

Nathalie Camelin · Yannick Estève
Carlos Martín-Vide (Eds.)

Statistical Language and Speech Processing

5th International Conference, SLSP 2017
Le Mans, France, October 23–25, 2017
Proceedings

Contents

Invited Paper

Author Profiling in Social Media: The Impact of Emotions on Discourse Analysis	3
<i>Paolo Rosso and Francisco Rangel</i>	

Language and Information Extraction

Neural Machine Translation by Generating Multiple Linguistic Factors	21
<i>Mercedes García-Martínez, Loïc Barrault, and Fethi Bougares</i>	
Analysis and Automatic Classification of Some Discourse Particles on a Large Set of French Spoken Corpora	32
<i>Denis Jouvet, Katarína Bartkova, Mathilde Dargnat, and Lou Lee</i>	
Learning Morphology of Natural Language as a Finite-State Grammar.	44
<i>Javad Nouri and Roman Yangarber</i>	
Incorporating Coreference to Automatic Evaluation of Coherence in Essays	58
<i>Michal Novák, Kateřina Rysová, Magdaléna Rysová, and Jiří Mirovský</i>	
Graph-Based Features for Automatic Online Abuse Detection.	70
<i>Etienne Papegnies, Vincent Labatut, Richard Dufour, and Georges Linarès</i>	
Exploring Temporal Analysis of Tweet Content from Cultural Events	82
<i>Mathias Quillot, Cassandre Ollivier, Richard Dufour, and Vincent Labatut</i>	
Towards a Relation-Based Argument Extraction Model for Argumentation Mining	94
<i>Gil Rocha and Henrique Lopes Cardoso</i>	

Post-processing and Applications of Automatic Transcriptions

Three Experiments on the Application of Automatic Speech Recognition in Industrial Environments	109
<i>Ferdinand Fuhrmann, Anna Maly, Christina Leitner, and Franz Graf</i>	
Enriching Confusion Networks for Post-processing	119
<i>Sahar Ghannay, Yannick Estève, and Nathalie Camelin</i>	

Attentional Parallel RNNs for Generating Punctuation in Transcribed Speech	131
<i>Alp Öktem, Mireia Farrús, and Leo Wanner</i>	
Lightweight Spoken Utterance Classification with CFG, tf-idf and Dynamic Programming	143
<i>Manny Rayner, Nikos Tsourakis, and Johanna Gerlach</i>	
Low Latency MaxEnt- and RNN-Based Word Sequence Models for Punctuation Restoration of Closed Caption Data	155
<i>Máté Ákos Tündik, Balázs Tarján, and György Szaszák</i>	

Speech: Paralinguistics and Synthesis

Unsupervised Speech Unit Discovery Using K-means and Neural Networks	169
<i>Céline Manenti, Thomas Pellegrini, and Julien Pinquier</i>	
Noise and Speech Estimation as Auxiliary Tasks for Robust Speech Recognition	181
<i>Gueorgui Pironkov, Stéphane Dupont, Sean U.N. Wood, and Thierry Dutoit</i>	
Unified Approach to Development of ASR Systems for East Slavic Languages	193
<i>Radek Safarik and Jan Nouza</i>	
A Regularization Post Layer: An Additional Way How to Make Deep Neural Networks Robust	204
<i>Jan Vaněk, Jan Zelinka, Daniel Soutner, and Josef Psutka</i>	

Speech Recognition: Modeling and Resources

Detecting Stuttering Events in Transcripts of Children's Speech	217
<i>Sadeen Alharbi, Madina Hasan, Anthony J.H. Simons, Shelagh Brumfitt, and Phil Green</i>	
Introducing AmuS: The Amused Speech Database	229
<i>Kevin El Haddad, Ilaria Torre, Emer Gilmartin, Hüseyin Çakmak, Stéphane Dupont, Thierry Dutoit, and Nick Campbell</i>	
Lexical Emphasis Detection in Spoken French Using F-BANKs and Neural Networks	241
<i>Abdelwahab Heba, Thomas Pellegrini, Tom Jorquera, Régine André-Obrecht, and Jean-Pierre Lorré</i>	

Speaker Change Detection Using Binary Key Modelling with Contextual Information	250
<i>Jose Patino, Héctor Delgado, and Nicholas Evans</i>	
Perception of Expressivity in TTS: Linguistics, Phonetics or Prosody?	262
<i>Marie Tahon, Gwénolé Lecorvé, Damien Lolive, and Raheel Qader</i>	
Author Index	275