

16th Annual Conference of the International Speech Communication Association (INTERSPEECH 2015)

Speech Beyond Speech Towards a Better
Understanding of the Most Important Biosignal

Dresden, Germany
6 – 10 September 2015

Volume 1 of 5

ISBN: 978-1-5108-1790-6

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2015) by the International Speech Communications Association
All rights reserved.

Printed by Curran Associates, Inc. (2016)

For permission requests, please contact the International Speech Communications Association
at the address below.

International Speech Communications Association
c/o Mme Emmanuelle FOXONET
4 Rue des Fauvettes - Lous Tourils
F-66390 Baixas, France

Phone: 49 228 735 639
Fax: 33 468 385 827

secretariat@isca-speech.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Oral Session 1: Feature Extraction and Modeling with Neural Networks

Chairs: Peter Bell, Bhuvana Ramabhadran

Room: Large Hall, Time 11:00 - 13:00, Monday, September 7, 2015

- PAGE 1
Mon-O-1-1
11:00 - 11:20
- Learning the Speech Front-End with Raw Waveform CLDNNs**
Tara N. Sainath, Ron J. Weiss, Andrew Senior, Kevin W. Wilson, Oriol Vinyals, Google, USA
- PAGE 6
Mon-O-1-2
11:20 - 11:40
- Architectures for Deep Neural Network Based Acoustic Models Defined Over Windowed Speech Waveforms**
Mayank Bhargava, Richard Rose, McGill University, Canada
- PAGE 11
Mon-O-1-3
11:40 - 12:00
- Analysis of CNN-Based Speech Recognition System Using Raw Speech as Input**
Dimitri Palaz¹, Mathew Magimai-Doss¹, Ronan Collobert²
¹Idiap Research Institute, Switzerland; ²Facebook, USA
- PAGE 16
Mon-O-1-4
12:00 - 12:20
- Bilinear Map of Filter-Bank Outputs for DNN-Based Speech Recognition**
Tetsuji Ogawa¹, Kenshiro Ueda¹, Kouichi Katsurada², Tetsunori Kobayashi¹, Tsuneo Nitta¹
¹Waseda University, Japan; ²Toyohashi University of Technology, Japan
- PAGE 21
Mon-O-1-5
12:20 - 12:40
- Speech Recognition with Temporal Neural Networks**
Payton Lin¹, Dau-Cheng Lyu², Yun-Fan Chang¹, Yu Tsao¹
¹Academia Sinica, Taiwan; ²ASUS, Taiwan
- PAGE 26
Mon-O-1-6
12:40 - 13:00
- Convolutional Neural Networks for Acoustic Modeling of Raw Time Signal in LVCSR**
Pavel Golik, Zoltán Tüske, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

Oral Session 2: Prosody

Chairs: Štefan Beňuš, Katalin Mády

Room: Conference 1, Time 11:00 - 13:00, Monday, September 7, 2015

- PAGE 31
Mon-O-2-1
11:00 - 11:20
- Stable and Unstable Intervals as a Basic Segmentation Procedure of the Speech Signal**
Ulrike Glavitsch¹, Lei He², Volker Dellwo²
¹EMPA, Switzerland; ²Universität Zürich, Switzerland
- PAGE 36
Mon-O-2-2
11:20 - 11:40
- Polysyllabic Shortening and Word-Final Lengthening in English**
Andreas Windmann¹, Juraj Šimko², Petra Wagner¹
¹Universität Bielefeld, Germany; ²University of Helsinki, Finland
- PAGE 41
Mon-O-2-3
11:40 - 12:00
- The Acoustics of Word Stress in English as a Function of Stress Level and Speaking Style**
Anders Eriksson, Mattias Heldner, Stockholm University, Sweden
- PAGE 46
Mon-O-2-4
12:00 - 12:20
- Pitch Accent Distribution in German Infant-Directed Speech**
Katharina Zahner, Muna Pohl, Bettina Braun, Universität Konstanz, Germany
- PAGE 51
Mon-O-2-5
12:20 - 12:40
- Acoustic Correlates of Perceived Syllable Prominence in German**
Hansjörg Mixdorff¹, Christian Cossio-Mercado², Angelika Hönemann³, Jorge Gurlekian², Diego Evin², Humberto Torres²
¹BHT Berlin, Germany; ²Universidad de Buenos Aires, Argentina; ³Universität Bielefeld, Germany
- PAGE 56
Mon-O-2-6
12:40 - 13:00
- Cross-Modality Matching of Linguistic and Emotional Prosody**
Simone Simonetti, Jeusun Kim, Chris Davis, University of Western Sydney, Australia

Oral Session 3: Speech Intelligibility Enhancement

Chairs: Paavo Alku, Jon Barker

Room: Conference 2+3, Time 11:00 – 13:00, Monday, September 7, 2015

- PAGE 60
Mon-O-3-1
11:00 – 11:20
- A Fast Algorithm for Improved Intelligibility of Speech-in-Noise Based on Frequency and Time Domain Energy Reallocation**
Tudor-Cătălin Zorilă, Yannis Stylianou, Toshiba Research Europe, UK
- PAGE 65
Mon-O-3-2
11:20 – 11:40
- Intelligibility Enhancement of Casual Speech for Reverberant Environments Inspired by Clear Speech Properties**
Maria Koutsogiannaki¹, Petko N. Petkov², Yannis Stylianou¹
¹University of Crete, Greece; ²Toshiba Research Europe, UK
- PAGE 70
Mon-O-3-3
11:40 – 12:00
- Intelligibility Enhancement of Vocal Announcements for Public Address Systems: A Design for All Through a Presbycusis Pre-Compensation Filter**
A. Ben Jemaa¹, N. Mechergui¹, G. Courtois², A. Mudry³, S. Djaziri-Larbi⁴, M. Turki⁴, H. Lissek², M. Jaidane⁴
¹Université de Tunis El Manar, Tunisia; ²EPFL, Switzerland; ³Stanford University, USA; ⁴Telnet, Tunisia
- PAGE 75
Mon-O-3-4
12:00 – 12:20
- Model-Based Integration of Reverberation for Noise-Adaptive Near-End Listening Enhancement**
Henning Schepker¹, David Hülsmeier², Jan RENNIES², Simon Doclo¹
¹Carl von Ossietzky Universität Oldenburg, Germany; ²Fraunhofer IDMT, Germany
- PAGE 80
Mon-O-3-5
12:20 – 12:40
- Online Lombard-Adaptation in Incremental Speech Synthesis**
Sebastian Rottschäfer, Hendrik Buschmeier, Herwin van Welbergen, Stefan Kopp, Universität Bielefeld, Germany
- PAGE 85
Mon-O-3-6
12:40 – 13:00
- Comparison of Gaussian Process Regression and Gaussian Mixture Models in Spectral Tilt Modelling for Intelligibility Enhancement of Telephone Speech**
Emma Jokinen, Ulpu Remes, Paavo Alku, Aalto University, Finland

Oral Session 4: Detecting and Predicting Mental and Social Disorders

Chairs: Julien Epps, Elizabeth Shriberg

Room: Conference 4+5, Time 11:00 – 13:00, Monday, September 7, 2015

- PAGE 90
Mon-O-4-1
11:00 – 11:20
- A Discriminative Reliability-Aware Classification Model with Applications to Intelligibility Classification in Pathological Speech**
Naveen Kumar, Shrikanth S. Narayanan, University of Southern California, USA
- PAGE 95
Mon-O-4-2
11:20 – 11:40
- Voiced/Unvoiced Transitions in Speech as a Potential Bio-Marker to Detect Parkinson's Disease**
J.R. Orozco-Arroyave¹, Florian Hönig², J.D. Arias-Londoño¹, J.F. Vargas-Bonilla¹, Sabine Skodda³, J. Ruzs⁴, Elmar Nöth²
¹Universidad de Antioquia, Colombia; ²FAU Erlangen-Nürnberg, Germany; ³Ruhr-Universität Bochum, Germany; ⁴Czech Technical University in Prague, Czech Republic
- PAGE 100
Mon-O-4-3
11:40 – 12:00
- Low-Frequency Components Analysis in Running Speech for the Automatic Detection of Parkinson's Disease**
T. Villa-Cañas¹, J.D. Arias-Londoño¹, J.R. Orozco-Arroyave¹, J.F. Vargas-Bonilla¹, Elmar Nöth²
¹Universidad de Antioquia, Colombia; ²FAU Erlangen-Nürnberg, Germany
- PAGE 105
Mon-O-4-4
12:00 – 12:20
- Automatic Detection of Parkinson's Disease from Continuous Speech Recorded in Non-Controlled Noise Conditions**
J.C. Vásquez-Correa¹, T. Arias-Vergara¹, J.R. Orozco-Arroyave¹, J.F. Vargas-Bonilla¹, J.D. Arias-Londoño¹, Elmar Nöth²
¹Universidad de Antioquia, Colombia; ²FAU Erlangen-Nürnberg, Germany

Mon-O-4 continued...

PAGE 110
Mon-O-4-5
12:20 - 12:40

Relevance Vector Machine for Depression Prediction

Nicholas Cummins¹, Vidhyasaharan Sethu¹, Julien Epps¹, Jarek Krajewski²
¹University of New South Wales, Australia; ²Universität Wuppertal, Germany

PAGE 115
Mon-O-4-6
12:40 - 13:00

Typicality and Emotion in the Voice of Children with Autism Spectrum Condition: Evidence Across Three Languages

Erik Marchi¹, Björn Schuller², Simon Baron-Cohen³, Ofer Golan⁴, Sven Bölte⁵, Prerna Arora⁶, Reinhold Häb-Umbach⁶
¹Technische Universität München, Germany; ²Imperial College London, UK; ³University of Cambridge, UK; ⁴Bar-Ilan University, Israel; ⁵Karolinska Institute, Sweden; ⁶Universität Paderborn, Germany

Oral Session 5: Spoken Language Understanding

Chairs: David Suendermann-Oeft, Hugo Van hamme

Room: Conference 6, Time 11:00 - 13:00, Monday, September 7, 2015

PAGE 120
Mon-O-5-1
11:00 - 11:20

Deep Contextual Language Understanding in Spoken Dialogue Systems

Chunxi Liu¹, Puyang Xu², Ruhi Sarikaya²
¹Johns Hopkins University, USA; ²Microsoft, USA

PAGE 125
Mon-O-5-2
11:20 - 11:40

RNN-Based Labeled Data Generation for Spoken Language Understanding

Yik-Cheung Tam, Yangyang Shi, Hunk Chen, Mei-Yuh Hwang, Microsoft, China

PAGE 130
Mon-O-5-3
11:40 - 12:00

Is it Time to Switch to Word Embedding and Recurrent Neural Networks for Spoken Language Understanding?

Vedran Vukotic¹, Christian Raymond¹, Guillaume Gravier²
¹INSA de Rennes, France; ²INRIA, France

PAGE 135
Mon-O-5-4
12:00 - 12:20

Recurrent Neural Network and LSTM Models for Lexical Utterance Classification

Suman Ravuri¹, Andreas Stolcke²
¹ICSI, USA; ²Microsoft, USA

PAGE 140
Mon-O-5-5
12:20 - 12:40

Semantic Retrieval of Personal Photos Using a Deep Autoencoder Fusing Visual Features with Speech Annotations Represented as Word/Paragraph Vectors

Hung-tsung Lu, Yuan-ming Liou, Hung-yi Lee, Lin-shan Lee, National Taiwan University, Taiwan

PAGE 145
Mon-O-5-6
12:40 - 13:00

A Comparison of Normalization Techniques Applied to Latent Space Representations for Speech Analytics

Mohamed Morchid, Richard Dufour, Driss Matrouf, LIA, France

Special Session 1: Active Perception in Human and Machine Speech Communication

Chairs: Martin Heckmann, Dorothea Kolossa

Room: Hall 1, Time 11:00 – 13:00, Monday, September 7, 2015

- PAGE 150
Mon-SP1-1
11:00 – 11:17
- The Effect of Soft, Modal and Loud Voice Levels on Entrainment in Noisy Conditions**
Éva Székely, Mark T. Keane, Julie Carson-Berndsen, University College Dublin, Ireland
- PAGE 155
Mon-SP1-2
11:17 – 11:34
- Does Voice Anthropomorphism Affect Lexical Alignment in Speech-Based Human-Computer Dialogue?**
Benjamin R. Cowan, Holly P. Branigan, University of Edinburgh, UK
- PAGE 160
Mon-SP1-3
11:34 – 11:51
- Exploiting Top-Down Source Models to Improve Binaural Localisation of Multiple Sources in Reverberant Environments**
Ning Ma, Guy J. Brown, Jose A. Gonzalez, University of Sheffield, UK
- PAGE 165
Mon-SP1-4
11:51 – 12:08
- Binaural Sound Source Localisation and Tracking Using a Dynamic Spherical Head Model**
Christopher Schymura¹, Fiete Winter², Dorothea Kolossa¹, Sascha Spors²
¹Ruhr-Universität Bochum, Germany; ²Universität Rostock, Germany
- PAGE 170
Mon-SP1-5
12:08 – 12:25
- The Role of Temporal Resolution in Modulation-Based Speech Segregation**
Tobias May, Thomas Bentsen, Torsten Dau, Technical University of Denmark, Denmark
- PAGE 175
Mon-SP1-6
12:25 – 12:42
- Improving Automatic Speech Recognition in Spatially-Aware Hearing Aids**
Hendrik Kayser, Constantin Spille, Daniel Marquardt, Bernd T. Meyer, Carl von Ossietzky Universität Oldenburg, Germany

Mon-SP1 continued...

- PAGE 180
Mon-SP1-7
12:42 – 12:59
- Dereverberation for Active Human-Robot Communication Robust to Speaker's Face Orientation**
Randy Gomez, Levko Ivanchuk, Keisuke Nakamura, Takeshi Mizumoto, Kazuhiro Nakadai, Honda Research Institute Japan, Japan

Poster Session 1: Speaker Recognition

Chair: Sandro Cumani

Room: Hall 2, Time 11:00 – 13:00, Monday, September 7, 2015

- PAGE 185
Mon-P-1-1 **Multi-Task Learning for Text-Dependent Speaker Verification**
Nanxin Chen, Yanmin Qian, Kai Yu, Shanghai Jiao Tong University, China
- PAGE 190
Mon-P-1-2 **JFA for Speaker Recognition with Random Digit Strings**
Themos Stafylakis¹, Patrick Kenny¹, Md. Jahangir Alam¹, Marcel Kockmann²
¹CRIM, Canada; ²VoiceTrust, Canada
- PAGE 195
Mon-P-1-3 **Structured Prediction for Speaker Identification in TV Series**
Elena Knyazeva, Guillaume Wisniewski, Hervé Bredin, François Yvon, LIMSI, France
- PAGE 200
Mon-P-1-4 **Speaker Recognition by Means of Acoustic and Phonetically Informed GMMs**
Sandro Cumani, Pietro Laface, Farzana Kulsoom, Politecnico di Torino, Italy
- PAGE 205
Mon-P-1-5 **A Fast Approach to Psychoacoustic Model Compensation for Robust Speaker Recognition in Additive Noise**
Ashish Panda, TCS Innovation Labs Mumbai, India
- PAGE 210
Mon-P-1-6 **Blind Score Normalization Method for PLDA Based Speaker Recognition**
Danila Doroshin, Nikolay Lubimov, Marina Nastasenko, Mikhail Kotov, Stel Computer Systems, Russia
- PAGE 214
Mon-P-1-7 **Non-Linear PLDA for i-Vector Speaker Verification**
Sergey Novoselov, Timur Pekhovsky, Oleg Kudashev, Valentin S. Mendeleev, Alexey Prudnikov, Speech Technology Center, Russia
- PAGE 219
Mon-P-1-8 **On the Need of Template Protection for Voice Authentication**
Carlos Vaquero, Patricia Rodríguez, AGNITiO, Spain

Mon-P-1 continued...

- PAGE 224
Mon-P-1-9 **Evaluation and Calibration of Short-Term Aging Effects in Speaker Verification**
Finnian Kelly, John H.L. Hansen, University of Texas at Dallas, USA
- PAGE 229
Mon-P-1-10 **Phone-Centric Local Variability Vector for Text-Constrained Speaker Verification**
Liping Chen¹, Kong Aik Lee², Bin Ma², Wu Guo¹, Haizhou Li², Li-Rong Dai¹
*¹USTC, China; ²A*STAR, Singapore*
- PAGE 234
Mon-P-1-11 **Cosine Distance Features for Robust Speaker Verification**
Kuruvachan K. George¹, C. Santhosh Kumar¹, Ramachandran K.I.¹, Ashish Panda²
¹Amrita University, India; ²TCS Innovation Labs Mumbai, India
- PAGE 239
Mon-P-1-12 **Voice Liveness Detection Algorithms Based on Pop Noise Caused by Human Breath for Automatic Speaker Verification**
Sayaka Shiota¹, Fernando Villavicencio², Junichi Yamagishi², Nobutaka Ono², Isao Echizen², Tomoko Matsui³
¹Tokyo Metropolitan University, Japan; ²NII, Japan; ³ISM, Japan
- PAGE 244
Mon-P-1-13 **Noise Robust Speaker Recognition with Convolutional Sparse Coding**
Antti Hurmalainen¹, Rahim Saeidi², Tuomas Virtanen¹
¹Tampere University of Technology, Finland; ²Aalto University, Finland
- PAGE 249
Mon-P-1-14 **Combining Amplitude and Phase-Based Features for Speaker Verification with Short Duration Utterances**
Md. Jahangir Alam, Patrick Kenny, Themos Stafylakis, CRIM, Canada

Poster Session 2: Speech Synthesis 1

Chair: Inma Hernaez

Room: Hall 2, Time 11:00 – 13:00, Monday, September 7, 2015

- PAGE 254
Mon-P-2-1 **Phase Perception of the Glottal Excitation of Vocoded Speech**
Tuomo Raitio¹, Lauri Juvela¹, Antti Suni², Martti Vainio², Paavo Alku¹
¹Aalto University, Finland; ²University of Helsinki, Finland
- PAGE 259
Mon-P-2-2 **Using Acoustics to Improve Pronunciation for Synthesis of Low Resource Languages**
Sunayana Sitaram, Serena Jeblee, Alan W. Black, Carnegie Mellon University, USA
- PAGE 264
Mon-P-2-3 **Sub-Band Text-to-Speech Combining Sample-Based Spectrum with Statistically Generated Spectrum**
Tadashi Inai¹, Sunao Hara¹, Masanobu Abe¹, Yusuke Ijima², Noboru Miyazaki², Hideyuki Mizuno²
¹Okayama University, Japan; ²NTT Corporation, Japan
- PAGE 269
Mon-P-2-4 **Pruning Redundant Synthesis Units Based on Static and Delta Unit Appearance Frequency**
Heng Lu¹, Wei Zhang¹, Xu Shao¹, Quan Zhou², Wenhui Lei², Hongbin Zhou², Andrew Breen³
¹Nuance Communications, USA; ²Nuance Communications, China; ³Nuance Communications, UK
- PAGE 274
Mon-P-2-5 **Emotional Transplant in Statistical Speech Synthesis Based on Emotion Additive Model**
Yamato Ohtani, Yu Nasu, Masahiro Morita, Masami Akamine, Toshiba Corporation, Japan

Mon-P-2 continued...

- PAGE 279
Mon-P-2-6 **Generalized Variable Parameter HMMs Based Acoustic-to-Articulatory Inversion**
Xurong Xie, Xunying Liu, Lan Wang, Rongfeng Su, Chinese Academy of Sciences, China
- PAGE 284
Mon-P-2-7 **Semi-Supervised Training of a Voice Conversion Mapping Function Using a Joint-Autoencoder**
Seyed Hamidreza Mohammadi, Alexander Kain, Oregon Health & Science University, USA
- PAGE 289
Mon-P-2-8 **On Glottal Source Shape Parameter Transformation Using a Novel Deterministic and Stochastic Speech Analysis and Synthesis System**
Stefan Huber, Axel Roebel, IRCAM, France
- PAGE 294
Mon-P-2-9 **Fluent Personalized Speech Synthesis with Prosodic Word-Level Spontaneous Speech Generation**
Yi-Chin Huang, Chung-Hsien Wu, Ming-Ge Shie, National Cheng-Kung University, Taiwan
- PAGE 299
Mon-P-2-10 **Non-Native Speech Synthesis Preserving Speaker Individuality Based on Partial Correction of Prosodic and Phonetic Characteristics**
Yuji Oshima, Shinnosuke Takamichi, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura, NAIST, Japan
- PAGE 304
Mon-P-2-11 **Evaluation of State Mapping Based Foreign Accent Conversion**
Markus Toman, Michael Pucher, FTW, Austria
- PAGE 309
Mon-P-2-12 **Minimum Trajectory Error Training for Deep Neural Networks, Combined with Stacked Bottleneck Features**
Zhizheng Wu, Simon King, University of Edinburgh, UK

Poster Session 3: Mining and Annotation of Spoken and Multimodal Resources

Chair: Murat Saraçlar

Room: Foyer, Time 11:00 - 13:00, Monday, September 7, 2015

- PAGE 314
Mon-P-3-1 **Anomaly-Based Annotation Errors Detection in TTS Corpora**
Jindřich Matoušek, Daniel Tihelka, University of West Bohemia, Czech Republic
- PAGE 319
Mon-P-3-2 **Analysing Automatic Descriptions of Intonation with ICARUS**
Katrin Schweitzer, Markus Gärtner, Arndt Riestler, Ina Rösiger, Kerstin Eckart, Jonas Kuhn, Grzegorz Dogil, Universität Stuttgart, Germany
- PAGE 324
Mon-P-3-3 **iCALL Corpus: Mandarin Chinese Spoken by Non-Native Speakers of European Descent**
*Nancy F. Chen, Rong Tong, Darren Wee, Peixuan Lee, Bin Ma, Haizhou Li, A*STAR, Singapore*
- PAGE 329
Mon-P-3-4 **Development of a Cantonese Dysarthric Speech Corpus**
Ka Ho Wong¹, Yu Ting Yeung¹, Edwin H.Y. Chan¹, Patrick C.M. Wong¹, Gina-Anne Levow², Helen Meng¹
¹Chinese University of Hong Kong, China; ²University of Washington, USA
- PAGE 334
Mon-P-3-5 **StyleX: A Corpus of Educational Videos for Research on Speaking Styles and Their Impact on Engagement and Learning**
Harish Arsikere, Sonal Patil, Ranjeet Kumar, Kundan Shrivastava, Om Deshmukh, Xerox Research Center India, India

Mon-P-3 continued...

- PAGE 339
Mon-P-3-6 **A Dialog Act Tagging Approach to Behavioral Coding: A Case Study of Addiction Counseling Conversations**
Doğan Can¹, David C. Atkins², Shrikanth S. Narayanan¹
¹University of Southern California, USA; ²University of Washington, USA
- PAGE 344
Mon-P-3-7 **Analysing Rhythm in Ritual Discourse in Yucatec Maya Using Automatic Speech Alignment**
Valentina Vapnarsky¹, Claude Barras², Cédric Becquey¹, David Doukhan³, Martine Adda-Decker², Lori Lamel²
¹LESC (UMR 7186), France; ²LIMSI, France; ³IRCAM, France
- PAGE 349
Mon-P-3-8 **Noise-Matched Training of CRF Based Sentence End Detection Models**
Madina Hasan, Rama Doddipatla, Thomas Hain, University of Sheffield, UK
- PAGE 354
Mon-P-3-9 **The Effect of Spectral Slope on Pitch Perception**
Jianjing Kuang, Mark Liberman, University of Pennsylvania, USA

Special Event: Speaker Comparison for Forensic and Investigative Applications

Chairs: Jean-François Bonastre, Joseph P. Campbell

Room: Large Hall, Time 14:30 - 16:30, Monday, September 7, 2015

Oral Session 6: Speech Production Data and Models

Chairs: Jianwu Dang, Pascal Perrier

Room: Conference 1, Time 14:30 – 16:30, Monday, September 7, 2015

- PAGE 359
Mon-O-6-1
14:30 – 14:50
- Combined Cine- and Tagged-MRI for Tracking Landmarks on the Tongue Surface**
Honghao Bao¹, Wenhuan Lu¹, Kiyoshi Honda¹, Jianguo Wei¹, Qiang Fang², Jianwu Dang¹
¹Tianjin University, China; ²Chinese Academy of Social Sciences, China
- PAGE 364
Mon-O-6-2
14:50 – 15:10
- Human Vocal Tract Growth: A Longitudinal Study of the Development of Various Anatomical Structures**
Guillaume Barbier¹, Louis-Jean Boë¹, Guillaume Captier², Rafael Laboissière³
¹GIPSA, France; ²Université de Montpellier, France; ³LPNC, France
- PAGE 369
Mon-O-6-3
15:10 – 15:30
- Analysis of Coarticulated Speech Using Estimated Articulatory Trajectories**
Ganesh Sivaraman¹, Vikramjit Mitra², Mark K. Tiede³, Elliot Saltzman⁴, Louis Goldstein⁵, Carol Espy-Wilson¹
¹University of Maryland, USA; ²SRI International, USA; ³Haskins Laboratories, USA; ⁴Boston University, USA; ⁵University of Southern California, USA
- PAGE 374
Mon-O-6-4
15:30 – 15:50
- Speech Planning in 4-Year-Old Children versus Adults: Acoustic and Articulatory Analyses**
Guillaume Barbier¹, Pascal Perrier¹, Lucie Ménard², Yohan Payan³, Mark K. Tiede⁴, Joseph S. Perkell⁵
¹GIPSA, France; ²Université du Québec à Montréal, Canada; ³TIMC-IMAG, France; ⁴Haskins Laboratories, USA; ⁵MIT, USA
- PAGE 379
Mon-O-6-5
15:50 – 16:10
- Morphological and Acoustic Analysis of the Vocal Tract Using a Multi-Speaker Volumetric MRI Dataset**
Tokihiko Kaburagi, Kyushu University, Japan

Mon-O-6 continued...

- PAGE 384
Mon-O-6-6
16:10 – 16:30
- Experimental Assessment of the Tongue Incompressibility Hypothesis During Speech Production**
Zisis Iason Skordilis¹, Vikram Ramanarayanan², Louis Goldstein¹, Shrikanth S. Narayanan¹
¹University of Southern California, USA; ²Educational Testing Service, USA

Oral Session 7: Deep Neural Networks in Language and Accent Recognition

Chairs: Ville Hautamäki, Michael Wagner

Room: Conference 2+3, Time 14:30 - 16:30, Monday, September 7, 2015

- PAGE 389
Mon-O-7-1
14:30 - 14:50
- Multilingual Bottleneck Features for Language Recognition**
Radek Fér, Pavel Matějka, František Grézl, Oldřich Plchot, Jan Černocký, Brno University of Technology, Czech Republic
- PAGE 394
Mon-O-7-2
14:50 - 15:10
- DNN Senone MAP Multinomial i-Vectors for Phonotactic Language Recognition**
Alan McCree, Daniel Garcia-Romero, Johns Hopkins University, USA
- PAGE 398
Mon-O-7-3
15:10 - 15:30
- Deep Bottleneck Network Based I-Vector Representation for Language Identification**
Yan Song, Xinhai Hong, Bing Jiang, Ruilian Cui, Ian McLoughlin, Li-Rong Dai, USTC, China
- PAGE 403
Mon-O-7-4
15:30 - 15:50
- An End-to-End Approach to Language Identification in Short Utterances Using Convolutional Neural Networks**
Alicia Lozano-Diez, Ruben Zazo-Candil, Javier Gonzalez-Dominguez, Doroteo T. Toledano, Joaquin Gonzalez-Rodriguez, Universidad Autónoma de Madrid, Spain
- PAGE 408
Mon-O-7-5
15:50 - 16:10
- Boosting Universal Speech Attributes Classification with Deep Neural Network for Foreign Accent Characterization**
Ville Hautamäki¹, Sabato Marco Siniscalchi², Hamid Behravan¹, Valerio Mario Salerno², Ivan Kukanov¹
¹University of Eastern Finland, Finland; ²Università di Enna Kore, Italy
- PAGE 413
Mon-O-7-6
16:10 - 16:30
- Multilingual Tandem Bottleneck Feature for Language Identification**
Wang Geng, Jie Li, Shanshan Zhang, Xinyuan Cai, Bo Xu, Chinese Academy of Sciences, China

Oral Session 8: Speech Transmission

Chairs: Tom Bäckström, Gernot Kubin

Room: Conference 4+5, Time 14:30 - 16:30, Monday, September 7, 2015

- PAGE 418
Mon-O-8-1
14:30 - 14:50
- On Compressibility of Neural Network Phonological Features for Low Bit Rate Speech Coding**
Afsaneh Asaei, Milos Cernak, Hervé Bourlard, Idiap Research Institute, Switzerland
- PAGE 423
Mon-O-8-2
14:50 - 15:10
- Robust and Accurate LSF Location with Laguerre Method**
Michał Lenarczyk, Polish Academy of Sciences, Poland
- PAGE 428
Mon-O-8-3
15:10 - 15:30
- Interactivity-Aware Payout Adaptation**
Jochen Issing¹, Nikolaus Färber², Reinhard German¹
¹FAU Erlangen-Nürnberg, Germany; ²Fraunhofer IIS, Germany
- PAGE 433
Mon-O-8-4
15:30 - 15:50
- Advanced Time Shrinking Using a Drop Classifier Based on Codec Features**
Jochen Issing¹, Nikolaus Färber², Reinhard German¹
¹FAU Erlangen-Nürnberg, Germany; ²Fraunhofer IIS, Germany
- PAGE 438
Mon-O-8-5
15:50 - 16:10
- Measuring and Monitoring Speech Quality for Voice Over IP with POLQA, ViSQOL and P.563**
Andrew Hines¹, Eoin Gillen², Naomi Harte²
¹Dublin Institute of Technology, Ireland; ²Trinity College Dublin, Ireland
- PAGE 443
Mon-O-8-6
16:10 - 16:30
- Towards the Prediction of Human Speaker Identification Performance from Measured Speech Quality**
Laura Fernández Gallardo¹, Sebastian Möller²
¹University of Canberra, Australia; ²T-Labs, Germany

Oral Session 9: Language Modeling for Conversational Speech

Chairs: Michael Levit, Hsin Min Wang

Room: Conference 6, Time 14:30 – 16:30, Monday, September 7, 2015

- PAGE 448
Mon-O-9-1
14:30 – 14:50
- Personalization of Word-Phrase-Entity Language Models**
M. Levit, Andreas Stolcke, R. Subba, S. Parthasarathy, S. Chang, S. Xie, T. Anastasakos, Benoit Dumoulin, Microsoft, USA
- PAGE 453
Mon-O-9-2
14:50 – 15:10
- Discriminative Bilinear Language Modeling for Broadcast Transcriptions**
Akio Kobayashi, Manon Ichiki, Takahiro Oku, Kazuo Onoe, Shoei Sato, NHK, Japan
- PAGE 458
Mon-O-9-3
15:10 – 15:30
- Recognize Foreign Low-Frequency Words with Similar Pairs**
Xi Ma, Xiaoxi Wang, Dong Wang, Zhiyong Zhang, Tsinghua University, China
- PAGE 463
Mon-O-9-4
15:30 – 15:50
- Combinations of Various Language Model Technologies Including Data Expansion and Adaptation in Spontaneous Speech Recognition**
Ryo Masumura¹, Taichi Asami¹, Takanobu Oba¹, Hirokazu Masataki¹, Sumitaka Sakauchi¹, Akinori Ito²
¹NTT Corporation, Japan; ²Tohoku University, Japan
- PAGE 468
Mon-O-9-5
15:50 – 16:10
- Bringing Contextual Information to Google Speech Recognition**
Petar Aleksic, Mohammadreza Ghodsi, Assaf Michaely, Cyril Allauzen, Keith Hall, Brian Roark, David Rybach, Pedro Moreno, Google, USA
- PAGE 473
Mon-O-9-6
16:10 – 16:30
- Sequence-Based Class Tagging for Robust Transcription in ASR**
Lucy Vasserman, Vlad Schogol, Keith Hall, Google, USA

Special Session 2a: Interspeech 2015 Computational Paralinguistics Challenge (ComParE): Degree of Nativeness, Parkinson's & Eating Condition

Chairs: Anton Batliner, Stefan Steidl

Room: Hall 1, Time 14:30 – 16:30, Monday, September 7, 2015

- PAGE 478
Mon-SP2a-1
14:30 – 14:40
- The INTERSPEECH 2015 Computational Paralinguistics Challenge: Nativeness, Parkinson's & Eating Condition**
Björn Schuller¹, Stefan Steidl², Anton Batliner², Simone Hantke³, Florian Höning², J.R. Orozco-Arroyave², Elmar Nöth², Yue Zhang³, Felix Weninger³
¹Imperial College London, UK; ²FAU Erlangen-Nürnberg, Germany; ³Technische Universität München, Germany
- Mon-SP2a-1
14:40 – 14:50
- The Degree of Nativeness Sub-Challenge: The Data**
Florian Höning, FAU Erlangen-Nürnberg, Germany
- PAGE 483
Mon-SP2a-2
14:50 – 15:00
- Phrase Accentuation Verification and Phonetic Variation Measurement for the Degree of Nativeness Sub-Challenge**
Claude Montacié, Marie-José Caraty, STIH (EA 4509), France
- PAGE 488
Mon-SP2a-3
15:00 – 15:10
- Combining Multiple Approaches to Predict the Degree of Nativeness**
Eugénio Ribeiro¹, Jaime Ferreira¹, Julia Olcoz², Alberto Abad¹, Helena Moniz¹, Fernando Batista¹, Isabel Trancoso¹
¹INESC-ID Lisboa, Portugal; ²Universidad de Zaragoza, Spain

Mon-SP2a continued...

PAGE 493
Mon-SP2a-4
15:10 - 15:20

Automated Evaluation of Non-Native English Pronunciation Quality: Combining Knowledge- and Data-Driven Features at Multiple Time Scales

Matthew P. Black, Daniel Bone, Zisis Iason Skordilis, Rahul Gupta, Wei Xia, Pavlos Papadopoulos, Sandeep Nallan Chakravarthula, Bo Xiao, Maarten Van Segbroeck, Jangwon Kim, Panayiotis G. Georgiou, Shrikanth S. Narayanan, University of Southern California, USA

Mon-SP2a-i2
15:20 - 15:30

The Parkinson's Condition Sub-Challenge: The Data

J.R. Orozco-Arroyave, Universidad de Antioquia, Colombia

PAGE 498
Mon-SP2a-5
15:30 - 15:40

Estimating the Severity of Parkinson's Disease from Speech Using Linear Regression and Database Partitioning

Dávid Sztahó, Gábor Kiss, Klára Vicsi, BME, Hungary

PAGE 503
Mon-SP2a-6
15:40 - 15:50

Random Forest-Based Prediction of Parkinson's Disease Progression Using Acoustic, ASR and Intelligibility Features

Alexander Zlotnik¹, Juan M. Montero¹, Rubén San-Segundo¹, Ascensión Gallardo-Antolín²

¹Universidad Politécnica de Madrid, Spain; ²Universidad Carlos III de Madrid, Spain

PAGE 508
Mon-SP2a-7
15:50 - 16:00

Automatic Recognition of Unified Parkinson's Disease Rating from Speech with Acoustic, i-Vector and Phonotactic Features

Guozhen An¹, David Guy Brizan¹, Min Ma¹, Michelle Morales¹, Ali Raza Syed¹, Andrew Rosenberg²

¹CUNY Graduate Center, USA; ²CUNY Queens College, USA

PAGE 513
Mon-SP2a-8
16:00 - 16:10

Parkinson's Condition Estimation Using Speech Acoustic and Inversely Mapped Articulatory Data

Seongjun Hahm, Jun Wang, University of Texas at Dallas, USA

Mon-SP2a continued...

PAGE 518
Mon-SP2a-9
16:10 - 16:20

Segment-Dependent Dynamics in Predicting Parkinson's Disease

James R. Williamson¹, Thomas F. Quatieri¹, Brian S. Helfer¹, Joseph Perricone¹, Satrajit S. Ghosh², Gregory Ciccarelli¹, Daryush D. Mehta¹

¹MIT Lincoln Laboratory, USA; ²MIT, USA

Mon-SP2a-D
16:20 - 16:30

Discussion

Poster Session 4: Pronunciation, Prosody and Audiovisual Features and Models

Chair: Anton Ragni

Room: Hall 2, Time 14:30 - 16:30, Monday, September 7, 2015

- PAGE 523
Mon-P-4-1 **Recognition of Voiced Sounds with a Continuous State HMM**
S.M. Houghton, Colin J. Champion, Philip Weber, University of Birmingham, UK
- PAGE 528
Mon-P-4-2 **Learning Speech Rate in Speech Recognition**
Xiangyu Zeng, Shi Yin, Dong Wang, Tsinghua University, China
- PAGE 533
Mon-P-4-3 **Pronunciation and Silence Probability Modeling for ASR**
Guoguo Chen, Hainan Xu, Minhua Wu, Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, USA
- PAGE 538
Mon-P-4-4 **Exploring Minimal Pronunciation Modeling for Low Resource Languages**
Marelle Davel¹, Etienne Barnard¹, Charl van Heerden¹, William Hartmann², Damianos Karakos², Richard Schwartz², Stavros Tsakalidis²
¹North-West University, South Africa; ²Raytheon BBN Technologies, USA
- PAGE 543
Mon-P-4-5 **Attribute Knowledge Integration for Speech Recognition Based on Multi-Task Learning Neural Networks**
Hao Zheng¹, Zhanlei Yang¹, Liwei Qiao², Jianping Li², Wenju Liu¹
¹Chinese Academy of Sciences, China; ²SGCC, China
- PAGE 548
Mon-P-4-6 **Detecting Audio-Visual Synchrony Using Deep Neural Networks**
Etienne Marcheret¹, Gerasimos Potamianos², Josef Vopicka³, Vaibhava Goel¹
¹IBM T.J. Watson Research Center, USA; ²University of Thessaly, Greece; ³IBM, Czech Republic

Mon-P-4 continued...

- PAGE 553
Mon-P-4-7 **Cross Database Training of Audio-Visual Hidden Markov Models for Phone Recognition**
Shahram Kalantari, David Dean, Houman Ghaemmaghami, Sridha Sridharan, Clinton Fookes, Queensland University of Technology, Australia
- PAGE 558
Mon-P-4-8 **Incorporating Visual Information for Spoken Term Detection**
Shahram Kalantari, David Dean, Sridha Sridharan, Queensland University of Technology, Australia
- PAGE 563
Mon-P-4-9 **Integration of Deep Bottleneck Features for Audio-Visual Speech Recognition**
Hiroshi Ninomiya¹, Norihide Kitaoka², Satoshi Tamura³, Yurie Iribe⁴, Kazuya Takeda¹
¹Nagoya University, Japan; ²University of Tokushima, Japan; ³Gifu University, Japan; ⁴Aichi Prefectural University, Japan
- PAGE 568
Mon-P-4-10 **Automatic Detection of Sentence Prominence in Speech Using Predictability of Word-Level Acoustic Features**
Sofoklis Kakouros, Okko Räsänen, Aalto University, Finland
- PAGE 573
Mon-P-4-11 **An Empirical Model of Emphatic Word Detection**
Milos Cernak, Pierre-Edouard Honnet, Idiap Research Institute, Switzerland
- PAGE 578
Mon-P-4-12 **Using Tilt for Automatic Emphasis Detection with Bayesian Networks**
Yishuang Ning¹, Zhiyong Wu¹, Xiaoyan Lou², Helen Meng¹, Jia Jia¹, Lianhong Cai¹
¹Tsinghua University, China; ²Beijing Samsung Telecom R&D Center, China

Poster Session 5: Speech Analysis and Representation

Chair: Jen-Tzung Chien

Room: Hall 2, Time 14:30 – 16:30, Monday, September 7, 2015

- PAGE 583
Mon-P-5-1 **Analysis of a Low-Dimensional Bottleneck Neural Network Representation of Speech for Modelling Speech Dynamics**
Linxue Bai, Peter Jančovič, Martin Russell, Philip Weber, University of Birmingham, UK
- PAGE 588
Mon-P-5-2 **Statistical Acoustic-to-Articulatory Mapping Unified with Speaker Normalization Based on Voice Conversion**
Hidetsugu Uchida, Daisuke Saito, Nobuaki Minematsu, Keikichi Hirose, University of Tokyo, Japan
- PAGE 593
Mon-P-5-3 **Analysis of Features from Analytic Representation of Speech Using MP-ABX Measures**
Raghavendra Reddy Pappagari, Karthika Vijayan, K. Sri Rama Murty, IIT Hyderabad, India
- PAGE 598
Mon-P-5-4 **Source-Filter Separation of Speech Signal in the Phase Domain**
Erfan Loweimi, Jon Barker, Thomas Hain, University of Sheffield, UK
- PAGE 603
Mon-P-5-5 **A Maximum Likelihood Approach to the Detection of Moments of Maximum Excitation and Its Application to High-Quality Speech Parameterization**
Ranniera Maia¹, Yannis Stylianou¹, Masami Akamine²
¹Toshiba Research Europe, UK; ²Toshiba Corporation, Japan
- PAGE 608
Mon-P-5-6 **SABR: Sparse, Anchor-Based Representation of the Speech Signal**
Christopher Liberatore, Sandesh Aryal, Zelun Wang, Seth Polsley, Ricardo Gutierrez-Osuna, Texas A&M University, USA

Mon-P-5 continued...

- PAGE 613
Mon-P-5-7 **Automatic Transformation of Irregular to Regular Voice by Residual Analysis and Synthesis**
Tamás Gábor Csapó, Géza Németh, BME, Hungary
- PAGE 618
Mon-P-5-8 **Optical Sensor Calibration for Electro-Optical Stomatography**
Simon Preuß, Peter Birkholz, Technische Universität Dresden, Germany
- PAGE 623
Mon-P-5-9 **From Text to Formants — Indirect Model for Trajectory Prediction Based on a Multi-Speaker Parallel Speech Database**
Kálmán Abari¹, Tamás Gábor Csapó², Bálint Pál Tóth², Gábor Olaszy²
¹University of Debrecen, Hungary; ²BME, Hungary
- PAGE 628
Mon-P-5-10 **Layered Nonnegative Matrix Factorization for Speech Separation**
Chung-Chien Hsu, Jen-Tzung Chien, Tai-Shih Chi, National Chiao Tung University, Taiwan
- PAGE 633
Mon-P-5-11 **Robust Tongue Tracking in Ultrasound Images: A Multi-Hypothesis Approach**
Catherine Laporte¹, Lucie Ménard²
¹École de Technologie Supérieure, Canada; ²Université du Québec à Montréal, Canada
- PAGE 638
Mon-P-5-12 **Objective Measures for Predicting the Intelligibility of Spectrally Smoothed Speech with Artificial Excitation**
Danny Websdale, Thomas Le Cornu, Ben Milner, University of East Anglia, UK

Poster Session 6: Speech Recognition — Technologies and Systems for New Applications

Chair: Daniel Willett

Room: Foyer, Time 14:30 – 16:30, Monday, September 7, 2015

- PAGE 643
Mon-P-6-1 **Mispronunciation Detection Without Nonnative Training Data**
Ann Lee, James Glass, MIT, USA
- PAGE 648
Mon-P-6-2 **Automatic Accentedness Evaluation of Non-Native Speech Using Phonetic and Sub-Phonetic Posterior Probabilities**
Ramya Rasipuram, Milos Cernak, Alexandre Nachen, Mathew Magimai-Doss, Idiap Research Institute, Switzerland
- PAGE 653
Mon-P-6-3 **Using F0 Contours to Assess Nativeness in a Sentence Repeat Task**
Min Ma¹, Keelan Evanini², Anastassia Loukina², Xinhao Wang², Klaus Zechner²
¹CUNY Graduate Center, USA; ²Educational Testing Service, USA
- PAGE 658
Mon-P-6-4 **Using Linguistic Indicators of Difficulty to Identify Mild Cognitive Impairment**
Rebecca Lunsford, Peter A. Heeman, Oregon Health & Science University, USA
- PAGE 663
Mon-P-6-5 **Automatic Intelligibility Measures Applied to Speech Signals Simulating Age-Related Hearing Loss**
Lionel Fontan¹, Jérôme Farinas¹, Isabelle Ferrané¹, Julien Pinquier¹, Xavier Aumont²
¹IRIT, France; ²LETRA (ANR-14-LAB4-0003-01), France
- PAGE 668
Mon-P-6-6 **Assessing Empathy Using Static and Dynamic Behavior Models Based on Therapist's Language in Addiction Counseling**
Sandeep Nallan Chakravarthula¹, Bo Xiao¹, Zac E. Imel², David C. Atkins³, Panayiotis G. Georgiou¹
¹University of Southern California, USA; ²University of Utah, USA; ³University of Washington, USA

Mon-P-6 continued...

- PAGE 673
Mon-P-6-7 **SVitchboard II and FiSVer I: High-Quality Limited-Complexity Corpora of Conversational English Speech**
Yuzong Liu, Rishabh Iyer, Katrin Kirchhoff, Jeff Bilmes, University of Washington, USA
- PAGE 678
Mon-P-6-8 **Fully Unsupervised Small-Vocabulary Speech Recognition Using a Segmental Bayesian Model**
Herman Kamper¹, Aren Jansen², Sharon Goldwater¹
¹University of Edinburgh, UK; ²Johns Hopkins University, USA
- PAGE 683
Mon-P-6-9 **LSTM for Punctuation Restoration in Speech Transcripts**
Ottokar Tilk, Tanel Alumäe, Tallinn University of Technology, Estonia
- PAGE 688
Mon-P-6-10 **Noise Robust Exemplar Matching for Speech Enhancement: Applications to Automatic Speech Recognition**
Emre Yilmaz, Deepak Baby, Hugo Van hamme, Katholieke Universiteit Leuven, Belgium
- PAGE 693
Mon-P-6-11 **A Study on Robust Detection of Pronunciation Erroneous Tendency Based on Deep Neural Network**
Yingming Gao, Yanlu Xie, Wen Cao, Jinsong Zhang, BLCU, China
- PAGE 697
Mon-P-6-12 **Vowel Mispronunciation Detection Using DNN Acoustic Models with Cross-Lingual Training**
Shrikant Joshi, Nachiket Deo, Preeti Rao, IIT Bombay, India
- PAGE 702
Mon-P-6-13 **Confidence-Features and Confidence-Scores for ASR Applications in Arbitration and DNN Speaker Adaptation**
Kshitiz Kumar, Ziad Al Bawab, Yong Zhao, Chaojun Liu, Benoit Dumoulin, Yifan Gong, Microsoft, USA

Mon-P-6 continued...

PAGE 707
Mon-P-6-14

Topic Modeling for Conference Analytics
Pengfei Liu¹, Shoaib Jameel¹, Wai Lam¹, Bin Ma², Helen Meng¹
*¹Chinese University of Hong Kong, China; ²A*STAR, Singapore*

PAGE 712
Mon-P-6-15

Sparse Coding Based Features for Speech Units Classification
Pulkit Sharma, Vinayak Abrol, A.D. Dileep, Anil Kumar Sao, IIT Mandi, India

Show and Tell Session 1

Chair: Siegfried Kunzmann

Room: Seminar 1, Time 14:30 - 16:30, Monday, September 7, 2015

PAGE 716
Show&Tell-1-1

Smarter Driving with IDA, the Intelligent Driving Assistant for Singapore
*Andreea I. Niculescu, Ngoc Thuy Huong Thai, Chongjia Ni, Boon Pang Lim, Kheng Hui Yeo, Rafael E. Banchs, A*STAR, Singapore*

PAGE 718
Show&Tell-1-2

Talk it Out: Adding Speech Interaction to Support Informational and Transactional Applications on Public Touch-Screen Kiosks
*Kheng Hui Yeo, Rafael E. Banchs, A*STAR, Singapore*

PAGE 720
Show&Tell-1-3

Conversational Agent and Management Tools for Conference and Tourism Domain
*Luis Fernando D'Haro, Seokhwan Kim, Rafael E. Banchs, A*STAR, Singapore*

PAGE 722
Show&Tell-1-4

Latvian Speech-to-Text Transcription Service
Askars Salimbajevs, Jevgenijs Strigins, Tilde, Latvia

PAGE 724
Show&Tell-1-5

System Supporting Speaker Identification in Emergency Call Center
Jakub Gałka, Joanna Grzybowska, Magdalena Igras, Paweł Jaciów, Kamil Wajda, Marcin Witkowski, Mariusz Ziółko, AGH UST, Poland

PAGE 726
Show&Tell-1-6

QAT² — The QCRI Advanced Transcription and Translation System
Ahmed Abdelali, Ahmed Ali, Francisco Guzmán, Felix Stahlberg, Stephan Vogel, Yifan Zhang, Hamad Bin Khalifa University, Qatar

PAGE 728
Show&Tell-1-7

Implementation of a Live Dialectal Media Subtitling System
Michael Stadtschnitzer, Christoph Schmidt, Fraunhofer IAIS, Germany

Show&Tell-1 continued...

PAGE 730
Show&Tell-1-8

A System for Automatic Broadcast News Summarisation, Geolocation and Translation
Peter Bell, Catherine Lai, Clare Llewellyn, Alexandra Birch, Mark Sinclair, University of Edinburgh, UK

PAGE 732
Show&Tell-1-9

Media Monitoring System for Latvian Radio and TV Broadcasts
Artūrs Znotiņš¹, Kaspars Polis², Roberts Darģis¹
¹University of Latvia, Latvia; ²LETA, Latvia

PAGE 734
Show&Tell-1-10

Meeting Assistant Application
Michel Assayag¹, Jonathan Huang², Jonathan Mamou¹, Oren Pereg¹, Saurav Sahay², Oren Shamir¹, Georg Stemmer³, Moshe Wasserblat¹
¹Intel, Israel; ²Intel, USA; ³Intel, Germany

Oral Session 10: Distant and Reverberant Speech Recognition

Chairs: John H. L. Hansen, Rita Singh

Room: Large Hall, Time 17:00 - 19:00, Monday, September 7, 2015

- PAGE 736
Mon-O-10-1
17:00 - 17:20
- Bayesian Integration of Sound Source Separation and Speech Recognition: A New Approach to Simultaneous Speech Recognition**
Kousuke Itakura, Izaya Nishimuta, Yoshiaki Bando, Katsutoshi Itoyama, Kazuyoshi Yoshii, Kyoto University, Japan
- PAGE 741
Mon-O-10-2
17:20 - 17:40
- Channel Selection in the Short-Time Modulation Domain for Distant Speech Recognition**
Ivan Himawan¹, Petr Motlicek¹, Sridha Sridharan², David Dean², Dian Tjondronegoro²
¹Idiap Research Institute, Switzerland; ²Queensland University of Technology, Australia
- PAGE 746
Mon-O-10-3
17:40 - 18:00
- A Multi-Channel Speech Enhancement Framework for Robust NMF-Based Speech Recognition for Speech-Impaired Users**
Gert Dekkers, Toon van Waterschoot, Bart Vanrumste, Bert Van Den Broeck, Jort F. Gemmeke, Hugo Van hamme, Peter Karsmakers, Katholieke Universiteit Leuven, Belgium
- PAGE 751
Mon-O-10-4
18:00 - 18:20
- Sound Source Separation Algorithm Using Phase Difference and Angle Distribution Modeling Near the Target**
Chanwoo Kim, Kean K. Chin, Google, USA
- PAGE 756
Mon-O-10-5
18:20 - 18:40
- Contaminated Speech Training Methods for Robust DNN-HMM Distant Speech Recognition**
Mirco Ravanelli, Maurizio Omologo, FBK, Italy
- PAGE 761
Mon-O-10-6
18:40 - 19:00
- Distance-Aware DNNs for Robust Speech Recognition**
Yajie Miao, Florian Metze, Carnegie Mellon University, USA
-
- INTERSPEECH 2015 Table of Contents
-

Oral Session 11: Speech Analysis and Representation 1

Chairs: Thomas Drugman, Panayiotis Georgiou

Room: Conference 1, Time 17:00 - 19:00, Monday, September 7, 2015

- PAGE 766
Mon-O-11-1
17:00 - 17:20
- Vocal Tremor Analysis via AM-FM Decomposition of Empirical Modes of the Glottal Cycle Length Time Series**
Christophe Mertens¹, Francis Grenez¹, François Viallet², Alain Ghio³, Sabine Skodda⁴, Jean Schoentgen¹
¹Université Libre de Bruxelles, Belgium; ²Centre Hospitalier du Pays d'Aix, France; ³LPL, France; ⁴Ruhr-Universität Bochum, Germany
- PAGE 771
Mon-O-11-2
17:20 - 17:40
- Estimating Lower Vocal Tract Features with Closed-Open Phase Spectral Analyses**
Elizabeth Godoy, Nicolas Malyska, Thomas F. Quatieri, MIT Lincoln Laboratory, USA
- PAGE 776
Mon-O-11-3
17:40 - 18:00
- Inductive Implementation of Segmental HMMs as CS-HMMs**
S.M. Houghton, Colin J. Champion, University of Birmingham, UK
- PAGE 781
Mon-O-11-4
18:00 - 18:20
- A Discriminative Analysis Within and Across Voiced and Unvoiced Consonants in Neutral and Whispered Speech in Multiple Indian Languages**
G. Nisha Meenakshi, Prasanta Kumar Ghosh, Indian Institute of Science, India
- PAGE 786
Mon-O-11-5
18:20 - 18:40
- Aligning Meeting Recordings via Adaptive Fingerprinting**
T.J. Tsai¹, Andreas Stolcke²
¹University of California at Berkeley, USA; ²Microsoft, USA
- PAGE 791
Mon-O-11-6
18:40 - 19:00
- On Representation Learning for Artificial Bandwidth Extension**
Matthias Zöhrer¹, Robert Peharz², Franz Pernkopf¹
¹Technische Universität Graz, Austria; ²Medizinische Universität Graz, Austria

Oral Session 12: L2 Speech Perception and Production

Chairs: Tarayuki Arai, Catherine Best

Room: Conference 2+3, Time 17:00 - 19:00, Monday, September 7, 2015

- PAGE 796
Mon-O-12-1
17:00 - 17:20
- Perception and Production of Vowel Contrasts in German Learners of English**
Helena Levy, Universität Freiburg, Germany
- PAGE 801
Mon-O-12-2
17:20 - 17:40
- Goodness of Tone (GOT) for Non-Native Mandarin Tone Recognition**
*Rong Tong, Nancy F. Chen, Bin Ma, Haizhou Li, A*STAR, Singapore*
- PAGE 806
Mon-O-12-3
17:40 - 18:00
- The Effect of High-Variability Training on the Perception and Production of French Stops by German Native Speakers**
Jeanin Jügler¹, Frank Zimmerer¹, Bernd Möbius¹, Christoph Draxler²
¹Universität des Saarlandes, Germany; ²LMU München, Germany
- PAGE 811
Mon-O-12-4
18:00 - 18:20
- Perception of Mandarin Tones by Native Tibetan Speakers**
Wenfu Bao, Hui Feng, Jianwu Dang, Zhilei Liu, Yang Yu, Siyu Wang, Tianjin University, China
- PAGE 815
Mon-O-12-5
18:20 - 18:40
- Study of Acoustic Correlates of English Lexical Stress Produced by Native (L1) Bengali Speakers Compared to Native (L1) English Speakers**
Shambhu Nath Saha, Shyamal Kr. Das Mandal, IIT Kharagpur, India
- PAGE 820
Mon-O-12-6
18:40 - 19:00
- Prosodic Phrasing Unique to the Acquisition of L2 Intonation — An Analysis of L2 Japanese Intonation by L1 Swedish Learners**
Yasuko Nagano-Madsen, University of Gothenburg, Sweden

Oral Session 13: Information and Metadata Extraction from Speech

Chairs: Pedro Moreno, Olivier Siohan

Room: Conference 4+5, Time 17:00 - 19:00, Monday, September 7, 2015

- PAGE 824
Mon-O-13-1
17:00 - 17:20
- Fusion of LVCSR and Posteriorgram Based Keyword Search**
Leda Sari, Batuhan Gündoğdu, Murat Saraçlar, Boğaziçi Üniversitesi, Turkey
- PAGE 829
Mon-O-13-2
17:20 - 17:40
- Improving Speech Recognition and Keyword Search for Low Resource Languages Using Web Data**
Gideon Mendels¹, Erica Cooper¹, Victor Soto¹, Julia Hirschberg¹, Mark J.F. Gales², Kate M. Knill², Anton Ragni², Haipeng Wang²
¹Columbia University, USA; ²University of Cambridge, UK
- PAGE 834
Mon-O-13-3
17:40 - 18:00
- Two-Step Spoken Term Detection Using SVM Classifier Trained with Pre-Indexed Keywords Based on ASR Result**
Kentaro Domoto¹, Takehito Utsuro¹, Naoki Sawada², Hiromitsu Nishizaki²
¹University of Tsukuba, Japan; ²University of Yamanashi, Japan
- PAGE 839
Mon-O-13-4
18:00 - 18:20
- Enhancing Low Resource Keyword Spotting with Automatically Retrieved Web Documents**
Le Zhang, Damianos Karakos, William Hartmann, Roger Hsiao, Richard Schwartz, Stavros Tsakalidis, Raytheon BBN Technologies, USA
- PAGE 844
Mon-O-13-5
18:20 - 18:40
- A Comparison Between a DNN and a CRF Disfluency Detection and Reconstruction System**
Dario Bertero, Linlin Wang, Ho Yin Chan, Pascale Fung, HKUST, China
- PAGE 849
Mon-O-13-6
18:40 - 19:00
- Recurrent Neural Networks for Incremental Disfluency Detection**
Julian Hough, David Schlangen, Universität Bielefeld, Germany

Oral Session 14: Deep Neural Networks for Speech Synthesis

Chairs: Raul Fernandez, Yoshihiko Nankaku

Room: Conference 6, Time 17:00 – 19:00, Monday, September 7, 2015

- PAGE 854
Mon-O-14-1
17:00 – 17:20
- Fusion of Multiple Parameterisations for DNN-Based Sinusoidal Speech Synthesis with Multi-Task Learning**
Qiong Hu¹, Zhizheng Wu¹, Korin Richmond¹, Junichi Yamagishi¹, Yannis Stylianou², Ranniery Maia²
¹University of Edinburgh, UK; ²Toshiba Research Europe, UK
- PAGE 859
Mon-O-14-2
17:20 – 17:40
- An Investigation of Recurrent Neural Network Architectures for Statistical Parametric Speech Synthesis**
Sivanand Achanta, Tejas Godambe, Suryakanth V. Gangashetty, IIIT Hyderabad, India
- PAGE 864
Mon-O-14-3
17:40 – 18:00
- Sequence Generation Error (SGE) Minimization Based Deep Neural Networks Training for Text-to-Speech Synthesis**
Yuchen Fan, Yao Qian, Frank K. Soong, Lei He, Microsoft, China
- PAGE 869
Mon-O-14-4
18:00 – 18:20
- Towards Minimum Perceptual Error Training for DNN-Based Speech Synthesis**
Cassia Valentini-Botinhao, Zhizheng Wu, Simon King, University of Edinburgh, UK
- PAGE 874
Mon-O-14-5
18:20 – 18:40
- Deep Neural Network-Based Statistical Parametric Speech Synthesis System Using Improved Time-Frequency Trajectory Excitation Model**
Eunwoo Song, Hong-Goo Kang, Yonsei University, Korea
- PAGE 879
Mon-O-14-6
18:40 – 19:00
- A Study of Speaker Adaptation for DNN-Based Speech Synthesis**
Zhizheng Wu, Pawel Swietojanski, Christophe Veaux, Steve Renals, Simon King, University of Edinburgh, UK

INTERSPEECH 2015 Table of Contents

Special Session 2b: Interspeech 2015 Computational Paralinguistics Challenge (ComParE): Degree of Nateness, Parkinson's & Eating Condition

Chairs: Anton Batliner, Stefan Steidl

Room: Hall 1, Time 17:00 – 19:00, Monday, September 7, 2015

- Mon-SP2b-i
17:00 – 17:10
- The Eating Condition Sub-Challenge: The Data**
Anton Batliner, FAU Erlangen-Nürnberg, Germany
- PAGE 884
Mon-SP2b-1
17:10 – 17:20
- Automatic Classification of Eating Conditions from Speech Using Acoustic Feature Selection and a Set of Hierarchical Support Vector Machine Classifiers**
Abhay Prasad, Prasanta Kumar Ghosh, Indian Institute of Science, India
- PAGE 889
Mon-SP2b-2
17:20 – 17:30
- Combining Hierarchical Classification with Frequency Weighting for the Recognition of Eating Conditions**
Johannes Wagner, Andreas Seiderer, Florian Lingensfelder, Elisabeth André, Universität Augsburg, Germany
- PAGE 894
Mon-SP2b-3
17:30 – 17:40
- Acoustic Group Feature Selection Using Wrapper Method for Automatic Eating Condition Recognition**
Dara Pir¹, Theodore Brown²
¹CUNY Graduate Center, USA; ²CUNY Queens College, USA
- PAGE 899
Mon-SP2b-4
17:40 – 17:50
- Comparing SVM, Softmax, and Shallow Neural Networks for Eating Condition Classification**
Thomas Pellegrini, IRIT, France
- PAGE 904
Mon-SP2b-5
17:50 – 18:00
- Using Representation Learning and Out-of-Domain Data for a Paralinguistic Speech Task**
Benjamin Milde, Chris Biemann, Technische Universität Darmstadt, Germany

Mon-SP2b continued...

PAGE 909
Mon-SP2b-6
18:00 – 18:10

Fisher Vectors with Cascaded Normalization for Paralinguistic Analysis

Heysem Kaya¹, Alexey A. Karpov², Albert Ali Salah¹

¹Boğaziçi Üniversitesi, Turkey; ²Russian Academy of Sciences, Russia

PAGE 914
Mon-SP2b-7
18:10 – 18:20

Automatic Estimation of Parkinson's Disease Severity from Diverse Speech Tasks

Jangwon Kim, Md. Nasir, Rahul Gupta, Maarten Van Segbroeck, Daniel Bone, Matthew P. Black, Zisis Iason Skordilis, Zhaojun Yang, Panayiotis G. Georgiou, Shrikanth S. Narayanan, University of Southern California, USA

PAGE 919
Mon-SP2b-8
18:20 – 18:30

Assessing the Degree of Nativeness and Parkinson's Condition Using Gaussian Processes and Deep Rectifier Neural Networks

Tamás Grósz¹, Róbert Busa-Fekete², Gábor Gosztolya³, László Tóth³

¹University of Szeged, Hungary; ²Universität Paderborn, Germany; ³MTA-SZTE RGAI, Hungary

Mon-SP2b-5
18:30 – 18:40

The INTERSPEECH 2015 Computational Paralinguistics Challenge: A Summary of Results

Stefan Steidl, FAU Erlangen-Nürnberg, Germany

Mon-SP2b-W
18:40 – 18:50

Wrapping Up: The Story of the Compare Challenges, What We Learned and Where to Go

Anton Batliner, FAU Erlangen-Nürnberg, Germany

Mon-SP2b-D
18:50 – 19:00

Discussion

Poster Session 7: Prosody 1

Chair: Hongwei Ding

Room: Hall 2, Time 17:00 – 19:00, Monday, September 7, 2015

PAGE 924
Mon-P-7-1

Pitch Scaling as a Perceptual Cue for Questions in German

Jan Michalsky, Carl von Ossietzky Universität Oldenburg, Germany

PAGE 929
Mon-P-7-2

Parameterization of Prosodic Headedness

Uwe D. Reichel¹, Katalin Mády², Štefan Beňuš³

¹LMU München, Germany; ²Hungarian Academy of Sciences, Hungary; ³UKF, Slovak Republic

PAGE 934
Mon-P-7-3

Detection of Mizo Tones

Biswajit Dev Sarma, Priyankoo Sarmah, Wendy Lalhminghlui, S.R. Mahadeva Prasanna, IIT Guwahati, India

PAGE 938
Mon-P-7-4

The Intonation of Echo *wh*-Questions

Sophie Repp, Lena Rosin, Humboldt-Universität zu Berlin, Germany

PAGE 943
Mon-P-7-5

Immediately Postverbal Questions in Urdu

Farhat Jabeen, Tina Bögel, Miriam Butt, Universität Konstanz, Germany

PAGE 948
Mon-P-7-6

Prosodic (Non-)Realisation of Broad, Narrow and Contrastive Focus in Hungarian: A Production and a Perception Study

Katalin Mády, Hungarian Academy of Sciences, Hungary

PAGE 953
Mon-P-7-7

F0 Discontinuity as a Marker of Prosodic Boundary Strength in Lombard Speech

Štefan Beňuš¹, Uwe D. Reichel², Juraj Šimko³

¹UKF, Slovak Republic; ²LMU München, Germany; ³University of Helsinki, Finland

Poster Session 8: Prosody 2

Chair: Hongwei Ding

Room: Hall 2, Time 17:00 – 19:00, Monday, September 7, 2015

- PAGE 958
Mon-P-8-1 **Comparing Journalistic and Spontaneous Speech: Prosodic and Spectral Analysis**
Cédric Gendrot, Martine Adda-Decker, Yaru Wu, LPP (UMR 7018), France
- PAGE 963
Mon-P-8-2 **Rhythm Influences the Tonal Realisation of Focus**
Nadja Schauffler, Katrin Schweitzer, Universität Stuttgart, Germany
- PAGE 968
Mon-P-8-3 **Linguistic Measures of Pitch Range in Slavic and Germanic Languages**
*Bistra Andreeva¹, Bernd Möbius¹, Grazyna Demenko², Frank Zimmerer¹,
Jeanin Jügler¹*
*¹Universität des Saarlandes, Germany; ²Adam Mickiewicz University in Poznań,
Poland*
- PAGE 973
Mon-P-8-4 **The Effect of Stress on Vowel Space in Daxi Hakka Chinese**
Chunan Qiu, Jie Liang, Tongji University, China
- PAGE 978
Mon-P-8-5 **Declination, Peak Height and Pitch Level in Declaratives and Questions of South
Connaught Irish**
Maria O'Reilly, Ailbhe Ní Chasaide, Trinity College Dublin, Ireland
- PAGE 983
Mon-P-8-6 **Contextual Variation of Tones in Mizo**
Priyankoo Sarmah, Leena Dihingia, Wendy Lalhminghlu, IIT Guwahati, India
- PAGE 987
Mon-P-8-7 **The Prosodic Marking of Rhetorical Questions in German**
*Daniela Wochner, Jana Schlegel, Nicole Dehé, Bettina Braun, Universität Konstanz,
Germany*

Poster Session 9: Speaker and Language Recognition

Chair: Kong Aik Lee

Room: Foyer, Time 17:00 – 19:00, Monday, September 7, 2015

- PAGE 992
Mon-P-9-1 **High-Resolution Acoustic Modeling and Compact Language Modeling of
Language-Universal Speech Attributes for Spoken Language Identification**
Yannan Wang¹, Jun Du¹, Li-Rong Dai¹, Chin-Hui Lee²
¹USTC, China; ²Georgia Institute of Technology, USA
- PAGE 997
Mon-P-9-2 **Phonemes Frequency Based PLLR Dimensionality Reduction for Language
Recognition**
*Saad Irtza¹, Vidhyasaharan Sethu¹, Phu Ngoc Le¹, Eliathamby Ambikairajah¹,
Haizhou Li²*
*¹University of New South Wales, Australia; ²A*STAR, Singapore*
- PAGE 1002
Mon-P-9-3 **Exploiting i-Vector Posterior Covariances for Short-Duration Language
Recognition**
Sandro Cumani¹, Oldřich Plchot², Radek Fér²
¹Politecnico di Torino, Italy; ²Brno University of Technology, Czech Republic
- PAGE 1007
Mon-P-9-4 **Using the Beat Histogram for Speech Rhythm Description and Language
Identification**
Athanasios Lykartsis, Stefan Weinzierl, Technische Universität Berlin, Germany
- PAGE 1012
Mon-P-9-5 **Speaker Recognition for Speech Under Face Cover**
*Rahim Saeidi¹, Tuija Niemi², Hanna Karppelin³, Jouni Pohjalainen¹,
Tomi Kinnunen⁴, Paavo Alku¹*
*¹Aalto University, Finland; ²National Bureau of Investigation, Finland; ³University of
Helsinki, Finland; ⁴University of Eastern Finland, Finland*

Mon-P-9 continued...

- PAGE 1017
Mon-P-9-6 **Dataset-Invariant Covariance Normalization for Out-Domain PLDA Speaker Verification**
Md. Hafizur Rahman, Ahilan Kanagasundaram, David Dean, Sridha Sridharan, Queensland University of Technology, Australia
- PAGE 1022
Mon-P-9-7 **Sparse Coding of Total Variability Matrix**
Longting Xu¹, Kong Aik Lee², Haizhou Li², Zhen Yang¹
*¹NJUPT, China; ²A*STAR, Singapore*
- PAGE 1027
Mon-P-9-8 **Duration Dependent Covariance Regularization in PLDA Modeling for Speaker Verification**
Weicheng Cai¹, Ming Li², Lin Li³, QingYang Hong³
¹JRI, China; ²Sun Yat-sen University, China; ³Xiamen University, China
- PAGE 1032
Mon-P-9-9 **Exploiting Supervector Structure for Speaker Recognition Trained on a Small Development Set**
Hagai Aronowitz, IBM Research Haifa, Israel
- PAGE 1037
Mon-P-9-10 **Modified-Prior PLDA and Score Calibration for Duration Mismatch Compensation in Speaker Recognition System**
QingYang Hong¹, Lin Li¹, Ming Li², Ling Huang¹, Lihong Wan¹, Jun Zhang¹
¹Xiamen University, China; ²Sun Yat-sen University, China
- PAGE 1042
Mon-P-9-11 **Speaker Verification Using Gaussian Posteriorgrams on Fixed Phrase Short Utterances**
Sarfaraz Jelil¹, Rohan Kumar Das², Rohit Sinha², S.R. Mahadeva Prasanna²
¹North-Eastern Hill University, India; ²IIT Guwahati, India

Mon-P-9 continued...

- PAGE 1047
Mon-P-9-12 **Importance of Intelligible Phonemes for Human Speaker Recognition in Different Channel Bandwidths**
Laura Fernández Gallardo¹, Sebastian Möller², Michael Wagner¹
¹University of Canberra, Australia; ²T-Labs, Germany
- PAGE 1052
Mon-P-9-13 **Denosing Autoencoder-Based Speaker Feature Restoration for Utterances of Short Duration**
Hitoshi Yamamoto, Takafumi Koshinaka, NEC Corporation, Japan
- PAGE 1057
Mon-P-9-14 **Full Multicondition Training for Robust i-Vector Based Speaker Recognition**
Dayana Ribas¹, Emmanuel Vincent², José Ramón Calvo¹
¹CENATAV, Cuba; ²INRIA, France

Show and Tell Session 2

Chair: Volker Fischer

Room: Seminar 1, Time 17:00 - 19:00, Monday, September 7, 2015

- PAGE 1062
Show&Tell-2-1 **SARMATA 2.0 Automatic Polish Language Speech Recognition System**
Bartosz Ziółko, Tomasz Jadczyk, Dawid Skurzok, Piotr Żelasko, Jakub Gałka, Tomasz Pędzimaż, Ireneusz Gawlik, Szymon Pałka, AGH UST, Poland
- PAGE 1064
Show&Tell-2-2 **Remeeting — Get More Out of Meetings**
Arlo Faria, Korbinian Riedhammer, Remeeting, USA
- PAGE 1066
Show&Tell-2-3 **Web Application System for Pronunciation Practice by Children with Disabilities and to Support Cooperation of Teachers and Medical Workers**
Ikuyo Masuda-Katsuse, Kindai University, Japan
- PAGE 1068
Show&Tell-2-4 **PATSY — It's All About Pronunciation!**
Caroline Kaufhold¹, Vadim Gamidov², Andreas Kiessling², Klaus Reinhard², Elmar Nöth¹
¹FAU Erlangen-Nürnberg, Germany; ²e.sigma Technology, Germany
- PAGE 1070
Show&Tell-2-5 **Real-Time Pitch Modification System for Speech and Singing Voice**
Elias Azarov, Maxim Vashkevich, Denis Likhachov, Alexander Petrovsky, BSUIR, Belarus
- PAGE 1072
Show&Tell-2-6 **Nao is Doing Humour in the CHIST-ERA JOKER Project**
Guillaume Dubuisson Duplessis¹, Lucile Béchade², Mohamed A. Sehili¹, Agnès Delaborde¹, Vincent Letard¹, Anne-Laure Ligozat¹, Paul Deléglise³, Yannick Estève³, Sophie Rosset¹, Laurence Devillers¹
¹LMSI, France; ²Université Paris Sud, France; ³LIUM, France
- PAGE 1074
Show&Tell-2-7 **ABIMS — Auditory Bewildered Interaction Measurement System**
Lisa Lange, Bartholomäus Pfeiffer, Daniel Duran, Universität Stuttgart, Germany

Oral Session 15: Neural Networks and Speaker Adaptation

Chairs: Penny Karanasou, Daniel Povey

Room: Large Hall, Time 09:00 - 11:00, Tuesday, September 8, 2015

- PAGE 1076
Tue-O-15-1
09:00 - 09:20 **Maximum a posteriori Adaptation of Network Parameters in Deep Models**
Zhen Huang¹, Sabato Marco Siniscalchi¹, I-Fan Chen¹, Jinyu Li², Jiadong Wu¹, Chin-Hui Lee¹
¹Georgia Institute of Technology, USA; ²Microsoft, USA
- PAGE 1081
Tue-O-15-2
09:20 - 09:40 **Regularized Sequence-Level Deep Neural Network Model Adaptation**
Yan Huang, Yifan Gong, Microsoft, USA
- PAGE 1086
Tue-O-15-3
09:40 - 10:00 **Modeling Speaker Variability Using Long Short-Term Memory Networks for Speech Recognition**
Xiangang Li, Xihong Wu, Peking University, China
- PAGE 1091
Tue-O-15-4
10:00 - 10:20 **Intermediate-Layer DNN Adaptation for Offline and Session-Based Iterative Speaker Adaptation**
Kshitiz Kumar, Chaojun Liu, Kaisheng Yao, Yifan Gong, Microsoft, USA
- PAGE 1096
Tue-O-15-5
10:20 - 10:40 **Speaker Adaptation of Convolutional Neural Network Using Speaker Specific Subspace Vectors of SGMM**
Murali Karthick B., Prateek Kolhar, S. Umesh, IIT Madras, India
- PAGE 1101
Tue-O-15-6
10:40 - 11:00 **On Speaker Adaptation of Long Short-Term Memory Recurrent Neural Networks**
Yajie Miao, Florian Metze, Carnegie Mellon University, USA

Oral Session 16: Brain- and Other Biosignal-based Spoken Communication

Chairs: Thomas F. Quatieri, Tanja Schultz

Room: Conference 1, Time 09:00 – 11:00, Tuesday, September 8, 2015

- PAGE 1106
Tue-O-16-1
09:00 – 09:20
- Automatic Identification of Received Language in MEG**
Emilio Parisotto¹, Youness A. Ghassabeh², Matt J. MacDonald³, Adelina Cozma⁴, Elizabeth W. Pang⁵, Frank Rudzicz¹
¹University of Toronto, Canada; ²UHN Toronto Rehabilitation Institute, Canada; ³SickKids Research Institute, Canada; ⁴University of Calgary, Canada; ⁵Hospital for Sick Children, Canada
- PAGE 1111
Tue-O-16-2
09:20 – 09:40
- Detection of Cardiovascular Reactivity in Speech**
Laurens van der Werff, Jón Guðnason, Kamilla Rún Jóhannsdóttir, Reykjavik University, Iceland
- PAGE 1116
Tue-O-16-3
09:40 – 10:00
- Lateralization in Emotional Speech Perception Following Transcranial Direct Current Stimulation**
Alex Francois-Nienaber¹, Jed A. Meltzer², Frank Rudzicz¹
¹University of Toronto, Canada; ²Rotman Research Institute, Canada
- PAGE 1121
Tue-O-16-4
10:00 – 10:20
- Speech Reconstruction from Human Auditory Cortex with Deep Neural Networks**
Minda Yang¹, Sameer A. Sheth², Catherine A. Schevon², Guy M. McKhann II², Nima Mesgarani¹
¹Columbia University, USA; ²Columbia University Medical Center, USA
- PAGE 1126
Tue-O-16-5
10:20 – 10:40
- Temporal Dynamics of the Speech Readiness Potential, and Its Use in a Neural Decoder of Speech-Motor Intention**
Jonathan S. Brumberg, Nichol Castro, Akshatha Rao, University of Kansas, USA
- PAGE 1131
Tue-O-16-6
10:40 – 11:00
- Continuous Speech Recognition from ECoG**
Dominic Heger¹, Christian Herff¹, Adriana de Pestors², Dominic Telaar¹, Peter Brunner², Gerwin Schalk², Tanja Schultz¹
¹KIT, Germany; ²New York State Department of Health, USA

Oral Session 17: Deep Neural Networks in Speaker Recognition

Chairs: Patrick Kenny, Alan McCree

Room: Conference 2+3, Time 09:00 – 11:00, Tuesday, September 8, 2015

- PAGE 1136
Tue-O-17-1
09:00 – 09:20
- Locally-Connected and Convolutional Neural Networks for Small Footprint Speaker Recognition**
Yu-hsin Chen, Ignacio Lopez-Moreno, Tara N. Sainath, Mirkó Visontai, Raziell Alvarez, Carolina Parada, Google, USA
- PAGE 1141
Tue-O-17-2
09:20 – 09:40
- Insights into Deep Neural Networks for Speaker Recognition**
Daniel Garcia-Romero, Alan McCree, Johns Hopkins University, USA
- PAGE 1146
Tue-O-17-3
09:40 – 10:00
- A Unified Deep Neural Network for Speaker and Language Recognition**
Fred Richardson¹, Douglas A. Reynolds¹, Najim Dehak²
¹MIT Lincoln Laboratory, USA; ²MIT, USA
- PAGE 1151
Tue-O-17-4
10:00 – 10:20
- Investigation of Bottleneck Features and Multilingual Deep Neural Networks for Speaker Verification**
Yao Tian, Meng Cai, Liang He, Jia Liu, Tsinghua University, China
- PAGE 1156
Tue-O-17-5
10:20 – 10:40
- Frequency Offset Correction in Single Sideband (SSB) Speech by Deep Neural Network for Speaker Verification**
Hua Xing, Gang Liu, John H.L. Hansen, University of Texas at Dallas, USA
- PAGE 1161
Tue-O-17-6
10:40 – 11:00
- Exploring Robustness of DNN/RNN for Extracting Speaker Baum-Welch Statistics in Mismatched Conditions**
Hao Zheng, Shanshan Zhang, Wenju Liu, Chinese Academy of Sciences, China

Oral Session 18: Speech Analysis and Representation 2

Chairs: Miloš Cerňak, Shinji Watanabe

Room: Conference 4+5, Time 09:00 - 11:00, Tuesday, September 8, 2015

- PAGE 1166
Tue-O-18-1
09:00 - 09:20
- AM-FM Based Filter Bank Analysis for Estimation of Spectro-Temporal Envelopes and Its Application for Speaker Recognition in Noisy Reverberant Environments**
Dhananjaya Gowda, Rahim Saeidi, Paavo Alku, Aalto University, Finland
- PAGE 1171
Tue-O-18-2
09:20 - 09:40
- Fast and Accurate Phase Unwrapping**
Thomas Drugman, Yanniss Stylianou, Toshiba Research Europe, UK
- PAGE 1176
Tue-O-18-3
09:40 - 10:00
- Sparse Representation with Temporal Max-Smoothing for Acoustic Event Detection**
Xugang Lu¹, Peng Shen¹, Yu Tsao², Chiori Hori¹, Hisashi Kawai¹
¹NICT, Japan; ²Academia Sinica, Taiwan
- PAGE 1181
Tue-O-18-4
10:00 - 10:20
- Estimation of Glottal Closure Instants from Telephone Speech Using a Group Delay-Based Approach That Considers Speech Signal as a Spectrum**
Anushiya Rachel G., Vijayalakshmi P., Nagarajan T., SSN, India
- PAGE 1186
Tue-O-18-5
10:20 - 10:40
- The Role of Prosody and Voice Quality in Text-Dependent Categories of Storytelling Across Languages**
Raúl Montaña, Francesc Alías, Universitat Ramon Llull, Spain
- PAGE 1191
Tue-O-18-6
10:40 - 11:00
- Neuromorphic Based Oscillatory Device for Incremental Syllable Boundary Detection**
Alexandre Hyafil¹, Milos Cernak²
¹Universitat Pompeu Fabra, Spain; ²Idiap Research Institute, Switzerland

Oral Session 19: Statistical Parametric Speech Synthesis

Chairs: Michael Pucher, Junichi Yamagishi

Room: Conference 6, Time 09:00 - 11:00, Tuesday, September 8, 2015

- PAGE 1196
Tue-O-19-1
09:00 - 09:20
- Simultaneous Optimization of Multiple Tree Structures for Factor Analyzed HMM-Based Speech Synthesis**
Takenori Yoshimura, Kei Hashimoto, Yoshihiko Nankaku, Keiichi Tokuda, Nagoya Institute of Technology, Japan
- PAGE 1201
Tue-O-19-2
09:20 - 09:40
- HMM Training Strategy for Incremental Speech Synthesis**
Maël Pouget¹, Thomas Hueber¹, Gérard Bailly¹, Timo Baumann²
¹GIPSA, France; ²Universität Hamburg, Germany
- PAGE 1206
Tue-O-19-3
09:40 - 10:00
- Modulation Spectrum-Constrained Trajectory Training Algorithm for HMM-Based Speech Synthesis**
Shinnosuke Takamichi¹, Tomoki Toda¹, Alan W. Black², Satoshi Nakamura¹
¹NAIST, Japan; ²Carnegie Mellon University, USA
- PAGE 1211
Tue-O-19-4
10:00 - 10:20
- Random Forests for Statistical Speech Synthesis**
Alan W. Black, Prasanna Kumar Muthukumar, Carnegie Mellon University, USA
- PAGE 1216
Tue-O-19-5
10:20 - 10:40
- Speaker Adaptation Using Relevance Vector Regression for HMM-Based Expressive TTS**
Doo Hwa Hong, Joun Yeop Lee, Se Young Jang, Nam Soo Kim, Seoul National University, Korea
- PAGE 1221
Tue-O-19-6
10:40 - 11:00
- Towards a Linear Dynamical Model Based Speech Synthesizer**
Vassilis Tsiaras¹, Ranniery Maia², Vassilis Diakouloukas¹, Yannis Stylianou², Vassilis Digalakis¹
¹Technical University of Crete, Greece; ²Toshiba Research Europe, UK

Special Session 3: Speech Science in End-user Applications

Chairs: Felix Burkhardt, Nils Lenke

Room: Hall 1, Time 09:00 - 11:00, Tuesday, September 8, 2015

PAGE 1226
Tue-SP3-1
09:00 - 09:20

Providing Objective Metrics of Team Communication Skills via Interpersonal Coordination Mechanisms

Céline De Looze¹, Brian Vaughan², Finnian Kelly³, Alison Kay¹

¹Trinity College Dublin, Ireland; ²Dublin Institute of Technology, Ireland; ³University of Texas at Dallas, USA

PAGE 1231
Tue-SP3-2
09:20 - 09:40

Dialog Act Modeling for Virtual Personal Assistant Applications Using a Small Volume of Labeled Data and Domain Knowledge

Donghyeon Lee, Jinsik Lee, Eun-Kyoung Kim, Jaewon Lee, Samsung Electronics, Korea

PAGE 1236
Tue-SP3-3
09:40 - 10:00

A Polyglot Domain Optimised Text-to-Speech System for Railway Station Announcements

Csaba Zainkó, Mátyás Bartalis, Géza Németh, Gábor Olaszy, BME, Hungary

PAGE 1241
Tue-SP3-4
10:00 - 10:20

Development of Hindi Speech Recognition System of Agricultural Commodities Using Deep Neural Network

Partho Mandal, Shalini Jain, Gaurav Ojha, Anupam Shukla, ABV-IIITM, India

PAGE 1246
Tue-SP3-5
10:20 - 10:40

Real-Time Audio Signal Enhancement for Hands-Free Speech Applications

Thomas Fehér, Michael Freitag, Christian Gruber, voice INTER connect, Germany

PAGE 1251
Tue-SP3-6
10:40 - 11:00

Personalized Synthetic Voices for Speaking Impaired: Website and App

D. Erro¹, Inma Hernaez¹, Agustin Alonso¹, D. García-Lorenzo¹, Eva Navas¹, J. Ye¹, H. Arzelus², Igor Jauk³, N.Q. Hy⁴, C. Magariños⁵, R. Pérez-Ramón¹, M. Sulír⁶, Xiaohai Tian⁴, X. Wang⁷

¹Universidad del País Vasco, Spain; ²VicomTech-IK4, Spain; ³Universitat Politècnica de Catalunya, Spain; ⁴NTU, Singapore; ⁵Universidade de Vigo, Spain; ⁶Technical University of Košice, Slovak Republic; ⁷USTC, China

Poster Session 10: Speech Recognition: Evaluation and Low-resource Languages

Chair: Jia Cui

Room: Hall 2, Time 09:00 - 11:00, Tuesday, September 8, 2015

- PAGE 1255
Tue-P-10-1 **Under-Resourced Speech Recognition Based on the Speech Manifold**
Reza Sahraeian¹, Dirk Van Compernelle¹, Febe de Wet²
¹Katholieke Universiteit Leuven, Belgium; ²CSIR, South Africa
- PAGE 1260
Tue-P-10-2 **Multilingual Features Based Keyword Search for Very Low-Resource Languages**
Pavel Golik, Zoltán Tüske, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany
- PAGE 1265
Tue-P-10-3 **Second Language Speech Recognition Using Multiple-Pass Decoding with Lexicon Represented by Multiple Reduced Phoneme Sets**
Xiaoyun Wang, Seiichi Yamamoto, Doshisha University, Japan
- PAGE 1270
Tue-P-10-4 **Using Resources from a Closely-Related Language to Develop ASR for a Very Under-Resourced Language: A Case Study for Iban**
Sarah Samson Juan¹, Laurent Besacier², Benjamin Lecouteux², Mohamed Dyab²
¹UNIMAS, Malaysia; ²LIG (UMR 5217), France
- PAGE 1275
Tue-P-10-5 **Prediction of Speech Recognition Accuracy for Utterance Classification**
Maxim L. Korenevsky, Andrey B. Smirnov, Valentin S. Mendeleev, ITMO University, Russia
- PAGE 1280
Tue-P-10-6 **Error Bounds for Context Reduction and Feature Omission**
Eugen Beck, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

Tue-P-10 continued...

- PAGE 1285
Tue-P-10-7 **A Metric for Evaluating Speech Recognizer Output Based on Human-Perception Model**
Nobuyasu Itoh¹, Gakuto Kurata¹, Ryuki Tachibana¹, Masafumi Nishimura²
¹IBM Research Tokyo, Japan; ²Shizuoka University, Japan
- PAGE 1289
Tue-P-10-8 **How to Evaluate ASR Output for Named Entity Recognition?**
Mohamed Ameer Ben Jannet¹, Olivier Galibert¹, Martine Adda-Decker², Sophie Rosset³
¹LNE, France; ²LPP (UMR 7018), France; ³LIMSI, France

Poster Session 11: Emotion

Chair: Anton Batliner

Room: Hall 2, Time 09:00 - 11:00, Tuesday, September 8, 2015

- PAGE 1294
Tue-P-11-1 **Acoustic-Prosodic Analysis of Attitudinal Expressions in German**
Hansjörg Mixdorff¹, Angelika Hönemann², Albert Rilliard³
¹BHT Berlin, Germany; ²Universität Bielefeld, Germany; ³LIMSI, France
- PAGE 1299
Tue-P-11-2 **Continuous Emotion Tracking Using Total Variability Space**
Hossein Khaki, Engin Erzin, Koç Üniversitesi, Turkey
- PAGE 1304
Tue-P-11-3 **An Analysis of the Relationship Between Signal-Derived Vocal Arousal Score and Human Emotion Production and Perception**
Chi-Chun Lee¹, Daniel Bone², Shrikanth S. Narayanan²
¹National Tsing Hua University, Taiwan; ²University of Southern California, USA
- PAGE 1309
Tue-P-11-4 **Morphology of Vocal Affect Bursts: Exploring Expressive Interjections in Japanese Conversation**
Hiroki Mori, Utsunomiya University, Japan
- PAGE 1314
Tue-P-11-5 **Emotion Clustering Based on Probabilistic Linear Discriminant Analysis**
Mahnoosh Mehrabani, Ozlem Kalinli, Ruxin Chen, Sony, USA
- PAGE 1319
Tue-P-11-6 **Objective Study of the Performance Degradation in Emotion Recognition Through the AMR-WB+ Codec**
Aaron Albin, Elliot Moore, Georgia Institute of Technology, USA
- PAGE 1324
Tue-P-11-7 **Analysis of Excitation Source Features of Speech for Emotion Recognition**
Sudarsana Reddy Kadiri, P. Gangamohan, Suryakanth V. Gangashetty, B. Yegnanarayana, IIT Hyderabad, India

Tue-P-11 continued...

- PAGE 1329
Tue-P-11-8 **An Investigation of Emotion Change Detection from Speech**
Zhaocheng Huang, Julien Epps, Eliathamby Ambikairajah, University of New South Wales, Australia
- PAGE 1334
Tue-P-11-9 **Crosslinguistic Comparison on the Perception of Mandarin Attitudinal Speech**
Wentao Gu¹, Ping Tang¹, Keikichi Hirose², Véronique Aubergé³
¹Nanjing Normal University, China; ²University of Tokyo, Japan; ³CNRS, France
- PAGE 1339
Tue-P-11-10 **Conflict Intensity Estimation from Speech Using Greedy Forward-Backward Feature Selection**
Gábor Gosztolya, MTA-SZTE RGAI, Hungary

Poster Session 12: Spoken Language Understanding

Chair: Ruhi Sarikaya

Room: Foyer, Time 09:00 - 11:00, Tuesday, September 8, 2015

- PAGE 1344
Tue-P-12-1 **Study of Entity-Topic Models for OOV Proper Name Retrieval**
Imran Sheikh, Irina Illina, Dominique Fohr, LORIA, France
- PAGE 1349
Tue-P-12-2 **Audio Quotation Marks for Natural Language Understanding**
Simon Boutin¹, Réal Tremblay², Patrick Cardinal¹, Doug Peters², Pierre Dumouchel¹
¹École de Technologie Supérieure, Canada; ²Nuance Communications, Canada
- PAGE 1353
Tue-P-12-3 **Using Word Confusion Networks for Slot Filling in Spoken Language Understanding**
Xiaohao Yang, Jia Liu, Tsinghua University, China
- PAGE 1358
Tue-P-12-4 **Distributed Representation-Based Spoken Word Sense Induction**
Justin Chiu, Yajie Miao, Alan W. Black, Alexander I. Rudnicky, Carnegie Mellon University, USA
- PAGE 1363
Tue-P-12-5 **Structuring Lectures in Massive Open Online Courses (MOOCs) for Efficient Learning by Linking Similar Sections and Predicting Prerequisites**
Sheng-syun Shen¹, Hung-yi Lee¹, Shang-wen Li², Victor Zue², Lin-shan Lee¹
¹National Taiwan University, Taiwan; ²MIT, USA
- PAGE 1368
Tue-P-12-6 **News Talk-Show Chaptering with Journalistic Genres**
Delphine Charlet, Géraldine Damnati, Jérémy Trione, Orange Labs, France
- PAGE 1373
Tue-P-12-7 **An Analysis of Time-Aggregated and Time-Series Features for Scoring Different Aspects of Multimodal Presentation Data**
Vikram Ramanarayanan, Lei Chen, Chee Wee Leong, Gary Feng, David Suendermann-Oeft, Educational Testing Service, USA

INTERSPEECH 2015 Table of Contents

Tue-P-12 continued...

- PAGE 1378
Tue-P-12-8 **Incorporating Prosodic Prominence Evidence into Term Weights for Spoken Content Retrieval**
David N. Racca, Gareth J.F. Jones, Dublin City University, Ireland
- PAGE 1383
Tue-P-12-9 **Leveraging Word Embeddings for Spoken Document Summarization**
Kuan-Yu Chen¹, Shih-Hung Liu¹, Hsin-Min Wang¹, Berlin Chen², Hsin-Hsi Chen³
¹Academia Sinica, Taiwan; ²National Taiwan Normal University, Taiwan; ³National Taiwan University, Taiwan
- PAGE 1388
Tue-P-12-10 **Mutually Exclusive Grounding for Weakly Supervised Non-Negative Matrix Factorisation**
Vincent Renkens, Hugo Van hamme, Katholieke Universiteit Leuven, Belgium
- PAGE 1393
Tue-P-12-11 **Using Semantic Maps for Robust Natural Language Interaction with Robots**
Emanuele Bastianelli¹, Danilo Croce¹, Roberto Basili¹, Daniele Nardi²
¹Università di Roma "Tor Vergata", Italy; ²Università di Roma "La Sapienza", Italy
- PAGE 1398
Tue-P-12-12 **Efficient Learning for Spoken Language Understanding Tasks with Word Embedding Based Pre-Training**
Yi Luan¹, Shinji Watanabe², Bret Harsham²
¹University of Washington, USA; ²MERL, USA
- PAGE 1403
Tue-P-12-13 **Zero-Shot Semantic Parser for Spoken Language Understanding**
Emmanuel Ferreira, Bassam Jabaian, Fabrice Lefèvre, LIA, France
- PAGE 1408
Tue-P-12-14 **Adapting Lexical Representation and OOV Handling from Written to Spoken Language with Word Embedding**
Jeremie Tafforeau, Thierry Artieres, Benoit Favre, Frederic Bechet, LIF (UMR 7279), France

Poster Session 13: Language Modeling for Speech Recognition

Chair: Xunying Liu

Room: Seminar 5+6, Time 09:00 – 11:00, Tuesday, September 8, 2015

- PAGE 1413
Tue-P-13-1 **Multi-Stream Long Short-Term Memory Neural Network Language Model**
Ebru Arisoy¹, Murat Saraçlar²
¹MEF Üniversitesi, Turkey; ²Boğaziçi Üniversitesi, Turkey
- PAGE 1418
Tue-P-13-2 **Composition-Based On-the-Fly Rescoring for Salient n-Gram Biasing**
Keith Hall, Eunjoon Cho, Cyril Allauzen, Françoise Beaufays, Noah Coccaro, Kaisuke Nakajima, Michael Riley, Brian Roark, David Rybach, Linda Zhang, Google, USA
- PAGE 1423
Tue-P-13-3 **Learning Phrase Patterns for ASR Name Error Detection Using Semantic Similarity**
Alex Marin¹, Mari Ostendorf², Ji He²
¹Microsoft, USA; ²University of Washington, USA
- PAGE 1428
Tue-P-13-4 **Sparse Non-Negative Matrix Language Modeling for Skip-Grams**
Noam Shazeer, Joris Pelemans, Ciprian Chelba, Google, USA
- PAGE 1433
Tue-P-13-5 **Pruning Sparse Non-Negative Matrix N-Gram Language Models**
Joris Pelemans, Noam Shazeer, Ciprian Chelba, Google, USA
- PAGE 1438
Tue-P-13-6 **Geo-Location for Voice Search Language Modeling**
Ciprian Chelba, Xuedong Zhang, Keith Hall, Google, USA
- PAGE 1443
Tue-P-13-7 **On Efficient Training of Word Classes and Their Application to Recurrent Neural Network Language Models**
Rami Botros, Kazuki Irie, Martin Sundermeyer, Hermann Ney, RWTH Aachen University, Germany

Tue-P-13 continued...

- PAGE 1448
Tue-P-13-8 **Deep Semantic Encodings for Language Modeling**
Ali Orkan Bayer, Giuseppe Ricciardi, Università di Trento, Italy
- PAGE 1453
Tue-P-13-9 **Learning OOV Through Semantic Relatedness in Spoken Dialog Systems**
Ming Sun, Yun-Nung Chen, Alexander I. Rudnicky, Carnegie Mellon University, USA
- PAGE 1458
Tue-P-13-10 **TDTO Language Modeling with Feedforward Neural Networks**
Tze Yuang Chong¹, Rafael E. Banchs², Eng Siong Chng¹, Haizhou Li¹
*¹TL@NTU, Singapore; ²A*STAR, Singapore*

Oral Session 20: Fast Efficient and Scalable Computing for Neural Nets

Chairs: Tara Sainath, Frank Seide

Room: Large Hall, Time 14:00 - 16:00, Tuesday, September 8, 2015

- PAGE 1463
Tue-O-20-1
14:00 - 14:20
- Improvements to the Pruning Behavior of DNN Acoustic Models**
Matthias Paulik, Apple, USA
- PAGE 1468
Tue-O-20-2
14:20 - 14:40
- Fast and Accurate Recurrent Neural Network Acoustic Models for Speech Recognition**
Haşim Sak, Andrew Senior, Kanishka Rao, Françoise Beaufays, Google, USA
- PAGE 1473
Tue-O-20-3
14:40 - 15:00
- Compressing Deep Neural Networks Using a Rank-Constrained Topology**
Preetum Nakkiran¹, Raziq Alvarez², Rohit Prabhavalkar², Carolina Parada²
¹University of California at Berkeley, USA; ²Google, USA
- PAGE 1478
Tue-O-20-4
15:00 - 15:20
- Convolutional Neural Networks for Small-Footprint Keyword Spotting**
Tara N. Sainath, Carolina Parada, Google, USA
- PAGE 1483
Tue-O-20-5
15:20 - 15:40
- Efficient GPU Implementation of Convolutional Neural Networks for Speech Recognition**
Ewout van den Berg¹, Daniel Brand², Rajesh Bordawekar², Leonid Rachevsky¹, Bhuvana Ramabhadran¹
¹IBM Watson, USA; ²IBM T.J. Watson Research Center, USA
- PAGE 1488
Tue-O-20-6
15:40 - 16:00
- Scalable Distributed DNN Training Using Commodity GPU Cloud Computing**
Nikko Strom, Amazon.com, USA

Oral Session 21: Source Separation and Computational Auditory Scene Analysis

Chairs: Rajesh M. Hegde, Chin-Hui Lee

Room: Conference 1, Time 14:00 - 16:00, Tuesday, September 8, 2015

- PAGE 1493
Tue-O-21-1
14:00 - 14:20
- Joint Source Localization and Separation in Spherical Harmonic Domain Using a Sparsity Based Method**
Sachin N. Kalkur, Sandeep Reddy C., Rajesh M. Hegde, IIT Kanpur, India
- PAGE 1498
Tue-O-21-2
14:20 - 14:40
- Regularized Non-Negative Matrix Factorization Using Alternating Direction Method of Multipliers and Its Application to Source Separation**
Shaofei Zhang¹, Dong-Yan Huang², Lei Xie¹, Eng Siong Chng³, Haizhou Li², Minghui Dong²
*¹Northwestern Polytechnical University, China; ²A*STAR, Singapore; ³NTU, Singapore*
- PAGE 1503
Tue-O-21-3
14:40 - 15:00
- Two-Stage Multi-Target Joint Learning for Monaural Speech Separation**
Shuai Nie¹, Shan Liang¹, Wei Xue¹, Xueliang Zhang², Wenju Liu¹, Like Dong³, Hong Yang³
¹Chinese Academy of Sciences, China; ²Inner Mongolia University, China; ³SGCC, China
- PAGE 1508
Tue-O-21-4
15:00 - 15:20
- Multi-Objective Learning and Mask-Based Post-Processing for Deep Neural Network Based Speech Enhancement**
Yong Xu¹, Jun Du¹, Zhen Huang², Li-Rong Dai¹, Chin-Hui Lee²
¹USTC, China; ²Georgia Institute of Technology, USA
- PAGE 1513
Tue-O-21-5
15:20 - 15:40
- Discriminative Nonnegative Matrix Factorization Using Cross-Reconstruction Error for Source Separation**
Kisoo Kwon¹, Jong Won Shin², Hyung Yong Kim¹, Nam Soo Kim¹
¹Seoul National University, Korea; ²GIST, Korea

Tue-O-21 continued...

PAGE 1517
Tue-O-21-6
15:40 - 16:00

Using Audio and Visual Information for Single Channel Speaker Separation
Faheem Khan, Ben Milner, University of East Anglia, UK

Oral Session 22: Emotion

Chairs: Carlos Busso Recabarren, Nick Campbell

Room: Conference 2+3, Time 14:00 - 16:00, Tuesday, September 8, 2015

PAGE 1522
Tue-O-22-1
14:00 - 14:20

Exploring Acoustic Differences Between Cantonese (Tonal) and English (Non-Tonal) Spoken Expressions of Emotions
Chee Seng Chong, Jeesun Kim, Chris Davis, University of Western Sydney, Australia

PAGE 1527
Tue-O-22-2
14:20 - 14:40

Valence, Arousal and Dominance Estimation for English, German, Greek, Portuguese and Spanish Lexica Using Semantic Models
Elisavet Palogiannidi¹, Elias Iosif², Polychronis Koutsakis¹, Alexandros Potamianos²
¹Technical University of Crete, Greece; ²Athena RC, Greece

PAGE 1532
Tue-O-22-3
14:40 - 15:00

Dimensionality Reduction for Speech Emotion Features by Multiscale Kernels
Xinzhou Xu¹, Jun Deng¹, Wenming Zheng², Li Zhao², Björn Schuller³
¹Technische Universität München, Germany; ²Southeast University, China; ³Imperial College London, UK

PAGE 1537
Tue-O-22-4
15:00 - 15:20

High-Level Feature Representation Using Recurrent Neural Network for Speech Emotion Recognition
Jinkyu Lee¹, Ivan Tashev²
¹Yonsei University, Korea; ²Microsoft, USA

PAGE 1541
Tue-O-22-5
15:20 - 15:40

Speech Emotion Classification Using Tree-Structured Sparse Logistic Regression
Myung Jong Kim, Joohong Yoo, Younggwon Kim, Hoirin Kim, KAIST, Korea

PAGE 1546
Tue-O-22-6
15:40 - 16:00

Annotators' Agreement and Spontaneous Emotion Classification Performance
Bogdan Vlasenko¹, Andreas Wendemuth²
¹Idiap Research Institute, Switzerland; ²Otto-von-Guericke-Universität Magdeburg, Germany

Oral Session 23: Computational Models of Human Speech Perception

Chairs: Martin Cooke, Okko Räsänen

Room: Conference 4+5, Time 14:00 - 16:00, Tuesday, September 8, 2015

- PAGE 1551
Tue-O-23-1
14:00 - 14:20
- On the Nature of the Features Generated in the Human Auditory Pathway for Phone Recognition**
Harald Höge, Universität der Bundeswehr München, Germany
- PAGE 1556
Tue-O-23-2
14:20 - 14:40
- How the Slope of the Speech Spectrum Affects the Perception of Speaker Size**
Kodai Yamamoto¹, Toshio Irino¹, Ryuichi Nisimura¹, Hideki Kawahara¹, Roy D. Patterson²
¹Wakayama University, Japan; ²University of Cambridge, UK
- PAGE 1561
Tue-O-23-3
14:40 - 15:00
- Weakly-Supervised Word Learning is Improved by an Active Online Algorithm**
Heikki Rasilo, Okko Räsänen, Aalto University, Finland
- PAGE 1566
Tue-O-23-4
15:00 - 15:20
- The Effect of Cochlear Implant Processing on Speaker Intelligibility: A Perceptual Study and Computer Model**
Lin Lin¹, Jon Barker², Guy J. Brown²
¹Jilin University, China; ²University of Sheffield, UK
- PAGE 1571
Tue-O-23-5
15:20 - 15:40
- Phonetic-Phonological Feature Emerges by Associating Phonetic with Semantic Information — A GSOM-Based Modeling Study**
Mengxue Cao¹, Aijun Li¹, Qiang Fang¹, Bernd J. Kröger²
¹Chinese Academy of Social Sciences, China; ²RWTH Aachen University, Germany
- PAGE 1576
Tue-O-23-6
15:40 - 16:00
- DIANA: Towards Computational Modeling Reaction Times in Lexical Decision in North American English**
L. ten Bosch¹, L. Boves¹, B. Tucker², M. Ernestus¹
¹Radboud Universiteit Nijmegen, The Netherlands; ²University of Alberta, Canada

INTERSPEECH 2015 Table of Contents

Oral Session 24: Prosody Modeling for Speech Synthesis

Chairs: Katrin Schweitzer, Ingmar Steiner

Room: Conference 6, Time 14:00 - 16:00, Tuesday, September 8, 2015

- PAGE 1581
Tue-O-24-1
14:00 - 14:20
- Automatic Phrase Boundary Labeling of Speech Synthesis Database Using Context-Dependent HMMs and N-Gram Prior Distributions**
Qian Chen¹, Zhen-Hua Ling¹, Chen-Yu Yang², Li-Rong Dai¹
*¹USTC, China; ²A*STAR, Singapore*
- PAGE 1586
Tue-O-24-2
14:20 - 14:40
- A Perceptual Investigation of Wavelet-Based Decomposition of *f*₀ for Text-to-Speech Synthesis**
Manuel Sam Ribeiro, Junichi Yamagishi, Robert A.J. Clark, University of Edinburgh, UK
- PAGE 1591
Tue-O-24-3
14:40 - 15:00
- Duration Prediction Using Multi-Level Model for GPR-Based Speech Synthesis**
Decha Moungsri, Tomoki Koriyama, Takao Kobayashi, Tokyo Institute of Technology, Japan
- PAGE 1596
Tue-O-24-4
15:00 - 15:20
- Data-Driven Foot-Based Intonation Generator for Text-to-Speech Synthesis**
Mahsa Sadat Elyasi Langarani, Jan van Santen, Seyed Hamidreza Mohammadi, Alexander Kain, Oregon Health & Science University, USA
- PAGE 1601
Tue-O-24-5
15:20 - 15:40
- Weighted Correlation Based Atom Decomposition Intonation Modelling**
Branislav Gerazov¹, Pierre-Edouard Honnet², Aleksandar Gjoreski¹, Philip N. Garner²
¹UKiM, Macedonia; ²Idiap Research Institute, Switzerland
- PAGE 1606
Tue-O-24-6
15:40 - 16:00
- Using Deep Bidirectional Recurrent Neural Networks for Prosodic-Target Prediction in a Unit-Selection Text-to-Speech System**
Raul Fernandez¹, Asaf Rendel², Bhuvana Ramabhadran¹, Ron Hoory²
¹IBM T.J. Watson Research Center, USA; ²IBM Research Haifa, Israel

Special Session 4: Speech and Language Processing of Children's Speech

Chairs: Kay Berkling, Diego Giuliani

Room: Hall 1, Time 14:00 - 16:00, Tuesday, September 8, 2015

Oral Presentations

PAGE 1611
Tue-SP4-1
14:00 - 14:20

Large Vocabulary Automatic Speech Recognition for Children

Hank Liao, Golan Pundak, Olivier Siohan, Melissa K. Carroll, Noah Coccaro, Qi-Ming Jiang, Tara N. Sainath, Andrew Senior, Françoise Beaufays, Michiel Bacchiani, Google, USA

PAGE 1616
Tue-SP4-2
14:20 - 14:40

Acoustic-Prosodic Correlates of 'Awkward' Prosody in Story Retellings from Adolescents with Autism

Daniel Bone¹, Matthew P. Black¹, Anil Ramakrishna¹, Ruth Grossman², Shrikanth S. Narayanan¹
¹University of Southern California, USA; ²Emerson College, USA

PAGE 1621
Tue-SP4-3
14:40 - 15:00

Evidence of Phonological Processes in Automatic Recognition of Children's Speech

Eva Fringi¹, Jill Fain Lehman², Martin Russell¹
¹University of Birmingham, UK; ²Disney Research, USA

Poster Presentations

PAGE 1625
Tue-SP4-4

Influence of Speaker Familiarity on Blind and Visually Impaired Children's Perception of Synthetic Voices in Audio Games (Poster)

Michael Pucher¹, Markus Toman¹, Dietmar Schabus¹, Cassia Valentini-Botinhao², Junichi Yamagishi², Bettina Zillinger³, Erich Schmid⁴
¹FTW, Austria; ²University of Edinburgh, UK; ³Fachhochschule Wiener Neustadt, Austria; ⁴Bundes-Blindenerziehungsinstitut, Austria
Tue-SP4 continued...

PAGE 1630
Tue-SP4-5

Low-Memory Fast On-Line Adaptation for Acoustically Mismatched Children's Speech Recognition (Poster)

S. Shahnawazuddin, Rohit Sinha, IIT Guwahati, India

PAGE 1635
Tue-SP4-6

Large Vocabulary Children's Speech Recognition with DNN-HMM and SGMM Acoustic Modeling (Poster)

Diego Giuliani¹, Bagher BabaAli²
¹FBK, Italy; ²University of Tehran, Iran

PAGE 1640
Tue-SP4-7

HMM Adaptation for Child Speech Synthesis (Poster)

Avashna Govender¹, Febe de Wet¹, Jules-Raymond Tapamo²
¹CSIR, South Africa; ²University of KwaZulu-Natal, South Africa

PAGE 1645
Tue-SP4-8

Vocal Turn-Taking Patterns in Groups of Children Performing Collaborative Tasks: An Exploratory Study (Poster)

Jaebok Kim, Khiet P. Truong, Vicky Charisi, Cristina Zaga, Manja Lohse, Dirk Heylen, Vanessa Evers, University of Twente, The Netherlands

PAGE 1650
Tue-SP4-9

Towards an Automated Screening Tool for Pediatric Speech Delay (Poster)

Roozbeh Sadeghian, Stephen A. Zahorian, Binghamton University, USA

PAGE 1655
Tue-SP4-10

Children's Reading Aloud Performance: A Database and Automatic Detection of Disfluencies (Poster)

Jorge Proença¹, Dirce Celorico¹, Sara Candeias², Carla Lopes¹, Fernando Perdigão¹
¹Instituto de Telecomunicações, Portugal; ²Microsoft, Portugal

PAGE 1660
Tue-SP4-11

Keyword Spotting in Multi-Player Voice Driven Games for Children (Poster)

Harshavardhan Sundar¹, Jill Fain Lehman², Rita Singh²
¹Carnegie Mellon University, USA; ²Disney Research, USA

Tue-SP4 continued...

- PAGE 1665
Tue-SP4-12 **Age-Dependent Height Estimation and Speaker Normalization for Children's Speech Using the First Three Subglottal Resonances (Poster)**
Jinxi Guo¹, Rohit Paturi¹, Gary Yeung¹, Steven M. Lulich², Harish Arsikere³, Abeer Alwan¹
¹University of California at Los Angeles, USA; ²Indiana University, USA; ³Xerox Research Center India, India

Poster Session 14: Syllables and Segments

Chairs: Shyam Agrawal, Marcin Włodarczak

Room: Hall 2, Time 14:00 - 16:00, Tuesday, September 8, 2015

- PAGE 1670
Tue-P-14-1 **The Effect of Speakers' Regional Varieties on Listeners' Decision-Making**
Adrian Leemann¹, Camilla Bernardasci², Francis Nolan¹
¹University of Cambridge, UK; ²Universität Zürich, Switzerland
- PAGE 1675
Tue-P-14-2 **Word-Initial Glottal Stop Insertion, Hiatus Resolution and Linking in British English**
Robert Fuchs, Westfälische Wilhelms-Universität Münster, Germany
- PAGE 1680
Tue-P-14-3 **Acoustic Analysis of Mandarin Affricates**
Shanpeng Li, Wentao Gu, Nanjing Normal University, China
- PAGE 1685
Tue-P-14-4 **Homophonous Phonotactic and Morphotactic Consonant Clusters in Word-Final Position**
Hannah Leykum, Sylvia Moosmüller, Wolfgang U. Dressler, Austrian Academy of Sciences, Austria
- PAGE 1690
Tue-P-14-5 **Consonant Duration and VOT as a Function of Syllable Complexity and Voicing in a Sub-Set of Spanish Clusters**
Mark Gibson¹, Ana María Fernández Planas², Adamantios Gafos³, Emily Ramirez¹
¹Rice University, USA; ²Universitat de Barcelona, Spain; ³Universität Potsdam, Germany
- PAGE 1695
Tue-P-14-6 **Hands-On Tool Producing Front Vowels for Phonetic Education: Aiming for Pronunciation Training with Tactile Sensation**
Takayuki Arai, Sophia University, Japan
- PAGE 1700
Tue-P-14-7 **Acoustics of Articulatory Constraints: Vowel Classification and Nasalization**
Indranil Dutta, Ayushi Pandey, EFL University, India

Tue-P-14 continued...

- PAGE 1705
Tue-P-14-8 **Voice-Conditioned Allophones of MOUTH and PRICE in Bahamian Creole**
Janina Kraus, LMU München, Germany
- PAGE 1710
Tue-P-14-9 **Analysis of Spatial Variation with App-Based Crowdsourced Audio Data**
Marie-José Kolly¹, Adrian Leemann², Florian Matter³
¹LIMSI, France; ²University of Cambridge, UK; ³University of Bern, Switzerland
- PAGE 1715
Tue-P-14-10 **Confusability in L2 Vowels: Analyzing the Role of Different Features**
Mátyás Jani¹, Catia Cucchiari², Roeland van Hout², Helmer Strik²
¹Pázmány Péter Catholic University, Hungary; ²Radboud Universiteit Nijmegen, The Netherlands
- PAGE 1720
Tue-P-14-11 **Perception of French Speakers' German Vowels**
Frank Zimmerer, Jürgen Trouvain, Universität des Saarlandes, Germany
- PAGE 1725
Tue-P-14-12 **Unintuitive Phonetic Behavior in Tswana Post-Nasal Stops**
Jagoda Bruni, Daniel Duran, Grzegorz Dogil, Universität Stuttgart, Germany

Poster Session 15: Speech Enhancement

Chair: Rainer Martin

Room: Hall 2, Time 14:00 – 16:00, Tuesday, September 8, 2015

- PAGE 1730
Tue-P-15-1 **Modeling Temporal Dependency for Robust Estimation of LP Model Parameters in Speech Enhancement**
Chun Hoy Wong, Tan Lee, Yu Ting Yeung, P.C. Ching, Chinese University of Hong Kong, China
- PAGE 1735
Tue-P-15-2 **Learning a Speech Manifold for Signal Subspace Speech Denoising**
Colin Vaz, Shrikanth S. Narayanan, University of Southern California, USA
- PAGE 1740
Tue-P-15-3 **An Iterative Speech Model-Based a priori SNR Estimator**
Samy Elshamy¹, Nilesch Madhu², Wouter Tirry², Tim Fingscheidt¹
¹Technische Universität Braunschweig, Germany; ²NXP Software, Belgium
- PAGE 1745
Tue-P-15-4 **Multi-Resolution Stacking for Speech Separation Based on Boosted DNN**
Xiao-Lei Zhang, DeLiang Wang, Ohio State University, USA
- PAGE 1750
Tue-P-15-5 **Least Squares Estimate of the Initial Phases in STFT Based Speech Enhancement**
Sidsel Marie Nørholm¹, Martin Krawczyk-Becker², Timo Gerkmann², Steven van de Par², Jesper Rindom Jensen¹, Mads Græsbøll Christensen¹
¹Aalborg University, Denmark; ²Carl von Ossietzky Universität Oldenburg, Germany
- PAGE 1755
Tue-P-15-6 **Enhancement of Non-Stationary Speech Using Harmonic Chirp Filters**
Sidsel Marie Nørholm, Jesper Rindom Jensen, Mads Græsbøll Christensen, Aalborg University, Denmark
- PAGE 1760
Tue-P-15-7 **Text-Informed Speech Enhancement with Deep Neural Networks**
Keisuke Kinoshita, Marc Delcroix, Atsunori Ogawa, Tomohiro Nakatani, NTT Corporation, Japan

Tue-P-15 continued...

- PAGE 1765
Tue-P-15-8 **Complex Tensor Factorization in Modulation Frequency Domain for Single-Channel Speech Enhancement**
Shogo Masaya, Masashi Unoki, JAIST, Japan
- PAGE 1770
Tue-P-15-9 **Systematic Integration of Acoustic Echo Canceller and Noise Reduction Modules for Voice Communication Systems**
Hyeonjoo Kang¹, JeeSok Lee¹, Soonho Baek², Hong-Goo Kang¹
¹Yonsei University, Korea; ²Samsung Electronics, Korea
- PAGE 1775
Tue-P-15-10 **DNN-Based Residual Echo Suppression**
Chul Min Lee¹, Jong Won Shin², Nam Soo Kim¹
¹Seoul National University, Korea; ²GIST, Korea
- PAGE 1780
Tue-P-15-11 **Codebook-Based Speech Enhancement Using Markov Process and Speech-Presence Probability**
Qi He, Changchun Bao, Feng Bao, Beijing University of Technology, China
- PAGE 1785
Tue-P-15-12 **On Optimal Smoothing in Minimum Statistics Based Noise Tracking**
Aleksej Chinaev, Reinhold Haeb-Umbach, Universität Paderborn, Germany
- PAGE 1790
Tue-P-15-13 **A Data-Driven Speech Enhancement Method Based on Modeled Long-Range Temporal Dynamics**
Yue Hao, Changchun Bao, Feng Bao, Feng Deng, Beijing University of Technology, China
- PAGE 1795
Tue-P-15-14 **Improved Phase Reconstruction in Single-Channel Speech Separation**
Florian Mayer, Pejman Mowlae, Technische Universität Graz, Austria

Poster Session 16: Spoken Language Understanding

Chair: Roger K. Moore

Room: Foyer, Time 14:00 - 16:00, Tuesday, September 8, 2015

- PAGE 1800
Tue-P-16-1 **Dialog State Tracking Using Long Short-Term Memory Neural Networks**
Xiaohao Yang, Jia Liu, Tsinghua University, China
- PAGE 1805
Tue-P-16-2 **Detecting Repetitions in Spoken Dialogue Systems Using Phonetic Distances**
*José Lopes¹, Giampiero Salvi¹, Gabriel Skantze¹, Alberto Abad², Joakim Gustafson¹,
Fernando Batista², Raveesh Meena¹, Isabel Trancoso²*
¹KTH, Sweden; ²INESC-ID Lisboa, Portugal
- PAGE 1810
Tue-P-16-3 **Multi-Language Hypotheses Ranking and Domain Tracking for Open Domain
Dialogue Systems**
Paul A. Crook, Jean-Philippe Robichaud, Ruhi Sarikaya, Microsoft, USA
- PAGE 1815
Tue-P-16-4 **Measuring Mimicry in Task-Oriented Conversations: Degree of Mimicry is Related
to Task Difficulty**
*Vijay Solanki, Alessandro Vinciarelli, Jane Stuart-Smith, Rachel Smith, University of
Glasgow, UK*
- PAGE 1820
Tue-P-16-5 **Auto-Imputing Radial Basis Functions for Neural-Network Turn-Taking Models**
Kornel Laskowski, Carnegie Mellon University, USA
- PAGE 1825
Tue-P-16-6 **Effect of Gender and Call Duration on Customer Satisfaction in Call Center Big
Data**
*Quim Llimona, Jordi Luque, Xavier Anguera, Zoraida Hidalgo, Souneil Park,
Nuria Oliver, Telefónica I+D, Spain*
- PAGE 1830
Tue-P-16-7 **Using Profile Similarity to Measure Agreement in Personality Perception**
Zoraida Callejas¹, David Griol²
¹Universidad de Granada, Spain; ²Universidad Carlos III de Madrid, Spain

INTERSPEECH 2015 Table of Contents

Tue-P-16 continued...

- PAGE 1835
Tue-P-16-8 **Relieving Mental Stress of Speakers Using a Tele-Operated Robot in Foreign
Language Speech Education**
*Shizuka Nakamura, Miki Watanabe, Yuichiro Yoshikawa, Kohei Ogawa,
Hiroshi Ishiguro, Osaka University, Japan*
- PAGE 1839
Tue-P-16-9 **Backward Mimicry and Forward Influence in Prosodic Contour Choice in Standard
American English**
Agustín Gravano¹, Štefan Beňuš², Rivka Levitan³, Julia Hirschberg⁴
*¹Universidad de Buenos Aires, Argentina; ²UKF, Slovak Republic; ³CUNY Brooklyn
College, USA; ⁴Columbia University, USA*
- PAGE 1844
Tue-P-16-10 **The Role of Speakers and Context in Classifying Competition in Overlapping
Speech**
*Shammur Absar Chowdhury, Morena Danieli, Giuseppe Riccardi, Università di Trento,
Italy*
- PAGE 1849
Tue-P-16-11 **Automatic Detection and Annotation of Disfluencies in Spoken French Corpora**
George Christodoulides¹, Mathieu Avanzi²
¹Université catholique de Louvain, Belgium; ²University of Cambridge, UK
- PAGE 1854
Tue-P-16-12 **Clustering Novel Intents in a Conversational Interaction System with Semantic
Parsing**
Dilek Hakkani-Tür, Yun-Cheng Ju, Geoffrey Zweig, Gokhan Tur, Microsoft, USA
- PAGE 1859
Tue-P-16-13 **Semantic Analysis of Spoken Input Using Markov Logic Networks**
Vladimir Despotovic¹, Oliver Walter², Reinhold Haeb-Umbach²
¹University of Belgrade, Serbia; ²Universität Paderborn, Germany

Tue-P-16 continued...

PAGE 1864
Tue-P-16-14

Hierarchical Discriminative Model for Spoken Language Understanding Based on Convolutional Neural Network

Jan Švec, Adam Chýlek, Luboš Šmídl, University of West Bohemia, Czech Republic

PAGE 1869
Tue-P-16-15

Learning Semantic Hierarchy with Distributed Representations for Unsupervised Spoken Language Understanding

Yun-Nung Chen, William Yang Wang, Alexander I. Rudnicky, Carnegie Mellon University, USA

Show and Tell Session 3

Chair: Hemant Patil

Room: Seminar 1, Time 14:00 - 16:00, Tuesday, September 8, 2015

PAGE 1874
Show&Tell-3-1

Phontasia — A Game for Training German Orthography

Kay Berkling¹, Nadine Pflaumer², Alexei Coyplove³

¹DHBW, Germany; ²Logopraxen, Germany; ³Inline Internet Online, Germany

PAGE 1876
Show&Tell-3-2

E-Commu-Book: An Assistive Technology for Users with Speech Impairments

Ka Ho Wong, Wai Kim Leung, Helen Meng, Chinese University of Hong Kong, China

PAGE 1878
Show&Tell-3-3

Swiss GraphoGame: Concept and Design Presentation of a Computerised Reading Intervention for Children with High Risk for Poor Reading Outcomes

Martina Röthlisberger¹, Iliana I. Karipidis¹, Georgette Pleisch¹, Volker Dellwo¹, Ulla Richardson², Silvia Brem¹

¹Universität Zürich, Switzerland; ²University of Jyväskylä, Finland

PAGE 1880
Show&Tell-3-4

Neolexon — A Therapy App for Patients with Aphasia

Jakob Pfab¹, Hanna Jakob², Mona Späth², Christoph Draxler¹

¹LMU München, Germany; ²Städtisches Klinikum München, Germany

PAGE 1882
Show&Tell-3-5

Acoustic Stress Detection for Improved Navigation of Educational Videos

Sonal Patil, Harish Arsikere, Om Deshmukh, Xerox Research Center India, India

PAGE 1884
Show&Tell-3-6

Multimodal Read-Aloud eBooks for Language Learning

Xavier Anguera, Sinkronigo, Spain

PAGE 1886
Show&Tell-3-7

Speech Technologies for African Languages: Example of a Multilingual Calculator for Education

Laurent Besacier¹, Elodie Gauthier¹, Mathieu Mangeot¹, Philippe Bretier²,

Paul Bagshaw², Olivier Rosec², Thierry Moudenc², François Pellegrino³,

Sylvie Voisin³, Egidio Marsico³, Pascal Nocera⁴

¹LIG (UMR 5217), France; ²Voxygen, France; ³DDL (UMR 5596), France; ⁴LIA, France

Special Event: Privacy Issues in Speech Data Collection and Usage

Chair: Michael Wagner

Room: Large Hall, Time 16:30 - 17:30, Tuesday, September 8, 2015

Oral Session 25: Phonetic Recognition: Novel Approaches and Understanding

Chairs: Lukáš Burget, Yifan Gong

Room: Large Hall, Time 09:00 - 11:00, Wednesday, September 9, 2015

PAGE 1888
Wed-O-25-1
09:00 - 09:20

Time-Frequency Kernel-Based CNN for Speech Recognition

*Tuo Zhao*¹, *Yunxin Zhao*¹, *Xin Chen*²

¹University of Missouri, USA; ²Pearson Knowledge Technologies, USA

PAGE 1893
Wed-O-25-2
09:20 - 09:40

Consonant Recognition with Continuous-State Hidden Markov Models and Perceptually-Motivated Features

Philip Weber, *Colin J. Champion*, *S.M. Houghton*, *Peter Jančovič*, *Martin Russell*,
University of Birmingham, UK

PAGE 1898
Wed-O-25-3
09:40 - 10:00

Investigating Factor Analysis Features for Deep Neural Networks in Noisy Speech Recognition

Sriram Ganapathy, *Samuel Thomas*, *Dimitrios Dimitriadis*, *Steven Rennie*, IBM T.J. Watson Research Center, USA

PAGE 1903
Wed-O-25-4
10:00 - 10:20

Ensemble of Gaussian Mixture Localized Neural Networks with Application to Phone Recognition

Ruchir Travadi, *Shrikanth S. Narayanan*, University of Southern California, USA

PAGE 1908
Wed-O-25-5
10:20 - 10:40

DNN Derived Filters for Processing of Modulation Spectrum of Speech

Jan Pešán, *Lukáš Burget*, *Hynek Hermansky*, *Karel Veselý*, Brno University of Technology, Czech Republic

PAGE 1912
Wed-O-25-6
10:40 - 11:00

Exploring How Deep Neural Networks Form Phonemic Categories

*Tasha Nagamine*¹, *Michael L. Seltzer*², *Nima Mesgarani*¹

¹Columbia University, USA; ²Microsoft, USA

INTERSPEECH 2015 Table of Contents

Oral Session 26: Varieties of Speech

Chairs: Mattias Heldner, Petra Wagner

Room: Conference 1, Time 09:00 - 11:00, Wednesday, September 9, 2015

PAGE 1917
Wed-O-26-1
09:00 - 09:20

Pronunciation Accuracy and Intelligibility of Non-Native Speech

Anastassia Loukina, *Melissa Lopez*, *Keelan Evanini*, *David Suendermann-Oeft*,
Alexei V. Ivanov, *Klaus Zechner*, Educational Testing Service, USA

PAGE 1922
Wed-O-26-2
09:20 - 09:40

Productions of /h/ in German: French vs. German Speakers

Frank Zimmerer, *Jürgen Trouvain*, Universität des Saarlandes, Germany

PAGE 1927
Wed-O-26-3
09:40 - 10:00

German Non-Native Realizations of French Voiced Fricatives in Final Position of a Group of Words

Anne Bonneau, *Martine Cadot*, LORIA, France

PAGE 1932
Wed-O-26-4
10:00 - 10:20

From Newcastle MOUTH to Aussie Ears: Australians' Perceptual Assimilation and Adaptation for Newcastle UK Vowels

*Catherine T. Best*¹, *Jason A. Shaw*¹, *Gerard Docherty*², *Bronwen G. Evans*³,
*Paul Foulkes*⁴, *Jennifer Hay*⁵, *Jalal Al-Tamimi*⁶, *Katharine Mair*³, *Karen E. Mulak*¹,
*Sophie Wood*⁴

¹University of Western Sydney, Australia; ²Griffith University, Australia; ³University College London, UK; ⁴University of York, UK; ⁵University of Canterbury, New Zealand; ⁶Newcastle University, UK

PAGE 1937
Wed-O-26-5
10:20 - 10:40

Wubuy Coronal Stop Perception by Speakers of Three Dialects of Bangla

*Rikke Louise Bundgaard-Nielsen*¹, *Brett Baker*², *Olga Maxwell*², *Janet Fletcher*²

¹University of Western Sydney, Australia; ²University of Melbourne, Australia

PAGE 1942
Wed-O-26-6
10:40 - 11:00

Using Melody Metrics to Compare English Speech Read by Native Speakers and by L2 Chinese Speakers from Shanghai

*Daniel Hirst*¹, *Hongwei Ding*²

¹LPL, France; ²Shanghai Jiao Tong University, China

Oral Session 27: Conversational Interaction

Chairs: Alan W. Black, Nigel G. Ward

Room: Conference 2+3, Time 09:00 - 11:00, Wednesday, September 9, 2015

PAGE 1947
Wed-O-27-1
09:00 - 09:20

Predicting Therapist Empathy in Motivational Interviews Using Language Features Inspired by Psycholinguistic Norms

James Gibson¹, Nikolaos Malandrakis¹, Francisco Romero¹, David C. Atkins², Shrikanth S. Narayanan¹

¹University of Southern California, USA; ²University of Washington, USA

PAGE 1952
Wed-O-27-2
09:20 - 09:40

Therapy Language Analysis Using Automatically Generated Psycholinguistic Norms

Nikolaos Malandrakis, Shrikanth S. Narayanan, University of Southern California, USA

PAGE 1957
Wed-O-27-3
09:40 - 10:00

A Dynamic Model for Behavioral Analysis of Couple Interactions Using Acoustic Features

Wei Xia¹, James Gibson¹, Bo Xiao¹, Brian Baucom², Panayiotis G. Georgiou¹

¹University of Southern California, USA; ²University of Utah, USA

PAGE 1962
Wed-O-27-4
10:00 - 10:20

Analysis and Modeling of the Role of Laughter in Motivational Interviewing Based Psychotherapy Conversations

Rahul Gupta¹, Theodora Chaspari¹, Panayiotis G. Georgiou¹, David C. Atkins², Shrikanth S. Narayanan¹

¹University of Southern California, USA; ²University of Washington, USA

PAGE 1967
Wed-O-27-5
10:20 - 10:40

The Discourse Value of Social Signals at Topic Change Moments

Francesca Bonin, Nick Campbell, Carl Vogel, Trinity College Dublin, Ireland

PAGE 1972
Wed-O-27-6
10:40 - 11:00

Automatic Detection of Uncertainty in Spontaneous German Dialogue

Tobias Schrank, Barbara Schuppler, Technische Universität Graz, Austria

INTERSPEECH 2015 Table of Contents

Oral Session 28: Speech and Audio Segmentation and Classification

Chairs: Haizhou Li, Koichi Shinoda

Room: Conference 4+5, Time 09:00 - 11:00, Wednesday, September 9, 2015

PAGE 1977
Wed-O-28-1
09:00 - 09:20

Face Reading from Speech — Predicting Facial Action Units from Audio Cues

Fabien Ringeval¹, Erik Marchi¹, Marc Mehu², Klaus Scherer³, Björn Schuller³

¹Technische Universität München, Germany; ²Webster Vienna Private University, Austria; ³Université de Genève, Switzerland

PAGE 1982
Wed-O-28-2
09:20 - 09:40

A New Front-End for Classification of Non-Speech Sounds: A Study on Human Whistle

Mahesh Kumar Nandwana, Hynek Bořil, John H.L. Hansen, University of Texas at Dallas, USA

PAGE 1987
Wed-O-28-3
09:40 - 10:00

Robust Features for Sonorant Segmentation in Continuous Speech

Sri Harsha Dumpala, Bhanu Teja Nellore, Raghu Ram Nevali, Suryakanth V. Gangashetty, B. Yegnanarayana, IIT Hyderabad, India

PAGE 1992
Wed-O-28-4
10:00 - 10:20

Reduction of Reverberation Effects in the MFCC Modulation Spectrum for Improved Classification of Acoustic Signals

Sebastian Gergen, Anil Nagathil, Rainer Martin, Ruhr-Universität Bochum, Germany

PAGE 1997
Wed-O-28-5
10:20 - 10:40

Spiking Neural Networks and the Generalised Hough Transform for Speech Pattern Detection

*Jonathan Dennis, Huy Dat Tran, Haizhou Li, A*STAR, Singapore*

PAGE 2002
Wed-O-28-6
10:40 - 11:00

Acoustic Event Recognition Using Dominant Spectral Basis Vectors

Woohyun Choi¹, Sangwook Park¹, David K. Han², Hanseok Ko¹

¹Korea University, Korea; ²ONR, USA

Oral Session 29: Spoken Dialogue Systems

Chairs: Satoshi Nakamura, Alexander Rudnicky

Room: Conference 6, Time 09:00 – 11:00, Wednesday, September 9, 2015

- PAGE 2007
Wed-O-29-1
09:00 – 09:20
- Learning from Real Users: Rating Dialogue Success with Neural Networks for Reinforcement Learning in Spoken Dialogue Systems**
Pei-Hao Su, David Vandyke, Milica Gašić, Dongho Kim, Nikola Mrkšić, Tsung-Hsien Wen, Steve Young, University of Cambridge, UK
- PAGE 2012
Wed-O-29-2
09:20 – 09:40
- A Framework to Develop Context-Aware Adaptive Dialogue System**
David Griol¹, Zoraida Callejas², Ramón López-Cózar²
¹Universidad Carlos III de Madrid, Spain; ²Universidad de Granada, Spain
- PAGE 2017
Wed-O-29-3
09:40 – 10:00
- A Proposal to Develop Domain and Subtask-Adaptive Dialog Management Models**
David Griol¹, Zoraida Callejas²
¹Universidad Carlos III de Madrid, Spain; ²Universidad de Granada, Spain
- PAGE 2022
Wed-O-29-4
10:00 – 10:20
- Hypotheses Ranking and State Tracking for a Multi-Domain Dialog System Using Multiple ASR Alternates**
Omar Zia Khan, Jean-Philippe Robichaud, Paul A. Crook, Ruhi Sarikaya, Microsoft, USA
- PAGE 2027
Wed-O-29-5
10:20 – 10:40
- An Entropy Minimization Framework for Goal-Driven Dialogue Management**
Ji Wu¹, Miao Li¹, Chin-Hui Lee²
¹Tsinghua University, China; ²Georgia Institute of Technology, USA
- PAGE 2032
Wed-O-29-6
10:40 – 11:00
- Context-Dependent Error Correction of Spoken Referring Expressions**
Ingrid Zukerman, Andisheh Partovi, Su Nam Kim, Monash University, Australia

Special Session 5: Automatic Speaker Verification Spoofing and Countermeasures (ASVspoof 2015)

Chairs: Tomi Kinnunen, Zhizheng Wu

Room: Hall 1, Time 09:00 – 11:00, Wednesday, September 9, 2015

- Wed-SP5-1
09:00 – 09:30
- Introductory Talk by the Organizers**
Zhizheng Wu¹, Tomi Kinnunen²
¹University of Edinburgh, UK; ²University of Eastern Finland, Finland
- Poster Presentations
- PAGE 2037
Wed-SP5-1
- ASVspoof 2015: The First Automatic Speaker Verification Spoofing and Countermeasures Challenge (Poster)**
Zhizheng Wu¹, Tomi Kinnunen², Nicholas Evans³, Junichi Yamagishi¹, Cemal Haniççi², Md. Sahidullah², Aleksandr Sizov²
¹University of Edinburgh, UK; ²University of Eastern Finland, Finland; ³EURECOM, France
- PAGE 2042
Wed-SP5-2
- The AHOLAB RPS SSD Spoofing Challenge 2015 Submission (Poster)**
Jon Sanchez, Ibon Saratxaga, Inma Hernaez, Eva Navas, D. Erro, Universidad del País Vasco, Spain
- PAGE 2047
Wed-SP5-3
- Human vs Machine Spoofing Detection on Wideband and Narrowband Data (Poster)**
Mirjam Wester, Zhizheng Wu, Junichi Yamagishi, University of Edinburgh, UK
- PAGE 2052
Wed-SP5-4
- Spoofing Speech Detection Using High Dimensional Magnitude and Phase Features: The NTU Approach for ASVspoof 2015 Challenge (Poster)**
Xiong Xiao¹, Xiaohai Tian², Steven Du¹, Haihua Xu¹, Eng Siong Chng¹, Haizhou Li¹
¹TL@NTU, Singapore; ²NTU, Singapore

Wed-SP5 continued...

- PAGE 2057
Wed-SP5-5 **Classifiers for Synthetic Speech Detection: A Comparison (Poster)**
Cemal Hanilçi, Tomi Kinnunen, Md. Sahidullah, Aleksandr Sizov, University of Eastern Finland, Finland
- PAGE 2062
Wed-SP5-6 **Combining Evidences from Mel Cepstral, Cochlear Filter Cepstral and Instantaneous Frequency Features for Detection of Natural vs. Spoofed Speech (Poster)**
Tanvina B. Patel, Hemant A. Patil, DA-IICT, India
- PAGE 2067
Wed-SP5-7 **Spoofing Detection with DNN and One-Class SVM for the ASVspoof 2015 Challenge (Poster)**
Jesús Villalba, Antonio Miguel, Alfonso Ortega, Eduardo Lleida, Universidad de Zaragoza, Spain
- PAGE 2072
Wed-SP5-8 **Development of CRIM System for the Automatic Speaker Verification Spoofing and Countermeasures Challenge 2015 (Poster)**
Md. Jahangir Alam, Patrick Kenny, Gautam Bhattacharya, Themis Stafylakis, CRIM, Canada
- PAGE 2077
Wed-SP5-9 **Spoofing Countermeasure Based on Analysis of Linear Prediction Error (Poster)**
Artur Janicki, Warsaw University of Technology, Poland
- PAGE 2082
Wed-SP5-10 **Simultaneous Utilization of Spectral Magnitude and Phase Information to Extract Supervectors for Speaker Verification Anti-Spoofing (Poster)**
Yi Liu¹, Yao Tian¹, Liang He¹, Jia Liu¹, Michael T. Johnson²
¹Tsinghua University, China; ²Marquette University, USA
- PAGE 2087
Wed-SP5-11 **A Comparison of Features for Synthetic Speech Detection (Poster)**
Md. Sahidullah, Tomi Kinnunen, Cemal Hanilçi, University of Eastern Finland, Finland

Wed-SP5 continued...

- PAGE 2092
Wed-SP5-12 **Relative Phase Information for Detecting Human Speech and Spoofed Speech (Poster)**
Longbiao Wang¹, Yohei Yoshida¹, Yuta Kawakami¹, Seiichi Nakagawa²
¹Nagaoka University of Technology, Japan; ²Toyohashi University of Technology, Japan
- PAGE 2097
Wed-SP5-13 **Robust Deep Feature for Spoofing Detection — The SJTU System for ASVspoof 2015 Challenge (Poster)**
Nanxin Chen, Yanmin Qian, Heinrich Dinkel, Bo Chen, Kai Yu, Shanghai Jiao Tong University, China
- Wed-SP5-D
10:30 – 11:00 **Open Discussion and Future Plans**
Junichi Yamagishi¹, Nicholas Evans²
¹University of Edinburgh, UK; ²EURECOM, France

Poster Session 17: Acoustic Modeling and Decoding Methods for Speech Recognition

Chair: Thomas Hain

Room: Hall 2, Time 09:00 - 11:00, Wednesday, September 9, 2015

- PAGE 2102
Wed-P-17-1 **Applying GPGPU to Recurrent Neural Network Language Model Based Fast Network Search in the Real-Time LVCSR**
Kyungmin Lee, Chiyoun Park, Ilhwan Kim, Namhoon Kim, Jaewon Lee, Samsung Electronics, Korea
- PAGE 2107
Wed-P-17-2 **Real-Time Integration of Dynamic Context Information for Improving Automatic Speech Recognition**
Youssef Oualil¹, Marc Schulder¹, Hartmut Helmke², Anna Schmidt¹, Dietrich Klakow¹
¹Universität des Saarlandes, Germany; ²DLR, Germany
- PAGE 2112
Wed-P-17-3 **Rapid Vocabulary Addition to Context-Dependent Decoder Graphs**
Cyril Allauzen, Michael Riley, Google, USA
- PAGE 2117
Wed-P-17-4 **Modeling Phonetic Context with Non-Random Forests for Speech Recognition**
Hainan Xu, Guoguo Chen, Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, USA
- PAGE 2122
Wed-P-17-5 **Ant Colony Algorithm Applied to Automatic Speech Recognition Graph Decoding**
Benjamin Lecouteux, Didier Schwab, LIG (UMR 5217), France
- PAGE 2127
Wed-P-17-6 **Garbage Modeling for On-Device Speech Recognition**
Christophe Van Gysel¹, Leonid Velikovich², Ian McGraw², Françoise Beaufays²
¹University of Amsterdam, The Netherlands; ²Google, USA

Wed-P-17 continued...

- PAGE 2132
Wed-P-17-7 **A Comparative Study of BNF and DNN Multilingual Training on Cross-Lingual Low-Resource Speech Recognition**
Haihua Xu, Van Hai Do, Xiong Xiao, Eng Siong Chng, TL@NTU, Singapore
- PAGE 2137
Wed-P-17-8 **Neural Higher-Order Factors in Conditional Random Fields for Phoneme Classification**
Martin Ratajczak¹, Sebastian Tschitschek², Franz Pernkopf¹
¹Technische Universität Graz, Austria; ²ETH Zürich, Switzerland
- PAGE 2142
Wed-P-17-9 **Stacked Auto-Encoder for ASR Error Detection and Word Error Rate Prediction**
Shahab Jalalvand, Daniele Falavigna, FBK, Italy

Poster Session 18: Speech Production Measurements and Analyses

Chairs: Prasanta Ghosh, Frank Rudzicz

Room: Hall 2, Time 09:00 – 11:00, Wednesday, September 9, 2015

- PAGE 2147
Wed-P-18-1 **Estimation of the Air-Tissue Boundaries of the Vocal Tract in the Mid-Sagittal Plane from Electromagnetic Articulograph Data**
Satyabrata Parida¹, Ashok Kumar Pattem², Prasanta Kumar Ghosh²
¹IIT Kharagpur, India; ²Indian Institute of Science, India
- PAGE 2152
Wed-P-18-2 **A New Italian Dataset of Parallel Acoustic and Articulatory Data**
Claudia Canevari, Leonardo Badino, Luciano Fadiqa, Istituto Italiano di Tecnologia, Italy
- PAGE 2157
Wed-P-18-3 **Error Analysis of Extracted Tongue Contours from 2D Ultrasound Images**
Tamás Gábor Csapó¹, Steven M. Lulich²
¹BME, Hungary; ²Indiana University, USA
- PAGE 2162
Wed-P-18-4 **Accuracy of a Markerless Acquisition Technique for Studying Speech Articulators**
Andrea Bandini¹, Slim Ouni², Piero Cosi³, Silvia Orlandi¹, Claudia Manfredi¹
¹Università di Firenze, Italy; ²LORIA, France; ³CNR-ISTC, Italy
- PAGE 2167
Wed-P-18-5 **Measuring Oral and Nasal Airflow in Production of Chinese Plosive**
Yujie Chi, Kiyoshi Honda, Jianguo Wei, Hui Feng, Jianwu Dang, Tianjin University, China
- PAGE 2172
Wed-P-18-6 **Enhanced Videokymographic Data Analysis Based on Vocal Folds Dynamics Modeling**
Carlo Drioli, Gian Luca Foresti, Università di Udine, Italy
- PAGE 2177
Wed-P-18-7 **Interpolation of Tongue Fleshpoint Kinematics from Combined EMA Position and Orientation Data**
Andrew J. Kolb, Michael T. Johnson, Jeffrey Berry, Marquette University, USA

INTERSPEECH 2015 Table of Contents

Wed-P-18 continued...

- PAGE 2182
Wed-P-18-8 **A New Technique for Assessing Glottal Dynamics in Speech and Singing by Means of Optical-Flow Computation**
Gustavo Andrade-Miranda¹, Nathalie Henrich Bernardoni², Juan Ignacio Godino-Llorente¹
¹Universidad Politécnica de Madrid, Spain; ²GIPSA, France
- PAGE 2187
Wed-P-18-9 **On the Incompatibility of Trilling and Palatalization: A Single-Subject Study of Sustained Apical and Uvular Trills**
Alexei Kochetov, Phil Howson, University of Toronto, Canada
- PAGE 2192
Wed-P-18-10 **Articulatory Movement Prediction Using Deep Bidirectional Long Short-Term Memory Based Recurrent Neural Networks and Word/Phone Embeddings**
Pengcheng Zhu, Lei Xie, Yunlin Chen, Northwestern Polytechnical University, China

Poster Session 19: Speech Synthesis 2

Chair: *Cassia Valentini-Botinhao*

Room: *Foyer, Time 09:00 - 11:00, Wednesday, September 9, 2015*

- PAGE 2197
Wed-P-19-1 **Combining Extreme Learning Machine and Decision Tree for Duration Prediction in HMM Based Speech Synthesis**
Yang Wang, Minghao Yang, Zhengqi Wen, Jianhua Tao, Chinese Academy of Sciences, China
- PAGE 2202
Wed-P-19-2 **F0 Parameterization of Glottalized Tones for HMM-Based Vietnamese TTS**
Duy Khanh Ninh, Yoichi Yamashita, Ritsumeikan University, Japan
- PAGE 2207
Wed-P-19-3 **Deep Neural Network Context Embeddings for Model Selection in Rich-Context HMM Synthesis**
Thomas Merritt, Junichi Yamagishi, Zhizheng Wu, Oliver Watts, Simon King, University of Edinburgh, UK
- PAGE 2212
Wed-P-19-4 **An Investigation of Context Clustering for Statistical Speech Synthesis with Deep Neural Network**
Bo Chen, Zhehuai Chen, Jiachen Xu, Kai Yu, Shanghai Jiao Tong University, China
- PAGE 2217
Wed-P-19-5 **Sentence-Level Control Vectors for Deep Neural Network Speech Synthesis**
Oliver Watts, Zhizheng Wu, Simon King, University of Edinburgh, UK
- PAGE 2222
Wed-P-19-6 **Micro-Structure of Disfluencies: Basics for Conversational Speech Synthesis**
Simon Betz, Petra Wagner, David Schlangen, Universität Bielefeld, Germany
- PAGE 2227
Wed-P-19-7 **Using Automatic Stress Extraction from Audio for Improved Prosody Modelling in Speech Synthesis**
*György Szaszák¹, András Beke², Gábor Olasz¹, Bálint Pál Tóth¹
¹BME, Hungary; ²Hungarian Academy of Sciences, Hungary*

Wed-P-19 continued...

- PAGE 2232
Wed-P-19-8 **Reconstructing Voices Within the Multiple-Average-Voice-Model Framework**
*Pierre Lanchantin¹, Christophe Veaux², Mark J.F. Gales¹, Simon King², Junichi Yamagishi²
¹University of Cambridge, UK; ²University of Edinburgh, UK*
- PAGE 2237
Wed-P-19-9 **HMM Based Myanmar Text to Speech System**
*Ye Kyaw Thu¹, Win Pa Pa², Jinfu Ni¹, Yoshinori Shiga¹, Andrew Finch¹, Chiori Hori¹, Hisashi Kawai¹, Eiichiro Sumita¹
¹NICT, Japan; ²UCSY, Myanmar*
- PAGE 2242
Wed-P-19-10 **Multiple Feed-Forward Deep Neural Networks for Statistical Parametric Speech Synthesis**
*Shinji Takaki¹, SangJin Kim², Junichi Yamagishi¹, JongJin Kim²
¹NII, Japan; ²Naver Corporation, Korea*

Special Event: Standardization Brown-Bag

Chair: David Suendermann-Oeft

Room: Large Hall, Time 12:30 - 14:00, Wednesday, September 9, 2015

INTERSPEECH 2015 Table of Contents

Oral Session 30: Spoken Translation & Speech-to-speech

Chairs: Roland Kuhn, Frank Seide

Room: Large Hall, Time 14:00 - 16:00, Wednesday, September 9, 2015

PAGE 2247
Wed-O-30-1
14:00 - 14:20

Adapting Machine Translation Models Toward Misrecognized Speech with Text-to-Speech Pronunciation Rules and Acoustic Confusability

Nicholas Ruiz¹, Qin Gao², William Lewis², Marcello Federico¹
¹FBK, Italy; ²Microsoft, USA

PAGE 2252
Wed-O-30-2
14:20 - 14:40

“Speech is Silver, but Silence is Golden”: Improving Speech-to-Speech Translation Performance by Slashing Users Input

Frederic Bechet, Benoit Favre, Mickael Rouvier, LIF (UMR 7279), France

PAGE 2257
Wed-O-30-3
14:40 - 15:00

A Study on the Stability and Effectiveness of Features in Quality Estimation for Spoken Language Translation

Raymond W.M. Ng, Kashif Shah, Lucia Specia, Thomas Hain, University of Sheffield, UK

PAGE 2262
Wed-O-30-4
15:00 - 15:20

Efficient Language Model Adaptation for Automatic Speech Recognition of Spoken Translations

Joris Pelemans¹, Tom Vanallemeersch¹, Kris Demuyne², Hugo Van hamme¹, Patrick Wambacq¹

¹Katholieke Universiteit Leuven, Belgium; ²Ghent University, Belgium

PAGE 2267
Wed-O-30-5
15:20 - 15:40

Speed or Accuracy? A Study in Evaluation of Simultaneous Speech Translation

Takashi Mieno, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura, NAIST, Japan

PAGE 2272
Wed-O-30-6
15:40 - 16:00

Large Scale Speech-to-Text Translation with Out-of-Domain Corpora Using Better Context-Based Models and Domain Adaptation

Marcin Junczys-Dowmunt¹, Paweł Przybyś¹, Arleta Staszuk¹, Eun-Kyoung Kim², Jaewon Lee²

¹Samsung Poland R&D Center, Poland; ²Samsung Electronics, Korea

Oral Session 31: Voice Activity Detection

Chairs: Ma Bin, Nam Soo Kim

Room: Conference 1, Time 14:00 – 16:00, Wednesday, September 9, 2015

- PAGE 2277
Wed-O-31-1
14:00 – 14:20
- A Statistical Model-Based Voice Activity Detection Using Multiple DNNs and Noise Awareness**
Inyoung Hwang, Jaeseong Sim, Sang-Hyeon Kim, Kwang-Sub Song, Joon-Hyuk Chang, Hanyang University, Korea
- PAGE 2282
Wed-O-31-2
14:20 – 14:40
- A Universal VAD Based on Jointly Trained Deep Neural Networks**
Qing Wang¹, Jun Du¹, Xiao Bao¹, Zi-Rui Wang¹, Li-Rong Dai¹, Chin-Hui Lee²
¹USTC, China; ²Georgia Institute of Technology, USA
- PAGE 2287
Wed-O-31-3
14:40 – 15:00
- Spectrographic Speech Mask Estimation Using the Time-Frequency Correlation of Speech Presence**
Ge Zhan, Zhaoqiong Huang, Dongwen Ying, Jielin Pan, Yonghong Yan, Chinese Academy of Sciences, China
- PAGE 2292
Wed-O-31-4
15:00 – 15:20
- Complete-Linkage Clustering for Voice Activity Detection in Audio and Visual Speech**
Houman Ghaemmaghami, David Dean, Shahram Kalantari, Sridha Sridharan, Clinton Fookes, Queensland University of Technology, Australia
- PAGE 2297
Wed-O-31-5
15:20 – 15:40
- A Model Based Voice Activity Detector for Noisy Environments**
Kaavya Sriskandaraja, Vidhyasaharan Sethu, Phu Ngoc Le, Eliathamby Ambikairajah, University of New South Wales, Australia
- PAGE 2302
Wed-O-31-6
15:40 – 16:00
- An Unsupervised Visual-Only Voice Activity Detection Approach Using Temporal Orofacial Features**
Fei Tao, John H.L. Hansen, Carlos Busso, University of Texas at Dallas, USA

Oral Session 32: Advances in iVector-based Speaker Verification

Chairs: Hagai Aronowitz, Douglas A. Reynolds

Room: Conference 2+3, Time 14:00 – 16:00, Wednesday, September 9, 2015

- PAGE 2307
Wed-O-32-1
14:00 – 14:20
- An I-Vector Backend for Speaker Verification**
Patrick Kenny¹, Themis Stafylakis¹, Md. Jahangir Alam¹, Marcel Kockmann²
¹CRIM, Canada; ²VoiceTrust, Canada
- PAGE 2312
Wed-O-32-2
14:20 – 14:40
- Multi-Channel Speaker Verification Based on Total Variability Modelling**
Joana Correia¹, Alessio Brutti², Alberto Abad¹
¹INESC-ID Lisboa, Portugal; ²FBK, Italy
- PAGE 2317
Wed-O-32-3
14:40 – 15:00
- SNR-Invariant PLDA Modeling for Robust Speaker Verification**
Na Li, Man-Wai Mak, Hong Kong Polytechnic University, China
- PAGE 2322
Wed-O-32-4
15:00 – 15:20
- Investigating In-Domain Data Requirements for PLDA Training**
Md. Hafizur Rahman, David Dean, Ahilan Kanagasundaram, Sridha Sridharan, Queensland University of Technology, Australia
- PAGE 2327
Wed-O-32-5
15:20 – 15:40
- Migrating i-Vectors Between Speaker Recognition Systems Using Regression Neural Networks**
Oldřej Glembek, Pavel Matějka, Oldřich Plchot, Jan Pešán, Lukáš Burget, Petr Schwarz, Brno University of Technology, Czech Republic
- PAGE 2332
Wed-O-32-6
15:40 – 16:00
- Improving PLDA Speaker Verification Using WMFD and Linear-Weighted Approaches in Limited Microphone Data Conditions**
Ahilan Kanagasundaram, David Dean, Sridha Sridharan, Queensland University of Technology, Australia

Oral Session 33: Voice Quality

Chairs: Paavo Alku, Carlos Ishi

Room: Conference 4+5, Time 14:00 - 16:00, Wednesday, September 9, 2015

- PAGE 2337
Wed-O-33-1
14:00 - 14:20
- The Relationship Between Voice Source Parameters and the Maxima Dispersion Quotient (MDQ)**
Christer Gobl, Irena Yanushevskaya, Ailbhe Ní Chasaide, Trinity College Dublin, Ireland
- PAGE 2342
Wed-O-33-2
14:20 - 14:40
- Glottal Inverse Filtering Based on Quadratic Programming**
Manu Airaksinen¹, Tom Bäckström², Paavo Alku¹
¹Aalto University, Finland; ²FAU Erlangen-Nürnberg, Germany
- PAGE 2347
Wed-O-33-3
14:40 - 15:00
- Automatic Detection of Creaky Voice Using Epoch Parameters**
N.P. Narendra, K. Sreenivasa Rao, IIT Kharagpur, India
- PAGE 2352
Wed-O-33-4
15:00 - 15:20
- Perception of Voicing in the Absence of Native Voicing Experience**
Rikke Louise Bundgaard-Nielsen¹, Brett Baker²
¹University of Western Sydney, Australia; ²University of Melbourne, Australia
- PAGE 2357
Wed-O-33-5
15:20 - 15:40
- The Relationship Between Acoustic and Perceived Intraspeaker Variability in Voice Quality**
Jody Kreiman, Soo Jin Park, Patricia A. Keating, Abeer Alwan, University of California at Los Angeles, USA
- PAGE 2361
Wed-O-33-6
15:40 - 16:00
- Perceptual Cues of Whispered Tones: Are They Really Special?**
Li Jiao¹, Qiuwu Ma¹, Ting Wang¹, Yi Xu²
¹Tongji University, China; ²University College London, UK

INTERSPEECH 2015 Table of Contents

Oral Session 34: Neural Networks for Language Modeling

Chairs: Ebru Arisoy, Tomas Mikolov

Room: Conference 6, Time 14:00 - 16:00, Wednesday, September 9, 2015

- PAGE 2366
Wed-O-34-1
14:00 - 14:20
- Multiscale Recurrent Neural Network Based Language Model**
Tsuyoshi Morioka¹, Tomoharu Iwata², Takaaki Hori², Tetsunori Kobayashi¹
¹Waseda University, Japan; ²NTT Corporation, Japan
- PAGE 2371
Wed-O-34-2
14:20 - 14:40
- Bag-of-Words Input for Long History Representation in Neural Network-Based Language Models for Speech Recognition**
Kazuki Irie, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany
- PAGE 2376
Wed-O-34-3
14:40 - 15:00
- Efficient Machine Translation Decoding with Slow Language Models**
Ahmad Emami, IBM T.J. Watson Research Center, USA
- PAGE 2380
Wed-O-34-4
15:00 - 15:20
- Latent Words Recurrent Neural Network Language Models**
Ryo Masumura¹, Taichi Asami¹, Takanobu Oba¹, Hirokazu Masataki¹, Sumitaka Sakauchi¹, Akinori Ito²
¹NTT Corporation, Japan; ²Tohoku University, Japan
- PAGE 2385
Wed-O-34-5
15:20 - 15:40
- Combining Multiple-Type Input Units Using Recurrent Neural Network for LVCSR Language Modeling**
Vataya Chunwijitra, Ananlada Chotimongkol, Chai Wutiw WATCHAI, NECTEC, Thailand
- PAGE 2390
Wed-O-34-6
15:40 - 16:00
- Prosodically-Enhanced Recurrent Neural Network Language Models**
Siva Reddy Gangireddy¹, Steve Renals¹, Yoshihiko Nankaku², Akinobu Lee²
¹University of Edinburgh, UK; ²Nagoya Institute of Technology, Japan

Special Session 6: Biosignal-based Spoken Communication

Chairs: Matthias Janke, Michael Wand

Room: Hall 1, Time 14:00 - 16:00, Wednesday, September 9, 2015

- Wed-SP6-i
14:00 - 14:05
- Welcome and Introduction**
Matthias Janke¹, Michael Wand²
¹Universität Bremen, Germany; ²IDSIA, Switzerland
- PAGE 2395
Wed-SP6-1
14:05 - 14:20
- A Comprehensive 3D Biomechanically-Driven Vocal Tract Model Including Inverse Dynamics for Speech Research**
Peter Anderson¹, Negar M. Harandi¹, Scott Moisi², Ian Stavness³, Sidney Fels¹
¹University of British Columbia, Canada; ²MPI for Psycholinguistics, The Netherlands; ³University of Saskatchewan, Canada
- PAGE 2400
Wed-SP6-2
14:20 - 14:35
- Low Frequency Ultrasonic Voice Activity Detection Using Convolutional Neural Networks**
Ian McLoughlin¹, Yan Song²
¹University of Kent, UK; ²USTC, China
- PAGE 2405
Wed-SP6-3
14:35 - 14:50
- Real-Time Control of a DNN-Based Articulatory Synthesizer for Silent Speech Conversion: A Pilot Study**
Florent Bocquelet¹, Thomas Hueber², Laurent Girin², Christophe Savariaux², Blaise Yvert¹
¹Clinatec, France; ²GIPSA, France
- PAGE 2410
Wed-SP6-4
14:50 - 15:05
- Tongue Tracking in Ultrasound Images Using EigenTongue Decomposition and Artificial Neural Networks**
Diandra Fabre, Thomas Hueber, Florent Bocquelet, Pierre Badin, GIPSA, France

INTERSPEECH 2015 Table of Contents

Wed-SP6 continued...

- PAGE 2415
Wed-SP6-5
15:05 - 15:20
- Speaker-Independent Silent Speech Recognition with Across-Speaker Articulatory Normalization and Speaker Adaptive Training**
Jun Wang, Seongjun Hahm, University of Texas at Dallas, USA
- PAGE 2420
Wed-SP6-6
15:20 - 15:35
- Codebook Clustering for Unit Selection Based EMG-to-Speech Conversion**
Lorenz Diener, Matthias Janke, Tanja Schultz, KIT, Germany
- PAGE 2425
Wed-SP6-7
15:35 - 15:50
- Flexible Tracking of Auditory Attention**
Majid Mirbagheri, Bradley Ekin, Les Atlas, Adrian K.C. Lee, University of Washington, USA
- Wed-SP6-P
15:50 - 16:00
- Panel and Discussion**
Matthias Janke¹, Michael Wand²
¹Universität Bremen, Germany; ²IDSIA, Switzerland

Poster Session 20: Robust Speech Recognition: Features, Far-field and Reverberation

Chair: Andrew Senior

Room: Hall 2, Time 14:00 - 16:00, Wednesday, September 9, 2015

- PAGE 2430
Wed-P-20-1 **A Study on Deep Neural Network Acoustic Model Adaptation for Robust Far-Field Speech Recognition**
Seyedmahdad Mirsamadi, John H.L. Hansen, University of Texas at Dallas, USA
- PAGE 2435
Wed-P-20-2 **Speech Dereverberation Using Long Short-Term Memory**
Masato Mimura, Shinsuke Sakai, Tatsuya Kawahara, Kyoto University, Japan
- PAGE 2440
Wed-P-20-3 **Reverberation Robust Acoustic Modeling Using i-Vectors with Time Delay Neural Networks**
Vijayaditya Peddinti, Guoguo Chen, Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, USA
- PAGE 2445
Wed-P-20-4 **Delta-MelSpectra Features for Noise Robustness to DNN-Based ASR Systems**
Kshitiz Kumar, Chaojun Liu, Yifan Gong, Microsoft, USA
- PAGE 2449
Wed-P-20-5 **Combating Reverberation in Large Vocabulary Continuous Speech Recognition**
Vikramjit Mitra, Julien Van Hout, Mitchell McLaren, Wen Wang, Martin Graciarena, Dimitra Vergyri, Horacio Franco, SRI International, USA
- PAGE 2454
Wed-P-20-6 **Three Ways to Adapt a CTS Recognizer to Unseen Reverberated Speech in BUT System for the ASPIRE Challenge**
Martin Karafiát, František Grézl, Lukáš Burget, Igor Szöke, Jan Černocký, Brno University of Technology, Czech Republic
- PAGE 2459
Wed-P-20-7 **Robust Parameter Estimation for Audio Declipping in Noise**
Mark J. Harvilla, Richard M. Stern, Carnegie Mellon University, USA

INTERSPEECH 2015 Table of Contents

Wed-P-20 continued...

- PAGE 2464
Wed-P-20-8 **Multi-Task Learning Deep Neural Networks for Speech Feature Denoising**
Bin Huang¹, Dengfeng Ke², Hao Zheng², Bo Xu², Yanyan Xu¹, Kaile Su³
¹Beijing Forestry University, China; ²Chinese Academy of Sciences, China; ³Griffith University, Australia
- PAGE 2469
Wed-P-20-9 **Time-Frequency Masking for Large Scale Robust Speech Recognition**
Yuxuan Wang¹, Ananya Misra², Kean K. Chin²
¹Ohio State University, USA; ²Google, USA
- PAGE 2474
Wed-P-20-10 **Efficient Use of DNN Bottleneck Features in Generalized Variable Parameter HMMs for Noise Robust Speech Recognition**
Rongfeng Su, Xurong Xie, Xunying Liu, Lan Wang, Chinese Academy of Sciences, China
- PAGE 2479
Wed-P-20-11 **Investigating Modulation Spectrogram Features for Deep Neural Network-Based Automatic Speech Recognition**
Deepak Baby, Hugo Van hamme, Katholieke Universiteit Leuven, Belgium
- PAGE 2484
Wed-P-20-12 **Deep Neural Network Based Spectral Feature Mapping for Robust Speech Recognition**
Kun Han, Yanzhang He, Deblin Bagchi, Eric Fosler-Lussier, DeLiang Wang, Ohio State University, USA

Poster Session 21: Social Signals, Assessment and Paralinguistics

Chairs: Laurence Devillers, Julia Hirschberg

Room: Hall 2, Time 14:00 - 16:00, Wednesday, September 9, 2015

- PAGE 2489
Wed-P-21-1 **Analyzing Speech Rate Entrainment and Its Relation to Therapist Empathy in Drug Addiction Counseling**
Bo Xiao¹, Zac E. Imel², David C. Atkins³, Panayiotis G. Georgiou¹, Shrikanth S. Narayanan¹
¹University of Southern California, USA; ²University of Utah, USA; ³University of Washington, USA
- PAGE 2494
Wed-P-21-2 **Agreement and Disagreement Utterance Detection in Conversational Speech by Extracting and Integrating Local Features**
Atsushi Ando, Taichi Asami, Manabu Okamoto, Hirokazu Masataki, Sumitaka Sakauchi, NTT Corporation, Japan
- PAGE 2499
Wed-P-21-3 **Still Together?: The Role of Acoustic Features in Predicting Marital Outcome**
Md. Nasir¹, Wei Xia¹, Bo Xiao¹, Brian Baucom², Shrikanth S. Narayanan¹, Panayiotis G. Georgiou¹
¹University of Southern California, USA; ²University of Utah, USA
- PAGE 2504
Wed-P-21-4 **On Evaluation Metrics for Social Signal Detection**
Gábor Gosztolya, MTA-SZTE RGAI, Hungary
- PAGE 2509
Wed-P-21-5 **Laughter and Filler Detection in Naturalistic Audio**
Lakshmish Kaushik, Abhijeet Sangwan, John H.L. Hansen, University of Texas at Dallas, USA
- PAGE 2514
Wed-P-21-6 **Automatic Formatted Transcripts for Videos**
Aasish Pappu, Amanda Stent, Yahoo!, USA

Wed-P-21 continued...

- PAGE 2519
Wed-P-21-7 **Does My Speech Rock? Automatic Assessment of Public Speaking Skills**
Lucas Azais¹, Adrien Payan¹, Tianjiao Sun¹, Guillaume Vidal¹, Tina Zhang¹, Eduardo Coutinho¹, Florian Eyben², Björn Schuller¹
¹Imperial College London, UK; ²audEERING, Germany
- PAGE 2524
Wed-P-21-8 **Verbal Intelligence Identification Based on Text Classification**
Roman Sergienko, Alexander Schmitt, Universität Ulm, Germany
- PAGE 2529
Wed-P-21-9 **A Multimodal Approach for Automatic Assessment of School Principals' Oral Presentation During Pre-Service Training Program**
Shan-Wen Hsiao¹, Hung-Ching Sun¹, Ming-Chuan Hsieh², Ming-Hsueh Tsai², Hsin-Chih Lin², Chi-Chun Lee¹
¹National Tsing Hua University, Taiwan; ²NAER, Taiwan
- PAGE 2534
Wed-P-21-10 **Are You TED Talk Material? Comparing Prosody in Professors and TED Speakers**
T.J. Tsai, University of California at Berkeley, USA
- PAGE 2539
Wed-P-21-11 **Detection of Cognitive States and Their Correlation to Speech Recognition Performance in Speech-to-Speech Machine Translation Systems**
Hayakawa Akira, Fasih Haider, Loredana Cerrato, Nick Campbell, Saturnino Luz, Trinity College Dublin, Ireland

Poster Session 22: Bandwidth Extension, Quality and Intelligibility Measures

Chair: *Tim Fingscheidt*

Room: *Foyer, Time 14:00 – 16:00, Wednesday, September 9, 2015*

- PAGE 2544
Wed-P-22-1 **Perceptual Speech Quality Dimensions in a Conversational Situation**
Friedemann Köster, Sebastian Möller, T-Labs, Germany
- PAGE 2549
Wed-P-22-2 **Multidimensional Evaluation and Predicting Overall Speech Quality**
Jens Berger, Anna Llagostera, SwissQual, Switzerland
- PAGE 2553
Wed-P-22-3 **On Speech Intelligibility Estimation of Phase-Aware Single-Channel Speech Enhancement**
Andreas Gaich, Pejman Mowlae, Technische Universität Graz, Austria
- PAGE 2558
Wed-P-22-4 **A Framework for the Evaluation of Microscopic Intelligibility Models**
Ricard Marxer¹, Martin Cooke², Jon Barker¹
¹University of Sheffield, UK; ²Ikerbasque, Spain
- PAGE 2563
Wed-P-22-5 **A Binaural Short Time Objective Intelligibility Measure for Noisy and Enhanced Speech**
Asger Heidemann Andersen¹, Jan Mark de Haan², Zheng-Hua Tan¹, Jesper Jensen¹
¹Aalborg University, Denmark; ²Oticon, Denmark
- PAGE 2568
Wed-P-22-6 **A Glimpse-Based Approach for Predicting Binaural Intelligibility with Single and Multiple Maskers in Anechoic Conditions**
Yan Tang¹, Martin Cooke², Bruno M. Fazenda¹, Trevor J. Cox¹
¹University of Salford, UK; ²Ikerbasque, Spain

INTERSPEECH 2015 Table of Contents

Wed-P-22 continued...

- PAGE 2573
Wed-P-22-7 **Improving the Prediction Power of the Speech Transmission Index to Account for Non-Linear Distortions Introduced by Noise-Reduction Algorithms**
Fei Chen, SUSTC, China
- PAGE 2578
Wed-P-22-8 **DNN-Based Speech Bandwidth Expansion and Its Application to Adding High-Frequency Missing Features for Automatic Speech Recognition of Narrowband Speech**
Kehuang Li¹, Zhen Huang¹, Yong Xu², Chin-Hui Lee¹
¹Georgia Institute of Technology, USA; ²USTC, China
- PAGE 2583
Wed-P-22-9 **Speech Quality Evaluation of Artificial Bandwidth Extension: Comparing Subjective Judgments and Instrumental Predictions**
Hannu Pulakka¹, Ville Myllylä¹, Anssi Rämö², Paavo Alku³
¹Microsoft, Finland; ²Nokia Research Center, Finland; ³Aalto University, Finland
- PAGE 2588
Wed-P-22-10 **Synchronous Overlap and Add of Spectra for Enhancement of Excitation in Artificial Bandwidth Extension of Speech**
M.A. Tuğtekin Turan, Engin Erzin, Koç Üniversitesi, Turkey
- PAGE 2593
Wed-P-22-11 **Speech Bandwidth Expansion Based on Deep Neural Networks**
Yingxue Wang¹, Shenghui Zhao¹, Wenbo Liu², Ming Li², Jingming Kuang¹
¹Beijing University of Technology, China; ²Sun Yat-sen University, China
- PAGE 2598
Wed-P-22-12 **A Novel Method of Artificial Bandwidth Extension Using Deep Architecture**
Bin Liu, Jianhua Tao, Zhengqi Wen, Ya Li, Danish Bukhari, Chinese Academy of Sciences, China

Show and Tell Session 4

Chair: Björn Hoffmeister

Room: Seminar 1, Time 14:00 - 16:00, Wednesday, September 9, 2015

- PAGE 2603
Show&Tell-4-1 **The RedDots Platform for Mobile Crowd-Sourcing of Speech Data**
*Kong Aik Lee, Guangsen Wang, Kam Pheng Ng, Hanwu Sun, Trung Hieu Nguyen, Ngoc Thuy Huong Thai, Bin Ma, Haizhou Li, A*STAR, Singapore*
- PAGE 2605
Show&Tell-4-2 **Two Extensions of Umeda and Teranishi's Physical Models of the Human Vocal Tract**
Takayuki Arai, Sophia University, Japan
- PAGE 2607
Show&Tell-4-3 **Collaborative Annotation for Person Identification in TV Shows**
Matheuz Budnik¹, Laurent Besacier¹, Johann Poignant², Hervé Bredin², Claude Barras², Mickael Stefas³, Pierrick Bruneau³, Thomas Tamisier³
¹LIG (UMR 5217), France; ²LIMSI, France; ³LIST, Luxembourg
- PAGE 2609
Show&Tell-4-4 **Phonetic/Linguistic Web Services at BAS**
Thomas Kisler, Florian Schiel, Uwe D. Reichel, Christoph Draxler, LMU München, Germany
- PAGE 2611
Show&Tell-4-5 **Managing Speech Databases with emuR and the EMU-webApp**
Raphael Winkelmann, LMU München, Germany
- PAGE 2613
Show&Tell-4-6 **Visual Comparison of Speaker Groups**
Sebastian Wankerl, Florian Hönig, Anton Batliner, J.R. Orozco-Arroyave, Elmar Nöth, FAU Erlangen-Nürnberg, Germany
- PAGE 2615
Show&Tell-4-7 **Tools for Rapid Customization of S2S Systems for Emergent Domains**
Rohit Kumar, Matthew E. Roy, Sanjika Hewavitharana, Dennis N. Mehay, Nina Zinovieva, Raytheon BBN Technologies, USA

Show&Tell-4 continued...

- PAGE 2617
Show&Tell-4-8 **The Speech Recognition Virtual Kitchen Turns One**
Florian Metze¹, Eric Riebling¹, Eric Fosler-Lussier², Andrew Plummer², Rebecca Bates³
¹Carnegie Mellon University, USA; ²Ohio State University, USA; ³Minnesota State University, USA
- PAGE 2619
Show&Tell-4-9 **Model-Based Adaptive Pre-Processing of Speech for Enhanced Intelligibility in Noise and Reverberation**
Jan RENNIES¹, Andreas Volgenandt¹, Henning Schepker², Simon Doclo¹
¹Fraunhofer IDMT, Germany; ²Carl von Ossietzky Universität Oldenburg, Germany
- PAGE 2621
Show&Tell-4-10 **Experiences with and New Application Ideas for the Interspeech App**
Sebastian Möller, Tilo Westermann, T-Labs, Germany
- PAGE 2623
Show&Tell-4-11 **Traditional IVR and Visual IVR — Killing Two Birds with One Stone**
Dmitry Sityaev, Praphul Kumar, Rajesh Ramchander, Genesys, USA

Oral Session 35: Discriminative Acoustic Training Methods for ASR

Chairs: Steve Renals, George Saon

Room: Large Hall, Time 16:30 - 18:30, Wednesday, September 9, 2015

- PAGE 2625
Wed-O-35-1
16:30 - 16:50 **Annotating Large Lattices with the Exact Word Error**
Rogier C. van Dalen, Mark J.F. Gales, University of Cambridge, UK
- PAGE 2630
Wed-O-35-2
16:50 - 17:10 **Semi-Supervised Maximum Mutual Information Training of Deep Neural Network Acoustic Models**
Vimal Manohar, Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, USA
- PAGE 2635
Wed-O-35-3
17:10 - 17:30 **Rectified Linear Neural Networks with Tied-Scalar Regularization for LVCSR**
Shiliang Zhang¹, Hui Jiang², Si Wei¹, Li-Rong Dai¹
¹USTC, China; ²York University, Canada
- PAGE 2640
Wed-O-35-4
17:30 - 17:50 **Segmental Conditional Random Fields with Deep Neural Networks as Acoustic Models for First-Pass Word Recognition**
Yanzhang He, Eric Fosler-Lussier, Ohio State University, USA
- PAGE 2645
Wed-O-35-5
17:50 - 18:10 **Distinct Triphone Acoustic Modeling Using Deep Neural Networks**
Dongpeng Chen, Brian Mak, HKUST, China
- PAGE 2650
Wed-O-35-6
18:10 - 18:30 **Minimum Word Error Training of RNN-Based Voice Activity Detection**
Gregory Gelly, Jean-Luc Gauvain, LIMSI, France

Oral Session 36: Syllables and Segments

Chairs: Anders Eriksson, Frank Zimmerer

Room: Conference 1, Time 16:30 - 18:30, Wednesday, September 9, 2015

- PAGE 2655
Wed-O-36-1
16:30 - 16:50 **Classification of Place-of-Articulation of Stop Consonants Using Temporal Analysis**
A.P. Prathosh¹, A.G. Ramakrishnan², T.V. Ananthapadmanabha³
¹Xerox Research Center India, India; ²Indian Institute of Science, India; ³Voice and Speech Systems, India
- PAGE 2660
Wed-O-36-2
16:50 - 17:10 **The Emergence of Nasal Velar Codas in Brazilian Portuguese: An rt-MRI Study**
Marissa Barlaz, Maojing Fu, Zhi-Pei Liang, Ryan Shosted, Brad Sutton, University of Illinois at Urbana-Champaign, USA
- PAGE 2665
Wed-O-36-3
17:10 - 17:30 **Salient Dimensions in Implicit Phonotactic Learning**
Elise Michon¹, Emmanuel Dupoux², Alejandrina Cristia²
¹Boğaziçi Üniversitesi, Turkey; ²LSCP, France
- PAGE 2670
Wed-O-36-4
17:30 - 17:50 **An Acoustic Examination of the Three-Way Sibilant Contrast in Lower Sorbian**
Phil Howson, University of Toronto, Canada
- PAGE 2675
Wed-O-36-5
17:50 - 18:10 **Investigating Consonant Reduction in Mandarin Chinese with Improved Forced Alignment**
Jiahong Yuan, Mark Liberman, University of Pennsylvania, USA
- PAGE 2679
Wed-O-36-6
18:10 - 18:30 **Durational Characteristics and Timing Patterns of Russian Onset Clusters at Two Speaking Rates**
Marianne Pouplier¹, Stefania Marin¹, Alexei Kochetov²
¹LMU München, Germany; ²University of Toronto, Canada

Oral Session 37: Topics in Paralinguistics

Chairs: Shrikanth Narayanan, Khiat Truong

Room: Conference 2+3, Time 16:30 - 18:30, Wednesday, September 9, 2015

PAGE 2684
Wed-O-37-1
16:30 - 16:50

Vocal Biomarkers to Discriminate Cognitive Load in a Working Memory Task
Thomas F. Quatieri¹, James R. Williamson¹, Christopher J. Smalt¹, Tejash Patel¹, Joseph Perricone¹, Daryush D. Mehta¹, Brian S. Helfer¹, Gregory Ciccarelli¹, Darrell Ricke¹, Nicolas Malyska¹, Jeff Palmer¹, Kristin Heaton², Marianna Eddy³, Joseph Moran³
¹MIT Lincoln Laboratory, USA; ²USARIEM, USA; ³NSRDEC, USA

PAGE 2689
Wed-O-37-2
16:50 - 17:10

I-Vector Based Physical Task Stress Detection with Different Fusion Strategies
Chunlei Zhang, Gang Liu, Chengzhu Yu, John H.L. Hansen, University of Texas at Dallas, USA

PAGE 2694
Wed-O-37-3
17:10 - 17:30

Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech Using ASR
László Tóth¹, Gábor Gosztolya¹, Veronika Vincze¹, Ildikó Hoffmann², Gréta Szatlóczi², Edit Biró², Fruzsina Zsura², Magdolna Pákáski², János Kálmán²
¹MTA-SZTE RGAI, Hungary; ²University of Szeged, Hungary

PAGE 2699
Wed-O-37-4
17:30 - 17:50

Contemporary Stochastic Feature Selection Algorithms for Speech-Based Emotion Recognition
Maxim Sidorov¹, Christina Brester², Alexander Schmitt¹
¹Universität Ulm, Germany; ²Siberian State Aerospace University, Russia

Wed-O-37 continued...

PAGE 2704
Wed-O-37-5
17:50 - 18:10

Effect of Different Jitter-Induced Glottal Pulse Shape Changes in Periodicity Perturbation Measures
Carlos A. Ferrer¹, Diana Torres¹, Eduardo González¹, José Ramón Calvo², Eduardo Castillo³
¹Universidad Central de Las Villas, Cuba; ²CENATAV, Cuba; ³University of New Brunswick, Canada

PAGE 2709
Wed-O-37-6
18:10 - 18:30

Automatic Audio Sentiment Extraction Using Keyword Spotting
Lakshmish Kaushik, Abhijeet Sangwan, John H.L. Hansen, University of Texas at Dallas, USA

Oral Session 38: Spoken Language Processing

Chairs: *Giuseppe Riccardi, Ingrid Zukerman*

Room: *Conference 4+5, Time 16:30 - 18:30, Wednesday, September 9, 2015*

- PAGE 2714
Wed-O-38-1
16:30 - 16:50
- Unsupervised Relation Detection Using Automatic Alignment of Query Patterns Extracted from Knowledge Graphs and Query Click Logs**
Panupong Pasupat¹, Dilek Hakkani-Tür²
¹Stanford University, USA; ²Microsoft, USA
- PAGE 2719
Wed-O-38-2
16:50 - 17:10
- A Latent Variable Model for Joint Pause Prediction and Dependency Parsing**
The Tung Nguyen, Graham Neubig, Hiroyuki Shindo, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura, NAIST, Japan
- PAGE 2724
Wed-O-38-3
17:10 - 17:30
- Extractive Meeting Summarization Through Speaker Zone Detection**
Mohammad Hadi Bokaei¹, Hossein Sameti¹, Yang Liu²
¹Sharif University of Technology, Iran; ²University of Texas at Dallas, USA
- PAGE 2729
Wed-O-38-4
17:30 - 17:50
- Positional Language Modeling for Extractive Broadcast News Speech Summarization**
Shih-Hung Liu¹, Kuan-Yu Chen¹, Berlin Chen², Hsin-Min Wang¹, Hsu-Chun Yen³, Wen-Lian Hsu¹
¹Academia Sinica, Taiwan; ²National Taiwan Normal University, Taiwan; ³National Taiwan University, Taiwan
- PAGE 2734
Wed-O-38-5
17:50 - 18:10
- Speech-Based Location Estimation of First Responders in a Simulated Search and Rescue Scenario**
Saeid Mokaram, Roger K. Moore, University of Sheffield, UK
- PAGE 2739
Wed-O-38-6
18:10 - 18:30
- Constructive Feedback, Thinking Process and Cooperation: Assessing the Quality of Classroom Interaction**
Tahir Sousa¹, Lucie Flekova², Margot Mieskes³, Iryna Gurevych²
¹University of Minnesota, USA; ²Technische Universität Darmstadt, Germany; ³Hochschule Darmstadt, Germany

INTERSPEECH 2015 Table of Contents

Oral Session 39: Voice Conversion

Chairs: *Masami Akamine, Yannis Stylianou*

Room: *Conference 6, Time 16:30 - 18:30, Wednesday, September 9, 2015*

- PAGE 2744
Wed-O-39-1
16:30 - 16:50
- A Real-Time Variable-Q Non-Stationary Gabor Transform for Pitch Shifting**
*Dong-Yan Huang, Minghui Dong, Haizhou Li, A*STAR, Singapore*
- PAGE 2749
Wed-O-39-2
16:50 - 17:10
- Many-to-Many Voice Conversion Based on Multiple Non-Negative Matrix Factorization**
Ryo Aihara, Testuya Takiguchi, Yasuo Arika, Kobe University, Japan
- PAGE 2754
Wed-O-39-3
17:10 - 17:30
- Statistical Singing Voice Conversion Based on Direct Waveform Modification with Global Variance**
Kazuhiro Kobayashi, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura, NAIST, Japan
- PAGE 2759
Wed-O-39-4
17:30 - 17:50
- System Fusion for High-Performance Voice Conversion**
Xiaohai Tian¹, Zhizheng Wu², Siu Wa Lee³, Quy Hy Nguyen¹, Minghui Dong³, Eng Siong Chng¹
¹NTU, Singapore; ²University of Edinburgh, UK; ³A*STAR, Singapore
- PAGE 2764
Wed-O-39-5
17:50 - 18:10
- Speaker Adaptation Using Only Vocalic Segments via Frequency Warping**
Agustin Alonso, D. Erro, Eva Navas, Inma Hernaez, Universidad del País Vasco, Spain
- PAGE 2769
Wed-O-39-6
18:10 - 18:30
- Non-Audible Murmur Enhancement Based on Statistical Conversion Using Air- and Body-Conductive Microphones in Noisy Environments**
Yusuke Tajiri, Kou Tanaka, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura, NAIST, Japan

Special Session 7: Advanced Crowdsourcing for Speech and Beyond

Chairs: Gina-Anne Levow, Tim Polzehl

Room: Hall 1, Time 16:30 – 18:30, Wednesday, September 9, 2015

- Wed-SP7-i
16:30 – 16:50
- Introduction by the Organizers**
Tim Polzehl¹, Gina-Anne Levow²
¹T-Labs, Germany; ²University of Washington, USA
- Oral Presentations
- PAGE 2774
Wed-SP7-1
16:50 – 17:05
- Transcribing Continuous Speech Using Mismatched Crowdsourcing**
Preethi Jyothi, Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign, USA
- PAGE 2779
Wed-SP7-2
17:05 – 17:20
- Selection and Aggregation Techniques for Crowdsourced Semantic Annotation Task**
Shammur Absar Chowdhury¹, Marcos Calvo², Arindam Ghosh¹, Evgeny A. Stepanov¹, Ali Orkan Bayer¹, Giuseppe Riccardi¹, Fernando García², Emilio Sanchis²
¹Università di Trento, Italy; ²Universidad Politécnica de Valencia, Spain
- PAGE 2784
Wed-SP7-3
17:20 – 17:35
- Controlling Quality and Handling Fraud in Large Scale Crowdsourcing Speech Data Collections**
Spencer Rothwell, Ahmad Elshenawy, Steele Carter, Daniela Braga, Faraz Romani, Michael Kennewick, Bob Kennewick, VoiceBox Technologies, USA
- PAGE 2789
Wed-SP7-4
17:35 – 17:50
- Data Collection and Annotation for State-of-the-Art NER Using Unmanaged Crowds**
Spencer Rothwell, Steele Carter, Ahmad Elshenawy, Vladislavs Dovgalecs, Safiyah Saleem, Daniela Braga, Bob Kennewick, VoiceBox Technologies, USA

INTERSPEECH 2015 Table of Contents

Wed-SP7 continued...

- PAGE 2794
Wed-SP7-5
17:50 – 18:05
- Robustness in Speech Quality Assessment and Temporal Training Expiry in Mobile Crowdsourcing Environments**
Tim Polzehl, Babak Naderi, Friedemann Köster, Sebastian Möller, T-Labs, Germany
- Poster Presentations
- PAGE 2799
Wed-SP7-6
- Effect of Trapping Questions on the Reliability of Speech Quality Judgments in a Crowdsourcing Paradigm (Poster)**
Babak Naderi, Tim Polzehl, Ina Wechsung, Friedemann Köster, Sebastian Möller, T-Labs, Germany
- PAGE 2804
Wed-SP7-7
- Voice Äpp: A Mobile App for Crowdsourcing Swiss German Dialect Data (Poster)**
Adrian Leemann¹, Marie-José Kolly², Jean-Philippe Goldman³, Volker Dellwo⁴, Ingrid Hove⁴, Ibrahim Almajai³, Sarah Grimm⁵, Sylvain Robert⁵, Daniel Wanitsch⁶
¹University of Cambridge, UK; ²LIMSI, France; ³Université de Genève, Switzerland; ⁴Universität Zürich, Switzerland; ⁵ETH Zürich, Switzerland; ⁶iBros.ch, Switzerland
- PAGE 2809
Wed-SP7-8
- Expert and Crowdsourced Annotation of Pronunciation Errors for Automatic Scoring Systems (Poster)**
Anastassia Loukina, Melissa Lopez, Keelan Evanini, David Suendermann-Oeft, Klaus Zechner, Educational Testing Service, USA
- PAGE 2814
Wed-SP7-9
- CapCap: An Output-Agreement Game for Video Captioning (Poster)**
Hernisa Kacorri¹, Kaoru Shinkawa², Shin Saito²
¹CUNY Graduate Center, USA; ²IBM Research Tokyo, Japan
- PAGE 2819
Wed-SP7-10
- Auris populi: Crowdsourced Native Transcriptions of Dutch Vowels Spoken by Adult Spanish Learners (Poster)**
Pepi Burgos, Eric Sanders, Catia Cucchiarini, Roeland van Hout, Helmer Strik, Radboud Universiteit Nijmegen, The Netherlands

Wed-SP7 continued...

- PAGE 2824
Wed-SP7-11 **Crowdsource a Little to Label a Lot: Labeling a Speech Corpus of Dialectal Arabic (Poster)**
Samantha Wray, Ahmed Ali, Hamad Bin Khalifa University, Qatar
- PAGE 2829
Wed-SP7-12 **Using Keyword Spotting to Help Humans Correct Captioning Faster (Poster)**
Yashesh Gaur, Florian Metze, Yajie Miao, Jeffrey P. Bigham, Carnegie Mellon University, USA
- PAGE 2834
Wed-SP7-13 **Validating and Optimizing a Crowdsourced Method for Gradient Measures of Child Speech (Poster)**
Tara McAllister Byun¹, Elaine Hitchcock², Daphna Harel¹
¹New York University, USA; ²Montclair State University, USA

Poster Session 23: Robust Speech Recognition: Adaptation

Chair: Najim Dehak

Room: Hall 2, Time 16:30 - 18:30, Wednesday, September 9, 2015

- PAGE 2839
Wed-P-23-1 **Joint Training of Speech Separation, Filterbank and Acoustic Model for Robust Automatic Speech Recognition**
Zhong-Qiu Wang, DeLiang Wang, Ohio State University, USA
- PAGE 2844
Wed-P-23-2 **Joint Environment and Speaker Normalization Using Factored Front-End CMLLR**
*Shakti Rath, Sunil Sivadas, Bin Ma, A*STAR, Singapore*
- PAGE 2849
Wed-P-23-3 **Robust Speech Recognition Using DNN-HMM Acoustic Model Combining Noise-Aware Training with Spectral Subtraction**
Akihiro Abe, Kazumasa Yamamoto, Seiichi Nakagawa, Toyohashi University of Technology, Japan
- PAGE 2854
Wed-P-23-4 **Robust i-Vector Extraction for Neural Network Adaptation in Noisy Environment**
Chengzhu Yu¹, Atsunori Ogawa², Marc Delcroix², Takuya Yoshioka², Tomohiro Nakatani², John H.L. Hansen¹
¹University of Texas at Dallas, USA; ²NTT Corporation, Japan
- PAGE 2858
Wed-P-23-5 **Spectrally Selective Dithering for Distorted Speech Recognition**
Michal Borsky, Petr Mizera, Petr Pollak, Czech Technical University in Prague, Czech Republic
- PAGE 2862
Wed-P-23-6 **Feature-Space Speaker Adaptation for Probabilistic Linear Discriminant Analysis Acoustic Models**
Liang Lu, Steve Renals, University of Edinburgh, UK
- PAGE 2867
Wed-P-23-7 **Speaker Adaptation Using the I-Vector Technique for Bottleneck Features**
Patrick Cardinal, Najim Dehak, Yu Zhang, James Glass, MIT, USA

Wed-P-23 continued...

- PAGE 2872
Wed-P-23-8 **I-Vector Estimation Using Informative Priors for Adaptation of Deep Neural Networks**
Penny Karanasou, Mark J.F. Gales, Philip C. Woodland, University of Cambridge, UK
- PAGE 2877
Wed-P-23-9 **Robust i-Vector Based Adaptation of DNN Acoustic Model for Speech Recognition**
Sri Garimella¹, Arindam Mandal², Nikko Strom², Bjorn Hoffmeister², Spyros Matsoukas², Sree Hari Krishnan Parthasarathi²
¹Amazon.com, India; ²Amazon.com, USA
- PAGE 2882
Wed-P-23-10 **GMM-Derived Features for Effective Unsupervised Adaptation of Deep Neural Network Acoustic Models**
Natalia Tomashenko¹, Yuri Khokhlov²
¹Speech Technology Center, Russia; ²STC-innovations, Russia
- PAGE 2887
Wed-P-23-11 **Unsupervised Adaptation for Deep Neural Network Using Linear Least Square Method**
Roger Hsiao, Tim Ng, Stavros Tsakalidis, Long Nguyen, Richard Schwartz, Raytheon BBN Technologies, USA
- PAGE 2892
Wed-P-23-12 **Ensemble Speaker Modeling Using Speaker Adaptive Training Deep Neural Network for Speaker Adaptation**
Sheng Li¹, Xugang Lu², Yuya Akita¹, Tatsuya Kawahara¹
¹Kyoto University, Japan; ²NICT, Japan
- PAGE 2897
Wed-P-23-13 **Data-Selective Transfer Learning for Multi-Domain Speech Recognition**
Mortaza Doulaty, Oscar Saz, Thomas Hain, University of Sheffield, UK

Poster Session 24: Speech and Audio Segmentation and Classification and Voice Activity Detection

Chair: Ming Li

Room: Hall 2, Time 16:30 - 18:30, Wednesday, September 9, 2015

- PAGE 2902
Wed-P-24-1 **Automatic Detection of Equipment Alarms in a Neonatal Intensive Care Unit Environment: A Knowledge-Based Approach**
Ganna Raboshchuk¹, Peter Jančovič², Climent Nadeu¹, Alex Peiró Lilja¹, Münevver Köküer², Blanca Muñoz Mahamud³, Ana Riverola de Veciana³
¹Universitat Politècnica de Catalunya, Spain; ²University of Birmingham, UK; ³Hospital Sant Joan de Déu Barcelona, Spain
- PAGE 2907
Wed-P-24-2 **"Multilingual" Deep Neural Network for Music Genre Classification**
Jia Dai¹, Wenju Liu¹, Chongjia Ni², Like Dong³, Hong Yang³
¹Chinese Academy of Sciences, China; ²SDUFE, China; ³SGCC, China
- PAGE 2912
Wed-P-24-3 **Accurate Endpointing with Expected Pause Duration**
Baiyang Liu, Bjorn Hoffmeister, Ariya Rastrow, Amazon.com, USA
- PAGE 2917
Wed-P-24-4 **Locality Constrained Transitive Distance Clustering on Speech Data**
Wenbo Liu¹, Zhiding Yu², Bhiksha Raj², Ming Li¹
¹Sun Yat-sen University, China; ²Carnegie Mellon University, USA
- PAGE 2922
Wed-P-24-5 **Feature Extraction Strategies in Deep Learning Based Acoustic Event Detection**
Miquel Espi, Masakiyo Fujimoto, Keisuke Kinoshita, Tomohiro Nakatani, NTT Corporation, Japan
- PAGE 2927
Wed-P-24-6 **An Acoustic Event Detection Framework and Evaluation Metric for Surveillance in Cars**
Peter Transfeld, Simon Receveur, Tim Fingscheidt, Technische Universität Braunschweig, Germany

Wed-P-24 continued...

PAGE 2932
Wed-P-24-7

Diachronic Semantic Cohesion for Topic Segmentation of TV Broadcast News
*Abdessalam Bouchekif¹, Géraldine Damnati¹, Yannick Estève², Delphine Charlet¹,
Nathalie Camelin²*
¹Orange Labs, France; ²LIUM, France

PAGE 2937
Wed-P-24-8

Comparison of Forced-Alignment Speech Recognition and Humans for Generating Reference VAD
Ivan Kraljevski¹, Zheng-Hua Tan², Maria Paola Bissiri³
¹voice INTER connect, Germany; ²Aalborg University, Denmark; ³Queen Margaret University, UK

PAGE 2942
Wed-P-24-9

Improving Voice Activity Detection in Movies
Bernhard Lehner, Gerhard Widmer, Reinhard Sonnleitner, Johannes Kepler Universität Linz, Austria

Poster Session 25: Speech and Hearing Disorders

Chair: Benjawan Kasisopa

Room: Foyer, Time 16:30 - 18:30, Wednesday, September 9, 2015

PAGE 2947
Wed-P-25-1

Language-Independent Method for Analysis of German Stuttering Recordings
Tomas Lustyk¹, Petr Bergl¹, Tino Haderlein², Elmar Nöth², Roman Cmejla¹
¹Czech Technical University in Prague, Czech Republic; ²FAU Erlangen-Nürnberg, Germany

PAGE 2952
Wed-P-25-2

An Investigation of MDVP Parameters for Voice Pathology Detection on Three Different Databases
Ahmed Al-nasheri, Zulfiqar Ali, Ghulam Muhammad, Mansour Alsulaiman, King Saud University, Saudi Arabia

PAGE 2957
Wed-P-25-3

Energy Distribution Analysis and Nonlinear Dynamical Analysis of Adductor Spasmodic Dysphonia
Jiantao Wu¹, Ping Yu², Nan Yan¹, Lan Wang¹, Xiaohui Yang³, Manwa L. Ng⁴
¹Chinese Academy of Sciences, China; ²PLA General Hospital, China; ³CUST, China; ⁴University of Hong Kong, China

PAGE 2962
Wed-P-25-4

Auditory-Visual Tone Perception in Hearing Impaired Thai Listeners
Benjawan Kasisopa¹, Nittayapa Klangpornkun², Denis Burnham¹
¹University of Western Sydney, Australia; ²Thammasat University, Thailand

PAGE 2967
Wed-P-25-5

Speech Intelligibility Decline in Individuals with Fast and Slow Rates of ALS Progression
Panying Rong¹, Yana Yunusova², Jordan R. Green¹
¹MGH Institute of Health Professions, USA; ²University of Toronto, Canada

Wed-P-25 continued...

- PAGE 2972
Wed-P-25-6 **Latency Analysis of Speech Shadowing Reveals Processing Differences in Japanese Adults Who Do and Do Not Stutter**
Rong Na A, Koichi Mori, Naomi Sakai, NRCD Research Institute, Japan
- PAGE 2977
Wed-P-25-7 **A Syllable-Based Analysis of Speech Temporal Organization: A Comparison Between Speaking Styles in Dysarthric and Healthy Populations**
Brigitte Bigi¹, Katarzyna Klessa², Laurianne Georgeton¹, Christine Meunier¹
¹LPL, France; ²Adam Mickiewicz University in Poznań, Poland
- PAGE 2982
Wed-P-25-8 **Autonomous Measurement of Speech Intelligibility Utilizing Automatic Speech Recognition**
Bernd T. Meyer, Birger Kollmeier, Jasper Ooster, Carl von Ossietzky Universität Oldenburg, Germany
- PAGE 2987
Wed-P-25-9 **Can You Hear Me? Acoustic Modifications in Speech Directed to Foreigners and Hearing-Impaired People**
Monja Angelika Knoll, Melissa Johnstone, Charlene Blakely, University of the West of Scotland, UK
- PAGE 2991
Wed-P-25-10 **Improving Automatic Forced Alignment for Dysarthric Speech Transcription**
Yu Ting Yeung, Ka Ho Wong, Helen Meng, Chinese University of Hong Kong, China

Poster Session 26: Speaker Recognition and Diarization

Chair: Anthony Larcher

Room: Seminar 5+6, Time 16:30 – 18:30, Wednesday, September 9, 2015

- PAGE 2996
Wed-P-26-1 **The RedDots Data Collection for Speaker Recognition**
Kong Aik Lee¹, Anthony Larcher², Guangsen Wang¹, Patrick Kenny³, Niko Brümmer⁴, David van Leeuwen⁵, Hagai Aronowitz⁶, Marcel Kockmann⁷, Carlos Vaquero⁸, Bin Ma¹, Haizhou Li¹, Themis Stafylakis³, Md. Jahangir Alam³, Albert Swart⁴, Javier Perez⁶
*¹A*STAR, Singapore; ²LIUM, France; ³CRIM, Canada; ⁴AGNITiO, South Africa; ⁵NovoLanguage, The Netherlands; ⁶IBM Research Haifa, Israel; ⁷VoiceTrust, Canada; ⁸AGNITiO, Spain*
- PAGE 3001
Wed-P-26-2 **Noise-Robust Speaker Recognition Based on Morphological Component Analysis**
Yongjun He¹, Chen Chen², Jiqing Han²
¹HUST, China; ²HIT, China
- PAGE 3006
Wed-P-26-3 **Analysis of Mutual Duration and Noise Effects in Speaker Recognition: Benefits of Condition-Matched Cohort Selection in Score Normalization**
Andreas Nautsch¹, Rahim Saeidi², Christian Rathgeb¹, Christoph Busch¹
¹Hochschule Darmstadt, Germany; ²Aalto University, Finland
- PAGE 3011
Wed-P-26-4 **Robustness to Additive Noise of Locally-Normalized Cepstral Coefficients in Speaker Verification**
Josué Fredes¹, José Novoa¹, Victor Poblete¹, Simon King², Richard M. Stern³, Néstor Becerra Yoma¹
¹Universidad de Chile, Chile; ²University of Edinburgh, UK; ³Carnegie Mellon University, USA
- PAGE 3016
Wed-P-26-5 **Probabilistic Linear Discriminant Analysis for Robust Speaker Identification in Co-Channel Speech**
Navid Shokouhi, John H.L. Hansen, University of Texas at Dallas, USA

Wed-P-26 continued...

- PAGE 3021
Wed-P-26-6 **Community Detection with Manifold Learning on Speaker i-Vector Space for Chinese**
Hongcui Wang¹, Di Jin¹, Lantian Li², Jianwu Dang¹
¹Tianjin University, China; ²Tsinghua University, China
- PAGE 3026
Wed-P-26-7 **A Comparison of Neural Network Feature Transforms for Speaker Diarization**
Sree Harsha Yella¹, Andreas Stolcke²
¹Idiap Research Institute, Switzerland; ²Microsoft, USA
- PAGE 3031
Wed-P-26-8 **Clustering Short *Push-to-Talk* Segments**
Ilya Shapiro, Neta Rabin, Irit Opher, Itshak Lapidot, Afeka, Israel
- PAGE 3036
Wed-P-26-9 **Exploring ANN Back-Ends for i-Vector Based Speaker Age Estimation**
Anna Fedorova¹, Ondřej Glembek², Tomi Kinnunen¹, Pavel Matějka²
¹University of Eastern Finland, Finland; ²Brno University of Technology, Czech Republic
- PAGE 3041
Wed-P-26-10 **Analysis of the Second Phase of the 2013–2014 i-Vector Machine Learning Challenge**
Désiré Bansé¹, George R. Doddington¹, Daniel Garcia-Romero², John J. Godfrey², Craig S. Greenberg¹, Jaime Hernández-Cordero³, John M. Howard¹, Alvin F. Martin¹, Lisa P. Mason³, Alan McCree², Douglas A. Reynolds⁴
¹NIST, USA; ²Johns Hopkins University, USA; ³DOD, USA; ⁴MIT Lincoln Laboratory, USA
- PAGE 3046
Wed-P-26-11 **NIST Language Recognition Evaluation — Plans for 2015**
Alvin F. Martin¹, Craig S. Greenberg¹, John M. Howard¹, Désiré Bansé¹, George R. Doddington¹, Jaime Hernández-Cordero², Lisa P. Mason²
¹NIST, USA; ²DOD, USA

Oral Session 40: Dialogue and Discourse

Chairs: Jens Edlund, Kornel Laskowski

Room: Conference 1, Time 09:00 – 11:00, Thursday, September 10, 2015

- PAGE 3051
Thu-O-40-1
09:00 – 09:20 **Communicative Needs and Respiratory Constraints**
Marcin Włodarczak¹, Mattias Heldner¹, Jens Edlund²
¹Stockholm University, Sweden; ²KTH, Sweden
- PAGE 3056
Thu-O-40-2
09:20 – 09:40 **Analysis and Classification of Cooperative and Competitive Dialogs**
Uwe D. Reichel¹, Nina Pörner¹, Dianne Nowack¹, Jennifer Cole²
¹LMU München, Germany; ²University of Illinois at Urbana-Champaign, USA
- PAGE 3061
Thu-O-40-3
09:40 – 10:00 **Towards Automatic Detection of Reported Speech in Dialogue Using Prosodic Cues**
Alessandra Cervone, Catherine Lai, Silvia Pareti, Peter Bell, University of Edinburgh, UK
- PAGE 3066
Thu-O-40-4
10:00 – 10:20 **Modeling Phrasing and Prominence Using Deep Recurrent Learning**
Andrew Rosenberg¹, Raul Fernandez², Bhuvana Ramabhadran²
¹CUNY Queens College, USA; ²IBM T.J. Watson Research Center, USA
- PAGE 3071
Thu-O-40-5
10:20 – 10:40 **Pitch Declination and Reset as a Function of Utterance Duration in Conversational Speech Data**
Céline De Looze, Irena Yanushevskaya, Andy Murphy, Eoghan O'Connor, Christer Gobl, Trinity College Dublin, Ireland
- PAGE 3076
Thu-O-40-6
10:40 – 11:00 **Investigating the Role of ‘yeah’ in Stance-Dense Conversation**
Valerie Freeman, Gina-Anne Levow, Richard Wright, Mari Ostendorf, University of Washington, USA

Oral Session 41: Speaker Diarization

Chairs: Jean-François Bonastre, Itshack Lapidot

Room: Conference 2+3, Time 09:00 - 11:00, Thursday, September 10, 2015

- PAGE 3081
Thu-O-41-1
09:00 - 09:20
- Factor Analysis for Speaker Segmentation and Improved Speaker Diarization**
Brecht Desplanques, Kris Demuynck, Jean-Pierre Martens, Ghent University, Belgium
- PAGE 3086
Thu-O-41-2
09:20 - 09:40
- Enhanced Speaker Diarization with Detection of Backchannels Using Eye-Gaze Information in Poster Conversations**
Koji Inoue¹, Yukoh Wakabayashi², Hiromasa Yoshimoto¹, Katsuya Takanashi¹, Tatsuya Kawahara¹
¹Kyoto University, Japan; ²Ritsumeikan University, Japan
- PAGE 3091
Thu-O-41-3
09:40 - 10:00
- Novel Clustering Selection Criterion for Fast Binary Key Speaker Diarization**
Héctor Delgado¹, Xavier Anguera², Corinne Fredouille³, Javier Serrano¹
¹Universidad Autónoma de Barcelona, Spain; ²Sinkronigo, Spain; ³LIA, France
- PAGE 3096
Thu-O-41-4
10:00 - 10:20
- Speaker Diarization with I-Vectors from DNN Senone Posteriors**
Gregory Sell, Daniel Garcia-Romero, Alan McCree, Johns Hopkins University, USA
- PAGE 3100
Thu-O-41-5
10:20 - 10:40
- Using Voice-Quality Measurements with Prosodic and Spectral Features for Speaker Diarization**
Abraham Woubie¹, Jordi Luque², Javier Hernando¹
¹Universitat Politècnica de Catalunya, Spain; ²Telefónica I+D, Spain
- PAGE 3105
Thu-O-41-6
10:40 - 11:00
- Integrating Online I-Vector Extractor with Information Bottleneck Based Speaker Diarization System**
Srikanth Madikeri, Ivan Himawan, Petr Motlicek, Marc Ferras, Idiap Research Institute, Switzerland

Oral Session 42: L1/L2 Speech Perception and Acquisition

Chairs: Gábor Pintér, Jürgen Trouvain

Room: Conference 4+5, Time 09:00 - 11:00, Thursday, September 10, 2015

- PAGE 3110
Thu-O-42-1
09:00 - 09:20
- Enhanced Processing of a Lost Language: Linguistic Knowledge or Linguistic Skill?**
Jiyoun Choi¹, Mirjam Broersma², Anne Cutler²
¹Hanyang University, Korea; ²MPI for Psycholinguistics, The Netherlands
- PAGE 3115
Thu-O-42-2
09:20 - 09:40
- Production Inconsistencies Delay Adaptation to Foreign Accents**
Ann-Kathrin Grohe¹, Gregory J. Poarch², Adriana Hanulíková³, Andrea Weber¹
¹Eberhard Karls Universität Tübingen, Germany; ²Westfälische Wilhelms-Universität Münster, Germany; ³Universität Freiburg, Germany
- PAGE 3120
Thu-O-42-3
09:40 - 10:00
- Acquisition of English Speech Rhythm by Monolingual Children**
Mikhail Ordin, Leona Polyanskaya, Universität Bielefeld, Germany
- PAGE 3125
Thu-O-42-4
10:00 - 10:20
- Durational Information in Word-Initial Lexical Embeddings in Spoken Dutch**
Odette Scharenborg, Radboud Universiteit Nijmegen, The Netherlands
- PAGE 3130
Thu-O-42-5
10:20 - 10:40
- The Development of Categorical Perception of Lexical Tones in Mandarin-Speaking Preschoolers**
Fei Chen, Nan Yan, Lan Wang, Tao Yang, Jiantao Wu, Han Zhao, Gang Peng, Chinese Academy of Sciences, China
- PAGE 3135
Thu-O-42-6
10:40 - 11:00
- Perception of Italian Liquids by Japanese Listeners: Comparisons to Spanish Liquids**
Tomohiko Ooigawa, Sophia University, Japan

Oral Session 43: LVCSR Systems and Applications

Chairs: Michiel Bacchiani, Florian Metzger

Room: Conference 6, Time 09:00 – 11:00, Thursday, September 10, 2015

PAGE 3140
Thu-O-43-1
09:00 – 09:20

The IBM 2015 English Conversational Telephone Speech Recognition System
George Saon, Hong-Kwang J. Kuo, Steven Rennie, Michael Picheny, IBM T.J. Watson Research Center, USA

PAGE 3145
Thu-O-43-2
09:20 – 09:40

The Cambridge University 2014 BOLT Conversational Telephone Mandarin Chinese LVCSR System for Speech Translation
Xunying Liu, Federico Flego, Linlin Wang, C. Zhang, Mark J.F. Gales, Philip C. Woodland, University of Cambridge, UK

PAGE 3150
Thu-O-43-3
09:40 – 10:00

The IBM BOLT Speech Transcription System
Samuel Thomas, George Saon, Hong-Kwang J. Kuo, Lidia Mangu, IBM T.J. Watson Research Center, USA

PAGE 3154
Thu-O-43-4
10:00 – 10:20

Improvements in RWTH LVCSR Evaluation Systems for Polish, Portuguese, English, Urdu, and Arabic
M. Ali Basha Shaik, Zoltán Tüske, M. Ali Tahir, Markus Nußbaum-Thom, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

PAGE 3159
Thu-O-43-5
10:20 – 10:40

Active Learning Based Data Selection for Limited Resource STT and KWS
Thiago Fraga-Silva¹, Jean-Luc Gauvain², Lori Lamel², Antoine Laurent¹, Viet-Bac Le¹, Abdel Messaoudi¹
¹Vocapia Research, France; ²LIMSI, France

PAGE 3164
Thu-O-43-6
10:40 – 11:00

Improved Hindi Broadcast ASR by Adapting the Language Model and Pronunciation Model Using a priori Syntactic and Morphophonemic Knowledge
Preethi Jyothi, Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign, USA

Special Session 8: Zero Resource Speech Technologies: Unsupervised Discovery of Linguistic Units

Chairs: Xavier Anguera, Emmanuel Dupoux, Maarten Versteegh

Room: Hall 1, Time 09:00 – 11:00, Thursday, September 10, 2015

PAGE 3169
Thu-SP8-1
09:00 – 09:10

The Zero Resource Speech Challenge 2015
Maarten Versteegh¹, Roland Thiollière¹, Thomas Schatz¹, Xuan Nga Cao¹, Xavier Anguera², Aren Jansen³, Emmanuel Dupoux¹
¹ENS, France; ²Telefónica I+D, Spain; ³Johns Hopkins University, USA

Thu-SP8-F
09:10 – 09:30

Poster Flash Presentation
Xavier Anguera¹, Emmanuel Dupoux², Maarten Versteegh³
¹Telefónica I+D, Spain; ²EHESS, France; ³ENS, France

Thu-SP8-D
09:30 – 10:00

Panel Discussion: Summary of Results, Discussion of Evaluation Metrics, and the Next Challenge
Xavier Anguera¹, Emmanuel Dupoux², Maarten Versteegh³
¹Telefónica I+D, Spain; ²EHESS, France; ³ENS, France

Poster Presentations

PAGE 3174
Thu-SP8-2

Discovering Discrete Subword Units with Binarized Autoencoders and Hidden-Markov-Model Encoders (Poster)
Leonardo Badino¹, Alessio Mereta¹, Lorenzo Rosasco²
¹Istituto Italiano di Tecnologia, Italy; ²Università di Genova, Italy

PAGE 3179
Thu-SP8-3

A Hybrid Dynamic Time Warping-Deep Neural Network Architecture for Unsupervised Acoustic Modeling (Poster)
Roland Thiollière, Ewan Dunbar, Gabriel Synnaeve, Maarten Versteegh, Emmanuel Dupoux, ENS, France

Thu-SP8 continued...

- PAGE 3184
Thu-SP8-4 **Automatic Segmentation and Clustering of Speech Using Sparse Coding and Metaheuristic Search (Poster)**
Wiehan Agenbag, Thomas Niesler, Stellenbosch University, South Africa
- PAGE 3189
Thu-SP8-5 **Parallel Inference of Dirichlet Process Gaussian Mixture Models for Unsupervised Acoustic Modeling: A Feasibility Study (Poster)**
Hongjie Chen¹, Cheung-Chi Leung², Lei Xie¹, Bin Ma², Haizhou Li²
*¹Northwestern Polytechnical University, China; ²A*STAR, Singapore*
- PAGE 3194
Thu-SP8-6 **Using Articulatory Features and Inferred Phonological Segments in Zero Resource Speech Processing (Poster)**
Pallavi Baljekar, Sunayana Sitaram, Prasanna Kumar Muthukumar, Alan W. Black, Carnegie Mellon University, USA
- PAGE 3199
Thu-SP8-7 **A Comparison of Neural Network Methods for Unsupervised Representation Learning on the Zero Resource Speech Challenge (Poster)**
Daniel Renshaw¹, Herman Kamper¹, Aren Jansen², Sharon Goldwater¹
¹University of Edinburgh, UK; ²Johns Hopkins University, USA
- PAGE 3204
Thu-SP8-8 **Unsupervised Word Discovery from Speech Using Automatic Segmentation into Syllable-Like Units (Poster)**
Okko Räsänen¹, Gabriel Doyle², Michael C. Frank²
¹Aalto University, Finland; ²Stanford University, USA
- PAGE 3209
Thu-SP8-9 **An Evaluation of Graph Clustering Methods for Unsupervised Term Discovery (Poster)**
Vince Lyzinski, Gregory Sell, Aren Jansen, Johns Hopkins University, USA

Poster Session 27: Neural Networks: Novel Architectures for LVCSR

Chair: Rogier Van Dalen

Room: Hall 2, Time 09:00 - 11:00, Thursday, September 10, 2015

- PAGE 3214
Thu-P-27-1 **A Time Delay Neural Network Architecture for Efficient Modeling of Long Temporal Contexts**
Vijayaditya Peddinti, Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, USA
- PAGE 3219
Thu-P-27-2 **Long Short-Term Memory Based Convolutional Recurrent Neural Networks for Large Vocabulary Speech Recognition**
Xiangang Li, Xihong Wu, Peking University, China
- PAGE 3224
Thu-P-27-3 **Parameterised Sigmoid and ReLU Hidden Activation Functions for DNN Acoustic Modelling**
C. Zhang, Philip C. Woodland, University of Cambridge, UK
- PAGE 3229
Thu-P-27-4 **Discriminative Template Learning in Group-Convolutional Networks for Invariant Speech Representations**
Chiyuan Zhang, Stephen Voinea, Georgios Evangelopoulos, Lorenzo Rosasco, Tomaso Poggio, MIT, USA
- PAGE 3234
Thu-P-27-5 **Investigation of Parametric Rectified Linear Units for Noise Robust Speech Recognition**
Sunil Sivadas¹, Zhenzhou Wu², Ma Bin¹
*¹A*STAR, Singapore; ²McGill University, Canada*
- PAGE 3239
Thu-P-27-6 **Multi-Softmax Deep Neural Network for Semi-Supervised Training**
Hang Su¹, Haihua Xu²
¹ICSI, USA; ²NTU, Singapore

Thu-P-27 continued...

- PAGE 3244
Thu-P-27-7 **A Multi-Region Deep Neural Network Model in Speech Recognition**
Jia Cui, George Saon, Bhuvana Ramabhadran, Brian Kingsbury, IBM T.J. Watson Research Center, USA
- PAGE 3249
Thu-P-27-8 **A Study of the Recurrent Neural Network Encoder-Decoder for Large Vocabulary Speech Recognition**
Liang Lu¹, Xingxing Zhang¹, Kyunghyun Cho², Steve Renals¹
¹University of Edinburgh, UK; ²Université de Montréal, Canada
- PAGE 3254
Thu-P-27-9 **Gaussian Free Cluster Tree Construction Using Deep Neural Network**
Linchen Zhu, Kevin Kilgour, Sebastian Stüker, Alex Waibel, KIT, Germany
- PAGE 3259
Thu-P-27-10 **Very Deep Convolutional Neural Networks for LVCSR**
Mengxiao Bi, Yanmin Qian, Kai Yu, Shanghai Jiao Tong University, China
- PAGE 3264
Thu-P-27-11 **Transferring Knowledge from a RNN to a DNN**
William Chan, Nan Rosemary Ke, Ian Lane, Carnegie Mellon University, USA
- PAGE 3269
Thu-P-27-12 **SVD-Based Universal DNN Modeling for Multiple Scenarios**
Changliang Liu¹, Jinyu Li², Yifan Gong²
¹Microsoft, China; ²Microsoft, USA
- PAGE 3274
Thu-P-27-13 **Speech Enhancement and Recognition Using Multi-Task Learning of Long Short-Term Memory Recurrent Neural Networks**
Zhuo Chen, Shinji Watanabe, Hakan Erdogan, John R. Hershey, MERL, USA

Poster Session 28: Speech and Music Analysis

Chair: Hideki Kawahara

Room: Hall 2, Time 09:00 - 11:00, Thursday, September 10, 2015

- PAGE 3279
Thu-P-28-1 **Speaker-Dependent Multipitch Tracking Using Deep Neural Networks**
Yuzhou Liu, DeLiang Wang, Ohio State University, USA
- PAGE 3284
Thu-P-28-2 **An Error Correction Scheme for GCI Detection Algorithms Using Pitch Smoothness Criterion**
Sujith P.¹, A.P. Prathosh², A.G. Ramakrishnan³, Prasanta Kumar Ghosh³
¹Ittiam Systems, India; ²Xerox Research Center India, India; ³Indian Institute of Science, India
- PAGE 3289
Thu-P-28-3 **Robust Pitch Estimation in Noisy Speech Using ZTW and Group Delay Function**
RaviShankar Prasad, B. Yegnanarayana, IIIT Hyderabad, India
- PAGE 3293
Thu-P-28-4 **Robust Localization of Single Sound Source Based on Phase Difference Regression**
Zhaoqiong Huang, Ge Zhan, Dongwen Ying, Yonghong Yan, Chinese Academy of Sciences, China
- PAGE 3298
Thu-P-28-5 **Frequency Map Selection Using a RBFN-Based Classifier in the MVDR Beamformer for Speaker Localization in Reverberant Rooms**
Daniele Salvati, Carlo Drioli, Gian Luca Foresti, Università di Udine, Italy
- PAGE 3302
Thu-P-28-6 **Exploiting Deep Neural Networks and Head Movements for Binaural Localisation of Multiple Speakers in Reverberant Conditions**
Ning Ma¹, Guy J. Brown¹, Tobias May²
¹University of Sheffield, UK; ²Technical University of Denmark, Denmark

Thu-P-28 continued...

- PAGE 3307
Thu-P-28-7 **Joint Optimization of Recurrent Networks Exploiting Source Auto-Regression for Source Separation**
Shuai Nie¹, Wei Xue¹, Shan Liang¹, Xueliang Zhang², Wenju Liu¹, Liwei Qiao³, Jianping Li³
¹Chinese Academy of Sciences, China; ²Inner Mongolia University, China; ³SGCC, China
- PAGE 3312
Thu-P-28-8 **Real-Time Audio-to-Score Alignment of Singing Voice Based on Melody and Lyric Information**
Rong Gong, Philippe Cuvillier, Nicolas Obin, Arshia Cont, IRCAM, France
- PAGE 3317
Thu-P-28-9 **Vocal Separation from Monaural Music Using Adaptive Auditory Filtering Based on Kernel Back-Fitting**
Jun-Yong Lee, Hye-Seung Cho, Hyoung-Gook Kim, Kwangwoon University, Korea
- PAGE 3321
Thu-P-28-10 **A Two-Stage Singing Voice Separation Algorithm Using Spectro-Temporal Modulation Features**
Frederick Z. Yen, Mao-Chang Huang, Tai-Shih Chi, National Chiao Tung University, Taiwan
- PAGE 3325
Thu-P-28-11 **Robust Sound Event Classification Using LBP-HOG Based Bag-of-Audio-Words Feature Representation**
Hyungjun Lim, Myung Jong Kim, Hoirin Kim, KAIST, Korea

Poster Session 29: Speech Synthesis 3

Chair: Gérard Bailly

Room: Foyer, Time 09:00 - 11:00, Thursday, September 10, 2015

- PAGE 3330
Thu-P-29-1 **Sequence-to-Sequence Neural Net Models for Grapheme-to-Phoneme Conversion**
Kaisheng Yao, Geoffrey Zweig, Microsoft, USA
- PAGE 3335
Thu-P-29-2 **Knowledge versus Data in TTS: Evaluation of a Continuum of Synthesis Systems**
Rosie Kay¹, Oliver Watts¹, Roberto Barra Chicote², Cassie Mayo¹
¹University of Edinburgh, UK; ²Universidad Politécnica de Madrid, Spain
- PAGE 3340
Thu-P-29-3 **Improving G2P from Wiktionary and Other (Web) Resources**
Steffen Eger, Goethe-Universität Frankfurt am Main, Germany
- PAGE 3345
Thu-P-29-4 **BLSTM Neural Networks for Speech Driven Head Motion Synthesis**
Chuang Ding, Pengcheng Zhu, Lei Xie, Northwestern Polytechnical University, China
- PAGE 3350
Thu-P-29-5 **Articulatory Controllable Speech Modification Based on Gaussian Mixture Models with Direct Waveform Modification Using Spectrum Differential**
Patrick Lumban Tobing, Kazuhiro Kobayashi, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura, NAIST, Japan
- PAGE 3355
Thu-P-29-6 **Reconstructing Intelligible Audio Speech from Visual Speech Features**
Thomas Le Cornu, Ben Milner, University of East Anglia, UK
- PAGE 3360
Thu-P-29-7 **Universal Grapheme-Based Speech Synthesis**
Sunayana Sitaram¹, Alok Parlikar¹, Gopala Krishna Anumanchipalli², Alan W. Black¹
¹Carnegie Mellon University, USA; ²University of California at San Francisco, USA

Thu-P-29 continued...

- PAGE 3365
Thu-P-29-8 **Artificial Personality and Disfluency**
Mirjam Wester¹, Matthew Aylett¹, Marcus Tomalin², Rasmus Dall¹
¹University of Edinburgh, UK; ²University of Cambridge, UK
- PAGE 3370
Thu-P-29-9 **Comparison of Chironomic Stylization versus Statistical Modeling of Prosody for Expressive Speech Synthesis**
Marc Evrard, Samuel Delalez, Christophe d'Alessandro, Albert Rilliard, LIMSI, France
- PAGE 3375
Thu-P-29-10 **A Multi-Layer F0 Model for Singing Voice Synthesis Using a B-Spline Representation with Intuitive Controls**
Luc Ardaillon, Gilles Degottex, Axel Roebel, IRCAM, France
- PAGE 3380
Thu-P-29-11 **Creating Expressive Synthetic Voices by Unsupervised Clustering of Audiobooks**
Igor Jauk¹, Antonio Bonafonte¹, Paula Lopez-Otero², Laura Docio-Fernandez²
¹Universitat Politècnica de Catalunya, Spain; ²Universidade de Vigo, Spain
- PAGE 3385
Thu-P-29-12 **Articulatory-Based Conversion of Foreign Accents with Deep Neural Networks**
Sandesh Aryal, Ricardo Gutierrez-Osuna, Texas A&M University, USA

Oral Session 44: Speech and Cognition in Adverse Conditions

Chairs: Louis ten Bosch, Bernd Meyer

Room: Large Hall, Time 14:00 - 16:00, Thursday, September 10, 2015

- PAGE 3390
Thu-O-44-1
14:00 - 14:20 **Action Planning and Congruency Effect Between Articulation and Grasping**
Mikko Tiainen, Lari Vainio, Kaisa Tiippana, Naeem Komeilipoor, Martti Vainio, University of Helsinki, Finland
- PAGE 3394
Thu-O-44-2
14:20 - 14:40 **Cognitive Workload and Vocabulary Sparseness: Theory and Practice**
Ron M. Hecht¹, Aharon Bar-Hillel², Stas Tiomkin¹, Hadar Levi¹, Omer Tsimhoni³, Naftali Tishby¹
¹Hebrew University of Jerusalem, Israel; ²Microsoft, Israel; ³General Motors, Israel
- PAGE 3399
Thu-O-44-3
14:40 - 15:00 **Counting Competing Speakers in a Timeframe — Human versus Computer**
Valentin Andrei, Horia Cucu, Andi Buzo, Corneliu Burileanu, Universitatea Politehnica din București, Romania
- PAGE 3404
Thu-O-44-4
15:00 - 15:20 **Segmental Contribution to the Intelligibility of Ideal Binary-Masked Sentences**
Fei Chen¹, Alexander Siu Tai Kwok²
¹SUSTC, China; ²University of Hong Kong, China
- PAGE 3408
Thu-O-44-5
15:20 - 15:40 **Perception of an Existing and Non-Existing L2 English Phoneme Behind Noise by Japanese Native Speakers**
Mako Ishida, Takayuki Arai, Sophia University, Japan
- PAGE 3412
Thu-O-44-6
15:40 - 16:00 **Viseme Comparison Based on Phonetic Cues for Varying Speech Accents**
Chitralekha Bhat, Sunil Kopparapu, TCS Innovation Labs Mumbai, India

Oral Session 45: Audio Signal Analysis and Representation

Chairs: Keisuke Kinoshita, Pejman Mowlae

Room: Conference 1, Time 14:00 – 16:00, Thursday, September 10, 2015

- PAGE 3417
Thu-O-45-1
14:00 – 14:20
- Quantifying Difference in Vocalizations of Bird Populations**
Colm O'Reilly, Nicola M. Marples, David J. Kelly, Naomi Harte, Trinity College Dublin, Ireland
- PAGE 3422
Thu-O-45-2
14:20 – 14:40
- Reverberation-Robust Acoustic Indoor Localization**
Jae Choi, Jeunghun Kim, Shin Jae Kang, Nam Soo Kim, Seoul National University, Korea
- PAGE 3426
Thu-O-45-3
14:40 – 15:00
- An Alternating Optimization Approach for Phase Retrieval**
Huaiping Ming¹, Dong-Yan Huang², Lei Xie¹, Haizhou Li², Minghui Dong²
*¹Northwestern Polytechnical University, China; ²A*STAR, Singapore*
- PAGE 3431
Thu-O-45-4
15:00 – 15:20
- Learning to Estimate Reverberation Time in Noisy and Reverberant Rooms**
Xiong Xiao¹, Shengkui Zhao², Xionghu Zhong³, Douglas L. Jones², Eng Siong Chng¹, Haizhou Li¹
¹TL@NTU, Singapore; ²Advanced Digital Sciences Center, Singapore; ³NTU, Singapore
- PAGE 3436
Thu-O-45-5
15:20 – 15:40
- Direction of Arrival Estimation Based on Reverberation Weighting and Noise Error Estimator**
Cheng Pang, Jie Zhang, Hong Liu, Peking University, China
- PAGE 3441
Thu-O-45-6
15:40 – 16:00
- Representing Nonspeech Audio Signals Through Speech Classification Models**
Huy Phan, Lars Hertel, Marco Maass, Radoslaw Mazur, Alfred Mertins, Universität zu Lübeck, Germany

Oral Session 46: Robustness in Speaker Recognition

Chairs: Joseph P. Campbell, Pietro Laface

Room: Conference 2+3, Time 14:00 – 16:00, Thursday, September 10, 2015

- PAGE 3446
Thu-O-46-1
14:00 – 14:20
- Mitigating the Effects of Non-Stationary Unseen Noises on Language Recognition Performance**
Luciana Ferrer¹, Mitchell McLaren², Aaron Lawson², Martin Graciarena²
¹Universidad de Buenos Aires, Argentina; ²SRI International, USA
- PAGE 3451
Thu-O-46-2
14:20 – 14:40
- An Information Theory Based Data-Homogeneity Measure for Voice Comparison**
Moez Ajili¹, Jean-François Bonastre¹, Solange Rossato², Juliette Kahn³, Itshak Lapidot⁴
¹LIA, France; ²LIG (UMR 5217), France; ³LNE, France; ⁴Afeka, Israel
- PAGE 3456
Thu-O-46-3
14:40 – 15:00
- The QUT-NOISE-SRE Protocol for the Evaluation of Noisy Speaker Recognition**
David Dean, Ahilan Kanagasundaram, Houman Ghaemmaghami, Md. Hafizur Rahman, Sridha Sridharan, Queensland University of Technology, Australia
- PAGE 3461
Thu-O-46-4
15:00 – 15:20
- Score Stabilization for Speaker Recognition Trained on a Small Development Set**
Hagai Aronowitz, IBM Research Haifa, Israel
- PAGE 3466
Thu-O-46-5
15:20 – 15:40
- Anti-Spoofing System: An Investigation of Measures to Detect Synthetic and Human Speech**
Abhinav Misra, Shivesh Ranjan, Chunlei Zhang, John H.L. Hansen, University of Texas at Dallas, USA
- PAGE 3471
Thu-O-46-6
15:40 – 16:00
- A Likelihood Ratio-Based Forensic Voice Comparison in Microphone vs. Mobile Mismatched Conditions Using Japanese /ai/**
Michael J. Carne, Australian National University, Australia

Oral Session 47: Evaluation of Speech Synthesis

Chairs: Alan W. Black, Rob Clark

Room: Conference 4+5, Time 14:00 - 16:00, Thursday, September 10, 2015

- PAGE 3476
Thu-O-47-1
14:00 - 14:20
- Are We Using Enough Listeners? No! — An Empirically-Supported Critique of Interspeech 2014 TTS Evaluations**
Mirjam Wester, Cassia Valentini-Botinhao, Gustav Eje Henter, University of Edinburgh, UK
- PAGE 3481
Thu-O-47-2
14:20 - 14:40
- How to Compare TTS Systems: A New Subjective Evaluation Methodology Focused on Differences**
Jonathan Chevelu¹, Damien Lolive¹, Sébastien Le Maguer², David Guennec¹
¹IRISA, France; ²Universität des Saarlandes, Germany
- PAGE 3486
Thu-O-47-3
14:40 - 15:00
- Double-Ended Prediction of the Naturalness Ratings of the Blizzard Challenge 2008–2013**
Lukas Latacz, Werner Verhelst, Vrije Universiteit Brussel, Belgium
- PAGE 3491
Thu-O-47-4
15:00 - 15:20
- Entropy-Based Sentence Selection for Speech Synthesis Using Phonetic and Prosodic Contexts**
Takashi Nose¹, Yusuke Arao², Takao Kobayashi², Komei Sugiura³, Yoshinori Shiga³, Akinori Ito¹
¹Tohoku University, Japan; ²Tokyo Institute of Technology, Japan; ³NICT, Japan
- PAGE 3496
Thu-O-47-5
15:20 - 15:40
- A Comparison of Speech Synthesis Systems Based on GPR, HMM, and DNN with a Small Amount of Training Data**
Tomoki Koriyama, Takao Kobayashi, Tokyo Institute of Technology, Japan
- PAGE 3501
Thu-O-47-6
15:40 - 16:00
- Objective Intelligibility Assessment of Text-to-Speech Systems Through Utterance Verification**
Raphael Ullmann, Ramya Rasipuram, Mathew Magimai-Doss, Hervé Bourlard, Idiap Research Institute, Switzerland

INTERSPEECH 2015 Table of Contents

Oral Session 48: Adaptive Methods for LVCSR

Chairs: Mark Gales, Brian Mak

Room: Conference 6, Time 14:00 - 16:00, Thursday, September 10, 2015

- PAGE 3506
Thu-O-48-1
14:00 - 14:20
- Continuous Word Representation Using Neural Networks for Proper Name Retrieval from Diachronic Documents**
Dominique Fohr, Irina Illina, LORIA, France
- PAGE 3511
Thu-O-48-2
14:20 - 14:40
- Recurrent Neural Network Language Model Adaptation for Multi-Genre Broadcast Speech Recognition**
X. Chen, T. Tan, Xunying Liu, Pierre Lanchantin, M. Wan, Mark J.F. Gales, Philip C. Woodland, University of Cambridge, UK
- PAGE 3516
Thu-O-48-3
14:40 - 15:00
- Paragraph Vector Based Topic Model for Language Model Adaptation**
Wengong Jin, Tianxing He, Yanmin Qian, Kai Yu, Shanghai Jiao Tong University, China
- PAGE 3521
Thu-O-48-4
15:00 - 15:20
- Personalized Speech Recognizer with Keyword-Based Personalized Lexicon and Language Model Using Word Vector Representations**
Ching-Feng Yeh, Yuan-ming Liou, Hung-yi Lee, Lin-shan Lee, National Taiwan University, Taiwan
- PAGE 3526
Thu-O-48-5
15:20 - 15:40
- Discriminative Data Selection for Lightly Supervised Training of Acoustic Model Using Closed Caption Texts**
Sheng Li, Yuya Akita, Tatsuya Kawahara, Kyoto University, Japan
- PAGE 3531
Thu-O-48-6
15:40 - 16:00
- Cross-Lingual Transfer Learning During Supervised Training in Low Resource Scenarios**
Amit Das, Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign, USA

Special Session 9: Robust Speech Processing Using Observation Uncertainty and Uncertainty Propagation

Chairs: Ramón F. Astudillo, Shinji Watanabe

Room: Hall 1, Time 14:00 - 16:00, Thursday, September 10, 2015

Thu-SP9-1
14:00 - 14:15

Session and Paper Overview

Ramón F. Astudillo¹, Shinji Watanabe², Ahmed Hussen Abdelaziz³,
Dorothea Kolossa³

¹INESC-ID Lisboa, Portugal; ²MERL, USA; ³Ruhr-Universität Bochum, Germany

Poster Presentations

PAGE 3536
Thu-SP9-1

Uncertainty Propagation for Noise Robust Speaker Recognition: The Case of NIST-SRE (Poster)

Dayana Ribas¹, Emmanuel Vincent², José Ramón Calvo¹

¹CENATAV, Cuba; ²INRIA, France

PAGE 3541
Thu-SP9-2

Uncertainty Training and Decoding Methods of Deep Neural Networks Based on Stochastic Representation of Enhanced Features (Poster)

Yuuki Tachioka¹, Shinji Watanabe²

¹Mitsubishi Electric, Japan; ²MERL, USA

PAGE 3546
Thu-SP9-3

Accounting for Uncertainty of i-Vectors in Speaker Recognition Using Uncertainty Propagation and Modified Imputation (Poster)

Rahim Saeidi, Paavo Alku, Aalto University, Finland

Thu-SP9 continued...

PAGE 3551
Thu-SP9-4

Autoencoder Based Multi-Stream Combination for Noise Robust Speech Recognition (Poster)

Sri Harish Mallidi¹, Tetsuji Ogawa², Karel Vesely³, Phani S. Nidadavolu¹,
Hynek Hermansky¹

¹Johns Hopkins University, USA; ²Waseda University, Japan; ³Brno University of
Technology, Czech Republic

PAGE 3556
Thu-SP9-5

Uncertainty Decoding for DNN-HMM Hybrid Systems Based on Numerical Sampling (Poster)

Christian Huemmer¹, Roland Maas¹, Andreas Schwarz¹, Ramón F. Astudillo²,
Walter Kellermann¹

¹FAU Erlangen-Nürnberg, Germany; ²INESC-ID Lisboa, Portugal

PAGE 3561
Thu-SP9-6

Uncertainty Propagation Through Deep Neural Networks (Poster)

Ahmed Hussen Abdelaziz¹, Shinji Watanabe², John R. Hershey², Emmanuel Vincent³,
Dorothea Kolossa¹

¹Ruhr-Universität Bochum, Germany; ²MERL, USA; ³INRIA, France

PAGE 3566
Thu-SP9-7

Handling Derivative Filterbank Features in Bounded-Marginalization-Based Missing Data Automatic Speech Recognition (Poster)

Marco Kühne, VoiceBox Technologies, Germany

PAGE 3571
Thu-SP9-8

Large-Scale, Sequence-Discriminative, Joint Adaptive Training for Masking-Based Robust ASR (Poster)

Arun Narayanan, Ananya Misra, Kean K. Chin, Google, USA

PAGE 3576
Thu-SP9-9

Integration of DNN Based Speech Enhancement and ASR (Poster)

Ramón F. Astudillo, Joana Correia, Isabel Trancoso, INESC-ID Lisboa, Portugal

Poster Session 30: Acoustic Model Adaptation and Training

Chair: Douglas O'Shaughnessy

Room: Hall 2, Time 14:00 - 16:00, Thursday, September 10, 2015

- PAGE 3581
Thu-P-30-1 **A General Artificial Neural Network Extension for HTK**
C. Zhang, Philip C. Woodland, University of Cambridge, UK
- PAGE 3586
Thu-P-30-2 **Audio Augmentation for Speech Recognition**
Tom Ko¹, Vijayaditya Peddinti², Daniel Povey², Sanjeev Khudanpur²
¹Huawei Technologies, China; ²Johns Hopkins University, USA
- PAGE 3590
Thu-P-30-3 **A Diversity-Penalizing Ensemble Training Method for Deep Learning**
Xiaohui Zhang, Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, USA
- PAGE 3595
Thu-P-30-4 **Deep Neural Network Training Emphasizing Central Frames**
Gakuto Kurata¹, Daniel Willett²
¹IBM Research Tokyo, Japan; ²Nuance Communications, Germany
- PAGE 3600
Thu-P-30-5 **Training Deep Bidirectional LSTM Acoustic Model for LVCSR by a Context-Sensitive-Chunk BPTT Approach**
Kai Chen¹, Zhi-Jie Yan², Qiang Huo²
¹USTC, China; ²Microsoft, China
- PAGE 3605
Thu-P-30-6 **Structured Output Layer with Auxiliary Targets for Context-Dependent Acoustic Modelling**
Pawel Swietojanski, Peter Bell, Steve Renals, University of Edinburgh, UK
- PAGE 3610
Thu-P-30-7 **Complementary Tasks for Context-Dependent Deep Neural Network Acoustic Models**
Peter Bell, Steve Renals, University of Edinburgh, UK

INTERSPEECH 2015 Table of Contents

Thu-P-30 continued...

- PAGE 3615
Thu-P-30-8 **Towards End-to-End Speech Recognition for Chinese Mandarin Using Long Short-Term Memory Recurrent Neural Networks**
Jie Li, Heng Zhang, Xinyuan Cai, Bo Xu, Chinese Academy of Sciences, China
- PAGE 3620
Thu-P-30-9 **Improving Deep Neural Networks Based Multi-Accent Mandarin Speech Recognition Using I-Vectors and Accent-Specific Top Layer**
Mingming Chen¹, Zhanlei Yang¹, Jizhong Liang², Yanpeng Li², Wenju Liu¹
¹Chinese Academy of Sciences, China; ²SGCC, China
- PAGE 3625
Thu-P-30-10 **Rapid Adaptation for Deep Neural Networks Through Multi-Task Learning**
Zhen Huang¹, Jinyu Li², Sabato Marco Siniscalchi¹, I-Fan Chen¹, Ji Wu¹, Chin-Hui Lee¹
¹Georgia Institute of Technology, USA; ²Microsoft, USA
- PAGE 3630
Thu-P-30-11 **fMLLR Based Feature-Space Speaker Adaptation of DNN Acoustic Models**
Sree Hari Krishnan Parthasarathi¹, Bjorn Hoffmeister¹, Spyros Matsoukas¹, Arindam Mandal¹, Nikko Strom¹, Sri Garimella²
¹Amazon.com, USA; ²Amazon.com, India
- PAGE 3635
Thu-P-30-12 **I-Vector Dependent Feature Space Transformations for Adaptive Speech Recognition**
Xiangang Li, Xihong Wu, Peking University, China
- PAGE 3640
Thu-P-30-13 **Unsupervised Domain Discovery Using Latent Dirichlet Allocation for Acoustic Modelling in Speech Recognition**
Mortaza Doulaty, Oscar Saz, Thomas Hain, University of Sheffield, UK
- PAGE 3645
Thu-P-30-14 **Training Data Selection for Acoustic Modeling via Submodular Optimization of Joint Kullback-Leibler Divergence**
Taichi Asami, Ryo Masumura, Hirokazu Masataki, Manabu Okamoto, Sumitaka Sakauchi, NTT Corporation, Japan

Poster Session 31: Spoken Term Detection, Spoken MT & Transliteration

Chair: Geoff Zweig

Room: Hall 2, Time 14:00 - 16:00, Thursday, September 10, 2015

- PAGE 3650
Thu-P-31-1 **Combination of NN and CRF Models for Joint Detection of Punctuation and Disfluencies**
Eunah Cho, Kevin Kilgour, Jan Niehues, Alex Waibel, KIT, Germany
- PAGE 3655
Thu-P-31-2 **Tunable Keyword-Aware Language Modeling and Context Dependent Fillers for LVCSR-Based Spoken Keyword Search**
Tze Siong Lau¹, I-Fan Chen², Chin-Hui Lee²
¹TL@NTU, Singapore; ²Georgia Institute of Technology, USA
- PAGE 3660
Thu-P-31-3 **Joint Decoding of Tandem and Hybrid Systems for Improved Keyword Spotting on Low Resource Languages**
Haipeng Wang, Anton Ragni, Mark J.F. Gales, Kate M. Knill, Philip C. Woodland, C. Zhang, University of Cambridge, UK
- PAGE 3665
Thu-P-31-4 **Preserving Word-Level Emphasis in Speech-to-Speech Translation Using Linear Regression HSMMs**
Quoc Truong Do, Shinnosuke Takamichi, Sakriani Sakti, Graham Neubig, Tomoki Toda, Satoshi Nakamura, NAIST, Japan
- PAGE 3670
Thu-P-31-5 **Phonology-Augmented Statistical Transliteration for Low-Resource Languages**
Hoang Gia Ngo¹, Nancy F. Chen², Binh Minh Nguyen¹, Bin Ma², Haizhou Li²
*¹NUS, Singapore; ²A*STAR, Singapore*

Thu-P-31 continued...

- PAGE 3675
Thu-P-31-6 **Evaluation of Re-Ranking by Prioritizing Highly Ranked Documents in Spoken Term Detection**
Kazuki Oouchi¹, Ryota Konno¹, Takahiro Akyu¹, Kazuma Konno¹, Kazunori Kojima¹, Kazuyo Tanaka², Shi-wook Lee³, Yoshiaki Itoh¹
¹Iwate Prefectural University, Japan; ²Tsukuba University, Japan; ³AIST, Japan
- PAGE 3680
Thu-P-31-7 **Distinctive Feature Based Representation of Speech for Query-by-Example Spoken Term Detection**
Abhijeet Saxena, B. Yegnanarayana, IIIT Hyderabad, India
- PAGE 3685
Thu-P-31-8 **Combination of Diverse Subword Units in Spoken Term Detection**
Shi-wook Lee¹, Kazuyo Tanaka², Yoshiaki Itoh³
¹AIST, Japan; ²Tsukuba University, Japan; ³Iwate Prefectural University, Japan
- PAGE 3690
Thu-P-31-9 **Sparse Modeling of Posterior Exemplars for Keyword Detection**
Dhananjay Ram, Afsaneh Asaei, Pranay Dighe, Hervé Bourlard, Idiap Research Institute, Switzerland

Poster Session 32: Stress, Load, and Pathologies

Chairs: John H. L. Hansen, Thomas F. Quatieri

Room: Foyer, Time 14:00 - 16:00, Thursday, September 10, 2015

- PAGE 3695
Thu-P-32-1 **Stress Level Detection Using Double-Layer Subband Filter**
*Tin Lay Nwe, Qianli Xu, Cuntai Guan, Bin Ma, A*STAR, Singapore*
- PAGE 3700
Thu-P-32-2 **Prosodic Characteristics of Read Speech Before and After Treadmill Running**
Jürgen Trouvain¹, Khiet P. Truong²
¹Universität des Saarlandes, Germany; ²University of Twente, The Netherlands
- PAGE 3705
Thu-P-32-3 **A Database for Analysis of Speech Under Physical Stress: Detection of Exercise Intensity While Running and Talking**
Khiet P. Truong¹, Arne Nieuwenhuys², Peter Beek³, Vanessa Evers¹
¹University of Twente, The Netherlands; ²Radboud Universiteit Nijmegen, The Netherlands; ³Vrije Universiteit Amsterdam, The Netherlands
- PAGE 3710
Thu-P-32-4 **Stressed Out: What Speech Tells Us About Stress**
Will Paul, Cecilia Ovesdotter Alm, Reynold Bailey, Joe Geigel, Linwei Wang, Rochester Institute of Technology, USA
- PAGE 3715
Thu-P-32-5 **Prediction of Heart Rate Changes from Speech Features During Interaction with a Misbehaving Dialog System**
Andreas Tsiartas, Andreas Kathol, Elizabeth Shriberg, Massimiliano de Zambotti, Adrian Willoughby, SRI International, USA
- PAGE 3720
Thu-P-32-6 **Acoustic Correlates for Perceived Effort Levels in Expressive Speech**
Mary Pietrowicz, Mark Hasegawa-Johnson, Karrie Karahalios, University of Illinois at Urbana-Champaign, USA

Thu-P-32 continued...

- PAGE 3725
Thu-P-32-7 **Pitch-Based Speech Perturbation Measures Using a Novel GCI Detection Algorithm: Application to Pathological Voice Classification**
Khalid Daoudi, Ashwini Jaya Kumar, INRIA, France
- PAGE 3729
Thu-P-32-8 **Speech-Based Assessment of PTSD in a Military Population Using Diverse Feature Classes**
Dimitra Vergyri¹, Bruce Knoth¹, Elizabeth Shriberg¹, Vikramjit Mitra¹, Mitchell McLaren¹, Luciana Ferrer¹, Pablo Garcia¹, Charles Marmar²
¹SRI International, USA; ²NYU Langone Medical Center, USA
- PAGE 3734
Thu-P-32-9 **Cognitive Impairment Prediction in the Elderly Based on Vocal Biomarkers**
Bea Yu¹, Thomas F. Quatieri¹, James R. Williamson¹, James C. Mundt²
¹MIT Lincoln Laboratory, USA; ²Center for Psych Consulting, USA
- PAGE 3739
Thu-P-32-10 **Automatic Age Detection in Normal and Pathological Voice**
J.-A. Gómez-García¹, L. Moro-Velázquez¹, Juan Ignacio Godino-Llorente¹, G. Castellanos-Domínguez²
¹Universidad Politécnica de Madrid, Spain; ²Universidad Nacional de Colombia, Colombia