

Seminário CLUL 2017

09 de junho, 10h30, Anf IV

Phonological measures for the assessment of text complexity

Marina Vigário, Carla Pires (iMed.U LISBOA), Afonso Cavaco

Measuring linguistic complexity is useful for multiple purposes (e.g. assessing L1 and L2 development; evaluating or building written and oral material targeting specific audiences). While the topic has long attracted the attention of scholars, most languages still lack tools for calculating linguistic complexity, and in general phonological units and patterns are little explored. In this talk, we report on recent work proposing and motivating a number of phonologically- and statistically-based measures for linguistic complexity assessment, and apply these measures to samples of oral and written material with expected differences in complexity/cost, in order to determine the potential of these indicators to distinguish texts in terms of degree of complexity/cost. Besides directly reflecting phonological complexity/cost, some of the proposed measures also reflect complexity in other areas of grammar (morphology, syntax and semantics). The method is applied to European Portuguese (EP), but is extendable to other languages as well. The measures investigated take into consideration inherent properties of phonological units or patterns and/or their frequency in the language. It is assumed that less common units or patterns contribute to higher levels of difficulty/cost. Almost all measures proposed showed statistical differences among three groups of texts. The results indicate that simple and automatically obtained measures coming from the distribution of phonological units and patterns are able to distinguish texts of variable complexity in the expected direction.