"997EMO" database design for affective speech research

Magdalena Oleśkowicz-Popiel, The Institute of Linguistics, Adam Mickiewicz University, Poznań

magda.jastrzebska@gmail.com

In studies on affect expressions a conflict between naturalness of speech and control over speech data acquisition is present. As pointed by Callejas&Cozar [1] the more control over the generated data the less spontaneity and naturalness of the expressed emotion and vice versa. There are three main approaches for collecting emotional speech samples: using actors for emotion portrayals, inducing emotions via various scenarios and procedures, recording natural speech reach in spontaneous affect expressions. Overviews of the existing database such as [2] show explicitly predominant number of acted and elicited emotion databases over spontaneous speech databases.

The present paper introduces acquisition and annotation of a "997EMO" database constructed for the needs of an ongoing project, which aims at establishing norms for dynamics of changes in expressive speech for Polish language in prosody and voice. To guarantee diversity of speech and quality of speech signal and to examine to what extent it is possible to control production of vocal emotion expressions, three types of speech material have been used in the database: spontaneous speech dialogs from 997 emergency call-center (Police), their imitations by professional actors and utterances with manipulated intonation contour. For multiple layer annotation of discreet and continuous features (including orthographic transcription, segmentation, context description, syntactic groups, dialog acts, emotion intensity and labels) Annotation System software has been employed [3].

Such a database design allows studying and comparing various speech parameters from both natural and artificial environment. Natural speech proves to be the most valuable data for measuring fundamental frequency and its derivatives (as the most prominent acoustic parameter that influences perception of emotions). Imitations of naturally occurring emotion expressions by actors allow evaluating ability to control expressiveness in human speech. Manipulations of intonation contour provide material for systematic study on the influence of pitch changes on perception of emotions.

Word count: 299

References:

[1]. Callejas, Z., López-Cózar, R., (2008). Influence of contextual information in emotion annotation for spoken dialogue systems. Speech Communication, vol. 50, pp. 416–433.

[2]. Ververidis D., Kotropoulos C. (2006). Emotional speech recognition: Resources, features, and methods. Speech Communication, 48 (9), pp. 1162-1181.

[3]. Klessa, K., Wagner, A., Oleśkowicz-Popiel, M. (2013) Using "Paralingua" database for investigation of affective states and paralinguistic features. Speech and Language Technology Vol. 16), 2013.