Prosody-based clustering of spoken corpus search results

Piotr Pęzik (University of Lodz, Poland)

The role of prosodic features of speech in the construction of meaning and discourse structure is an active area of linguistic research (Wichmann et al. 2009). The work described in this paper addresses the need for corpus-based tools for studying prosodic phenomena in conversational and read speech. More specifically, the paper presents an experimental prosody-based clustering functionality of the Spokes online search engine available for spoken Polish and English data (see: spokes.clarin-pl.eu, pelcra.clarin-pl.eu/SpokesBNC/ and pelcra.clarin-pl.eu/spokes2-web). The basic goal of this technique is to facilitate the analysis of concordance results according to selected features of the speech signal associated with word spans matching user queries. For example, given a simple query for the word *hello*, it is possible to automatically label pitch patterns detected for this surface form and thus distinguish between rising vs. falling pitch contours or relatively high- vs. relatively low-pitched instances of *hello* in the corpus. The assumption behind this approach is that even a simple sorting of search results by such labels may lead to the discovery of prosody-marked functions of words and multiword expressions.

Furthermore, normalized annotations of prosody can also be used as more complex feature vectors in order to cluster concordance result sets in a multidimensional space. The characteristics of speech considered in this experimental module of Spokes are a) duration of the matching search spans, b) pitch contours and c) voice intensity measurements. The normalization of pitch and intensity contours is performed using the MOMEL technique (Hirst & Espesser 1993) for simplifying contours and the INSINT scheme for encoding intonation patterns. The applicability of this new data exploration method to corpus-based studies of spoken language is evaluated on two use cases a) detecting potential prosodic distinctions between instances of ditropic expressions (cf. Van Lancker et al. 1981) and b) identifying functionally different but morphologically identical discourse and pragmatic markers.

References

Hirst, D. and Espesser, R. (1993). Automatic modeling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de phonétique d'Aix-en-Provence,* 15, 71-85.

Van Lancker, D., G. J. Canter, and D. Terbeek. 1981. Disambiguation of Ditropic Sentences: Acoustic and Phonetic Cues. *Journal of Speech and Hearing Research*, 24, no. 3, 330–35.

Wichmann, A., Dehé, N. and Barth-Weingarten D. 2009. Where Prosody Meets Pragmatics: Research at the Interface. In *Where Prosody Meets Pragmatics*, 1–20. Brill.