Editorial

Introduction to the Special Issue on Next Generation Computational Paralinguistics

Björn Schuller[1,2], Stefan Steidl[3], Anton Batliner[2,3], Alessandro Vinciarelli[4,5], Felix Burkhardt[6], Rob van Son[7,8]

[1] *Imperial College London, Department of Computing, England*
[2] *Technische Universität München, Machine Intelligence & Signal Processing Group, MMK, Germany*
[3] *Friedrich-Alexander-Universiät Erlangen-Nürnberg, Pattern Recognition Lab, Germany*
[4] *University of Glasgow, Scotland*
[5] *IDIAP, Switzerland*
[6] *Deutsche Telekom AG, Germany*
[7] *University of Amsterdam, The Netherlands*
[8] *Netherlands Cancer Institute, The Netherlands*
Corresponding Author: *Björn Schuller*
Imperial College London, Department of Computing, England
bjoern.schuller@imperial.ac.uk, telephone: +44-207-59-48357

Computational Paralinguistics recently reached a level of maturity allowing for first real-life applications in interaction, coaching, media retrieval, robotics, surveillance, and manifold further domains. In particular, an increasing level of realism is recently faced by coping with speaker independent analysis of highly naturalistic data in narrow-bandwidth, noisy, or reverberated conditions. At the same time, the richness of the range of speaker states and traits analysed computationally is increasingly widening up. This includes in particular also the degree of subjectivity faced with tasks such as perceived speaker personality, likability, or intelligibility, to name a few. Both these aspects require additional experience on the interplay of states and traits in speech, singing, and language. Further, with the integration in applications, novel aspects arise such as efficiency, reliability, self-learning, mobility, multi-cultural and multi-lingual aspects, handling groups of speaker or singers, standardisation, and user experience with such systems.

The INTERSPEECH 2012 Speaker Trait Challenge, which has been organised by the guest editors and Elmar Nöth, provided the first forum for comparison of results for perceived speaker traits under realistic conditions. In this special issue, we first summarise the findings from this Challenge: The introducing article *"A Survey on Perceived Speaker Traits: Personality, Likability, Pathology, and the First Challenge"* by the organisers of the INTERSPEECH 2012 Speaker State Challenge together with the further authors of the Challenge's introductory paper, namely Tobias Bocklet, Florian Eyben, Gelareh Mohammadi, Benjamin Weiss, and Felix Weninger, aims at providing a broad overview on the state of the art in perceived speaker trait recognition, as exemplified by speaker's perceived personality, likability, and intelligibility of pathological speakers. It further summarises the Challenge and provides deeper analysis of participants' results. This introductory article was handled in an independent review process by another editor. Then, two entries of the Challenge are described in more detail starting with the winning contribution of the Pathology Sub-Challenge. This is followed by contributions outside the Challenge touching upon Computational Paralinguistics in general by dealing with Parkinson's disease and its recognition from speech, and further dealing with recognition and manifestation of affect in speech and singing.

In the first article, *"Automatic intelligibility classification of sentence-level pathological speech"*, Jangwon Kim, Naveen Kumar, Andreas Tsiartas, Ming Li, and Shrikanth Narayanan propose novel sentence-level features for modelling atypicality, as well as a post-classification posterior smoothing scheme. Together with

feature-level fusion and subsystem decision fusion, these proved to be effective for the binary intelligibility classification of pathological speech.

Next, Jouni Pohjalainen, Okko Räsänen, and Serdar Kadioglu deal in *"Feature Selection Methods and Their Combinations in High-Dimensional Classification of Speaker Likability, Intelligibility and Personality Traits"* with feature selection and new supervised and unsupervised methods, using all three challenge corpora and the standard k-nearest-neighbors rule as classification algorithm. Best methods yield improved performance with fewer features.

Broadening from contributions dealing with the Challenge follows the article *"Fully Automated Assessment of the Severity of Parkinson's Disease from Speech"* by Alireza Bayestehtashk, Meysam Asgari, Izhak Shafran, and James McNames. For a corpus with speech from patients with Parkinson's disease, the authors refine feature extraction and demonstrate the effectiveness of different tasks for eliciting speech as well as a fully automated data collection and inference.

Houwei Cao, Ragini Verma, and Ani Nenkova improve in *"Speaker-sensitive Emotion Recognition via Ranking: Studies on Acted and Spontaneous Speech"* emotion recognition for acted and spontaneous speech databases by using ranking models and, especially for the spontaneous data, by combining ranking and standard classifiers.

With a fourth author, Houwei Cao, Arman Savran, Ragini Verma, and Ani Nenkova show in *"Acoustic and Lexical Representations for Affect Prediction in Spontaneous Conversations"* improved recognition of emotional dimensions when employing corpus-dependent bag-of-word representations and acoustic utterance-dependent regions of interest, using the AVEC 2012 challenge database.

As the only contribution dealing with singing rather than speaking, *"Comparing the acoustic expression of emotion in the speaking and the singing voice"* by Klaus R. Scherer, Johan Sundberg, Lucas Tamarit, and Gláucia L. Salomão closes this issue. For emotional singing, loudness, tempo, spectral balance and perturbation differentiate emotions most; in comparison to spoken, especially high arousal emotions (database with professional actors), singers tend to rely more on voice perturbation, especially vibrato.

Apart from the opening article by the guest editors, out of the many submissions received for this special issue, six were accepted divided into two Challenge-related participants and four general topics in Computational Paralinguistics. Four of the accepted papers underwent two revisions, the other three one.

1. Acknowledgement