

**12th Annual Conference of the
International Speech Communication
Association 2011**

(INTERSPEECH 2011)

**Florence, Italy
27-31 August 2011**

Volume 1 of 5

ISBN: 978-1-61839-270-1

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© (2011) by the International Speech Communications Association
All rights reserved.

Printed by Curran Associates, Inc. (2011)

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

International Speech Communications Association
c/o Emmanuelle Foxonet
Lous Tourils
F-66390 Baixas, France

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

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

TABLE OF CONTENTS

VOLUME 1

KEYNOTE SESSIONS

KEYNOTE 1

Chair: Isabel Trancoso

Speaking More Like You: Entrainment in Conversational Speech	1
<i>Julia Hirschberg</i>	

KEYNOTE 2

Chair: Giuseppe Riccardi

Neural Representations of Word Meanings	2
<i>Tom M. Mitchell</i>	

KEYNOTE 3

Chair: Giuseppe Riccardi

Honest Signals	3
<i>Alex (Sandy) Pentland</i>	
Signals and Speech	4
<i>Alex (Sandy) Pentland</i>	

KEYNOTE 4

Chair: Giuseppe Riccardi

Round Table: Future and Applications of Speech and Language Technologies for the Good Health of Society	8
<i>Gabriele Miceli, Björn Granström, Hiroshi Ishiguro</i>	

REGULAR ORAL SESSIONS

SPEAKER RECOGNITION – MODELING

Chair: Andrea Paoloni

Skew Gaussian Mixture Models for Speaker Recognition	12
<i>Avi Matza, Yuval Bistriz</i>	
Towards Goat Detection in Text-Dependent Speaker Verification	16
<i>Orith Toledo-Ronen, Hagai Aronowitz, Ron Hoory, Jason Pelecanos, David Nahamoo</i>	
Speaker Modeling Using Local Binary Decisions	20
<i>Jean-Francois Bonastre, Xavier Anguera, Gabriel H. Sierra, Pierre-Michel Bousquet</i>	
New Developments in Voice Biometrics for User Authentication	24
<i>Hagai Aronowitz, Ron Hoory, Jason Pelecanos, David Nahamoo</i>	
Evaluation of i-vector Speaker Recognition Systems for Forensic Application	28
<i>Miranti Indar Mandasari, Mitchell McLaren, David A. van Leeuwen</i>	
Mixture of PLDA Models in i-vector Space for Gender-Independent Speaker Recognition	32
<i>Mohammed Senoussaoui, Patrick Kenny, Niko Brummer, Edward de Villiers, Pierre Dumouchel</i>	

SPEECH PERCEPTION – SPEECH INTELLIGIBILITY

Chair: Anne Cutler

Segregation of Whispered Speech Interleaved with Noise or Speech Maskers	36
<i>Nandini Iyer, Douglas S. Brungart, Brian D. Simpson</i>	
Monaural Azimuth Localization Using Spectral Dynamics of Speech	40
<i>Roi Kliper, Hendrik Kayser, Daphna Weinshall, Israel Nelken, Jorn Anemuller</i>	
Prediction of Binaural Intelligibility Level Differences in Reverberation	44
<i>Jan Rannies, Thomas Brand, Birger Kollmeier</i>	
Let's All Speak Together! Exploring the Impact of Various Languages on the Comprehension of Speech in Multi-Linguistic Babble	48
<i>Aurore Gautreau, Michel Hoen, Fanny Meunier</i>	

Cross-Rate Variation in the Intelligibility of Dual-Rate Gated Speech in Older Listeners	52
<i>Valeriy Shafiro, Stanley Sheft, Robert Risley</i>	
An Effort-Inspired Auditory Model Front-End for Speech Recognition	56
<i>Chia-ying Lee, James Glass, Oded Ghitza</i>	

SPEECH REPRESENTATION AND MODELLING

Chair: Yannis Stylianou

A Long-Term Harmonic Plus Noise Model for Speech Signals	60
<i>Faten Ben Ali, Laurent Girin, Sonia Djaziri Larbi</i>	
A Frequency Domain Approach to ARX-LF Voiced Speech Parameterization and Synthesis	64
<i>Alan O Cinneide, David Dorran, Mikel Gainza, Eugene Coyle</i>	
Automatic Data-Driven Learning of Articulatory Primitives from Real-Time MRI Data Using Convolutional NMF with Sparseness Constraints	68
<i>Vikram Ramanarayanan, Athanasios Katsamanis, Shrikanth Narayanan</i>	
Online Pattern Learning for Non-Negative Convolutional Sparse Coding	72
<i>Dong Wang, Ravichander Vippera, Nicholas Evans</i>	
Sinewave Representations of Nonmodality	76
<i>Nicolas Malyska, Thomas F. Quatieri, Robert Dunn</i>	
Time-Varying Signal Adaptive Transform and IHT Recovery of Compressive Sensed Speech	80
<i>C. Srikanth Raj, T.V. Sreenivas</i>	

EMOTION, SPEAKING STYLE, AND SOCIAL BEHAVIOR

Chair: Anton Batliner

Acoustic-Linguistic Recognition of Interest in Speech with Bottleneck-BLSTM Nets	84
<i>Martin Wollmer, Felix Weninger, Florian Eyben, Bjorn Schuller</i>	
Automatic Detection of Anger in Human-Human Call Center Dialogs	88
<i>Mustafa Erden, Levent M. Arslan</i>	
Improved Classification of Speaking Styles for Mental Health Monitoring Using Phoneme Dynamics	92
<i>Keng-hao Chang, Howard Lei, John Canny</i>	
“You made me do it”: Classification of Blame in Married Couples' Interactions by Fusing Automatically Derived Speech and Language Information	96
<i>Matthew P. Black, Panayiotis G. Georgiou, Athanasios Katsamanis, Brian R. Baucom, Shrikanth Narayanan</i>	
Context and Priming Effects in the Recognition of Emotion of Old and Young Listeners	100
<i>Martijn Goudbeek, Marie Nilsenova</i>	
Acoustic and Prosodic Correlates of Social Behavior	104
<i>Agustin Gravano, Rivka Levitan, Laura Willson, Stefan Benus, Julia Hirschberg, Ani Nenkova</i>	

HMM-BASED SPEECH SYNTHESIS I

Chair: Keiichi Tokuda

Decision Tree-Based Clustering with Outlier Detection for HMM-Based Speech Synthesis	108
<i>Kyung Hwan Oh, June Sig Sung, Doo Hwa Hong, Nam Soo Kim</i>	
Prediction of Voice Aperiodicity Based on Spectral Representations in HMM Speech Synthesis	112
<i>Hanna Silen, Elina Helander, Moncef Gabbouj</i>	
A Perceptual Expressivity Modeling Technique for Speech Synthesis Based on Multiple-Regression HMM	116
<i>Takashi Nose, Takao Kobayashi</i>	
Multi-Speaker Modeling with Shared Prior Distributions and Model Structures for Bayesian Speech Synthesis	120
<i>Kei Hashimoto, Yoshihiko Nankaku, Keiichi Tokuda</i>	
Feature-Space Transform Tying in Unified Acoustic-Articulatory Modelling for Articulatory Control of HMM-Based Speech Synthesis	124
<i>Zhen-Hua Ling, Korin Richmond, Junichi Yamagishi</i>	
The Effect of Using Normalized Models in Statistical Speech Synthesis	128
<i>Matt Shannon, Heiga Zen, William Byrne</i>	

SPEAKER RECOGNITION – MODELING, AUTOMATIC PROCEDURES, ANALYSIS I

Chair: Luciano Romito

Restoring the Residual Speaker Information in Total Variability Modeling for Speaker Verification	132
<i>Ce Zhang, Rong Zheng, Bo Xu</i>	

New Developments in Joint Factor Analysis for Speaker Verification	136
<i>Hagai Aronowitz, Oren Barkan</i>	
Speaker Recognition Using Temporal Contours in Linguistic Units: The Case of Formant and Formant-Bandwidth Trajectories	140
<i>Joaquin Gonzalez-Rodriguez</i>	
Discriminatively Trained i-vector Extractor for Speaker Verification	144
<i>Ondrej Glembek, Lukas Burget, Niko Brummer, Oldrich Plchot, Pavel Matejka</i>	
Constrained Cepstral Speaker Recognition Using Matched UBM and JFA Training	148
<i>Michelle Hewlett Sanchez, Luciana Ferrer, Elizabeth Shriberg, Andreas Stolcke</i>	
A New Perspective on GMM Subspace Compensation Based on PCCA and Wiener Filtering	152
<i>Alan McCree, Douglas Sturim, Douglas Reynolds</i>	

SPEECH PERCEPTION – PERCEPTUAL LEARNING AND CROSS-LANGUAGE PERCEPTION

Chair: Catia Cucchiari

Perceptual Learning of Liquids	156
<i>Odette Scharenborg, Holger Mitterer, James M. McQueen</i>	
The Efficiency of Cross-Dialectal Word Recognition	160
<i>Annelie Tuinman, Holger Mitterer, Anne Cutler</i>	
Estimation of Perceptual Spaces for Speaker Identities Based on the Cross-Lingual Discrimination Task	164
<i>Minoru Tsuzaki, Keiichi Tokuda, Hisashi Kawai, Jinfu Ni</i>	
The Relation Between Perception and Production in L2 Phonological Processing	168
<i>Sharon Peperkamp, Camillia Bouchon</i>	
The Role of Word-Initial Glottal Stops in Recognizing English Words	172
<i>Maria Paola Bissiri, Maria Luisa Garcia Lecumberri, Martin Cooke, Jan Volin</i>	
Effect of Language Experience on the Categorical Perception of Cantonese Vowel Duration	176
<i>Caicai Zhang, Gang Peng, William S.-Y. Wang</i>	

SPEECH ANALYSIS

Chair: Thomas F. Quatieri

Adaptive Estimation of Zeros of Time-Varying Z-Transforms	180
<i>C.F. Pedersen, O. Andersen, P. Dalsgaard</i>	
Identifying Regions of Non-Modal Phonation Using Features of the Wavelet Transform	184
<i>John Kane, Christer Gobl</i>	
Acoustic Analysis of Whispered Speech for Phoneme and Speaker Dependency	188
<i>Xing Fan, Keith W. Godin, John H.L. Hansen</i>	
Multi-Party Speech Recovery Exploiting Structured Sparsity Models	192
<i>Afsaneh Asaei, Mohammad J. Taghizadeh, Herve Bourlard, Volkan Cevher</i>	
Modulation Spectrum Analysis for Recognition of Reverberant Speech	196
<i>Sri Harish Mallidi, Sriram Ganapathy, Hynek Hermansky</i>	
Discrete Choice Models for Non-Intrusive Quality Assessment	200
<i>Petko N. Petkov, W. Bastiaan Kleijn, Bert de Vries</i>	

SPEECH ENHANCEMENT AND DEREVERBERATION

Chair: Peter Kabal

Single Channel Dereverberation Using Example-Based Speech Enhancement with Uncertainty Decoding Technique	204
<i>Keisuke Kinoshita, Mehrez Souden, Marc Delcroix, Tomohiro Nakatani</i>	
A Statistical Room Impulse Response Model with Frequency Dependent Reverberation Time for Single-Microphone Late Reverberation Suppression	208
<i>Jan S. Erkelens, Richard Heusdens</i>	
An Assessment of the Improvement Potential of Time-Frequency Masking for Speech Dereverberation	212
<i>Chenxi Zheng, Tiago H. Falk, Wai-Yip Chan</i>	
Perceptual Improvement of a Two-Stage Algorithm for Speech Dereverberation	216
<i>Thiago de M. Prego, Amaro A. de Lima, Sergio L. Netto</i>	
A Model-Based Spectral Envelope Wiener Filter for Perceptually Motivated Speech Enhancement	220
<i>Najib Hadir, Friedrich Faubel, Dietrich Klakow</i>	

Binaural Noise-Reduction Method Based on Blind Source Separation and Perceptual Post Processing	224
<i>Jorge I. Marin-Hurtado, Devangi N. Parikh, David V. Anderson</i>	

ASR – FEATURE EXTRACTION II

Chair: Dong Yu

Region Dependent Transform on MLP Features for Speech Recognition	228
<i>Tim Ng, Bing Zhang, Spyros Matsoukas, Long Nguyen</i>	
Discriminant Sub-Space Projection of Spectro-Temporal Speech Features Based on Maximizing Mutual Information	232
<i>Martin Heckmann, Claudius Glaser</i>	
Combining Feature Space Discriminative Training with Long-Term Spectro-Temporal Features for Noise-Robust Speech Recognition	236
<i>Takashi Fukuda, Osamu Ichikawa, Masafumi Nishimura</i>	
Combining Frame and Segment Level Processing via Temporal Pooling for Phonetic Classification	240
<i>Sumit Chopra, Patrick Haffner, Dimitrios Dimitriadis</i>	
Improved Bottleneck Features Using Pretrained Deep Neural Networks	244
<i>Dong Yu, Michael L. Seltzer</i>	
Minimum Classification Error Based Spectro-Temporal Feature Extraction for Robust Audio Classification	248
<i>Yuan-Fu Liao, Chia-Hsing Lin, We-Der Fang</i>	

SPEAKER RECOGNITION – MODELING, AUTOMATIC PROCEDURES, ANALYSIS II

Chair: Kornel Laskowski

Data-Driven Gaussian Component Selection for Fast GMM-Based Speaker Verification	252
<i>Ce Zhang, Rong Zheng, Bo Xu</i>	
Analysis of i-vector Length Normalization in Speaker Recognition Systems	256
<i>Daniel Garcia-Romero, Carol Y. Espy-Wilson</i>	
An Analysis Framework Based on Random Subspace Sampling for Speaker Verification	260
<i>Weiwu Jiang, Zhifeng Li, Helen Meng</i>	
Factor Analysis Back Ends for MLLR Transforms in Speaker Recognition	264
<i>Nicolas Scheffer, Yun Lei, Luciana Ferrer</i>	
Report on Performance Results in the NIST 2010 Speaker Recognition Evaluation	268
<i>Craig S. Greenberg, Alvin F. Martin, Bradford N. Barr, George R. Doddington</i>	
iVector Fusion of Prosodic and Cepstral Features for Speaker Verification	272
<i>Marcel Kockmann, Luciana Ferrer, Lukas Burget, Jan Cernocky</i>	

SPEECH PRODUCTION – ARTICULATORY MEASUREMENTS

Chair: Paavo Alku

Visualization of Vocal Tract Shape Using Interleaved Real-Time MRI of Multiple Scan Planes	276
<i>Yoon-Chul Kim, Michael Proctor, Shrikanth Narayanan, Krishna S. Nayak</i>	
Biomechanical Tongue Models: An Approach to Studying Inter-Speaker Variability	280
<i>Ralf Winkler, Susanne Fuchs, Pascal Perrier, Mark Tiede</i>	
Quantifying Articulatory Distinctiveness of Vowels	284
<i>Jun Wang, Jordan R. Green, Ashok Samal, David B. Marx</i>	
Direct Estimation of Articulatory Kinematics from Real-Time Magnetic Resonance Image Sequences	288
<i>Michael Proctor, Adam Lammert, Athanasios Katsamanis, Louis Goldstein, Christina Hagedorn, Shrikanth Narayanan</i>	
Combined Optical Distance Sensing and Electropalatography to Measure Articulation	292
<i>Peter Birkholz, Christiane Neuschaefer-Rube</i>	
Simulating Post-L F0 Bouncing by Modeling Articulatory Dynamics	296
<i>Santitham Prom-on, Yi Xu, Fang Liu</i>	

ACOUSTIC EVENT DETECTION

Chair: Dirk van Compernelle

Learning New Acoustic Events in an HMM-Based System Using MAP Adaptation	300
<i>Jurgen T. Geiger, Mohamed Anouar Lakhal, Bjorn Schuller, Gerhard Rigoll</i>	
Alternative Frequency Scale Cepstral Coefficient for Robust Sound Event Recognition	304
<i>Yi Ren Leng, Huy Dat Tran, Norihide Kitaoka, Haizhou Li</i>	

Evaluation of Abnormal Sound Detection using Multi-Stage GMM in Various Environments	308
<i>Akinori Ito, Akihito Aiba, Masashi Ito, Shozo Makino</i>	
Unsupervised Learning of Acoustic Events Using Dynamic Time Warping and Hierarchical K-Means++ Clustering	312
<i>Joerg Schmalenstroerer, Markus Bartek, Reinhold Haeb-Umbach</i>	
Feature Extraction Assessment for an Acoustic-Event Classification Task Using the Entropy Triangle	316
<i>David Mejia-Navarrete, Ascension Gallardo-Antolin, Carmen Pelaez-Moreno, Francisco J. Valverde-Albacete</i>	
Unsupervised Audio Analysis for Categorizing Heterogeneous Consumer Domain Videos	320
<i>Pradeep Natarajan, Stavros Tsakalidis, Vasant Manohar, Rohit Prasad, Premkumar Natarajan</i>	

SPEECH SYNTHESIS – UNIT SELECTION AND HYBRID APPROACHES

Chair: Junichi Yamagishi

Enriching Text-to-Speech Synthesis Using Automatic Dialog Act Tags	324
<i>V. K. Rangarajan Sridhar, Ann Syrdal, Alistair D. Conkie, Srinivas Bangalore</i>	
Joint Target and Joint Cost Weight Training for Unit Selection Synthesis	328
<i>Lukas Latacz, Wesley Mattheyses, Werner Verhelst</i>	
Prominence-Based Prosody Prediction for Unit Selection Speech Synthesis	332
<i>Andreas Windmann, Igor Jauk, Fabio Tamburini, Petra Wagner</i>	
Evaluating the Meaning of Synthesized Listener Vocalizations	336
<i>Sathish Pammi, Marc Schroder</i>	
A Hybrid TTS Approach for Prosody and Acoustic Modules	340
<i>Inaki Sainz, Daniel Erro, Eva Navas, Inma Hernaez</i>	
Uniform Speech Parameterization for Multi-Form Segment Synthesis	344
<i>Alexander Sorin, Slava Shechtman, Vincent Pollet</i>	

SPEECH ENHANCEMENT ANALYSIS AND EVALUATION

Chair: Doug O'Shaughnessy

Theoretical Analysis of Musical Noise and Speech Distortion in Structure-Generalized Parametric Blind Spatial Subtraction Array	348
<i>Ryoichi Miyazaki, Hiroshi Saruwatari, Kiyohiro Shikano</i>	
Subjective and Objective Evaluation of Speech Intelligibility Enhancement Under Constant Energy and Duration Constraints	352
<i>Yan Tang, Martin Cooke</i>	
A Risk-Estimation-Based Comparison of Mean Square Error and Itakura-Saito Distortion Measures for Speech Enhancement	356
<i>Nagarjuna Reddy Muraka, Chandra Sekhar Seelamantula</i>	
On Noise Tracking for Noise Floor Estimation	360
<i>Mahdi Triki</i>	
Maximum a posteriori Estimation of Noise from Non-Acoustic Reference Signals in Very Low Signal-to-Noise Ratio Environments	364
<i>Ben Milner</i>	
Blind Speech Prior Estimation for Generalized Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator	368
<i>Ryo Wakisaka, Hiroshi Saruwatari, Kiyohiro Shikano, Tomoya Takatani</i>	

SPEAKER RECOGNITION – ANALYSIS AND STATISTICS I

Chair: David Van Leeuwen

Harmonic Structure Transform for Speaker Recognition	372
<i>Kornel Laskowski, Qin Jin</i>	
Combining Evidence from Spectral and Source-Like Features for Person Recognition from Humming	376
<i>Hemant A. Patil, Maulik C. Madhavi, Keshab K. Parhi</i>	
Improvements in Speaker Characterization Using Spectral Subband Energy Based on Harmonic plus Noise Model	380
<i>Yanhua Long, Zhi-Jie Yan, Frank K. Soong, Li-Rong Dai, Wu Guo</i>	
Implicit Segmentation in Two-Wire Speaker Recognition	384
<i>Yosef A. Solewicz, Hagai Aronowitz</i>	
Boosting Speaker Recognition Performance with Compact Representations	388
<i>Sibel Yaman, Jason Pelecanos, Mohamed Kamal Omar</i>	
Partitioning of Two-Speaker Conversation Datasets	392
<i>Carlos Vaquero, Alfonso Ortega, Eduardo Lleida</i>	

SPEECH PRODUCTION – COARTICULATION AND SPEECH TIMING

Chair: *Wim van Dommelen*

Jaw Movement in Vowels and Liquids Forming the Syllable Nucleus	396
<i>Stefan Benus, Marianne Pouplier</i>	
Coarticulation Across Prosodic Domains in Italian: An Ultrasound Investigation	400
<i>Barbara Gili Fivela, Antonio Stella, Sonia D'Apolito, Francesco Sigona</i>	
Investigating the Stability of Intergestural Timing Relations	404
<i>Juraj Simko, Fred Cummins, Stefan Benus</i>	
Speech Timing Organization for the Phonological Length Contrast in Italian Consonants	408
<i>Claudio Zmarich, Barbara Gili Fivela, Pascal Perrier, Christophe Savariaux, Graziano Tisato</i>	
Timing in Italian VNC Sequences at Different Speech Rates	412
<i>Chiara Celata, Silvia Calamai</i>	
Automatic Analysis of Singleton and Geminate Consonant Articulation Using Real-Time Magnetic Resonance Imaging	416
<i>Christina Hagedorn, Michael Proctor, Louis Goldstein</i>	

SPEECH SEGMENTATION

Chair: *Daniele Falavigna*

A Two-Stage Sample-Based Phone Boundary Detector Using Segmental Similarity Features	420
<i>Yih-Ru Wang</i>	
Iterative Improvement of Speaker Segmentation in a Noisy Environment Using High-Level Knowledge	424
<i>Qiang Huang, Stephen J. Cox</i>	
Hierarchical Audio Segmentation with HMM and Factor Analysis in Broadcast News Domain	428
<i>Diego Castan, Carlos Vaquero, Alfonso Ortega, David Martinez, Jesus Villalba, Eduardo Lleida</i>	
Syllable Segmentation of Continuous Speech Using Auditory Attention Cues	432
<i>Ozlem Kalinli</i>	
Exploiting Phone-Class Specific Landmarks for Refinement of Segment Boundaries in TTS Databases	436
<i>Vijayaditya Peddinti, Kishore Prahallad</i>	
Phoneme-Level Text to Audio Synchronization on Speech Signals with Background Music	440
<i>Agnes Pedone, Juan Jose Burred, Simon Maller, Pierre Leveau</i>	

ASR – ACOUSTIC MODELS II

Chair: *Frank Seide*

Conversational Speech Transcription Using Context-Dependent Deep Neural Networks	444
<i>Frank Seide, Gang Li, Dong Yu</i>	
Sequential Classification Criteria for NNs in Automatic Speech Recognition	448
<i>Guangsen Wang, Khe Chai Sim</i>	
Grapheme-Based Automatic Speech Recognition Using KL-HMM	452
<i>Mathew Magimai-Doss, Ramya Rasipuram, Guillermo Aradilla, Herve Boulard</i>	
Direct Error Rate Minimization of Hidden Markov Models	456
<i>Joseph Keshet, Chih-Chieh Cheng, Mark Stoehr, David McAllester</i>	
On the Effectiveness of Statistical Modeling Based Template Matching Approach for Continuous Speech Recognition	460
<i>Xie Sun, Xin Chen, Yunxin Zhao</i>	
Comparison of Smoothing Techniques for Robust Context Dependent Acoustic Modelling in Hybrid NN/HMM Systems	464
<i>Guangsen Wang, Khe Chai Sim</i>	

ROBUST SPEECH RECOGNITION II

Chair: *Maurizio Omologo*

Propagation of Uncertainty Through Multilayer Perceptrons for Robust Automatic Speech Recognition	468
<i>Ramon Fernandez Astudillo, Joao Paulo da Silva Neto</i>	
Mapping Sparse Representation to State Likelihoods in Noise-Robust Automatic Speech Recognition	472
<i>Katariina Mahkonen, Antti Hurmalainen, Tuomas Virtanen, Jort F. Gemmeke</i>	
Uncertainty Measures for Improving Exemplar-Based Source Separation	476
<i>Heikki Kallasjoki, Ulpu Remes, Jort F. Gemmeke, Tuomas Virtanen, Kalle J. Palomaki</i>	

Maximum Confidence Measure Based Interaural Phase Difference Estimation for Noise Masking in Dual-Microphone Robust Speech Recognition	480
<i>Hsien-Cheng Liao, Yuan-Fu Liao, Chin-Hui Lee</i>	
A Performance Monitoring Approach to Fusing Enhanced Spectrogram Channels in Robust Speech Recognition	484
<i>Shirin Badiezedegan, Richard Rose</i>	
Generalized Variable Parameter HMMs for Noise Robust Speech Recognition	488
<i>Ning Cheng, X. Liu, Lan Wang</i>	

SPEAKER RECOGNITION – ANALYSIS AND STATISTICS II

Chair: Mohammed Senoussaoui

Intersession Compensation and Scoring Methods in the i-vectors Space for Speaker Recognition	492
<i>Pierre-Michel Bousquet, Driss Matrouf, Jean-Francois Bonastre</i>	
Kernel Alignment Maximization for Speaker Recognition Based on High-Level Features	496
<i>Szymon Drgas, Adam Dabrowski</i>	
Kernel Partial Least Squares for Speaker Recognition	500
<i>Balaji Vasan Srinivasan, Daniel Garcia-Romero, Dmitry N. Zotkin, Ramani Duraiswami</i>	
Conversational-Side-Specific Inter-Session Variability Compensation	504
<i>Mohamed Kamal Omar, Jason Pelecanos</i>	
A Speaker Line-Up for the Likelihood Ratio	508
<i>David A. van Leeuwen, Niko Brummer</i>	
Towards Fully Bayesian Speaker Recognition: Integrating Out the Between-Speaker Covariance	512
<i>Jesus Villalba, Niko Brummer</i>	

PHYSIOLOGY AND PATHOLOGY OF SPOKEN LANGUAGE

Chair: Tim Bunnell

Novel VTEO Based Mel Cepstral Features for Classification of Normal and Pathological Voices	516
<i>Hemant A. Patil, Pallavi N. Baljekar</i>	
Temporal Performance of Dysarthric Patients in Speech and Tapping Tasks	520
<i>Eiji Shimura, Kazuhiko Kakehi</i>	
A Comparative Acoustic Study on Speech of Glossectomy Patients and Normal Subjects	524
<i>Xinhui Zhou, Maureen Stone, Carol Y. Espy-Wilson</i>	
Dysperiodicity Analysis of Perceptually Assessed Synthetic Speech Stimuli	528
<i>Ali Alpan, Francis Grenez, Jean Schoentgen</i>	
Is the Perception of Voice Quality Language-Dependant? A Comparison of French and Italian Listeners and Dysphonic Speakers	532
<i>Alain Ghio, Frederique Weisz, Giovanna Baracca, Giovanna Cantarella, Daniele Robert, Virginie Woisard, Franco Fussi, Antoine Giovanni</i>	
Automatic Selection of Acoustic and Non-Linear Dynamic Features in Voice Signals for Hypernasality Detection	536
<i>J.R. Orozco-Arroyave, S. Murillo-Rendon, A.M. Alvarez-Meza, J.D. Arias-Londono, E. Delgado-Trejos, J.F. Vargas-Bonilla, C.G. Castellanos-Dominguez</i>	

ASR – LEXICAL, PROSODIC AND MULTI-LINGUAL MODELS

Chair: Murat Saraclar

Learning from Mistakes: Expanding Pronunciation Lexicons Using Word Recognition Errors	540
<i>Sravana Reddy, Evandro Gouvea</i>	
Improving Non-Native ASR Through Stochastic Multilingual Phoneme Space Transformations	544
<i>David Imseng, Herve Bourlard, John Dines, Philip N. Garner, Mathew Magimai-Doss</i>	
Unsupervised Arabic Dialect Adaptation with Self-Training	548
<i>Scott Novotney, Rich Schwartz, Sanjeev Khudanpur</i>	
Template-Based Automatic Speech Recognition Meets Prosody	552
<i>Dino Seppi, Kris Demuyne, Dirk Van Compernelle</i>	
Pronunciation Learning from Continuous Speech	556
<i>Ibrahim Badr, Ian McGraw, James Glass</i>	
State-Level Data Borrowing for Low-Resource Speech Recognition Based on Subspace GMMs	560
<i>Yanmin Qian, Daniel Povey, Jia Liu</i>	

SOURCE SEPARATION

Chair: Tomohiro Nakatani

Blind Speech Separation in Multiple Environments Using a Frequency Oriented PCA Method for Convolutional Mixtures	564
<i>Y. Benabderrahmane, Sid-Ahmed Selouani, Douglas O'Shaughnessy</i>	
Blind Speech Separation in Time-Domain Using Block-Toeplitz Structure of Reconstructed Signal Matrices	568
<i>Zbynek Koldovsky, Jiri Malek, Petr Tichavsky</i>	
Generalized Method for Solving the Permutation Problem in Frequency-Domain Blind Source Separation of Convolved Speech Signals	572
<i>Auxiliadora Sarmiento, Ivan Duran, Sergio Cruces, Pablo Aguilera</i>	
Adaptation of Speaker-Specific Bases in Non-Negative Matrix Factorization for Single Channel Speech-Music Separation	576
<i>Emad M. Grais, Hakan Erdogan</i>	
An Informed Source Separation System for Speech Signals	580
<i>Shuhua Zhang, Laurent Girin</i>	
Adaptive Blocking Beamformer for Speech Separation	584
<i>Ngoc Thuy Tran, William Cowley, Andre Pollok</i>	

MULTIMODAL SIGNAL PROCESSING

Chair: Keikichi Hirose

Asynchronous Multimodal Text Entry Using Speech and Gesture Keyboards	588
<i>Per Ola Kristensson, Keith Vertanen</i>	
Robust Bimodal Person Identification Using Face and Speech with Limited Training Data and Corruption of Both Modalities	592
<i>Niall McLaughlin, Ji Ming, Danny Crookes</i>	
Toward a Multi-Speaker Visual Articulatory Feedback System	596
<i>Atef Ben Youssef, Thomas Hueber, Pierre Badin, Gerard Bailly</i>	
Statistical Mapping Between Articulatory and Acoustic Data for an Ultrasound-Based Silent Speech Interface	600
<i>Thomas Hueber, Elie-Laurent Benaroya, Bruce Denby, Gerard Chollet</i>	
Unsupervised Geometry Calibration of Acoustic Sensor Networks Using Source Correspondences	604
<i>Joerg Schmalenstroeer, Florian Jacob, Reinhold Haeb-Umbach, Marius H. Hennecke, Gernot A. Fink</i>	
Investigations on Speaking Mode Discrepancies in EMG-Based Speech Recognition	608
<i>Michael Wand, Matthias Janke, Tanja Schultz</i>	

ASR – LANGUAGE MODELS II

Chair: Stephan Kanthak

Empirical Evaluation and Combination of Advanced Language Modeling Techniques	612
<i>Tomas Mikolov, Anoop Deoras, Stefan Kombrink, Lukas Burget, Jan Cernocky</i>	
Personalizing Model M for Voice-Search	616
<i>Geoffrey Zweig, Shuangyu Chang</i>	
Sentence Selection by Direct Likelihood Maximization for Language Model Adaptation	620
<i>Takahiro Shinzaki, Yu Kubota, Sadaoki Furui, Eiji Utsunomiya, Yasutaka Shindoh</i>	
Feature Combination Approaches for Discriminative Language Models	624
<i>Ebru Arisoy, Bhuvana Ramabhadran, Hong-Kwang Jeff Kuo</i>	
On-Line Language Model Biasing for Multi-Pass Automatic Speech Recognition	628
<i>S. Ananthakrishnan, Stavros Tsakalidis, Rohit Prasad, Premkumar Natarajan</i>	
Mandarin Word-Character Hybrid-Input Neural Network Language Model	632
<i>Moonyoung Kang, Tim Ng, Long Nguyen</i>	

PHONOLOGY AND PHONETICS

Chair: Mark Hasegawa-Johnson

Laryngealization and Breathiness in Persian	636
<i>Vahid Sadeghi</i>	
Age-Dependent Differences in the Neutralization of the Intervocalic Voicing Contrast: Evidence from an Apparent-Time Study on East Franconian	640
<i>Viola Muller, Jonathan Harrington, Felicitas Kleber, Ulrich Reubold</i>	

Comparing Syllable Frequencies in Corpora of Written and Spoken Language	644
<i>Barbara Samlowski, Bernd Mobius, Petra Wagner</i>	
Sylli: Automatic Phonological Syllabification for Italian	648
<i>Luca Iacononi, Renata Savy</i>	
A Preliminary Study on the Production of Signs in Brazilian Sign Language when One of the Manual Articulators is Unavailable	652
<i>Andre N. Xavier, Plinio A. Barbosa</i>	
Electroglottograph and Acoustic Cues for Phonation Contrasts in Taiwan Min Falling Tones	656
<i>Ho-hsien Pan, Mao-Hsu Chen, Shao-Ren Lyu</i>	

VOICE CONVERSION

Chair: Junichi Yamagishi

One-to-Many Voice Conversion Based on Tensor Representation of Speaker Space	660
<i>Daisuke Saito, Keisuke Yamamoto, Nobuaki Minematsu, Keiichi Hirose</i>	
A Study on Bag of Gaussian Model with Application to Voice Conversion	664
<i>Yu Qiao, Tong Tong, Nobuaki Minematsu</i>	
A Bayesian Approach to Voice Conversion Based on GMMs Using Multiple Model Structures	668
<i>Lei Li, Yoshihiko Nankaku, Keiichi Tokuda</i>	
Quality Improvement of Voice Conversion Systems Based on Trellis Structured Vector Quantization	672
<i>Mahdi Eslami, Hamid Sheikhzadeh, Abolghasem Sayadiyan</i>	
Voice Conversion Using GMM with Enhanced Global Variance	676
<i>Hadas Benisty, David Malah</i>	

VOLUME 2

Spectral Envelope Transformation Using DFW and Amplitude Scaling for Voice Conversion with Parallel or Nonparallel Corpora	680
<i>Elizabeth Godoy, Olivier Rosec, Thierry Chonavel</i>	

ROBUST SPEECH RECOGNITION III

Chair: Richard Stern

Sinusoidal Approach for the Single-Channel Speech Separation and Recognition Challenge	684
<i>P. Mowlae, R. Saeidi, Zheng-Hua Tan, M.G. Christensen, Tomi Kinnunen, P. Franti, S.H. Jensen</i>	
Semi-Supervised Single-Channel Speech-Music Separation for Automatic Speech Recognition	688
<i>Cemil Demir, A. Taylan Cemgil, Murat Saraclar</i>	
A Level-Dependent Auditory Filter-Bank for Speech Recognition in Reverberant Environments	692
<i>HariKrishna Maganti, Marco Matassoni</i>	
A Multichannel Feature-Based Processing for Robust Speech Recognition	696
<i>Mehrez Souden, Keisuke Kinoshita, Marc Delcroix, Tomohiro Nakatani</i>	
Feature Normalization Using Structured Full Transforms for Robust Speech Recognition	700
<i>Xiong Xiao, Jinyu Li, Eng Siong Chng, Haizhou Li</i>	
A Robust Estimation Method of Noise Mixture Model for Noise Suppression	704
<i>Masakiyo Fujimoto, Shinji Watanabe, Tomohiro Nakatani</i>	

SPOKEN LANGUAGE UNDERSTANDING

Chair: Ruhi Sarikaya

Multi-Task Learning for Spoken Language Understanding with Shared Slots	708
<i>Xiao Li, Ye-Yi Wang, Gokhan Tur</i>	
Learning Weighted Entity Lists from Web Click Logs for Spoken Language Understanding	712
<i>Dustin Hillard, Asli Celikyilmaz, Dilek Hakkani-Tur, Gokhan Tur</i>	
Bootstrapping Domain Detection Using Query Click Logs for New Domains	716
<i>Dilek Hakkani-Tur, Gokhan Tur, Larry Heck, Elizabeth Shriberg</i>	
Approximate Inference for Domain Detection in Spoken Language Understanding	720
<i>Asli Celikyilmaz, Dilek Hakkani-Tur, Gokhan Tur</i>	
Speech Indexing Using Semantic Context Inference	724
<i>Chien-Lin Huang, Bin Ma, Haizhou Li, Chung-Hsien Wu</i>	
Automatically Optimizing Utterance Classification Performance without Human in the Loop	728
<i>Yun-Cheng Ju, Jasha Droppo</i>	

DIALECT AND ACCENT IDENTIFICATION

Chair: David Martinez

In Search of Cues Discriminating West-African Accents in French	732
<i>Philippe Boula de Mareuil, Jean-Luc Rouas, Manuela Yapomo</i>	
Computer and Human Recognition of Regional Accents of British English	736
<i>Abualsoud Hanani, Martin Russell, Michael J. Carey</i>	
Target-Aware Lattice Rescoring for Dialect Recognition	740
<i>Rong Tong, Bin Ma, Haizhou Li, Eng Siong Chng</i>	
Effective Arabic Dialect Classification Using Diverse Phonotactic Models	744
<i>Murat Akbacak, Dimitra Vergyri, Andreas Stolcke, Nicolas Scheffer, Arindam Mandal</i>	
Characterizing Deletion Transformations Across Dialects Using a Sophisticated Tying Mechanism	748
<i>Nancy F. Chen, Wade Shen, Joseph P. Campbell</i>	
Dialect and Accent Recognition Using Phonetic-Segmentation Supervectors	752
<i>Fadi Biadisy, Julia Hirschberg, Daniel P.W. Ellis</i>	

FIRST LANGUAGE ACQUISITION

Chair: Cinzia Avesani

The Multi Timescale Phoneme Acquisition Model of the Self-Organizing Based on the Dynamic Features	756
<i>Kouki Miyazawa, Hideaki Miura, Hideaki Kikuchi, Reiko Mazuka</i>	
The Time-Course of Talker-Specificity Effects for Newly-Learned Pseudowords: Evidence for a Hybrid Model of Lexical Representation	760
<i>Helen Brown, M. Gareth Gaskell</i>	
A Parametric Approach to Intonation Acquisition Research: Validation on Child-Directed Speech Data	764
<i>Britta Lintfert, Antje Schweitzer, Bernd Mobius</i>	
Modelling Novelty Preference in Word Learning	768
<i>Maarten Versteegh, Louis ten Bosch, Lou Boves</i>	
Using Imitation to Learn Infant-Adult Acoustic Mappings	772
<i>G. Ananthakrishnan, Giampiero Salvi</i>	
Thresholding Word Activations for Response Scoring --- Modelling Psycholinguistic Data	776
<i>Christina Bergmann, Louis ten Bosch, Lou Boves</i>	

ASR – ACOUSTIC MODELS III

Chair: Ralf Schlüter

Generalized Baum-Welch Algorithm and its Implication to a New Extended Baum-Welch Algorithm	780
<i>Roger Hsiao, Tanja Schultz</i>	
Word Boundary Modelling and Full Covariance Gaussians for Arabic Speech-to-Text Systems	784
<i>F. Diehl, M.J.F. Gales, X. Liu, M. Tomalin, P.C. Woodland</i>	
A Fully Automated Derivation of State-Based Eigentriphones for Triphone Modeling with No Tied States Using Regularization	788
<i>Tom Ko, Brian Mak</i>	
Reducing Computational Complexities of Exemplar-Based Sparse Representations with Applications to Large Vocabulary Speech Recognition	792
<i>Tara N. Sainath, Bhuvana Ramabhadran, David Nahamoo, Dimitri Kanevsky</i>	
An i-vector Based Approach to Training Data Clustering for Improved Speech Recognition	796
<i>Yu Zhang, Jian Xu, Zhi-Jie Yan, Qiang Huo</i>	
Rapid Training of Acoustic Models Using Graphics Processing Unit	800
<i>Senaka Buthpitiya, Ian Lane, Jike Chong</i>	

SPOKEN DIALOGUE SYSTEMS I

Chair: Olivier Pietquin

User Study of Spoken Decision Support System	804
<i>Teruhisa Misu, Kiyonori Ohtake, Chiori Hori, Hisashi Kawai, Satoshi Nakamura</i>	
Efficient Probabilistic Tracking of User Goal and Dialog History for Spoken Dialog Systems	808
<i>Antoine Raux, Yi Ma</i>	
Tackling a Shilly-Shally Classifier for Predicting Task Success in Spoken Dialogue Interaction	812
<i>Alexander Schmitt, Alexander Zgorzelski, Wolfgang Minker</i>	

Evaluation of Listening-Oriented Dialogue Control Rules Based on the Analysis of HMMs	816
<i>Toyomi Meguro, Yasuhiro Minami, Ryuichiro Higashinaka, Kohji Dohsaka</i>	
Large-Scale Experiments on Data-Driven Design of Commercial Spoken Dialog Systems	820
<i>D. Suendermann, J. Liscombe, J. Bloom, G. Li, R. Pieraccini</i>	
Comparing System-Driven and Free Dialogue in In-Vehicle Interaction	824
<i>Fredrik Kronlid, Jessica Villing, Alexander Berman, Staffan Larsson</i>	

SPOKEN LANGUAGE RESOURCES, EVALUATION AND STANDARDIZATION II

Chair: Paolo Baggia

Rapid Evaluation of Speech Representations for Spoken Term Discovery	828
<i>Michael A. Carlin, Samuel Thomas, Aren Jansen, Hynek Hermansky</i>	
Phonemic Similarity Metrics to Compare Pronunciation Methods	832
<i>Ben Hixon, Eric Schneider, Susan L. Epstein</i>	
Investigating the Effect of Number of Interlocutors on the Quality of Experience for Multi-Party Audio Conferencing	836
<i>Janto Skowronek, Alexander Raake</i>	
On Development of Consistently Punctuated Speech Corpora	840
<i>Jachym Kolar, Lori Lamel</i>	
A Multimodal Real-Time MRI Articulatory Corpus for Speech Research	844
<i>Shrikanth Narayanan, Erik Bresch, Prasanta Kumar Ghosh, Louis Goldstein, Athanasios Katsamanis, Yoon Kim, Adam Lammert, Michael Proctor, Vikram Ramanarayanan, Yinghua Zhu</i>	
Building an Audio-Visual Corpus of Australian English: Large Corpus Collection with an Economical Portable and Replicable Black Box	848
<i>Denis Burnham, Dominique Estival, Steven Fazio, Jette Viethen, Felicity Cox, Robert Dale, Steve Cassidy, Julien Epps, Roberto Togneri, Michael Wagner, Yuko Kinoshita, Roland Gocke, Joanne Arciuli, Marc Onslow, Trent Lewis, Andrew Butcher, John Hajek</i>	

LANGUAGE IDENTIFICATION

Chair: Philippe Boula de Mareuil

Data-Driven UBM Generation via Tied Gaussians for GMM-Supervector Based Accent Identification	852
<i>Rong Zheng, Ce Zhang, Bo Xu</i>	
I3A Language Recognition System for Albayzin 2010 LRE	856
<i>David Martinez, Jesus Villalba, Antonio Miguel, Alfonso Ortega, Eduardo Lleida</i>	
Dimensionality Reduction for Using High-Order n-Grams in SVM-Based Phonotactic Language Recognition	860
<i>Mikel Penagarikano, Amparo Varona, Luis Javier Rodriguez-Fuentes, German Bordel</i>	
Language Recognition via i-vectors and Dimensionality Reduction	864
<i>Najim Dehak, Pedro A. Torres-Carrasquillo, Douglas Reynolds, Reda Dehak</i>	
Language Recognition in iVectors Space	868
<i>David Martinez, Oldrich Plchot, Lukas Burget, Ondrej Glembek, Pavel Matejka</i>	

SECOND LANGUAGE ACQUISITION, DEVELOPMENT AND LEARNING II

Chair: Robert Fox

On Mispronunciation Lexicon Generation Using Joint-Sequence Multigrams in Computer-Aided Pronunciation Training (CAPT)	872
<i>Xiaojun Qian, Helen Meng, Frank K. Soong</i>	
Validating a Second Language Perception Model for Classroom Context --- A Longitudinal Study within the Perceptual Assimilation Model	876
<i>Bianca Sisinni, Mirko Grimaldi</i>	
The Role of Variability in Non-Native Perceptual Learning of a Japanese Geminata-Singleton Fricative Contrast	880
<i>Makiko Sadakata, James M. McQueen</i>	
Fluency Changes with General Progress in L2 Proficiency	884
<i>Jared Bernstein, Jian Cheng, Masanori Suzuki</i>	
Tongue Gestures Awareness and Pronunciation Training	888
<i>Slim Ouni</i>	
Impact of Speaker Variability on Speech Perception in Non-Native Listeners	892
<i>Wim A. van Dommelen, Valerie Hazan</i>	

ASR – SEARCH, KEYWORD SPOTTING AND CONFIDENCE MEASURES II

Chair: *Geoffrey Zweig*

A Template Based Voice Trigger System Using Bhattacharyya Edit Distance	896
<i>Evelyn Kurniawati, Samsudin Ng, Karthik Muralidhar, Sapna George</i>	
Acoustic Look-Ahead for More Efficient Decoding in LVCSR	900
<i>D. Nolden, Ralf Schluter, Hermann Ney</i>	
A New Epsilon Filter for Efficient Composition of Weighted Finite-State Transducers	904
<i>Frank Duckhorn, Matthias Wolff, Rudiger Hoffmann</i>	
A Bottom-Up Stepwise Knowledge-Integration Approach to Large Vocabulary Continuous Speech Recognition Using Weighted Finite State Machines	908
<i>Sabato Marco Siniscalchi, Torbjorn Svendsen, Chin-Hui Lee</i>	
Combining Information Sources for Confidence Estimation with CRF Models	912
<i>M.S. Seigel, P.C. Woodland</i>	
Evaluation of Fast Spoken Term Detection Using a Suffix Array	916
<i>Kouichi Katsurada, Shinta Sawada, Shigeki Teshima, Yurie Iribe, Tsuneo Nitta</i>	

SLP FOR INFORMATION EXTRACTION AND RETRIEVAL I

Chair: *Haizhou Li*

Latent Topic Modeling for Audio Corpus Summarization	920
<i>Timothy J. Hazen</i>	
Investigation of Spontaneous Speech Characterization Applied to Speaker Role Recognition	924
<i>Richard Dufour, Yannick Esteve, Paul Deleglise</i>	
Zero-Resource Audio-Only Spoken Term Detection Based on a Combination of Template Matching Techniques	928
<i>Armando Muscariello, Guillaume Gravier, Frederic Bimbot</i>	
Automatic Learning in Content Indexing Service Using Phonetic Alignment	932
<i>Yeon-Jun Kim, David C. Gibbon</i>	
Leveraging Relevance Cues for Improved Spoken Document Retrieval	936
<i>Pei-Ning Chen, Kuan-Yu Chen, Berlin Chen</i>	
Spoken Lecture Summarization by Random Walk over a Graph Constructed with Automatically Extracted Key Terms	940
<i>Yun-Nung Chen, Yu Huang, Ching-Feng Yeh, Lin-shan Lee</i>	

SPEAKER DIARIZATION I

Chair: *Janez Zibert*

Speaker Diarization Using a priori Acoustic Information	944
<i>Hagai Aronowitz</i>	
Improved Overlapped Speech Handling for Speaker Diarization	948
<i>Kofi Boakye, Oriol Vinyals, Gerald Friedland</i>	
Exploiting Intra-Conversation Variability for Speaker Diarization	952
<i>Stephen Shum, Najim Dehak, Ekapol Chuangsuwanich, Douglas Reynolds, James Glass</i>	
Speaker Clustering Based on Non-Negative Matrix Factorization	956
<i>Masafumi Nishida, Seiichi Yamamoto</i>	
Information Bottleneck Features for HMM/GMM Speaker Diarization of Meetings Recordings	960
<i>Sree Harsha Yella, Fabio Valente</i>	
Cross Likelihood Ratio Based Speaker Clustering Using Eigenvoice Models	964
<i>D. Wang, Robbie Vogt, Sridha Sridharan, David Dean</i>	

PROSODY I

Chair: *Gerard Bailly*

A Quantitative Investigation of the Prosody of Verum Focus in Italian	968
<i>Giuseppina Turco, Michele Gubian, Jessamyn Schertz</i>	
Effects of Focus on f_0 and Duration in Irish (Gaelic) Declaratives	972
<i>Amelie Dorn, Ailbhe Ni Chasaide</i>	
The Phonology and Phonetics of Perceived Prosody: What do Listeners Imitate?	976
<i>Jennifer Cole, Stefanie Shattuck-Hufnagel</i>	
Uncovering the Effect of Imitation on Tonal Patterns of French Accentual Phrases	980
<i>Amandine Michelas, Noel Nguyen</i>	

Crossmodal Prosodic and Gestural Contribution to the Perception of Contrastive Focus	984
<i>Pilar Prieto, Cecilia Pugliesi, Joan Borrás-Comes, Ernesto Arroyo, Josep Blat</i>	
Temporal Relationship Between Auditory and Visual Prosodic Cues	988
<i>Erin Cvejic, Jeeseun Kim, Chris Davis</i>	

ASR – NEW PARADIGMS

Chair: John Hansen

New Methods for Template Selection and Compression in Continuous Speech Recognition	992
<i>Xie Sun, Yunxin Zhao</i>	
Structured Support Vector Machines for Noise Robust Continuous Speech Recognition	996
<i>Shi-Xiong Zhang, M.J.F. Gales</i>	
Continuous Digits Recognition Leveraging Invariant Structure	1000
<i>Masayuki Suzuki, Gakuto Kurata, Masafumi Nishimura, Nobuaki Minematsu</i>	
Convergence of Line Search A-Function Methods	1004
<i>Dimitri Kanevsky, David Nahamoo, Tara N. Sainath, Bhuvana Ramabhadran</i>	
Hidden Boosted MMI and Hierarchical State Posterior Feature for Automatic Speech Recognition Based on Hidden Conditional Neural Fields	1008
<i>Yasuhisa Fujii, Kazumasa Yamamoto, Seiichi Nakagawa</i>	
Recognition and Real Time Performances of a Lightweight Ultrasound Based Silent Speech Interface Employing a Language Model	1012
<i>Jun Cai, Bruce Denby, Pierre Roussel, Gerard Dreyfus, Lise Crevier-Buchman</i>	

SPOKEN DIALOGUE SYSTEMS II

Chair: Steve Young

Optimizing Situated Dialogue Management in Unknown Environments	1016
<i>Heriberto Cuayahuitl, Nina Dethlefs</i>	
Acoustic-Similarity Based Technique to Improve Concept Recognition	1020
<i>Om D. Deshmukh, Shajith Iqbal, Ashish Verma, Etienne Marcheret</i>	
Dialog Methods for Improved Alphanumeric String Capture	1024
<i>Doug Peters, Peter Stubbley</i>	
Detecting the Status of a Predictive Incremental Speech Understanding Model for Real-Time Decision-Making in a Spoken Dialogue System	1028
<i>David DeVault, Kenji Sagae, David Traum</i>	
User Simulation in Dialogue Systems Using Inverse Reinforcement Learning	1032
<i>Senthilkumar Chandramohan, Matthieu Geist, Fabrice Lefevre, Olivier Pietquin</i>	
Lossless Value Directed Compression of Complex User Goal States for Statistical Spoken Dialogue Systems	1036
<i>Paul A. Crook, Oliver Lemon</i>	

SPEAKER DIARIZATION II

Chair: Hagai Aronowitz

Prosodic and Phonetic Features for Speaker Clustering in Speaker Diarization Systems	1040
<i>Janez Zibert, France Mihelic</i>	
Diarization-Based Speaker Retrieval for Broadcast Television Archives	1044
<i>Marijn Huijbregts, David A. van Leeuwen</i>	
The Detection of Overlapping Speech with Prosodic Features for Speaker Diarization	1048
<i>Martin Zelenak, Javier Hernando</i>	
LP Residual Features for Robust, Privacy-Sensitive Speaker Diarization	1052
<i>Sree Hari Krishnan Parthasarathi, Herve Boudlard, Daniel Gatica-Perez</i>	
Extending the Task of Diarization to Speaker Attribution	1056
<i>Houman Ghaemmaghami, David Dean, Robbie Vogt, Sridha Sridharan</i>	
Comparing Multi-Stage Approaches for Cross-Show Speaker Diarization	1060
<i>Viet-Anh Tran, Viet Bac Le, Claude Barras, Lori Lamel</i>	

PROSODY II

Chair: Pilar Prieto

Analysing the Correspondence Between Automatic Prosodic Segmentation and Syntactic Structure	1064
<i>Gyorgy Szaszak, Katalin Nagy, Andras Beke</i>	

Long-Distance Rhythmic Dependencies and their Application to Automatic Language Identification	1068
<i>Joseph Tepperman, Emily Nava</i>	
Symbolic and Direct Sequential Modeling of Prosody for Classification of Speaking-Style and Nativeness	1072
<i>Andrew Rosenberg</i>	
Prosodic Analysis and Perception of Mandarin Utterances Conveying Attitudes	1076
<i>Wentao Gu, Ting Zhang, Hiroya Fujisaki</i>	
Predicting Taiwan Mandarin Tone Shapes from their Duration	1080
<i>Chierh Cheng, Michele Gubian</i>	
Variation of Accent Type and of Context --- Influences on Pragmatic Focus Interpretation	1084
<i>Charlotte Wollermann, Ulrich Schade, Bernhard Schroder</i>	

ADAPTATION FOR ASR

Chair: Phil Woodland

Model Adaptation for Automatic Speech Recognition Based on Multiple Time Scale Evolution	1088
<i>Shinji Watanabe, Atsushi Nakamura, Biing-Hwang Juang</i>	
Integrated Online Speaker Clustering and Adaptation	1092
<i>C. Breslin, K.K. Chin, M.J.F. Gales, Kate Knill</i>	
A Study on Speaker Normalized MLP Features in LVCSR	1096
<i>Zoltan Tuske, Christian Plahl, Ralf Schluter</i>	
Matrix-Variate Distribution of Training Models for Robust Speaker Adaptation	1100
<i>Yongwon Jeong, Young Kuk Kim</i>	
Separating Speaker and Environmental Variability Using Factored Transforms	1104
<i>Michael L. Seltzer, Alex Acero</i>	
Your Mobile Virtual Assistant Just Got Smarter!	1108
<i>Mazin Gilbert, Iker Arizmendi, Enrico Bocchieri, Diamantino Caseiro, Vincent Goffin, Andrej Ljolje, Mike Phillips, Chao Wang, Jay Wilpon</i>	

SLP FOR INFORMATION EXTRACTION AND RETRIEVAL II

Chair: Pascale Fung

Topic Segmentation of TV-Streams by Mathematical Morphology and Vectorization	1112
<i>Vincent Claveau, Sebastien Lefevre</i>	
Probabilistic Latent Semantic Analysis for Broadcast News Story Segmentation	1116
<i>Mimi Lu, Cheung-Chi Leung, Lei Xie, Bin Ma, Haizhou Li</i>	
Hybrid Speech Recognition for Voice Search: A Comparative Study	1120
<i>Evandro Gouvea</i>	
A New Phonetic Candidate Generator for Improving Search Query Efficiency	1124
<i>Bo Peng, Yao Qian, Frank K. Soong, Bo Zhang</i>	
Towards Voice-Input Symbolic Pattern Retrieval Using Parameter-Based Search	1128
<i>Yukiko Suzuki, Kiyooki Aikawa</i>	
A Language Independent Approach to Audio Search	1132
<i>Vikram Gupta, Jitendra Ajmera, Arun Kumar, Ashish Verma</i>	

REGULAR POSTER SESSIONS

SECOND LANGUAGE ACQUISITION, DEVELOPMENT AND LEARNING I

Chair: Joaquim Llisterra

Acquisition of Timing Patterns in Second Language	1136
<i>Mikhail Ordin, Leona Polyanskaya, Christiane Ulbrich</i>	
Context-Dependent Duration Modeling with Backoff Strategy and Look-Up Tables for Pronunciation Assessment and Mispronunciation Detection	1140
<i>Hongyan Li, Shen Huang, Shijin Wang, Bo Xu</i>	
Perceptual Training of Vowel Length Contrast of Japanese by L2 Listeners: Effects of an Isolated Word versus a Word Embedded in Sentences	1144
<i>Mee Somu, Keiichi Tajima, Hiroaki Kato, Yoshinori Sagisaka</i>	
Similar Vowels in L1/L2 Production: Confused or Discerned in Early L2 English Learners with Different Amount of Exposure	1148
<i>E-Chin Wu</i>	
Production and Perception of Estonian Vowels by Native and Non-Native Speakers	1152
<i>Lya Meister, Einar Meister</i>	

New Feature Parameters for Pronunciation Evaluation in English Presentations at International Conferences	1156
<i>Hiroshi Kibishi, Seiichi Nakagawa</i>	
Synchronous Reading: Learning French Orthography by Audiovisual Training	1160
<i>Gerard Bailly, Will Barbour</i>	
Phoneme Level Non-Native Pronunciation Analysis by an Auditory Model-Based Native Assessment Scheme	1164
<i>Christos Koniaris, Olov Engwall</i>	
The Open Front Vowel /æ/ in the Production and Perception of Czech Students of English	1168
<i>Pavel Sturm, Radek Skarnitzl</i>	
Error Selection for ASR-Based English Pronunciation Training in 'My Pronunciation Coach'	1172
<i>Catia Cucchiari, Henk van den Heuvel, Eric Sanders, Helmer Strik</i>	
An Experimental Analysis of Pitch Patterns in Japanese Speakers of English with Verification by Speech Re-Synthesis	1176
<i>Tomoko Nariai, Kazuyo Tanaka</i>	
An Analysis of Word Duration in Native Speakers and Japanese Speakers of English	1180
<i>Tomoko Nariai, Kazuyo Tanaka, Yoshiaki Ito</i>	

SPEECH ENHANCEMENT

Chair: Dietrich Klakow

Evaluating Artificial Bandwidth Extension by Conversational Tests in Car Using Mobile Devices with Integrated Hands-Free Functionality	1184
<i>Laura Laaksonen, Ville Myllyla, Riitta Niemisto</i>	
Low-Frequency Bandwidth Extension of Telephone Speech Using Sinusoidal Synthesis and Gaussian Mixture Model	1188
<i>Hannu Pulakka, Ulpu Remes, Santeri Yrttiaho, Kalle J. Palomaki, Mikko Kurimo, Paavo Alku</i>	
Memory-Based Approximation of the Gaussian Mixture Model Framework for Bandwidth Extension of Narrowband Speech	1192
<i>Amr H. Nour-Eldin, Peter Kabal</i>	
Speech Enhancement by Reconstruction from Cleaned Acoustic Features	1196
<i>Philip Harding, Ben Milner</i>	
A Soft Decision-Based Speech Enhancement Using Acoustic Noise Classification	1200
<i>Jae-Hun Choi, Sang-Kyun Kim, Joon-Hyuk Chang</i>	
A Noise Estimation Method Based on Speech Presence Probability and Spectral Sparseness	1204
<i>Chao Li, Wenju Liu</i>	
Improved <i>a posteriori</i> Speech Presence Probability Estimation Based on Cepstro-Temporal Smoothing and Time-Frequency Correlation	1208
<i>Chao Li, Wenju Liu</i>	
A Rapid Adaptation Algorithm for Tracking Highly Non-Stationary Noises based on Bayesian Inference for On-Line Spectral Change Point Detection	1212
<i>Md Foezur Rahman Chowdhury, Sid-Ahmed Selouani, Douglas O'Shaughnessy</i>	
Single Channel Speech Enhancement Using MMSE Estimation of Short-Time Modulation Magnitude Spectrum	1216
<i>Kuldip Paliwal, Belinda Schwerin, Kamil Wojcicki</i>	
Speech Enhancement Using Masking Properties in Adverse Environments	1220
<i>Atanu Saha, Tetsuya Shimamura</i>	
Phoneme-Dependent NMF for Speech Enhancement in Monaural Mixtures	1224
<i>Bhiksha Raj, Rita Singh, Tuomas Virtanen</i>	
Kernel PCA for Speech Enhancement	1228
<i>Christina Leitner, Franz Pernkopf, Gernot Kubin</i>	
Objective Intelligibility Prediction of Speech by Combining Correlation and Distortion Based Techniques	1232
<i>Angel M. Gomez, Belinda Schwerin, Kuldip Paliwal</i>	

ASR – FEATURE EXTRACTION I

Chair: Fabio Brugnara

Integrating Recent MLP Feature Extraction Techniques into TRAP Architecture	1236
<i>Frantisek Grezl, Martin Karafiat</i>	
Feature Frame Stacking in RNN-Based Tandem ASR Systems --- Learned vs. Predefined Context	1240
<i>Martin Wollmer, Bjorn Schuller, Gerhard Rigoll</i>	

Improved Acoustic Feature Combination for LVCSR by Neural Networks	1244
<i>Christian Plahl, Ralf Schluter, Hermann Ney</i>	
Hierarchical Tandem Features for ASR in Mandarin	1248
<i>Joel Pinto, Mathew Magimai-Doss, Herve Bourlard</i>	
Analysis and Comparison of Recent MLP Features for LVCSR Systems	1252
<i>Fabio Valente, Mathew Magimai-Doss, Wen Wang</i>	
Deep Learning of Speech Features for Improved Phonetic Recognition	1256
<i>Jaehyung Lee, Soo-Young Lee</i>	
Globality-Locality Consistent Discriminant Analysis for Phone Classification	1260
<i>Heyun Huang, Yang Liu, Jort F. Gemmeke, Louis ten Bosch, Bert Cranen, Lou Boves</i>	
Front-End Compensation Methods for LVCSR Under Lombard Effect	1264
<i>Hynek Boril, Frantisek Grezl, John H.L. Hansen</i>	
Classification of Fricatives Using Feature Extrapolation of Acoustic-Phonetic Features in Telephone Speech	1268
<i>Jung-Won Lee, Jeung-Yoon Choi, Hong-Goo Kang</i>	
Noise Robust Feature Extraction Based on Extended Weighted Linear Prediction in LVCSR	1272
<i>Sami Keronen, Jouni Pohjalainen, Paavo Alku, Mikko Kurimo</i>	
Comparing Different Flavors of Spectro-Temporal Features for ASR	1276
<i>Bernd T. Meyer, Suman V. Ravuri, Marc Rene Schadler, Nelson Morgan</i>	
VTLN in the MFCC Domain: Band-Limited versus Local Interpolation	1280
<i>Ehsan Variani, Thomas Schaaf</i>	
Multistream Bandpass Modulation Features for Robust Speech Recognition	1284
<i>Sridhar Krishna Nemala, Kailash Patil, Mounya Elhilali</i>	
An Analysis of Automatic Speech Recognition with Multiple Microphones	1288
<i>Davide Marino, Thomas Hain</i>	

SPOKEN DIALOGUE & SPOKEN LANGUAGE UNDERSTANDING SYSTEMS

Chair: Steve Renals

Multi-View Approach for Speaker Turn Role Labeling in TV Broadcast News Shows	1292
<i>Geraldine Damnati, Delphine Charlet</i>	
Evaluation of an Integrated Authoring Tool for Building Advanced Question-Answering Characters	1296
<i>Sudeep Gandhe, Michael Rushforth, Priti Aggarwal, David Traum</i>	
Towards Unsupervised Spoken Language Understanding: Exploiting Query Click Logs for Slot Filling	1300
<i>Gokhan Tur, Dilek Hakkani-Tur, Dustin Hillard, Asli Celikyilmaz</i>	
Web-Enhanced Content Retrieval for Information Access Dialogue System	1304
<i>Donghyeon Lee, Cheongjae Lee, Minwoo Jeong, Kyungduk Kim, Seokhwan Kim, Junhwi Choi, Gary Geunbae Lee</i>	
Uncertainty Management for On-Line Optimisation of a POMDP-Based Large-Scale Spoken Dialogue System	1308
<i>Lucie Daubigney, Milica Gasic, Senthilkumar Chandramohan, Matthieu Geist, Olivier Pietquin, Steve Young</i>	
Detection of Task-Incomplete Dialogs Based on Utterance-and-Behavior Tag N-Gram for Spoken Dialog Systems	1312
<i>Sunao Hara, Norihide Kitaoka, Kazuya Takeda</i>	
Shrinkage-Based Features for Natural Language Call Routing	1316
<i>Ruhi Sarikaya, Stanley F. Chen, Bhuvana Ramabhadran</i>	
Clustering with Modified Cosine Distance Learned from Constraints	1320
<i>Leonid Rachevsky, Dimitri Kanevsky, Ruhi Sarikaya, Bhuvana Ramabhadran</i>	
Computer-Assisted Disfluency Counts for Stuttered Speech	1324
<i>Peter A. Heeman, Andy McMillin, J. Scott Yaruss</i>	
Using Speaker ID to Discover Repeat Callers of a Spoken Dialog System	1328
<i>Andrew Fandrianto, Brian Langner, Alan W. Black</i>	
Semantic Graph Clustering for POMDP-Based Spoken Dialog Systems	1332
<i>Florian Pinault, Fabrice Lefevre</i>	
Learning Place-Names from Spoken Utterances and Localization Results by Mobile Robot	1336
<i>Ryo Taguchi, Yuji Yamada, Koosuke Hattori, Taizo Umezaki, Masahiro Hoguro, Naoto Iwahashi, Kotaro Funakoshi, Mikio Nakano</i>	
Active Learning for Dialogue Act Classification	1340
<i>Bjorn Gambäck, Fredrik Olsson, Oscar Tackstrom</i>	
Speaker Role Recognition Using Question Detection and Characterization	1344
<i>Thierry Bazillon, Benjamin Maza, Michael Rouvier, Frederic Bechet, Alexis Nasr</i>	

Learning Score Structure from Spoken Language for a Tennis Game	1348
<i>Qiang Huang, Stephen J. Cox</i>	
Semi-Automated Classifier Adaptation for Natural Language Call Routing	1352
<i>Silke M. Witt</i>	
Interactional Style Detection for Versatile Dialogue Response Using Prosodic and Semantic Features	1356
<i>Wei-Bin Liang, Chung-Hsien Wu, Chih-Hung Wang, Jhing-Fa Wang</i>	
Quality Aspects of Multimodal Dialog Systems: Identity, Stimulation and Success	1360
<i>Christine Kuhnel, Benjamin Weiss, Matthias Schulz, Sebastian Moller</i>	

PROSODIC STRUCTURE

Chair: Elizabeth Schriberg

Where Should Pitch Accents and Phrase Breaks Go? A Syntax Tree Transducer Solution	1364
<i>Joseph Tepperman, Emily Nava</i>	
Phrasal Prominences do not need Pitch Movements: Postfocal Phrasal Heads in Italian	1368
<i>Giuliano Bocci, Cinzia Avesani</i>	

VOLUME 3

Intonation of Left Dislocated Topics in Modern Greek	1372
<i>David Le Gac, Hiyon Yoo</i>	
Phrases, Pitch and Perceived Prominence in Maori	1376
<i>Laura Thompson, Catherine I. Watson, Ray Harlow, Jeanette King, Margaret Maclagan, Helen Charters, Peter Keegan</i>	
Perceptual Sensitivity to Prenuclear and Nuclear Intonational Patterns	1380
<i>Tomas Dubeda</i>	
Tonal Alignment Defined: The Case of Southern Irish English	1384
<i>Raya Kalaldehy</i>	
Using Mutual Information to Identify Regions of Analysis for Prosodic Analysis	1388
<i>Andrew Rosenberg</i>	
Prosodic Highlights in Mandarin Continuous Speech --- Cross-Genre Attributes and Implications	1392
<i>Chiu-yu Tseng, Chao-yu Su, Chi-Feng Huang</i>	
When Two Newly-Acquired Words are One: New Words Differing in Stress Alone are not Automatically Represented Differently	1396
<i>Simone Sulpizio, James M. McQueen</i>	
Automatic Determination of the Standard Chinese Prosodic Phrase Boundaries by F₀ Generation Model	1400
<i>Shehui Bu, Zhenjie Zhuo, Lingling Yang, Shuichi Itahashi</i>	
Measuring Speakers' Similarity in Speech by Means of Prosodic Cues: Methods and Potential	1404
<i>Celine De Looze, Stephane Rauzy</i>	
Tonal Variations in Mandarin: New Evidence from Spontaneous and Read Speech	1408
<i>Li-chiung Yang</i>	

LANGUAGE PROCESSING

Chair: Frederic Bechet

Accounting for Prosodic Information to Improve ASR-Based Topic Tracking for TV Broadcast News	1412
<i>Camille Guinaudeau, Julia Hirschberg</i>	
Morpheme Conversion for Connecting Speech Recognizer and Language Analyzers in Unsegmented Languages	1416
<i>Kenji Imamura, Tomoko Izumi, Kugatsu Sadamitsu, Kuniko Saito, Satoshi Kobashikawa, Hirokazu Masataki</i>	
Emotion Detection Based on Concept Inference and Spoken Sentence Analysis for Customer Service	1420
<i>Ren-Ying Fang, Bo-Wei Chen, Jhing-Fa Wang, Chung-Hsien Wu</i>	
Commas Recovery with Syntactic Features in French and in Czech	1424
<i>Christophe Cerisara, Pavel Kral, Claire Gardent</i>	
Redundancy Reduction in ASR of Spontaneous Speech Through Statistical Machine Translation	1428
<i>Daniele Falavigna</i>	
From Interview to News Text: A Study of Taiwan TV Political Interviews in Newspaper Reports	1432
<i>Chin-Chih Chiang</i>	

ASR – LANGUAGE MODELS I

Chair: Michael Riley

Unary Data Structures for Language Models	1436
<i>Jeffrey Sorensen, Cyril Allauzen</i>	
Bayesian Language Model Interpolation for Mobile Speech Input	1440
<i>Cyril Allauzen, Michael Riley</i>	
On the Estimation of Discount Parameters for Language Model Smoothing	1444
<i>Martin Sundermeyer, Ralf Schluter, Hermann Ney</i>	
N-Grams for Conditional Random Fields or a Failure-Transition (Φ) Posterior for Acyclic FSTs	1448
<i>Patrick Lehnen, Stefan Hahn, Hermann Ney</i>	
Hybrid Language Models Using Mixed Types of Sub-Lexical Units for Open Vocabulary German LVCSR	1452
<i>M. Ali Basha Shaik, Amr El-Desoky Mousa, Ralf Schluter, Hermann Ney</i>	
Morpheme Based Factored Language Models for German LVCSR	1456
<i>Amr El-Desoky Mousa, M. Ali Basha Shaik, Ralf Schluter, Hermann Ney</i>	
Compound Word Recombination for German LVCSR	1460
<i>Markus Nussbaum-Thom, Amr El-Desoky Mousa, Ralf Schluter, Hermann Ney</i>	
Lattice-Based Risk Minimization Training for Unsupervised Language Model Adaptation	1464
<i>Akio Kobayashi, Takahiro Oku, Shinichi Homma, Toru Imai, Seiichi Nakagawa</i>	
Similarity Language Model	1468
<i>Christian Gillot, Christophe Cerisara</i>	
Data Sampling and Dimensionality Reduction Approaches for Reranking ASR Outputs Using Discriminative Language Models	1472
<i>Erinc Dikici, Murat Semerci, Murat Saraclar, Ethem Alpaydin</i>	
Training a Language Model Using Webdata for Large Vocabulary Japanese Spontaneous Speech Recognition	1476
<i>Ryo Masumura, Seongjun Hahm, Akinori Ito</i>	
Large Vocabulary SOUL Neural Network Language Models	1480
<i>Hai-Son Le, Ilya Oparin, Abdel Messaoudi, Alexandre Allauzen, Jean-Luc Gauvain, Francois Yvon</i>	
Improved Spoken Query Transcription Using Co-Occurrence Information	1484
<i>Jonathan Mamou, Abhinav Sethy, Bhuvana Ramabhadran, Ron Hoory, Paul Vozila</i>	
Unsupervised Latent Speaker Language Modeling	1488
<i>Yik-Cheung Tam, Paul Vozila</i>	

SPOKEN LANGUAGE RESOURCES, EVALUATION AND STANDARDIZATION I

Chair: Sebastian Moeller

Measurement of Objective Intelligibility of Japanese Accented English Using ERJ (English Read by Japanese) Database	1492
<i>Nobuaki Minematsu, Koji Okabe, Keisuke Ogaki, Keikichi Hirose</i>	
From Single-Call to Multi-Call Quality: A Study on Long-Term Quality Integration in Audio-Visual Speech Communication	1496
<i>Sebastian Moller, Chihuy Bang, Teele Tamme, Markus Vaalgamaa, Benjamin Weiss</i>	
Optimal Selection of Limited Vocabulary Speech Corpora	1500
<i>Hui Lin, Jeff Bilmes</i>	
Open Source Multi-Language Audio Database for Spoken Language Processing Applications	1504
<i>Stephen A. Zahorian, Jiang Wu, Montri Karnjanadecha, Chandra SekharVootkuri, Brian Wong, Andrew Hwang, Eldar Tokhtamyshev</i>	
The USC CARE Corpus: Child-Psychologist Interactions of Children with Autism Spectrum Disorders	1508
<i>Matthew P. Black, Daniel Bone, Marian E. Williams, Phillip Gorrindo, Pat Levitt, Shrikanth Narayanan</i>	
Towards a Versatile Multi-Layered Description of Speech Corpora Using Algebraic Relations	1512
<i>Nelly Barbot, Vincent Barreaud, Olivier Boeffard, Laure Charonnat, Arnaud Delhay, Sebastien Le Maguer, Damien Lolive</i>	
Announcing the Electromagnetic Articulography (Day 1) Subset of the Family mngu0 Articulatory Corpus	1516
<i>Korin Richmond, Phil Hoole, Simon King</i>	
A Pitch Tracking Corpus with Evaluation on Multipitch Tracking Scenario	1520
<i>Gregor Pirker, Michael Wohlmayr, Stefan Petrik, Franz Pernkopf</i>	

On Building and Evaluating a Broadcast-News Audio Segmentation System	1524
<i>Taras Butko, Climent Nadeu</i>	
Time- and Acoustic-Mediated Alignment Algorithms for Speech Recognition Evaluation	1528
<i>Simon Dobrisesk, France Mihelic</i>	
Effects of Shortening Speech Prompts of In-Car Voice User Interfaces on Users Mental Models	1532
<i>Julia Niemann, Kati Schulz, Ina Wechsung</i>	
Speech Transcript Evaluation for Information Retrieval	1536
<i>Laurens van der Werff, Wessel Kraaij, Franciska de Jong</i>	
The Albayzin 2010 Language Recognition Evaluation	1540
<i>Luis Javier Rodriguez-Fuentes, Mikel Penagarikano, Amparo Varona, Mireia Diez, German Bordel</i>	
Progress and Prospects for Speech Technology: Results from Three Sexennial Surveys	1544
<i>Roger K. Moore</i>	
Painless WFST Cascade Construction for LVCSR - Transducersaurus	1548
<i>Josef R. Novak, Nobuaki Minematsu, Keikichi Hirose</i>	

PARALINGUISTIC INFORMATION – CLASSIFICATION AND DETECTION

Chair: Julia Hirschberg

On the Use of Multimodal Cues for the Prediction of Degrees of Involvement in Spontaneous Conversation	1552
<i>Catharine Oertel, Stefan Scherer, Nick Campbell</i>	
Anger Recognition in Spoken Dialog Using Linguistic and Para-Linguistic Information	1556
<i>Narichika Nomoto, Masafumi Tamoto, Hirokazu Masataki, Osamu Yoshioka, Satoshi Takahashi</i>	
Recognition of Personality Traits from Human Spoken Conversations	1560
<i>A.V. Ivanov, G. Riccardi, A.J. Sporka, J. Franc</i>	
Using Multiple Databases for Training in Emotion Recognition: To Unite or to Vote?	1564
<i>Bjorn Schuller, Zixing Zhang, Felix Weninger, Gerhard Rigoll</i>	
Would You Buy a Car from Me?' - On the Likability of Telephone Voices	1568
<i>Felix Burkhardt, Bjorn Schuller, Benjamin Weiss, Felix Weninger</i>	
Automatic Identification of Salient Acoustic Instances in Couples' Behavioral Interactions Using Diverse Density Support Vector Machines	1572
<i>James Gibson, Athanasios Katsamanis, Matthew P. Black, Shrikanth Narayanan</i>	
Predicting Speaker Changes and Listener Responses with and without Eye-Contact	1576
<i>Daniel Neiberg, Joakim Gustafson</i>	
Emotion Classification Using Inter- and Intra-Subband Energy Variation	1580
<i>Senaka Amarakeerthi, Tin Lay Nwe, Liyanage C. De Silva, Michael Cohen</i>	
Emotion Classification of Infants' Cries Using Duration Ratios of Acoustic Segments	1584
<i>K. Kitahara, S. Michiwiki, M. Sato, S. Matsunaga, M. Yamashita, K. Shinohara</i>	
Vowels Formants Analysis Allows Straightforward Detection of High Arousal Acted and Spontaneous Emotions	1588
<i>Bogdan Vlasenko, Dmytro Prylipko, David Philippou-Hubner, Andreas Wendemuth</i>	
Intra-, Inter-, and Cross-Cultural Classification of Vocal Affect	1592
<i>Daniel Neiberg, Petri Laukka, Hillary Anger Elfenbein</i>	

APPLICATIONS FOR LEARNING, EDUCATION, AGED AND HANDICAPPED PERSONS

Chair: Roberto Gretter

Verifying Human Users in Speech-Based Interactions	1596
<i>Sajad Shiral-Shahreza, Yashar Ganjali, Ravin Balakrishnan</i>	
Automatic Assessment of Prosody in High-Stakes English Tests	1600
<i>Jian Cheng</i>	
Improvement of Segmental Mispronunciation Detection with Prior Knowledge Extracted from Large L2 Speech Corpus	1604
<i>Dean Luo, Xuesong Yang, Lan Wang</i>	
Off-Topic Detection in Automated Speech Assessment Applications	1608
<i>Jian Cheng, Jianqiang Shen</i>	
Towards Context-Dependent Phonetic Spelling Error Correction in Children's Freely Composed Text for Diagnostic and Pedagogical Purposes	1612
<i>Sebastian Stuker, Johanna Fay, Kay Berkling</i>	
Factored Translation Models for Improving a Speech into Sign Language Translation System	1616
<i>V. Lopez-Ludena, R. San-Segundo, R. Cordoba, J. Ferreiros, J.M. Montero, J.M. Pardo</i>	

Formant Maps in Hungarian Vowels --- Online Data Inventory for Research, and Education	1620
<i>Kalman Abari, Zsuzsanna Zsofia Racz, Gabor Olaszy</i>	
Automatic Subtitling of the Basque Parliament Plenary Sessions Videos	1624
<i>German Bordel, Silvia Nieto, Mikel Penagarikano, Luis Javier Rodriguez-Fuentes, Amparo Varona</i>	
Generating Animated Pronunciation from Speech Through Articulatory Feature Extraction.....	1628
<i>Yurie Iribe, Silasak Manosavanh, Kouichi Katsurada, Ryoko Hayashi, Chunyue Zhu, Tsuneo Nitta</i>	
A Tale of Two Tasks: Detecting Children's Off-Task Speech in a Reading Tutor.....	1632
<i>Wei Chen, Jack Mostow</i>	
Problems Encountered by Japanese EL2 with English Short Vowels as Illustrated on a 3D Vowel Chart.....	1636
<i>Toshiko Isei-Jaakkola, Takatoshi Naka, Keikichi Hirose</i>	
Automatic Generation of Listening Comprehension Learning Material in European Portuguese	1640
<i>Thomas Pellegrini, Rui Correia, Isabel Trancoso, Jorge Baptista, Nuno Mamede</i>	
Candidate Generation for ASR Output Error Correction Using a Context-Dependent Syllable Cluster-Based Confusion Matrix.....	1644
<i>Chao-Hong Liu, Chung-Hsien Wu, David Sarwono, Jhing-Fa Wang</i>	
Semi-Supervised Tree Support Vector Machine for Online Cough Recognition	1648
<i>Thai Hoa Huynh, Vu An Tran, Huy Dat Tran</i>	

ROBUST SPEECH RECOGNITION I

Chair: Pietro Laface

A Versatile Gaussian Splitting Approach to Non-Linear State Estimation and its Application to Noise-Robust ASR.....	1652
<i>Volker Leutnant, Alexander Krueger, Reinhold Haeb-Umbach</i>	
Generalized-Log Spectral Mean Normalization for Speech Recognition.....	1656
<i>Hilman F. Pardede, Koichi Shinoda</i>	
Zero-Crossing-Based Channel Attentive Weighting of Cepstral Features for Robust Speech Recognition: The ETRI 2011 CHiME Challenge System.....	1660
<i>Young-Ik Kim, Hoon-Young Cho, Sang-Hun Kim</i>	
Feature Compensation for Speech Recognition in Severely Adverse Environments Due to Background Noise and Channel Distortion.....	1664
<i>Wooil Kim, John H.L. Hansen</i>	
Binaural Cues for Fragment-Based Speech Recognition in Reverberant Multisource Environments.....	1668
<i>Ning Ma, Jon Barker, Heidi Christensen, Phil D. Green</i>	
Sub-Band Level Histogram Equalization for Robust Speech Recognition	1672
<i>Vikas Joshi, Raghavendra Bilgi, S. Umesh, L. Garcia, C. Benitez</i>	
GMM-Based Missing-Feature Reconstruction on Multi-Frame Windows.....	1676
<i>Ulpu Remes, Yoshihiko Nankaku, Keiichi Tokuda</i>	
Improvements of a Dual-Input DBN for Noise Robust ASR.....	1680
<i>Yang Sun, Jort F. Gemmeke, Bert Cranen, Louis ten Bosch, Lou Boves</i>	
Denosing Using Optimized Wavelet Filtering for Automatic Speech Recognition.....	1684
<i>Randy Gomez, Tatsuya Kawahara</i>	
Noise Robust Speaker-Independent Speech Recognition with Invariant-Integration Features Using Power-Bias Subtraction.....	1688
<i>Florian Muller, Alfred Mertins</i>	

ASR – ACOUSTIC MODELS I

Chair: Lori Lamel

Semi-Automatic Acoustic Model Generation from Large Unsynchronized Audio and Text Chunks.....	1692
<i>Michele Alessandrini, Giorgio Biagetti, Alessandro Curzi, Claudio Turchetti</i>	
Unsupervised Testing Strategies for ASR.....	1696
<i>Brian Strope, Doug Beeferman, Alexander Gruenstein, Xin Lei</i>	
Acoustic Model Training with Detecting Transcription Errors in the Training Data.....	1700
<i>Gakuto Kurata, Nobuyasu Itoh, Masafumi Nishimura</i>	
Towards Unsupervised Training of Speaker Independent Acoustic Models.....	1704
<i>Aren Jansen, Kenneth Church</i>	
Acoustic Modeling with Bootstrap and Restructuring Based on Full Covariance.....	1708
<i>Xiaodong Cui, Xin Chen, Jian Xue, Peder A. Olsen, John R. Hershey, Bowen Zhou</i>	
An i-vector Based Approach to Acoustic Sniffing for Irrelevant Variability Normalization Based Acoustic Model Training and Speech Recognition	1712
<i>Jian Xu, Yu Zhang, Zhi-Jie Yan, Qiang Huo</i>	

Log-Linear Optimization of Second-Order Polynomial Features with Subsequent Dimension Reduction for Speech Recognition	1716
<i>Muhammad Ali Tahir, Ralf Schluter, Hermann Ney</i>	
Genre Categorization and Modeling for Broadcast Speech Transcription	1720
<i>Qingqing Zhang, Lori Lamel, Jean-Luc Gauvain</i>	
Individual Error Minimization Learning Framework and its Applications to Speech Recognition and Utterance Verification	1724
<i>Sunghwan Shin, Ho-Young Jung, Biing-Hwang Juang</i>	
Effective Triphone Mapping for Acoustic Modeling in Speech Recognition	1728
<i>Sakhia Darjaa, Milos Cernak, Marian Trnka, Milan Rusko, Robert Sabo</i>	
Analysis of Dialectal Influence in Pan-Arabic ASR	1732
<i>Udhyakumar Nallasamy, Michael Garbus, Florian Metze, Qin Jin, Thomas Schaaf, Tanja Schultz</i>	
Connected Digit Recognition by Means of Reservoir Computing	1736
<i>Azarakhsh Jalalvand, Fabian Triefenbach, David Verstraeten, Jean-Pierre Martens</i>	
Large Margin - Minimum Classification Error Using Sum of Shifted Sigmoids as the Loss Function	1740
<i>Madhavi V. Ratnagiri, Biing-Hwang Juang, Lawrence Rabiner</i>	
Representing Phonological Features Through a Two-Level Finite State Model	1744
<i>Javier M. Olaso, M. Ines Torres, Raquel Justo</i>	
Optimization of the Gaussian Mixture Model Evaluation on GPU	1748
<i>Jan Vanek, Jan Trmal, Josef V. Psutka, Josef Psutka</i>	

SOURCE SEPARATION AND SPEECH ENHANCEMENT

Chair: Marco Matassoni

Monaural Voiced Speech Segregation Based on Pitch and Comb Filter	1752
<i>Xueliang Zhang, Wenju Liu</i>	
Fast and Simple Iterative Algorithm of Lp-Norm Minimization for Under-Determined Speech Separation	1756
<i>Yasuharu Hirasawa, Naoki Yasuraoka, Toru Takahashi, Tetsuya Ogata, Hiroshi G. Okuno</i>	
Monaural Speech Separation Based on a 2D Processing and Harmonic Analysis	1760
<i>Azam Rabiee, Saeed Setayeshi, Soo-Young Lee</i>	
Underdetermined Blind Source Separation with Fuzzy Clustering for Arbitrarily Arranged Sensors	1764
<i>Ingrid Jafari, Serajul Haque, Roberto Togneri, Sven Nordholm</i>	
On Initial Seed Selection for Frequency Domain Blind Speech Separation	1768
<i>Dang Hai Tran Vu, Reinhold Haeb-Umbach</i>	
Spatial Filter Calibration Based on Minimization of Modified LSD	1772
<i>Nobuaki Tanaka, Tetsuji Ogawa, Tetsunori Kobayashi</i>	
Probabilistic Spectrum Envelope: Categorized Audio-Features Representation for NMF-Based Sound Decomposition	1776
<i>Toru Nakashika, Tetsuya Takiguchi, Yasuo Ariki</i>	
A High Resolution Multiple Source Localization Based on Generalized Cumulant Structure (GCS) Matrix	1780
<i>Jinho Choi, Chang D. Yoo</i>	
Single Channel Speech Music Separation Using Nonnegative Matrix Factorization with Sliding Windows and Spectral Masks	1784
<i>Emad M. Grais, Hakan Erdogan</i>	
Perceptually-Inspired Processing for Multichannel Wiener Filter	1788
<i>Jorge I. Marin-Hurtado, David V. Anderson</i>	
Speech Recognition in Mixed Sound of Speech and Music Based on Vector Quantization and Non-Negative Matrix Factorization	1792
<i>Shoichi Nakano, Kazumasa Yamamoto, Seiichi Nakagawa</i>	
Reduction of Highly Nonstationary Ambient Noise by Integrating Spectral and Locational Characteristics of Speech and Noise for Robust ASR	1796
<i>Tomohiro Nakatani, Shoko Araki, Marc Delcroix, Takuya Yoshioka, Masakiyo Fujimoto</i>	
Voice Processing by Dynamic Glottal Models with Applications to Speech Enhancement	1800
<i>Carlo Drioli, Andrea Calanca</i>	
Supervised Sparse Coding Strategy in Cochlear Implants	1804
<i>Jinqiu Sang, Guoping Li, Hongmei Hu, Mark E. Lutman, Stefan Bleeck</i>	

HMM-BASED SPEECH SYNTHESIS II

Chair: Enrico Zovato

Continuous Control of the Degree of Articulation in HMM-Based Speech Synthesis	1808
<i>Benjamin Picart, Thomas Drugman, Thierry Dutoit</i>	
Estimation of Window Coefficients for Dynamic Feature Extraction for HMM-Based Speech Synthesis	1812
<i>Ling-Hui Chen, Yoshihiko Nankaku, Heiga Zen, Keiichi Tokuda, Zhen-Hua Ling, Li-Rong Dai</i>	
Inverse Filtering Based Harmonic Plus Noise Excitation Model for HMM-Based Speech Synthesis	1816
<i>Zhengqi Wen, Jianhua Tao</i>	
Improved HNM-Based Vocoder for Statistical Synthesizers	1820
<i>Daniel Erro, Inaki Sainz, Eva Navas, Inma Hernaez</i>	
A Statistical Phrase/Accent Model for Intonation Modeling	1824
<i>Gopala Krishna Anumanchipalli, Luis C. Oliveira, Alan W. Black</i>	
Intermediate-State HMMs to Capture Continuously-Changing Signal Features	1828
<i>Gustav Eje Henter, W. Bastiaan Kleijn</i>	
Automatic Sentence Selection from Speech Corpora Including Diverse Speech for Improved HMM-TTS Synthesis Quality	1832
<i>Norbert Braunschweiler, Sabine Buchholz</i>	
Phonological Knowledge Guided HMM State Mapping for Cross-Lingual Speaker Adaptation	1836
<i>Hui Liang, John Dines</i>	
Reformulating Prosodic Break Model into Segmental HMMs and Information Fusion	1840
<i>Nicolas Obin, Pierre Lanchantin, Anne Lacheret, Xavier Rodet</i>	
Multipulse Sequences for Residual Signal Modeling	1844
<i>Ranniery Maia, Heiga Zen, Kate Knill, M.J.F. Gales, Sabine Buchholz</i>	
Can Objective Measures Predict the Intelligibility of Modified HMM-Based Synthetic Speech in Noise?	1848
<i>Cassia Valentini-Botinhao, Junichi Yamagishi, Simon King</i>	
Speech Synthesis Based on Articulatory-Movement HMMs with Voice-Source Codebooks	1852
<i>Tsuneo Nitta, Takayuki Onoda, Masashi Kimura, Yurie Iribe, Kouichi Katsurada</i>	
Large-Scale Subjective Evaluations of Speech Rate Control Methods for HMM-Based Speech Synthesizers	1856
<i>Tsuneo Kato, Makoto Yamada, Nobuyuki Nishizawa, Keiichiro Oura, Keiichi Tokuda</i>	
HMM-Based Emphatic Speech Synthesis Using Unsupervised Context Labeling	1860
<i>Yu Maeno, Takashi Nose, Takao Kobayashi, Yusuke Ijima, Hideharu Nakajima, Hideyuki Mizuno, Osamu Yoshioka</i>	

PHONETICS AND PHONOLOGY, STRESS, ACCENT, RHYTHM

Chair: Bernd Mobius

Chinese and Italian Speech Rhythm: Normalization and the CCI Algorithm	1864
<i>Chiara Bertini, Pier Marco Bertinetto, Na Zhi</i>	
Rhythm Metrics on Syllables and Feet do not Work as Expected	1868
<i>Paolo Mairano, Antonio Romano</i>	
Applying Rhythm Features to Automatically Assess Non-Native Speech	1872
<i>Lei Chen, Klaus Zechner</i>	
Prosodic Synchrony in Co-Operative Task-Based Dialogues: A Measure of Agreement and Disagreement	1876
<i>Brian Vaughan</i>	
Low and High, Short and Long by Crook or by Hook?	1880
<i>Oliver Niebuhr, Astrid Wolf</i>	
Estimating Speaking Rate by Means of Rhythmicity Parameters	1884
<i>Christian Heinrich, Florian Schiel</i>	
Comparing Word and Syllable Prominence Rated by Naive Listeners	1888
<i>Denis Arnold, Bernd Mobius, Petra Wagner</i>	
L1/L2 Perception of Lexical Stress with F0 Peak-Delay: Effect of an Extra Syllable Added	1892
<i>Shinichi Tokuma, Yi Xu</i>	
Letter-to-Phoneme Conversion Based on Two-Stage Neural Network Focusing on Letter and Phoneme Contexts	1896
<i>Kheang Seng, Yurie Iribe, Tsuneo Nitta</i>	
An International English Speech Corpus for Longitudinal Study of Accent Development	1900
<i>Rosemary Orr, Hugo Quene, Roeland van Beek, Thari Diefenbach, David A. van Leeuwen, Marijn Huijbregts</i>	

A Corpus-Based Study of English Pronunciation Variations	1904
<i>Sunhee Kim, Kyuwghan Lee, Minhwa Chung</i>	
Long Term Average Speech Spectra in Yolngu Matha and Pitjantjatjara Speaking Females and Males	1908
<i>Hywel Stoakes, Andrew Butcher, Janet Fletcher, Marija Tabain</i>	
Context and Speaker Dependency in the Relation of Vowel Formants and Subglottal Resonances - Evidence from Hungarian	1912
<i>Tekla Etelka Graczi, Steven M. Lulich, Tamas Gabor Csapo, Andras Beke</i>	

ASR – SEARCH, KEYWORD SPOTTING AND CONFIDENCE MEASURES I

Chair: Mark Gales

Event Selection from Phone Posteriorgrams Using Matched Filters	1916
<i>Keith Kintzley, Aren Jansen, Hynek Hermansky</i>	
A Piecewise Aggregate Approximation Lower-Bound Estimate for Posteriorgram-Based Dynamic Time Warping	1920
<i>Yaodong Zhang, James Glass</i>	
OOV Detection and Recovery Using Hybrid Models with Different Fragments	1924
<i>Long Qin, Ming Sun, Alexander Rudnicky</i>	
AUC Optimization Based Confidence Measure for Keyword Spotting	1928
<i>Haiyang Li, Jiqing Han, Tieran Zheng</i>	
An Empirical Study of Multilingual Spoken Term Detection	1932
<i>Zejun Ma, Xiaorui Wang, Bo Xu</i>	
Fusing Multiple Confidence Measures for Chinese Spoken Term Detection	1936
<i>Zejun Ma, Xiaorui Wang, Bo Xu</i>	
Response Probability Based Decoding Algorithm for Large Vocabulary Continuous Speech Recognition	1940
<i>Zhanlei Yang, Hao Chao, Wenju Liu</i>	
Combining Lattice-Based Language Dependent and Independent Approaches for Out-of-Language Detection in LVCSR	1944
<i>Yuxiang Shan, Yan Deng, Jia Liu</i>	
Evaluation of Tree-Trellis Based Decoding in Over-Million LVCSR	1948
<i>Naoaki Ito, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda</i>	
Lattice Based Discriminative Model Combination Using Automatically Induced Phonetic Contexts	1952
<i>Hao Huang, Bing Hu Li</i>	
Predicting Human Perceived Accuracy of ASR Systems	1956
<i>Taniya Mishra, Andrej Ljolje, Mazin Gilbert</i>	
Cross-Lingual Study of ASR Errors: On the Role of the Context in Human Perception of Near-Homophones	1960
<i>I. Vasilescu, D. Yahia, N. Snoeren, Martine Adda-Decker, Lori Lamel</i>	
Performance Prediction of Speech Recognition Using Average-Voice-Based Speech Synthesis	1964
<i>Tatsuhiko Saito, Takashi Nose, Takao Kobayashi, Yohei Okato, Akio Horii</i>	
Confidence Measures for Turkish Call Center Conversations	1968
<i>Ali Haznedaroglu, Levent M. Arslan</i>	
Spoken Document Confidence Estimation Using Contextual Coherence	1972
<i>Taichi Asami, Narichika Nomoto, Satoshi Kobashikawa, Yoshikazu Yamaguchi, Hirokazu Masataki, Satoshi Takahashi</i>	

PITCH PROCESSING – SINGING VOICE ANALYSIS

Chair: Thomas Drugman

Fundamental Frequency Estimation Using Modified Higher Order Moments and Multiple Windows	1976
<i>Alipah Pawi, Saeed Vaseghi, Ben Milner, Seyed Ghorshi</i>	
EM-Based Gain Adaptation for Probabilistic Multipitch Tracking	1980
<i>Michael Wohlmayr, Franz Pernkopf</i>	
Joint Robust Voicing Detection and Pitch Estimation Based on Residual Harmonics	1984
<i>Thomas Drugman, Abeer Alwan</i>	
Epoch Extraction in High Pass Filtered Speech Using Hilbert Envelope	1988
<i>D. Govind, S.R.M. Prasanna, Debadatta Pati</i>	
Robust HNR-Based Closed-Loop Pitch and Harmonic Parameters Estimation	1992
<i>Alexander Pavlovets, Alexander Petrovsky</i>	
Exploring Bessel Features for Detection of Glottal Closure Instants	1996
<i>Chetana Prakash, Dhananjaya N., Suryakanth V. Gangashetty</i>	

Evaluation of Glottal Epoch Detection Algorithms on Different Voice Types	2000
<i>Joao P. Cabral, John Kane, Christer Gobl, Julie Carson-Berndsen</i>	
A Divide et impera Algorithm for Optimal Pitch Stylization	2004
<i>Antonio Origlia, Giovanni Abete, Francesco Cutugno, Iolanda Alfano, Renata Savy, Bogdan Ludusan</i>	
Singing Voice Analysis Using Relative Harmonic Delays	2008
<i>Ricardo Sousa, Anibal Ferreira</i>	
Singing Voice Synthesis: Singer-Dependent Vibrato Modeling and Coherent Processing of Spectral Envelope	2012
<i>S.W. Lee, Minghui Dong</i>	
Chorus Digitalis: Experiments in Chironomic Choir Singing	2016
<i>Sylvain Le Beux, Lionel Feugere, Christophe d'Alessandro</i>	

PROSODIC MODELING

Chair: Hiroya Fujisaki

Prominence Model for Prosodic Features in Automatic Lexical Stress and Pitch Accent Detection	2020
<i>Kun Li, Shuang Zhang, Mingxing Li, Wai-Kit Lo, Helen Meng</i>	
Hierarchical Stress Modeling in Mandarin Text-to-Speech	2024
<i>Ya Li, Jianhua Tao, Xiaoying Xu</i>	
Automatic Prosodic Events Detection by Using Syllable-Based Acoustic, Lexical and Syntactic Features	2028
<i>Chong-Jia Ni, Wenju Liu, Bo Xu</i>	
Using Dynamic Time Warping to Compute Prosodic Similarity Measures	2032
<i>Albert Rilliard, Alexandre Allauzen, Philippe Boula de Mareuil</i>	
Applying the Quantitative Target Approximation Model (qTA) to German and Brazilian Portuguese	2036
<i>Plinio A. Barbosa, Hansjorg Mixdorff, Sandra Madureira</i>	
Stylization and Trajectory Modelling of Short and Long Term Speech Prosody Variations	2040
<i>Nicolas Obin, Anne Lacheret, Xavier Rodet</i>	
Toward a Continuous Modeling of French Prosodic Structure: Using Acoustic Features to Predict Prominence Location and Prominence Degree	2044
<i>Mathieu Avanzi, Nicolas Obin, Anne Lacheret-Dujour, Bernard Victorri</i>	

VOLUME 4

Optimal Models of Prosodic Prominence Using the Bayesian Information Criterion	2048
<i>Tim Mahrt, Jui-Ting Huang, Yoonsook Mo, Margaret Fleck, Mark Hasegawa-Johnson, Jennifer Cole</i>	
Quantitative Analysis of Tone Coarticulation in Mandarin	2052
<i>Hussein Hussein, Hansjorg Mixdorff, Hue San Do, Rudiger Hoffmann</i>	
Tracking Pitch Contours Using Minimum Jerk Trajectories	2056
<i>Daniel Neiberg, G. Ananthkrishnan, Joakim Gustafson</i>	

DISCOURSE AND DIALOGUE

Chair: Patrick Ehlen

On the Use of Linguistic Features in an Automatic System for Speech Analytics of Telephone Conversations	2060
<i>Benjamin Maza, Marc El-Beze, Georges Linares, Renato De Mori</i>	
Determining what Questions to Ask, with the Help of Spectral Graph Theory	2064
<i>Abe Kazemzadeh, Sungbok Lee, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	
Are You Sure You're Paying Attention? - 'Uh-Huh' Communicating Understanding as a Marker of Attentiveness	2068
<i>Hendrik Buschmeier, Zofia Malisz, Marcin Wlodarczak, Stefan Kopp, Petra Wagner</i>	
Projectability of Transition-Relevance Places Using Prosodic Features in Japanese Spontaneous Conversation	2072
<i>Yuichi Ishimoto, Mika Enomoto, Hitoshi Iida</i>	
Measuring Final Lengthening for Speaker-Change Prediction	2076
<i>Anna Hjalmarsson, Kornel Laskowski</i>	
Incremental Learning and Forgetting in Stochastic Turn-Taking Models	2080
<i>Kornel Laskowski, Jens Edlund, Mattias Heldner</i>	
Reinforcement Learning of Argumentation Dialogue Policies in Negotiation	2084
<i>Kallirroi Georgila, David Traum</i>	

Topic Switching Strategies for Spoken Dialogue Systems	2088
<i>Tobias Heinroth, Savina Koleva, Wolfgang Minker</i>	
Unsupervised Clustering of Utterances Using Non-Parametric Bayesian Methods	2092
<i>Ryuichiro Higashinaka, Noriaki Kawamae, Kugatsu Sadamitsu, Yasuhiro Minami, Toyomi Meguro, Kohji Dohsaka, Hirohito Inagaki</i>	

SLP FOR SPEECH TRANSLATION, INFORMATION EXTRACTION

AND RETRIEVAL

Chair: Dekai Wu

OOV Sensitive Named-Entity Recognition in Speech	2096
<i>Carolina Parada, Mark Dredze, Frederick Jelinek</i>	
Speech Translation with Grammar Driven Probabilistic Phrasal Bilexica Extraction	2100
<i>Markus Saers, Dekai Wu, Chi-kiu Lo, Karteek Addanki</i>	
An Efficient Unified Extraction Algorithm for Bilingual Data	2104
<i>Christoph Tillmann, Sanjika Hewavitharana</i>	
Using Features from Topic Models to Alleviate Over-Generation in Hierarchical Phrase-Based Translation	2108
<i>Songfang Huang, Bowen Zhou</i>	
An Empirical Study on Improving Hierarchical Phrase-Based Translation Using Alignment Features	2112
<i>Songfang Huang, Bowen Zhou</i>	
Robust Speech Translation by Domain Adaptation	2116
<i>Xiaodong He, Li Deng</i>	
Enhancements to the Training Process of Classifier-Based Speech Translator via Topic Modeling	2120
<i>Emil Ettelaie, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	
A Scalable Approach to Building a Parallel Corpus from the Web	2124
<i>V. K. Rangarajan Sridhar, Luciano Barbosa, Srinivas Bangalore</i>	
Spoken Term Detection Results Using Plural Subword Models by Estimating Detection Performance for Each Query	2128
<i>Yoshiaki Itoh, Kohei Iwata, Masaaki Ishigame, Kazuyo Tanaka, Shi-wook Lee</i>	
SpeechForms: From Web to Speech and Back	2132
<i>Luciano Barbosa, Diamantino Caseiro, Giuseppe Di Fabrizio, Amanda Stent</i>	
Image Processing Filters for Line Detection-Based Spoken Term Detection	2136
<i>Kazuyuki Noritake, Hiroaki Nanjo, Takehiko Yoshimi</i>	
Using Latent Topic Features for Named Entity Extraction in Search Queries	2140
<i>Joe Polifroni, Francois Mairesse</i>	
Language Model Expansion Using Webdata for Spoken Document Retrieval	2144
<i>Ryo Masumura, Seongjun Hahm, Akinori Ito</i>	
Effects of Query Expansion for Spoken Document Passage Retrieval	2148
<i>Tomoyosi Akiba, Koichiro Honda</i>	
Unsupervised Hidden Markov Modeling of Spoken Queries for Spoken Term Detection without Speech Recognition	2152
<i>Chun-an Chan, Lin-shan Lee</i>	
Topic Identification from Audio Recordings Using Rich Recognition Results and Neural Network Based Classifiers	2156
<i>Roberto Gemello, Franco Mana, Pier Domenico Batsu</i>	

SPEECH SYNTHESIS – SELECTED TOPICS

Chair: Tomoki Toda

A Grammar Based Approach to Style Specific Phrase Prediction	2160
<i>Alok Parlikar, Alan W. Black</i>	
Unsupervised Features from Text for Speech Synthesis in a Speech-to-Speech Translation System	2164
<i>Oliver Watts, Bowen Zhou</i>	
Unsupervised Continuous-Valued Word Features for Phrase-Break Prediction without a Part-of-Speech Tagger	2168
<i>Oliver Watts, Junichi Yamagishi, Simon King</i>	
Albayzin 2010: A Spanish Text to Speech Evaluation	2172
<i>Francisco Campillo, Francisco Mendez, Montserrat Arza, Laura Docio, Antonio Bonafonte, Eva Navas, Inaki Sainz</i>	
Combining Active and Semi-Supervised Learning for Homograph Disambiguation in Mandarin Text-to-Speech Synthesis	2176
<i>Binbin Shen, Zhiyong Wu, Yongxin Wang, Lianhong Cai</i>	

Automatically Creating a Diphone Set from a Speech Database	2180
<i>Thomas Ewender, Beat Pfister</i>	
Automatic Viseme Clustering for Audiovisual Speech Synthesis	2184
<i>Wesley Mattheyses, Lukas Latacz, Werner Verhelst</i>	
Perceptual Quality Dimensions of Text-to-Speech Systems	2188
<i>Florian Hinterleitner, Sebastian Moller, Christoph Norrenbrock, Ulrich Heute</i>	
A Pointwise Approach to Pronunciation Estimation for a TTS Front-End	2192
<i>Shinsuke Mori, Graham Neubig</i>	
Correlating Text with Prosody	2196
<i>Mohamed Abou-Zleikha, Julie Carson-Berndsen</i>	
What is... Dengue Fever? - Modeling and Predicting Pronunciation Errors in a Text-to-Speech System	2200
<i>Andrew Rosenberg, Raul Fernandez, Bhuvana Ramabhadran</i>	
Aperiodicity Analysis for Quality Estimation of Text-to-Speech Signals	2204
<i>Christoph Norrenbrock, Ulrich Heute, Florian Hinterleitner, Sebastian Moller</i>	

HUMAN SPEECH AND SOUND PERCEPTION I

Chair: Denis Burnham

Parallels in Infants' Attention to Speech Articulation and to Physical Changes in Speech-Unrelated Objects	2208
<i>Eeva Klintfors, Ellen Marklund, Francisco Lacerda</i>	
Speech Events are Recoverable from Unlabeled Articulatory Data: Using an Unsupervised Clustering Approach on Data Obtained from Electromagnetic Midsagittal Articulography (EMA)	2212
<i>Daniel Duran, Jagoda Bruni, Grzegorz Dogil, Hinrich Schutze</i>	
Children's Recognition of their own Voice: Influence of Phonological Impairment	2216
<i>Sofia Strombergsson</i>	
Evaluation of Bone-Conducted Ultrasonic Hearing-Aid Regarding Transmission of Speaker Discrimination Information	2220
<i>Takayuki Kagomiya, Seiji Nakagawa</i>	
Impact of Different Feedback Mechanisms in EMG-Based Speech Recognition	2224
<i>Christian Herff, Matthias Janke, Michael Wand, Tanja Schultz</i>	
Phonotactic Constraints and the Segmentation of Cantonese Speech	2228
<i>Michael C.W. Yip</i>	
Reaction Time and Decision Difficulty in the Perception of Intonation	2232
<i>Katrin Schneider, Grzegorz Dogil, Bernd Mobius</i>	
Processing of Stress Related Acoustic Cues as Indexed by ERPs	2236
<i>Ferenc Honbolygo, Valeria Csepe</i>	
On the Relationship Between Perceived Accentness, Acoustic Similarity, and Processing Difficulty in Foreign-Accented Speech	2240
<i>Marijt J. Witteman, Andrea Weber, James M. McQueen</i>	
The Perception Boundary Between Single and Geminate Stops in 3- and 4-Mora Japanese Words	2244
<i>Shigeaki Amano, Yukari Hirata</i>	
Correlation Analysis of Acoustic Features with Perceptual Voice Quality Similarity for Similar Speaker Selection	2248
<i>Yusuke Ijima, Mitsuaki Isogai, Hideyuki Mizuno</i>	

MULTILINGUAL AND MULTIMODAL APPROACHES TO SPOKEN LANGUAGE

Chair: Michael Johnston

Can Audio-Visual Speech Recognition Outperform Acoustically Enhanced Speech Recognition in Automotive Environment?	2252
<i>Rajitha Navarathna, Tristan Kleinschmidt, David Dean, Sridha Sridharan, Patrick Lucey</i>	
A Multimodal Approach to Dictation of Handwritten Historical Documents	2256
<i>Vicent Alabau, Veronica Romero, Antonio-L. Lagarda, Carlos-D. Martinez-Hinarejos</i>	
Weight Optimization for Bimodal Unit-Selection Talking Head Synthesis	2260
<i>Asterios Toutios, Utpala Musti, Slim Ouni, Vincent Colotte</i>	
Modality Selection and Perceived Mental Effort in a Mobile Application	2264
<i>Stefan Schaffer, Benjamin Jockel, Ina Wechsung, Robert Schleicher, Sebastian Moller</i>	
A Cross-Lingual Spoken Content Search System	2268
<i>Jitendra Ajmera, Ashish Verma</i>	
NeMo: A Platform for Multilingual News Monitoring	2272
<i>C. Girardi, Roberto Gretter, Daniele Falavigna, Fabio Brugnara, Diego Giuliani, M. Federico</i>	

Unsupervised Learning of Acoustic Unit Descriptors for Audio Content Representation and Classification	2276
<i>Sourish Chaudhuri, Mark Harvilla, Bhiksha Raj</i>	
Conditioned Hidden Markov Model Fusion for Multimodal Classification	2280
<i>Michael Glodek, Stefan Scherer, Friedhelm Schwenker</i>	
Distant Speech Recognition in a Smart Home: Comparison of Several Multisource ASRs in Realistic Conditions	2284
<i>Benjamin Lecouteux, Michel Vacher, Francois Portet</i>	
A Robust Approach to Mining Repeated Sequence in Audio Stream	2288
<i>Jiansong Chen, Lei Zhu, Bailan Feng, Peng Ding, Bo Xu</i>	

ASR – NEW PARADIGMS AND OTHER TOPICS

Chair: Chin Hui Lee

Accelerated Parallelizable Neural Network Learning Algorithm for Speech Recognition	2292
<i>Dong Yu, Li Deng</i>	
Deep Convex Net: A Scalable Architecture for Speech Pattern Classification	2296
<i>Li Deng, Dong Yu</i>	
Modeling Broad Context for Tone Recognition with Conditional Random Fields	2300
<i>Siwei Wang, Gina-Anne Levow</i>	
Improved Tonal Language Speech Recognition by Integrating Spectro-Temporal Evidence and Pitch Information with Properly Chosen Tonal Acoustic Units	2304
<i>Shang-wen Li, Yow-bang Wang, Liang-che Sun, Lin-shan Lee</i>	
Kullback-Leibler Divergence-Based ASR Training Data Selection	2308
<i>Evandro Gouvea, Marelle H. Davel</i>	
Articulatory Feature Classification Using Nearest Neighbors	2312
<i>Arild Brandrud Naess, Karen Livescu, Rohit Prabhavalkar</i>	
Continuous Episodic Memory Based Speech Recognition Using Articulatory Dynamics	2316
<i>Sebastien Demange, Slim Ouni</i>	
Grapheme Model Interpolation and Arabic Pronunciation Generation	2320
<i>T. Li, P.C. Woodland, F. Diehl, M.J.F. Gales</i>	
Grapheme-to-Phoneme Conversion Using Conditional Random Fields	2324
<i>Irina Illina, Dominique Fohr, Denis Jouvet</i>	
Bilingual Acoustic Model Adaptation by Unit Merging on Different Levels and Cross-Level Integration	2328
<i>Ching-Feng Yeh, Chao-Yu Huang, Lin-shan Lee</i>	
A Qualitative Evaluation of Phoneme-to-Phoneme Technology	2332
<i>Marijn Schraagen, Gerrit Bloothoof</i>	
Cheap Bootstrap of Multi-Lingual Hidden Markov Models	2336
<i>Daniele Falavigna, Roberto Gretter</i>	
Adaptive Stream Fusion in Multistream Recognition of Speech	2340
<i>Nima Mesgarani, Samuel Thomas, Hynek Hermansky</i>	
Unsupervised Audio Patterns Discovery Using HMM-Based Self-Organized Units	2344
<i>Man-hung Siu, Herbert Gish, Steve Lowe, Arthur Chan</i>	
Nearest Neighbors with Learned Distances for Phonetic Frame Classification	2348
<i>John Labiak, Karen Livescu</i>	

SPEAKER RECOGNITION – MODELING, AUTOMATIC PROCEDURES, ANALYSIS III

Chair: Daniel Garcia-Romero

i-vector Based Speaker Recognition on Short Utterances	2352
<i>Ahilan Kanagasundaram, Robbie Vogt, David Dean, Sridha Sridharan, Michael Mason</i>	
Study of Overlapped Speech Detection for NIST SRE Summed Channel Speaker Recognition	2356
<i>Hanwu Sun, Bin Ma</i>	
Super-Dirichlet Mixture Models Using Differential Line Spectral Frequencies for Text-Independent Speaker Identification	2360
<i>Zhanyu Ma, Arne Leijon</i>	
Comparison of Voice Activity Detectors for Interview Speech in NIST Speaker Recognition Evaluation	2364
<i>Hon-Bill Yu, Man-Wai Mak</i>	

Eigen-Voice Based Anchor Modeling System for Speaker Identification Using MLLR Super-Vector	2368
<i>A.K. Sarkar, S. Umesh</i>	
Automatic Detection of Speaker Attributes Based on Utterance Text	2372
<i>Wen Wang, Andreas Kathol, Harry Bratt</i>	
Comparison of Speaker Recognition Approaches for Real Applications	2376
<i>Sandro Cumani, Pier Domenico Batzu, Daniele Colibro, Claudio Vair, Pietro Laface, Vasileios Vasilakakis</i>	
Modeling Speaker Personality Using Voice	2380
<i>Tim Polzehl, Sebastian Moller, Florian Metze</i>	
Structural Joint Factor Analysis for Speaker Recognition	2384
<i>Marc Ferras, Koichi Shinoda, Sadaoki Furui</i>	
Acoustic Forest for SMAP-Based Speaker Verification	2388
<i>Sangeeta Biswas, Marc Ferras, Koichi Shinoda, Sadaoki Furui</i>	
Mixture of Auto-Associative Neural Networks for Speaker Verification	2392
<i>G.S.V.S. Sivaram, Samuel Thomas, Hynek Hermansky</i>	

SPEECH AUDIO ANALYSIS AND CLASSIFICATION

Chair: Olivier Rosec

Stop Consonant Recognition by Temporal Fine Structure of Burst	2396
<i>Seppo Fagerlund, Unto K. Laine</i>	
Phonetic Classification Using Controlled Random Walks	2400
<i>Katrin Kirchhoff, Andrei Alexandrescu</i>	
Keyphrase Cloud Generation of Broadcast News	2404
<i>Luis Marujo, Marcio Viveiros, Joao P. Neto</i>	
Optimized Feature Extraction and HMMs in Subword Detectors	2408
<i>Alfonso M. Canterla, Magne H. Johnsen</i>	
Real-World Speech/Non-Speech Audio Classification Based on Sparse Representation Features and GPCs	2412
<i>Ziqiang Shi, Jiqing Han, Tieran Zheng</i>	
Privacy Preserving Speaker Verification Using Adapted GMMs	2416
<i>Manas A. Pathak, Bhiksha Raj</i>	
Clustering Expressive Speech Styles in Audiobooks Using Glottal Source Parameters	2420
<i>Eva Szekeby, Joao P. Cabral, Peter Cahill, Julie Carson-Berndsen</i>	
On the Use of the Rhythmogram for Automatic Syllabic Prominence Detection	2424
<i>Bogdan Ludusan, Antonio Origlia, Francesco Cutugno</i>	
Speech Modulation Features for Robust Nonnative Speech Accent Detection	2428
<i>Sethserey Sam, Xiong Xiao, Laurent Besacier, Eric Castelli, Haizhou Li, Eng Siong Chng</i>	
Frame-Level Vocal Effort Likelihood Space Modeling for Improved Whisper-Island Detection	2432
<i>Chi Zhang, John H.L. Hansen</i>	
Speaker Identification for Whispered Speech Using a Training Feature Transformation from Neutral to Whisper	2436
<i>Xing Fan, John H.L. Hansen</i>	
An Accurate and Robust Gender Identification Algorithm	2440
<i>Andrea DeMarco, Stephen J. Cox</i>	
Deep Belief Networks for Automatic Music Genre Classification	2444
<i>Xiaohong Yang, Qingcai Chen, Shusen Zhou, Xiaolong Wang</i>	
Image Representation of the Subband Power Distribution for Robust Sound Classification	2448
<i>Jonathan Dennis, Huy Dat Tran, Haizhou Li</i>	
Acoustic and Visual Cues of Turn-Taking Dynamics in Dyadic Interactions	2452
<i>Bo Xiao, Viktor Rozgic, Athanasios Katsamanis, Brian R. Baucom, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	

HUMAN SPEECH AND SOUND PERCEPTION II

Chair: Holger Mitterer

Pointing Gestures do not Influence the Perception of Lexical Stress	2456
<i>Alexandra Jesse, Holger Mitterer</i>	
Relationships Between Phonetic Features and Speech Perception - A Statistical Investigation from a Large Anechoic British English Corpus	2460
<i>Ian R. Cushing, Francis F. Li, Ken Worrall, Tim Jackson</i>	
The Representation of Speech in a Nonlinear Auditory Model: Time-Domain Analysis of Simulated Auditory-Nerve Firing Patterns	2464
<i>Guy J. Brown, Tim Jurgens, Ray Meddis, Matthew Robertson, Nicholas R. Clark</i>	

An Automatic Voice Pleasantness Classification System Based on Prosodic and Acoustic Patterns of Voice Preference	2468
<i>Luis Coelho, Daniela Braga, Miguel Sales-Dias, Carmen Garcia-Mateo</i>	
Contributions of F1 and F2 (F2') to the Perception of Plosive Consonants	2472
<i>Rene Carre, Pierre Divenyi, Willy Serniclaes, Emmanuel Ferragne, Egidio Marsico, Viet-Son Nguyen</i>	
Auditory Speech Processing is Affected by Visual Speech in the Periphery	2476
<i>Jeesun Kim, Chris Davis</i>	
Visual Speech Speeds Up Auditory Identification Responses	2480
<i>Tim Paris, Jeesun Kim, Chris Davis</i>	
Agglomerative Hierarchical Clustering of Emotions in Speech Based on Subjective Relative Similarity	2484
<i>Ryoichi Takashima, Tohru Nagano, Ryuki Tachibana, Masafumi Nishimura</i>	
Optimal Syllabic Rates and Processing Units in Perceiving Mandarin Spoken Sentences	2488
<i>Guangting Mai, Gang Peng</i>	
Cross-Lingual Speaker Discrimination Using Natural and Synthetic Speech	2492
<i>Mirjam Wester, Hui Liang</i>	

SPEECH AUDIO ANALYSIS

Chair: Toshiaki Fukada

Robust Audio Fingerprinting Based on Local Spectral Luminance Maxima Scheme	2496
<i>Yong-zhe Shi, Wei-Qiang Zhang, Jia Liu</i>	
Entropy-Rate Driven Inference of Stochastic Grammars	2500
<i>Unto K. Laine</i>	
An Efficient Pre-Processing Scheme to Improve the Sound Source Localization System in Noisy Environment	2504
<i>Sheng-Chieh Lee, K. Bharanitharan, Bo-Wei Chen, Jhing-Fa Wang, Chung-Hsien Wu, Min-Jian Liao</i>	
A Study on Auditory Feature Spaces for Speech-Driven Lip Animation	2508
<i>Guylaine Le-Jan, Yannick Benezeth, Guillaume Gravier, Frederic Bimbot</i>	
Phase-Only Speech Reconstruction Using Very Short Frames	2512
<i>Erfan Loweimi, Seyed Mohammad Ahadi, Hamid Sheikhzadeh</i>	
Frequency-Warped and Stabilized Time-Varying Cepstral Coefficients	2516
<i>Trond Skogstad, Torbjorn Svendsen</i>	
Using Human Perception for Automatic Accent Assessment	2520
<i>Freddy William, Abhijeet Sangwan, John H.L. Hansen</i>	
A Study of the Effectiveness of Articulatory Strokes for Phonemic Recognition	2524
<i>Carlos Molina, Sungbok Lee, Shrikanth Narayanan, Nestor Becerra Yoma</i>	
Auditory Filterbank Improves Voice Morphing	2528
<i>Erika Okamoto, Toshio Irino, Ryuichi Nisimura, Hideki Kawahara</i>	
Monaural Sound Localization	2532
<i>Anna Katharina Fuchs, Christian Feldbauer, Michael Stark</i>	

SPEECH CODING

Chair: Alan McCree

Dual-Mode AVQ Coding Based on Spectral Masking and Sparseness Detection for ITU-T G.711.1/G.722 Super-Wideband Extensions	2536
<i>Masahiro Fukui, Shigeaki Sasaki, Yusuke Hiwasaki, Kurihara Sachiko, Yoichi Haneda</i>	
Phone Impact Based Speech Transmission Technique for Reliable Speech Recognition in Poor Wireless Network Conditions	2540
<i>Azar Taufique, Kumaran Vijayasankar, Wool Kim, John H.L. Hansen, Marco Tacca, Andrea Fumagalli</i>	
Automatic Speech Codec Identification with Applications to Tampering Detection of Speech Recordings	2544
<i>Jingting Zhou, Daniel Garcia-Romero, Carol Y. Espy-Wilson</i>	
A Hybrid Quasi-Harmonic/CELP Wideband Speech Coding Scheme for Unit Selection TTS Synthesis	2548
<i>Chang-Heon Lee, Olivier Rosec, Yannis Stylianou</i>	
Voice Quality Characterization of IETF Opus Codec	2552
<i>Anssi Ramo, Henri Toukoma</i>	
Leja Ordering LSFs for Accurate Estimation of Predictor Coefficients	2556
<i>C.F. Pedersen</i>	
Improved Quality for Conversational VoIP Using Path Diversity	2560
<i>Qipeng Gong, Peter Kabal</i>	
Tree Encoding for the ITU-T G.711.1 Speech Coder	2564
<i>Abdul Hannan Khan, Peter Kabal</i>	

Parallel and Hierarchical Decision Making for Sparse Coding in Speech Recognition	2568
<i>Dong Wang, Ravichander Vipperla, Nicholas Evans</i>	
A New Model-Based Mandarin-Speech Coding System	2572
<i>Chen-Yu Chiang, Jyh-Her Yang, Ming-Chieh Liu, Yih-Ru Wang, Yuan-Fu Liao, Sin-Horn Chen</i>	

ROBUSTNESS AND ADAPTATION FOR ASR

Chair: Vivek Kumar

Using Unsupervised Feature-Based Speaker Adaptation for Improved Transcription of Spoken Archives	2576
<i>Petr Cerva, Karel Palecek, Jan Silovsky, Jan Nouza</i>	
Online Speaker Adaptation with Pre-Computed FMLLR Transformations	2580
<i>Volker Fischer, Siegfried Kunzmann</i>	
Instantaneous Speaker Adaptation Through Selection and Combination of fMLLR Transformation Matrices	2584
<i>Diego Giuliani, Fabio Brugnara</i>	
Joint Bilinear Transformation Space Based Maximum a posteriori Linear Regression Adaptation Using Prior with Variance Function	2588
<i>Hwa Jeon Song, Yunkeun Lee, Hyung Soon Kim</i>	
A Study on Combining VTLN and SAT to Improve the Performance of Automatic Speech Recognition	2592
<i>D.R. Sanand, Mikko Kurimo</i>	
Incorporating Regional Information to Enhance MAP-Based Stochastic Feature Compensation for Robust Speech Recognition	2596
<i>Yu Tsao, Paul R. Dixon, Chiori Hori, Hisashi Kawai</i>	
A Study on the Effect of Pitch on LPCC and PLPC Features for Children's ASR in Comparison to MFCC	2600
<i>Shweta Ghai, Rohit Sinha</i>	
About Handling Boundary Uncertainty in a Speaking Rate Dependent Modeling Approach	2604
<i>Denis Jouvet, Dominique Fohr, Irina Illina</i>	
An Active Learning Approach to Task Adaptation	2608
<i>Ji Wu, Zhiyang He, Ping Lv</i>	
Efficient Speaker and Noise Normalization for Robust Speech Recognition	2612
<i>Vikas Joshi, Raghavendra Bilgi, S. Umesh, C. Benitez, L. Garcia</i>	
How Realistic is Artificially Added Noise?	2616
<i>Thomas Winkler</i>	

VOICE ACTIVITY DETECTION

Chair: Abeer Alwan

Voice Activity Detection in MTF-Based Power Envelope Restoration	2620
<i>Masashi Unoki, Xugang Lu, Rico Petrick, Shota Morita, Masato Akagi, Rudiger Hoffmann</i>	
Using Spectral Fluctuation of Speech in Multi-Feature HMM-Based Voice Activity Detection	2624
<i>Miquel Espi, Shigeki Miyabe, Takuya Nishimoto, Nobutaka Ono, Shigeki Sagayama</i>	
Linear Dynamic Models for Voice Activity Detection	2628
<i>Kannu Mehta, Chau Khoa Pham, Eng Siong Chng</i>	
Detection of Shouted Speech in the Presence of Ambient Noise	2632
<i>Jouni Pohjalainen, Tuomo Raitio, Paavo Alku</i>	
Breath-Detection-Based Telephony Speech Phrasing	2636
<i>Takashi Fukuda, Osamu Ichikawa, Masafumi Nishimura</i>	
Multi-Channel Voice Activity Detection Based on Conic Constraints	2640
<i>Gibak Kim</i>	
Multi-Sensor Voice Activity Detection Based on Multiple Observation Hypothesis Testing	2644
<i>Theodoros Petsatodis, Fotios Talantzis, Christos Boukis, Zheng-Hua Tan, Ramjee Prasad</i>	
Online Speech Activity Detection in Broadcast News	2648
<i>Chao Gao, Guruprasad Saikumar, Saurabh Khanwalkar, Avi Herscovici, Anoop Kumar, Amit Srivastava, Premkumar Natarajan</i>	
A Real-Time Speech Command Detector for a Smart Control Room	2652
<i>Daniel Reich, Felix Putze, Dominic Heger, Joris Ijsselmuiden, Rainer Stiefelhagen, Tanja Schultz</i>	
Robust Voice Activity Detector for Real World Applications Using Harmonicity and Modulation Frequency	2656
<i>Ekapol Chuangsuwanich, James Glass</i>	

On Noise Robust Voice Activity Detection	2660
<i>Tomas Dekens, Werner Verhelst</i>	
Adaptive Regularization Framework for Robust Voice Activity Detection	2664
<i>Xugang Lu, Masashi Unoki, Ryosuke Isotani, Hisashi Kawai, Satoshi Nakamura</i>	

HUMAN SPEECH PRODUCTION I

Chair: Francesco Cutugno

On the Use of Extended Context for HMM-Based Spontaneous Conversational Speech Synthesis	2668
<i>Tomoki Koriyama, Takashi Nose, Takao Kobayashi</i>	
Predicting Tongue Positions from Acoustics and Facial Features	2672
<i>Asterios Toutios, Slim Ouni</i>	
Assessing Acoustic Reduction: Exploiting Local Structure in Speech	2676
<i>Louis ten Bosch, Annika Hamalainen, Mirjam Ernestus</i>	
The "Fortis-Lenis" Distinction in Bulgarian and German	2680
<i>Bistra Andreeva, Magdalena Wolska</i>	
Acoustic Correlates of Glottal Gaps	2684
<i>Gang Chen, Jody Kreiman, Yen-Liang Shue, Abeer Alwan</i>	
Using a Genetic Algorithm to Estimate Parameters of a Coarticulation Model	2688
<i>Brian O. Bush, John-Paul Hosom, Alexander Kain, Akiko Amano-Kusumoto</i>	
Synthesis of Breathly, Normal, and Pressed Phonation Using a Two-Mass Model with a Triangular Glottis	2692
<i>Peter Birkholz, Bernd J. Kroger, Christiane Neuschaefer-Rube</i>	
Analysis of Inter-Articulator Correlation in Acoustic-to-Articulatory Inversion Using Generalized Smoothness Criterion	2696
<i>Prasanta Kumar Ghosh, Shrikanth Narayanan</i>	
Frequency-Domain Representation of Source-Filter Coupling and its Effect in the Production of Voice	2700
<i>Tokihiko Kaburagi</i>	
Method for Speech Inversion with Large Scale Statistical Evaluation	2704
<i>Heikki Rasilo, Unto K. Laine, Okko Rasanen, Toomas Altsaar</i>	
Italian in the No-Man's Land Between Stress-Timing and Syllable-Timing? Speakers are More Stress-Timed than Listeners	2708
<i>Bettina Braun, Sabine Geiselmann</i>	
The Lombard Effect in Spontaneous Dialog Speech	2712
<i>Laura Folk, Florian Schiel</i>	

SPEAKER RECOGNITION – ANALYSIS AND STATISTICS III

Chair: Pierre-Michel Bousquet

Variational Bayesian Model Selection for GMM-Speaker Verification Using Universal Background Model	2716
<i>Timur Pekhovsky, Alexandra Lokhanova</i>	

VOLUME 5

To Weight or Not to Weight: Source-Normalised LDA for Speaker Recognition Using i-vectors	2720
<i>Mitchell McLaren, David A. van Leeuwen</i>	
Maximum Entropy Based Data Selection for Speaker Recognition	2724
<i>Chien-Lin Huang, Bin Ma</i>	
Addressing the Data-Imbalance Problem in Kernel-Based Speaker Verification via Utterance Partitioning and Speaker Comparison	2728
<i>Wei Rao, Man-Wai Mak</i>	
Single-Channel Head Orientation Estimation Based on Discrimination of Acoustic Transfer Function	2732
<i>Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Arika</i>	
Maximum Likelihood i-vector Space Using PCA for Speaker Verification	2736
<i>Zhenchun Lei, Yingchun Yang</i>	
Speaker Verification Using Sparse Representations on Total Variability i-vectors	2740
<i>Ming Li, Xiang Zhang, Yonghong Yan, Shrikanth Narayanan</i>	

Robust Speaker Recognition in Non-Stationary Room Environments Based on Empirical Mode Decomposition	2744
<i>Taufiq Hasan, John H.L. Hansen</i>	
Range Based Multi Microphone Array Fusion for Speaker Activity Detection in Small Meetings	2748
<i>Jani Even, Panikos Heracleous, Carlos T. Ishi, Norihiro Hagita</i>	
Speaker Verification Robust to Talking Style Variation Using Multiple Kernel Learning Based on Conditional Entropy Minimization	2752
<i>Tetsuji Ogawa, Hideitsu Hino, Noboru Murata, Tetsunori Kobayashi</i>	
Regularized Logistic Regression Fusion for Speaker Verification	2756
<i>Ville Hautamaki, Kong Aik Lee, Tomi Kinnunen, Bin Ma, Haizhou Li</i>	
A Longest Matching Segment Approach with Bayesian Adaptation - Application to Noise-Robust Speaker Recognition	2760
<i>Ayeh Jafari, Ramji Srinivasan, Danny Crookes, Ji Ming</i>	
Data Selection with Kurtosis and Nasality Features for Speaker Recognition	2764
<i>Howard Lei, Nikki Mirghafori</i>	
Use of the Harmonic Phase in Speaker Recognition	2768
<i>Inma Hernaez, Ibon Saratxaga, Jon Sanchez, Eva Navas, Iker Luengo</i>	

VOICE CONVERSION AND SPEECH SYNTHESIS

Chair: Alan Black

Gaussian Process Experts for Voice Conversion	2772
<i>Nicholas C.V. Pilkington, Heiga Zen, M.J.F. Gales</i>	
Intonation Conversion from Neutral to Expressive Speech	2776
<i>Christophe Veaux, Xavier Rodet</i>	
Speaker-Adaptive Speech Synthesis Based on Eigenvoice Conversion and Language-Dependent Prosodic Conversion in Speech-to-Speech Translation	2780
<i>Nobuhiko Hattori, Tomoki Toda, Hisashi Kawai, Hiroshi Saruwatari, Kiyohiro Shikano</i>	
Adding Glottal Source Information to Intra-Lingual Voice Conversion	2784
<i>Javier Perez, Antonio Bonafonte</i>	
Formant-Controlled HMM-Based Speech Synthesis	2788
<i>Ming Lei, Junichi Yamagishi, Korin Richmond, Zhen-Hua Ling, Simon King, Li-Rong Dai</i>	
Analysis of HMM-Based Lombard Speech Synthesis	2792
<i>Tuomo Raitio, Antti Suni, Martti Vainio, Paavo Alku</i>	
Discrete/Continuous Modelling of Speaking Style in HMM-Based Speech Synthesis: Design and Evaluation	2796
<i>Nicolas Obin, Pierre Lanchantin, Anne Lacheret, Xavier Rodet</i>	
Factored MLLR Adaptation for Singing Voice Generation	2800
<i>June Sig Sung, Doo Hwa Hong, Shin Jae Kang, Nam Soo Kim</i>	
Adaptation of Prosody in Speech Synthesis by Changing Command Values of the Generation Process Model of Fundamental Frequency	2804
<i>Keikichi Hirose, Keiko Ochi, Ryusuke Mihara, Hiroya Hashimoto, Daisuke Saito, Nobuaki Minematsu</i>	
Prosody Conversion for Emotional Mandarin Speech Synthesis Using the Tone Nucleus Model	2808
<i>Miaomiao Wen, Miaomiao Wang, Keikichi Hirose, Nobuaki Minematsu</i>	
Rapid Adaptation of Foreign-Accented HMM-Based Speech Synthesis	2812
<i>Reima Karhila, Mirjam Wester</i>	
The Effects of Phoneme Errors in Speaker Adaptation for HMM Speech Synthesis	2816
<i>Balint Toth, Tibor Fegyö, Geza Nemeth</i>	

HUMAN SPEECH PRODUCTION II

Chair: Francis Grenez

Articulatory Reduction in Mandarin Chinese Words	2820
<i>Jeffrey Berry, Sunjing Ji, Ian Fasel, Diana Archangeli</i>	
Morphological Variation in the Adult Vocal Tract: A Modeling Study of its Potential Acoustic Impact	2824
<i>Adam Lammert, Michael Proctor, Athanasios Katsamanis, Shrikanth Narayanan</i>	
Analysis and Automatic Estimation of Children's Subglottal Resonances	2828
<i>Steven M. Lulich, Harish Arsitkere, John R. Morton, Gary K.F. Leung, Abeer Alwan, Mitchell S. Sommers</i>	
Acceleration Sensor Based Estimates of Subglottal Resonances: Short vs. Long Vowels	2832
<i>Wolfgang Wokurek, Andreas Madsack</i>	
Comparison of Nasalance Measurements from Accelerometers and Microphones and Preliminary Development of Novel Features	2836
<i>Nicolas Audibert, Angélique Amélot</i>	

The Effect of Seeing the Interlocutor on Speech Production in Different Noise Types	2840
<i>Michael Fitzpatrick, Jeeseun Kim, Chris Davis</i>	
Conversing in the Presence of a Competing Conversation: Effects on Speech Production	2844
<i>Vincent Aubanel, Martin Cooke, Julian Villegas, Maria Luisa Garcia Lecumberri</i>	
Very Short Utterances and Timing in Turn-Taking	2848
<i>Mattias Heldner, Jens Edlund, Anna Hjalmarsson, Kornel Laskowski</i>	
Validating rt-MRI Based Articulatory Representations via Articulatory Recognition	2852
<i>Athanasios Katsamanis, Erik Bresch, Vikram Ramanarayanan, Shrikanth Narayanan</i>	
An Electropalatographic and Acoustic Study on Anticipatory Coarticulation in V1#C2V2 Sequences in Standard Chinese	2856
<i>Yinghao Li, Jiangping Kong</i>	
Final /t/ Reduction in Dutch Past-Participles: The Role of Word Predictability and Morphological Decomposability	2860
<i>Iris Hanique, Mirjam Ernestus</i>	
Parametrising Degree of Articulator Movement from Dynamic MRI Data	2864
<i>Zeynab Raeesy, Ladan Baghai-Ravary, John Coleman</i>	

SYSTEMS FOR LVCSR AND RICH TRANSCRIPTION

Chair: Diego Giuliani

Improving LVCSR System Combination Using Neural Network Language Model Cross Adaptation	2868
<i>X. Liu, M.J.F. Gales, P.C. Woodland</i>	
Towards High Performance LVCSR in Speech-to-Speech Translation System on Smart Phones	2872
<i>Jian Xue, Xiaodong Cui, Gregg Daggett, Etienne Marcheret, Bowen Zhou</i>	
Deploying Google Search by Voice in Cantonese	2876
<i>Yun-Hsuan Sung, Martin Jansche, Pedro J. Moreno</i>	
An Investigation in Speech Recognition for Colloquial Arabic	2880
<i>Sarah Al-Shareef, Thomas Hain</i>	
A Multithreaded Implementation of Viterbi Decoding on Recursive Transition Networks	2884
<i>Fabio Brugnara</i>	
Recurrent Neural Network Based Language Modeling in Meeting Recognition	2888
<i>Stefan Kombrink, Tomas Mikolov, Martin Karafiat, Lukas Burget</i>	
Ad-Hoc Meeting Transcription on Clusters of Mobile Devices	2892
<i>Michele Cossalter, Priya Sundararajan, Ian Lane</i>	
ROVER Enhancement with Automatic Error Detection	2896
<i>Kacem Abida, Fakhri Karray</i>	
Automatic Comma Insertion of Lecture Transcripts Based on Multiple Annotations	2900
<i>Yuya Akita, Tatsuya Kawahara</i>	

LANGUAGE, DIALECT IDENTIFICATION AND SPEAKER DIARIZATION

Chair: Nancy Chen

Study on the Relevance Factor of Maximum a Posteriori with GMM for Language Recognition	2904
<i>Chang Huai You, Haizhou Li, Kong Aik Lee</i>	
Improving Multiband Position-Pitch Algorithm for Localization and Tracking of Multiple Concurrent Speakers by Using a Frequency Selective Criterion	2908
<i>Tania Habib, Harald Romsdorfer</i>	
On the Use of Lattices of Time-Synchronous Cross-Decoder Phone Co-Occurrences in a SVM-Phonotactic Language Recognition System	2912
<i>Amparo Varona, Mikel Penagarikano, Luis Javier Rodriguez-Fuentes, German Bordel</i>	
Speaker Clustering Based on Utterance-Oriented Dirichlet Process Mixture Model	2916
<i>Naohiro Tawara, Shinji Watanabe, Tetsuji Ogawa, Tetsunori Kobayashi</i>	
PLDA-Based Clustering for Speaker Diarization of Broadcast Streams	2920
<i>Jan Silovsky, Jan Prazak, Petr Cerva, Jindrich Zdansky, Jan Nouza</i>	
iVector Approach to Phonotactic Language Recognition	2924
<i>Mehdi Souffar, Marcel Kockmann, Lukas Burget, Oldrich Plchot, Ondrej Glembek, Torbjorn Svendsen</i>	
Discriminative Features for Language Identification	2928
<i>Chris Alberti, Michiel Bacchiani</i>	
Perceptual Sensitivity to Dialectal and Generational Variations in Vowels	2932
<i>Robert Allen Fox, Ewa Jacewicz</i>	
Investigation of Cross-Show Speaker Diarization	2936
<i>Qian Yang, Qin Jin, Tanja Schultz</i>	

Language Identification for Text Chats	2940
<i>Vesa Siivola, Bryan Pellom, Meagan Sills</i>	
Spoken Language Recognition in the Latent Topic Simplex	2944
<i>Kong Aik Lee, Chang Huai You, Ville Hautamaki, Anthony Larcher, Haizhou Li</i>	

PARALINGUISTIC INFORMATION – ANALYSIS AND TOOLS

Chair: Shri Narayanan

Investigating Robustness of Spectral Moments on Normal- and High-Effort Speech	2948
<i>Frederike Gottsmann, Corinna Harwardt</i>	
Comparing the Impact of Raised Vocal Effort on Various Spectral Parameters	2952
<i>Corinna Harwardt</i>	
Vowel Context and Speaker Interactions Influencing Glottal Open Quotient and Formant Frequency Shifts in Physical Task Stress	2956
<i>Keith W. Godin, John H.L. Hansen</i>	
Prosodic Correlates of Individual Physiological Response to Stress	2960
<i>Serguei Pakhomov, Michael Kotlyar</i>	
The Vocal Effort of Dominance in Scenario Meetings	2964
<i>Marcela Charfuelan, Marc Schroder</i>	
A Preliminary Model of Emotional Prosody Using Multidimensional Scaling	2968
<i>Sona Patel, Rahul Shrivastav</i>	
An Exploratory Study of the Relations Between Perceived Emotion Strength and Articulatory Kinematics	2972
<i>Jangwon Kim, Sungbok Lee, Shrikanth Narayanan</i>	
Improved Acoustic Characterization of Breathy and Whispery Voices	2976
<i>Carlos T. Ishi, Hiroshi Ishiguro, Norihiro Hagita</i>	
Neutral to Target Emotion Conversion Using Source and Suprasegmental Information	2980
<i>D. Govind, S.R.M. Prasanna, B. Yegnanarayana</i>	
A Multimodal Analysis of Vocal and Visual Backchannels in Spontaneous Dialogs	2984
<i>Khiet P. Truong, Ronald Poppe, Iwan de Kok, Dirk Heylen</i>	
Kernel Models for Affective Lexicon Creation	2988
<i>Nikos Malandrakis, Alexandros Potamianos, Elias Iosif, Shrikanth Narayanan</i>	

SPECIAL SESSIONS

SPEECH AND LANGUAGE PROCESSING-BASED ASSISTIVE TECHNOLOGIES AND HEALTH APPLICATIONS I

Chairs: Tobias Bocklet, Gokhan Tur

Automatic Detection of Depression in Speech Using Gaussian Mixture Modeling with Factor Analysis	2992
<i>Douglas Sturim, Pedro A. Torres-Carrasquillo, Thomas F. Quatieri, Nicolas Malyska, Alan McCree</i>	
Utterance Verification for Automating the Hearing in Noise Test (HINT)	2996
<i>H. Timothy Bunnell, Jason Lilley, Sigfrid D. Soli, Ivan Pal</i>	
Analyzing the Nature of ECA Interactions in Children with Autism	3000
<i>Emily Mower, Chi-Chun Lee, James Gibson, Theodora Chaspari, Marian E. Williams, Shrikanth Narayanan</i>	

SPEECH AND LANGUAGE PROCESSING-BASED ASSISTIVE TECHNOLOGIES AND HEALTH APPLICATIONS II

Chairs: Shri Narayanan, Elmar Noeth

Incorporating Speech Recognition Engine into an Intelligent Assistive Reading System for Dyslexic Students	3004
<i>Theologos Athanaselis, Stelios Bakamidis, Ioannis Dologlou, Evmorfia N. Argyriou, Antonis Symvonis</i>	
An Investigation of Depressed Speech Detection: Features and Normalization	3008
<i>Nicholas Cummins, Julien Epps, Michael Breakspear, Roland Goecke</i>	
Using Prosodic and Spectral Features in Detecting Depression in Elderly Males	3012
<i>Michelle Hewlett Sanchez, Dimitra Vergyi, Luciana Ferrer, Colleen Richey, Pablo Garcia, Bruce Knoth, William Jarrold</i>	
Combining Phonological and Acoustic ASR-Free Features for Pathological Speech Intelligibility Assessment	3016
<i>Catherine Middag, Tobias Bocklet, Jean-Pierre Martens, Elmar Noth</i>	

Speech Synthesis Parameter Generation for the Assistive Silent Speech Interface MVOCA	3020
<i>Robin Hofe, Stephen R. Ell, Michael J. Fagan, James M. Gilbert, Phil D. Green, Roger K. Moore, Sergey I. Rybchenko</i>	
Spectral Features for Automatic Blind Intelligibility Estimation of Spastic Dysarthric Speech	3024
<i>Richard Hummel, Wai-Yip Chan, Tiago H. Falk</i>	
Extraction of Narrative Recall Patterns for Neuropsychological Assessment	3028
<i>Emily T. Prud'hommeaux, Brian Roark</i>	
Gesture Design of Hand-to-Speech Converter Derived from Speech-to-Hand Converter Based on Probabilistic Integration Model	3032
<i>Aki Kunikoshi, Yu Qiao, Daisuke Saito, Nobuaki Minematsu, Keikichi Hirose</i>	
Powered Wheelchair Control Using Acoustic-Based Recognition of Head Gesture Accompanying Speech	3036
<i>Akira Sasou</i>	
Analyzing Training Dependencies and Posterior Fusion in Discriminant Classification of Apnea Patients Based on Sustained and Connected Speech	3040
<i>Jose Luis Blanco, Ruben Fernandez, Doroteo Torre, F. Javier Caminero, Eduardo Lopez</i>	

CROWDSOURCING FOR SPEECH PROCESSING I

Chairs: Maxine Eskenazi, David Suendermann, Gina-Anne Levow

Speaking to the Crowd: Looking at Past Achievements in Using Crowdsourcing for Speech and Predicting Future Challenges	3044
<i>Gabriel Parent, Maxine Eskenazi</i>	

CROWDSOURCING FOR SPEECH PROCESSING II

Chairs: Maxine Eskenazi, David Suendermann, Gina-Anne Levow

A Transcription Task for Crowdsourcing with Automatic Quality Control	3048
<i>Chia-ying Lee, James Glass</i>	
Reliability-Weighted Acoustic Model Adaptation Using Crowd Sourced Transcriptions	3052
<i>Kartik Audhkhasi, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	
Crowdsourcing for Word Recognition in Noise	3056
<i>Martin Cooke, Jon Barker, Maria Luisa Garcia Lecumberri, Krzysztof Wasilewski</i>	
Crowdsourcing Preference Tests, and How to Detect Cheating	3060
<i>Sabine Buchholz, Javier Latorre</i>	
Growing a Spoken Language Interface on Amazon Mechanical Turk	3064
<i>Ian McGraw, James Glass, Stephanie Seneff</i>	
Real User Evaluation of Spoken Dialogue Systems Using Amazon Mechanical Turk	3068
<i>F. Jurcicek, S. Keizer, Milica Gasic, Francois Mairesse, B. Thomson, K. Yu, Steve Young</i>	
Quality Assessment of Crowdsourcing Transcriptions for African Languages	3072
<i>Hadrien Gelas, Solomon Teferra Abate, Laurent Besacier, Francois Pellegrino</i>	
Using Crowdsourcing to Provide Prosodic Annotations for Non-Native Speech	3076
<i>Keelan Evanini, Klaus Zechner</i>	
PodCastle: Recent Advances of a Spoken Document Retrieval Service Improved by Anonymous User Contributions	3080
<i>Masataka Goto, Jun Ogata</i>	

SPOKEN LANGUAGE PROCESSING OF HUMAN-HUMAN CONVERSATIONS I

Chair: Dilek Hakkani-Tur

Language-Independent Socio-Emotional Role Recognition in the AMI Meetings Corpus	3084
<i>Fabio Valente, Alessandro Vinciarelli</i>	
Measuring Acoustic-Prosodic Entrainment with Respect to Multiple Levels and Dimensions	3088
<i>Rivka Levitan, Julia Hirschberg</i>	
Automatic Call Quality Monitoring Using Cost-Sensitive Classification	3092
<i>Youngja Park</i>	

SPOKEN LANGUAGE PROCESSING OF HUMAN-HUMAN CONVERSATIONS II

Chair: Dilek Hakkani-Tur

Learning Influences from Word Use in Polylogue	3096
<i>Tomoharu Iwata, Shinji Watanabe</i>	

Identifying Agreement/Disagreement in Conversational Speech: A Cross-Lingual Study	3100
<i>Wen Wang, Kristin Precoda, Colleen Richey, Geoffrey Raymond</i>	
A Dual Channel Coupled Decoder for Fillers and Feedback	3104
<i>Daniel Neiberg, Joakim Gustafson</i>	
An Analysis of PCA-Based Vocal Entrainment Measures in Married Couples' Affective Spoken Interactions	3108
<i>Chi-Chun Lee, Athanasios Katsamanis, Matthew P. Black, Brian R. Baucom, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	

SPEECH AND AUDIO PROCESSING FOR HUMAN-ROBOT INTERACTION I

Chair: Laurence Devillers

Using Prominence Detection to Generate Acoustic Feedback in Tutoring Scenarios	3112
<i>Lars Schillingmann, Petra Wagner, Christian Munier, Britta Wrede, Katharina Rohlfing</i>	
Bayesian Extension of MUSIC for Sound Source Localization and Tracking	3116
<i>Takuma Otsuka, Kazuhiro Nakadai, Tetsuya Ogata, Hiroshi G. Okuno</i>	
Speech-Based Non-Prototypical Affect Recognition for Child-Robot Interaction in Reverberated Environments	3120
<i>Martin Wollmer, Felix Weninger, Stefan Steidl, Anton Batliner, Bjorn Schuller</i>	

SPEECH AND AUDIO PROCESSING FOR HUMAN-ROBOT INTERACTION II

Chair: Alex Rudnicky

Blind Source Separation for Robot Audition Using Fixed Beamforming with HRTFs	3124
<i>Mounira Maazaoui, Yves Grenier, Karim Abed-Meraim</i>	
Real-Life Emotion Detection from Speech in Human-Robot Interaction: Experiments Across Diverse Corpora with Child and Adult Voices	3128
<i>Marie Tahon, Agnes Delaborde, Laurence Devillers</i>	
Weighted Ordered Classes - Nearest Neighbors: A New Framework for Automatic Emotion Recognition from Speech	3132
<i>Yazid Attabi, Pierre Dumouchel</i>	
Prosodic Analysis of a Corpus of Tales	3136
<i>David Doukhan, Albert Rilliard, Sophie Rosset, Martine Adda-Decker, Christophe d'Alessandro</i>	
Analysis of Acoustic-Prosodic Features Related to Paralinguistic Information Carried by Interjections in Dialogue Speech	3140
<i>Carlos T. Ishi, Hiroshi Ishiguro, Norihiro Hagita</i>	
Robust Intonation Pattern Classification in Human Robot Interaction	3144
<i>Martin Heckmann, Kazuhiro Nakadai, Hirofumi Nakajima</i>	
ASR for Human-Symbiotic Robot ``EMIEW2'' with Mechanical Noise and Floor-Level Noise Reduction	3148
<i>Takashi Sumiyoshi, Masahito Togami, Yasunari Obuchi</i>	

SPEECH TECHNOLOGY FOR UNDER-RESOURCED LANGUAGES I

Chairs: Alexey Karpov, Laurent Besacier

Rapid Building of an ASR System for Under-Resourced Languages Based on Multilingual Unsupervised Training	3152
<i>Ngoc Thang Vu, Franziska Kraus, Tanja Schultz</i>	
Places and Manner of Articulation of Bangla Consonants: A EPG Based Study	3156
<i>Shyamal K. Das Mandal, Somnath Chandra, Swaran Lata, A.K. Datta</i>	
Efficient Harvesting of Internet Audio for Resource-Scarce ASR	3160
<i>Marelle H. Davel, Charl van Heerden, Neil Kleynhans, Etienne Barnard</i>	

SPEECH TECHNOLOGY FOR UNDER-RESOURCED LANGUAGES II

Chairs: Laurent Besacier, Alexey Karpov

Automatic Prosody Generation for Serbo-Croatian Speech Synthesis Based on Regression Trees	3164
<i>Milan Secujski, Darko Pekar, Niksa Jakovljevic</i>	
Very Large Vocabulary ASR for Spoken Russian with Syntactic and Morphemic Analysis	3168
<i>Alexey Karpov, Irina Kipyatkova, Audrey Ronzhin</i>	
Cross-Language Phone Recognition when the Target Language Phoneme Inventory is not Known	3172
<i>Timothy Kempton, Roger K. Moore, Thomas Hain</i>	

A Paradigm for Limited Vocabulary Speech Recognition Based on Redundant Spectro-Temporal Feature Sets	3176
<i>Sourish Chaudhuri, Bhiksha Raj, Tony Ezzat</i>	
GorUp: An Ontology-Driven Audio Information Retrieval System that Suits the Requirements of Under-Resourced Languages	3180
<i>N. Barroso, K. Lopez de Ipina, A. Ezeiza, C. Hernandez, N. Ezeiza, O. Barroso, U. Susperregi, S. Barroso</i>	
Woefzela - An Open-Source Platform for ASR Data Collection in the Developing World	3184
<i>Nic J. de Vries, Jaco Badenhorst, Marelle H. Davel, Etienne Barnard, Alta de Waal</i>	
A Study on the Perception of Tone and Intonation in Sesotho	3188
<i>Hansjorg Mixdorff, Lehlohonolo Mohasi, 'Malillo Machobane, Thomas Niesler</i>	
Developing a Broadband Automatic Speech Recognition System for Afrikaans	3192
<i>Febe de Wet, Alta de Waal, Gerhard B. van Huyssteen</i>	
Multi-Accent Speech Recognition of Afrikaans, Black and White Varieties of South African English	3196
<i>Herman Kamper, Thomas Niesler</i>	
Perceptual Representation of Consonant Sounds in Thai	3200
<i>C. Tantibundhit, C. Onsuwan, T. Saimai, N. Saimai, S. Thatphithakkul, P. Chootrakool, K. Kosawat, N. Thatphithakkul</i>	
A Cross-Lingual Approach to the Development of an HMM-Based Speech Synthesis System for Malay	3204
<i>Mumtaz B. Mustafa, Raja N. Aion, Roziati Zainuddin, Zuraidah M. Don, Gerry Knowles</i>	

SPECIAL EVENTS

SPEAKER STATE CHALLENGE – INTOXICATION AND SLEEPINESS I

Chair: Bjoern Schuller

The INTERSPEECH 2011 Speaker State Challenge	3208
<i>Bjorn Schuller, Stefan Steidl, Anton Batliner, Florian Schiel, Jarek Krajewski</i>	
Combining Multiple Phoneme-Based Classifiers with Audio Feature-Based Classifier for the Detection of Alcohol Intoxication	3212
<i>Claude Montacie, Marie-Jose Caraty</i>	
Intoxication Detection Using Phonetic, Phonotactic and Prosodic Cues	3216
<i>Fadi Biadisy, William Yang Wang, Andrew Rosenberg, Julia Hirschberg</i>	
Drink and Speak: On the Automatic Classification of Alcohol Intoxication by Acoustic, Prosodic and Text-Based Features	3220
<i>Tobias Bocklet, Korbinian Riedhammer, Elmar Noth</i>	
Intoxicated Speech Detection by Fusion of Speaker Normalized Hierarchical Features and GMM Supervectors	3224
<i>Daniel Bone, Matthew P. Black, Ming Li, Angeliki Metallinou, Sungbok Lee, Shrikanth Narayanan</i>	
Attention, Sobriety Checkpoint! Can Humans Determine by Means of Voice, if Someone is Drunk... and Can Automatic Classifiers Compete?	3228
<i>Stefan Ultes, Alexander Schmitt, Wolfgang Minker</i>	
Does it Groove or does it Stumble - Automatic Classification of Alcoholic Intoxication using Prosodic Features	3232
<i>Florian Honig, Anton Batliner, Elmar Noth</i>	

SPEECH PROCESSING TOOLS

Chair: Christoph Draxler

Speech Processing Tools - An Introduction to Interoperability	3236
<i>Christoph Draxler, Toomas Altsaar, Sadaoki Furui, Mark Liberman, Peter Wittenburg</i>	
EasyAlign: An Automatic Phonetic Alignment Tool Under Praat	3240
<i>Jean-Philippe Goldman</i>	
Mtrans: A Multi-Channel, Multi-Tier Speech Annotation Tool	3244
<i>Julian Villegas, Martin Cooke, Vincent Aubanel, Marco A. Piccolino-Boniforti</i>	
The JSafran Platform for Semi-Automatic Speech Processing	3248
<i>Christophe Cerisara, Claire Gardent</i>	
The Social Signal Interpretation Framework (SSI) for Real Time Signal Processing and Recognition	3252
<i>Johannes Wagner, Florian Lingenfeller, Elisabeth Andre</i>	
ELAN - Aspects of Interoperability and Functionality	3256
<i>Han Sloetjes, Peter Wittenburg, Aarthy Somasundaram</i>	
Open Source Voice Creation Toolkit for the MARY TTS Platform	3260
<i>Marc Schroder, Marcela Charfuelan, Sathish Pammi, Ingmar Steiner</i>	

Java Visual Speech Components for Rapid Application Development of GUI Based Speech Processing Applications.....	3264
<i>Stefan Steidl, Korbinian Riedhammer, Tobias Bocklet, Florian Honig, Elmar Noth</i>	
mTalk - A Multimodal Browser for Mobile Services.....	3268
<i>Michael Johnston, Giuseppe Di Fabbrizio, Simon Urbanek</i>	
Web-Based Automatic Speech Recognition Service - webASR.....	3272
<i>Stuart N. Wrigley, Thomas Hain</i>	
A Web Based Speech Transcription Workplace.....	3276
<i>Markus Klehr, Andreas Ratzka, Thomas Ross</i>	
WinPitch: A Multimodal Tool for Speech Analysis of Endangered Languages.....	3280
<i>Philippe Martin</i>	
Recording Caregiver Interactions for Machine Acquisition of Spoken Language Using the KLAIR Virtual Infant.....	3284
<i>Mark Huckvale</i>	

SPEAKER STATE CHALLENGE – INTOXICATION AND SLEEPINESS II

Chair: Anton Batliner

Perception of Alcoholic Intoxication in Speech.....	3288
<i>Florian Schiel</i>	
Detecting Sleepiness by Fusing Classifiers Trained with Novel Acoustic Features.....	3292
<i>Tauhidur Rahman, Soroosh Mariooryad, Shalini Keshavamurthy, Gang Liu, John H.L. Hansen, Carlos Busso</i>	
An HMM-Based Approach to the INTERSPEECH 2011 Speaker State Challenge.....	3296
<i>Albino Nogueiras Rodriguez</i>	
RANSAC-Based Training Data Selection for Speaker State Recognition.....	3300
<i>Elif Bozkurt, Engin Erzin, Cigdem Eroglu Erdem, A. Tanju Erdem</i>	
University of Ljubljana System for Interspeech 2011 Speaker State Challenge.....	3304
<i>Rok Gajsek, Simon Dobrisek, France Mihelic</i>	
Speaker State Classification Based on Fusion of Asymmetric SIMPLS and Support Vector Machines.....	3308
<i>Dong-Yan Huang, Shuzhi Sam Ge, Zhengchen Zhang</i>	

SHOW & TELL SESSIONS

SHOW & TELL DEMONSTRATION – SPEECH SYSTEMS AND APPLICATIONS

Chair: Dimitrios Dimitriadis

An Affective Spoken Storyteller.....	3312
<i>Felix Burkhardt</i>	
Text Driven 3D Photo-Realistic Talking Head.....	3314
<i>Lijuan Wang, Wei Han, Frank K. Soong, Qiang Huo</i>	
Physical Models Producing Vowels with Pitch Variation.....	3316
<i>Takayuki Arai</i>	
An Engine-Independent Text-to-Speech Workplace.....	3318
<i>Margot Mieskes</i>	
An Application to Test the Emotion Conveyed by Vocal and Musical Signals.....	3320
<i>Simone Carcone, Carlo Giovannella</i>	
Automatic Speech Recognition System Dedicated for Polish.....	3322
<i>Mariusz Ziolkowski, Jakub Galka, Bartosz Ziolkowski, Tomasz Jadczyk, Dawid Skurzok, Mariusz Masior</i>	
Joint Application of Speech and Speaker Recognition for Automation and Security in Smart Home.....	3324
<i>Kong Aik Lee, Anthony Larcher, Helen Thai, Bin Ma, Haizhou Li</i>	
Adding a Speech Cursor to a Multimodal Dialogue System.....	3326
<i>Staffan Larsson, Alexander Berman, Jessica Villing</i>	
Prosody Toolkit: Integrating HTK, Praat and WEKA.....	3328
<i>S. Thomas Christie, Serguei Pakhomov</i>	
Collecting Life Logs for Experience-Based Corpora.....	3330
<i>F. Francesconi, A. Ghosh, G. Riccardi, M. Ronchetti, A. Vagin</i>	

SHOW & TELL DEMONSTRATION – MOBILITY AND WEB-SERVICES

Chair: Mazin Gilbert

Making an Automatic Speech Recognition Service Freely Available on the Web.....	3332
<i>Stuart N. Wrigley, Thomas Hain</i>	

AT&T VoiceBuilder: A Cloud-Based Text-to-Speech Voice Builder Tool	3334
<i>Yeon-Jun Kim, Thomas Okken, Alistair D. Conkie, Giuseppe Di Fabbrizio</i>	
Extending Audio Notetaker to Browse WebASR Transcriptions	3336
<i>Roger Tucker, Dan Fry, Vincent Wan, Stuart N. Wrigley, Thomas Hain</i>	
A Web-Based Tool for Developing Multilingual Pronunciation Lexicons	3338
<i>Samantha Ainsley, Linne Ha, Martin Jansche, Ara Kim, Masayuki Nanzawa</i>	
Speak4it and the Multimodal Semantic Interpretation System	3340
<i>Michael Johnston, Patrick Ehlen</i>	
TSAB - Web Interface for Transcribed Speech Collections	3342
<i>Tanel Alumae, Ahti Kitsik</i>	
Visual Voice Mail to Text on the iPhone/iPad	3344
<i>Andrej Ljolje, Vincent Goffin, Diamantino Caseiro, Taniya Mishra, Mazin Gilbert</i>	
Percy - An HTML5 Framework for Media Rich Web Experiments on Mobile Devices	3346
<i>Christoph Draxler</i>	
The KLAIR Toolkit for Recording Interactive Dialogues with a Virtual Infant	3348
<i>Mark Huckvale</i>	
Real-Time Prototype for Integration of Blind Source Extraction and Robust Automatic Speech Recognition	3350
<i>Francesco Nesta, Marco Matassoni, HariKrishna Maganti</i>	
Author Index	