpolation methods for the representation of distinctions across varieties by combining two different quantitative rhythmic variables (%V and Δ%C).

Among the existing spatial interaction models, Huff was chosen for intonation and phrasing since it is one of the most used in the Geography field to analyze and predict patterns of spatial interaction (Cliquet 2006; Haynes & Fotheringham 1984) and to generate areas of influence/attractiveness of a given spatial point. As for rhythm, the IDW method was preferred because it is a robust method, thus a good estimator. In previous work (Oliveira *et al.* 2014), an index was manually computed to convert two dependent variables into a single one, using reference values for %V and Δ%C (Ramus *et al.* 1999, Frota & Vigário 2001), as well as their relative weight and impact on the rhythmic classification of each language/variety. However, in order to avoid a manual calculation of the index each time data are added to the model, a (automatic) solution was found thus reducing possible errors and improving data management. Thus, a Linear Regression Model was run to establish the equation behind rhythm, with rhythmic class as the dependent variable, and %V and Δ%C as predictors. The outputs of the statistical model provide the relative weight of each quantitative variable to the rhythmic classification. The resulting equation [$rhythm_i$=-1.99+(0.09*%V)+(-0.05*Δ%C)] was combined with the IDW method and the results are promising (mirroring the previous non-automatic findings), as shown in Fig.1. For the phrasing representation, a Huff model was used with an added constraint – accessibility, given by the principal road network – in order to test the possible effect of population mobility in the geographical representation of language variation (Fig. 2).

The current study was enriched with more data points to strengthen the geospatial analysis (Huff and IDW), with methods of automatic computation of complex variables and with population variables towards the major goal of achieving more accurate geographic representation of prosodic variation.

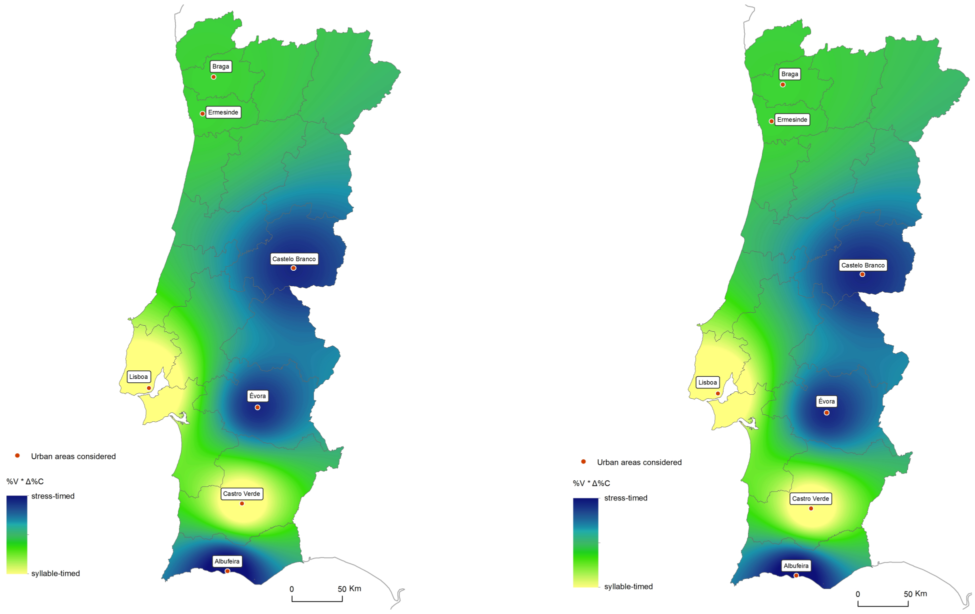

Figure 1. Rhythmic variation in EP, on the basis of two quantitative measures (%V and Δ%C) calculated for seven urban regions, labeled in the map. **Left side:** map obtained by means of a manually calculated index. **Right side:** map obtained by means of an automatic index, based on the equation behind rhythm.

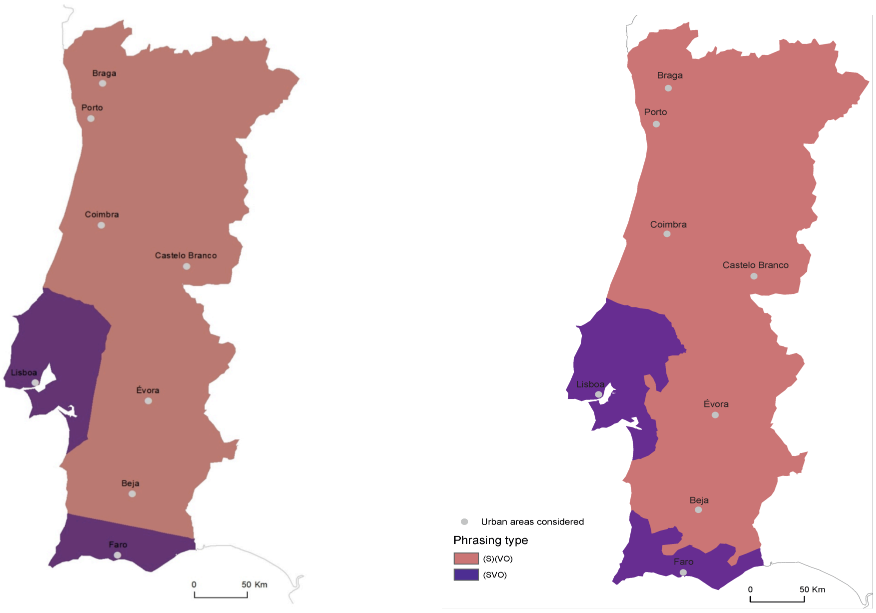

Figure 2. Phrasing pattern across EP varieties variation calculated for eight urban regions, labeled in the map. **Left side:** simple map, without geographical constraints. **Right side:** map including a geographical constraint – accessibility, given by the principal road network.

Selected references

Cliquet, G. 2006. Geomarketing: Methods and Strategies in Spatial Marketing. *GIS – Geographical Information Systems Series*. France: Wiley-ISTE.

Crespo-Sendra, V., M. Cruz, J. Silva & S. Frota. 2014. Asking questions across Portuguese varieties. Talk presented at the *6th International Conference on Tone and Intonation in Europe* (TIE). Netherlands: University of Utrecht.

Cruz, M. & S. Frota. 2014. Rhythm in central-southern varieties of European Portuguese: production and perception. In A. Moreno, F. Silva, I. Falé, I. Pereira & J. Veloso (org.). *Textos Selecionados do XXIX Encontro Nacional da Associação Portuguesa de Linguística*. Porto: APL, pp. 213-231.

Elordieta, G., S. Frota & M. Vigário. 2005. Subjects, objects and intonational phrasing in Spanish and Portuguese. *Studia Linguistica 59*(2-3), 110-143.

Frota, S., M. Cruz, F. Fernandes-Svartman, G. Collischonn, A. Fonseca, C. Serra, P. Oliveira & M. Vigário. 2015. Intonational variation in Portuguese: European and Brazilian varieties. In S. Frota & P. Prieto (eds). *Intonation in Romance*. Oxford: Oxford University Press, pp. 235-283.

Maguire, W. & A. McMahon. 2011. Quantifying relations between dialects. In W. Maguire & A. McMahon (eds.) *Analysing variation in English*. Cambridge: Cambridge University Press, pp. 93-120.

Oliveira, P., P. Palma, M. Cruz, B. Neto & S. Frota. 2014. Sistemas SIG e variação dialectal no Português Europeu: métodos geográficos inovadores no mapeamento da prosódia. Poster presented at the *EUE 2014 – 12º Encontro de Utilizadores Esri Portugal*, 21-22 May.