Annotating paralinguistic features of speech in the context of basic research and documentary linguistics

Katarzyna Klessa, Adam Mickiewicz University in Poznań, Poland

Kolloquium Linguistik, Bielefeld, 12.07.2017

Paralinguistic or non-linguistic features of speech continue to raise questions not only with regard to annotation tasks but also to the features' basic definitions and classifications. It is not always obvious where (or whether) to draw a line between paralinguistic and prosodic features. A range of approaches and practical solutions have been proposed and reported in the literature (see e.g., the numerous contributions to the *INTERSPEECH Computational Parlinguistics Challenge* series of events). As far as the annotation tasks are concerned, the questions that still persist include for example the choice of the feature specification and choice or the type of rating scales. Similar dilemmas may be faced in the case of perception-based experiments designed either for the needs of applications, including those strictly technological (e.g., naturally-sounding speech synthesis, automatic speaker and speech recognition), or research focused on psychology of interpersonal communication.

In this talk, I will present selected features of Annotation Pro software tool (annotationpro.org) and illustrate them with some of the most recent studies conducted with its use for the investigation of paralinguistic features.

Annotation Pro is a tool for speech annotation and annotation mining. It is extends the potential of a typical multi-layer annotation environment with a new component based on a graphical representation of feature space that supports annotation based on continuous rating scales. Apart from the annotation options, the program provides a simple perception experiment framework. The tool has so far been used within several projects, including analyses of emotions in speech (e.g., Oleśkowicz-Popiel & Bachan, 2017), inter-speaker convergence in the domains of speech and gesture (Karpiński et al., 2014; Czoska et al., 2015), or timing pattern variability (Klessa & Gibbon, 2014). Among others, current applications of the program comprise exploration of annotations of speech recordings in endangered and/or under-resourced languages (e.g., Beermann et al., 2015, Klessa et al., 2017).

The presented study examples will include data from chosen perception experiments conducted with the use of continuous and discrete rating scales, as well as preliminary observations from an on-going annotation mining work based on a multimodal Polish-German corpus of task-oriented dialogues (Karpiński & Klessa, submitted).

Selected references

Beermann D., Klessa K., Nyampong B.O. (2015). Using TypeCraft and Annotation Pro for multilayer annotation and analysis of Tense and Aspect in Krio, In Vetulani Z., Mariani J. (Eds.). *Proceedings of the 7th Language & Technology Conference*, Poznań.

Czoska, A., Klessa, K., Karpiński, M., Nowikow-Jarmołowicz, E. (2015). Prosody and gesture in dialogue: Cross-modal interactions. *Proceedings of 4th Gesture and Speech in Interaction (GESPIN) Conference*, Nantes, France, 83-88.

Karpiński, M., Klessa, K. (submitted). Methods, tools and techniques for multimodal analysis of accommodation in intercultural communication.

Karpiński, M., Klessa, K., Czoska, A. (2014). Local and global convergence in the temporal domain in Polish task-oriented dialogue, *Proceedings of the 7th Speech Prosody Conference*, 20-23 May 2014, Dublin, Ireland.

Klessa, K. & Gibbon, D. (2014). Annotation Pro + TGA: automation of speech timing analysis, *Proceedings of the 9th Language Resources and Evaluation Conference (LREC)*, Reykjavik, Iceland. ISBN 978-2-9517408-8-4, 1499-1505.

Klessa, K., Nau, N., Orlovs, O. (2017, in press). Timing patterns variability in Latgalian read speech. In van Dommelen, W. Koreman, J. (Eds.) *Nordic Prosody XII*. Peter Lang.

Klessa, K., Wagner, A., Oleśkowicz-Popiel, M. & Karpiński, M. (2013). 'Paralingua' – a new speech corpus for the studies of paralinguistic features. In Chelo Vargas-Sierra (Ed.). *Corpus Resources for Descriptive and Applied Studies. Current Challenges and Future Directions (CILC2013). Procedia – Social and Behavioral Science. Vol. 95*, 48-58. Elsevier, ISSN: 1877-0428.

Oleśkowicz-Popiel, M., Bachan, J. (2017). Manipulations of f0 contours in affective speech analysis. *28th Konferenz Elektronische Sprachsignalverarbeitung 2017*, Saarbrücken, 9-16.

Schuller, B., Steidl, S., Batliner, A., Hirschberg, J., Burgoon, J. K., Baird, A., ... & Evanini, K. (2016). The interspeech 2016 computational paralinguistics challenge: Deception, sincerity & native language. In *Proceedings of Interspeech*, San Francisco.

Acknowledgement

The creation of the Polish-German corpus of multimodal data has been created with the support of the Polish Ministry of Science and Higher Education within "The National Programme for the Development of Humanities" in the years 2014-2016 (for more see: http://borderland.amu.edu.pl/).