

Conference

Panorama of Experimental Prosody Research

Plínio A. Barbosa

Speech Prosody Studies Group/Linguistics Dep./Instituto de Estudos da Linguagem/Unicamp & CNPq

pabarbosa.unicampbr@gmail.com

This lecture starts by defining and contextualising experimental speech prosody as regards theory (observation, hypothesis testing, modeling) and the hypothetico-deductive-based process of experimentation. Issues involving the relation between the subject, the experimenter and the elicited corpus within an experimental setting will be raised and discussed in the context of the alleged lab speech vs spontaneous speech dichotomy. We will argue for a gradient account of a corpus characterisation within a space formed by two dimensions, text genre and degree of elicitation control by the experimenter. The selection of linguistic and paralinguistic variables for statistical analysis will be illustrated with four examples presenting independent, dependent and control variables in prosodic research. Key concepts of experimental research on speech rhythm and intonation in the prosodic literature are presented and illustrated with recent experiments. The first experiment reviews a method for automatically detecting changes in readiness for action from speech of twelve participants on a radio show with differing degrees of signal-to-noise ratio. In this illustration, three classic statistical analyses were used: Principal Component Analysis (PCA) allowed the reduction of the dimensionality of the judges' responses for three affective primitives (activation, valence, and involvement) to two factors explaining 97 % of the variance of the responses, ANOVA and post-hoc analyses assessed the distinctiveness of readiness for action (factor 1) among the twelve participants, and multiple regression analysis allowed the inference of judges' responses from acoustics (three prosodic parameters, mean of rate of fundamental frequency change, mean and standard deviation of spectral emphasis, explain 67 % of the variance of readiness for action). This setting was also used for relating content derived from Discourse Segments Purpose theory to acoustic information derived from twelve prosodic descriptors. The second example presents a way for evaluating the degree of word prominence and boundary strength by a

procedure that minimises the listener variation in the response by allowing yes/no responses only and, then, associating the z statistical test for proportions to obtain different degrees for the two aforementioned prosodic functions. Produced and perceived saliences were then compared. The relation between prominence and boundary is also reported as being highly co-dependent for the listener, although the two functions can be associated to different domains (parts) within the word. The third illustration proposes a methodology for studying the relation of rhythmic and intonational structures for the utterances by correlating normalised duration contours with F0 traces. A script for superposing the two traces in Praat will be demonstrated. The last illustration presents a method for advancing knowledge in speech rhythm research by going beyond typological studies by focussing on the investigation of rhythmic differences between utterances. The reading and storytelling of six Brazilian Portuguese (BP) speakers is used to assess perceived differences in way of speaking (*modo de falar*, in Portuguese) by a group of ten listeners. Perceived differences are inferred from eleven possible rhythmic descriptors by using different models of multiple linear and non-linear regression analyses. Only four parameters were found significant for explaining up to 71 % of the total variance of the listeners' responses: speech rate (in syllables per second), duration-related salience rate (in normalised duration peaks per second), non-prominent syllable rate (which excludes silent pauses and prominent syllables), and mean of duration-related degree of prominence/strength of boundary. F0 peak rate was found non-significant as was descriptors related to variation of syllable-sized or stress-group-size duration. This means that, at least for BP, listeners rely more on rate of syllable-size and stress-group-size units to judge differences in way of speaking.