

# **14th Annual Conference of the International Speech Communication Association (INTERSPEECH 2013)**

## **Speech in Life Sciences and Human Societies**

**Lyon, France  
25-29 August 2013**

**Volume 1 of 5**

**Editors:**

**F. Bimbot  
C. Cerisara  
C. Fougeron  
G. Gravier**

**L. Lamel  
F. Pellegrino  
P. Perrier**

**ISBN: 978-1-62993-443-3  
ISSN: 2308-457X**

**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© (2013) by the International Speech Communications Association  
All rights reserved.

Printed by Curran Associates, Inc. (2014)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# TABLE OF CONTENTS

## VOLUME 1

### ORAL SESSION 1: SYSTEMS FOR SEARCH/RETRIEVAL OF SPEECH DOCUMENTS

*Chairs: Martha Larson, Stavros Tsakalidis*

<b>Information Retrieval-Based Dynamic Time Warping</b> .....	1
<i>Xavier Anguera</i>	
<b>On the Computation of Document Frequency Statistics from Spoken Corpora Using Factor Automata</b> .....	6
<i>Dogan Can, Shrikanth Narayanan</i>	
<b>Acceleration of Spoken Term Detection Using a Suffix Array by Assigning Optimal Threshold Values to Sub-Keywords</b> .....	11
<i>Kouichi Katsurada, Seiichi Miura, Kheang Seng, Yurie Iribe, Tsuneo Nitta</i>	
<b>Strategies for High Accuracy Keyword Detection in Noisy Channels</b> .....	15
<i>Arindam Mandal, Julien van Hout, Yik-Cheung Tam, Vikramjit Mitra, Yun Lei, Jing Zheng, Dimitra Vergyri, Luciana Ferrer, Martin Graciarena, Andreas Kathol, Horacio Franco</i>	
<b>On the Calibration and Fusion of Heterogeneous Spoken Term Detection Systems</b> .....	20
<i>Alberto Abad, Luis Javier Rodriguez-Fuentes, Mikel Penagarikano, Amparo Varona, German Bordel</i>	
<b>Intensive Acoustic Models Constructed by Integrating Low-Occurrence Models for Spoken Term Detection</b> .....	25
<i>Shiro Narumi, Kazuma Konno, Takuya Nakano, Yoshiaki Itoh, Kazunori Kojima, Masaaki Ishigame, Kazuyo Tanaka, Shi-wook Lee</i>	

### ORAL SESSION 2: SPEECH ANALYSIS I

*Chairs: Masami Akamine, Kawasaki Henrich, Nathalie Henrich*

<b>Using Phonetic Feature Extraction to Determine Optimal Speech Regions for Maximising the Effectiveness of Glottal Source Analysis</b> .....	29
<i>John Kane, Irena Yanushevskaya, John Dalton, Christer Gobl, Ailbhe Ni Chasaide</i>	
<b>Beyond Bandlimited Sampling of Speech Spectral Envelope Imposed by the Harmonic Structure of Voiced Sounds</b> .....	34
<i>Hideki Kawahara, Masanori Morise, Tomoki Toda, Ryuichi Nisimura, Toshio Irino</i>	
<b>A Source-Filter Based Adaptive Harmonic Model and its Application to Speech Prosody Modification</b> .....	39
<i>JeeSok Lee, Frank K. Soong, Hong-Goo Kang</i>	
<b>Detection of Glottal Opening Instants Using Hilbert Envelope</b> .....	44
<i>K. Ramesh, S.R.M. Prasanna, D. Govind</i>	
<b>Robust Formant Detection Using Group Delay Function and Stabilized Weighted Linear Prediction</b> .....	49
<i>Dhananjaya Gowda, Jouni Pohjalainen, Mikko Kurimo, Paavo Alku</i>	
<b>A Source-Filter Separation Algorithm for Voiced Sounds Based on an Exact Anticausal/Causal Pole Decomposition for the Class of Periodic Signals</b> .....	54
<i>Thomas Hezard, Thomas Helie, Boris Doval</i>	

### ORAL SESSION 3: LANGUAGE AND DIALECT RECOGNITION

*Chairs: Kay Berklings, Karlsruhe Van Leeuwen, David Van Leeuwen*

<b>Parallel Absolute-Relative Feature Based Phonotactic Language Recognition</b> .....	59
<i>Weiwei Liu, Wei-Qiang Zhang, Zhiyi Li, Jia Liu</i>	
<b>Dimensionality Reduction of Phone Log-Likelihood Ratio Features for Spoken Language Recognition</b> .....	64
<i>Mireia Diez, Amparo Varona, Mikel Penagarikano, Luis Javier Rodriguez-Fuentes, German Bordel</i>	
<b>Improvements in Language Identification on the RATS Noisy Speech Corpus</b> .....	69
<i>Jeff Ma, Bing Zhang, Spyros Matsoukas, Sri Harish Mallidi, Feipeng Li, Hynek Hermansky</i>	
<b>Regularized Subspace n-Gram Model for Phonotactic iVector Extraction</b> .....	74
<i>Mehdi Souffar, Lukas Burget, Oldrich Plchot, Sandro Cumani, Jan Cernocky</i>	
<b>Foreign Accent Detection from Spoken Finnish Using i-Vectors</b> .....	79
<i>Hamid Behravan, Ville Hautamaki, Tomi Kinnunen</i>	

<b>Adaptive Gaussian Backend for Robust Language Identification</b> .....	84
<i>Mitchell McLaren, Aaron Lawson, Yun Lei, Nicolas Scheffer</i>	

#### **ORAL SESSION 4: ASR — NEURAL NETWORKS**

*Chairs: Hynek Hermansky, Alexander Waibel*

<b>Lattice-Based Training of Bottleneck Feature Extraction Neural Networks</b> .....	89
<i>Matthias Paulik</i>	
<b>Modular Combination of Deep Neural Networks for Acoustic Modeling</b> .....	94
<i>Jonas Gehring, Wonkyum Lee, Kevin Kilgour, Ian Lane, Yajie Miao, Alex Waibel</i>	
<b>Informative Spectro-Temporal Bottleneck Features for Noise-Robust Speech Recognition</b> .....	99
<i>Shuo-Yiin Chang, Nelson Morgan</i>	
<b>A Scalable Approach to Using DNN-Derived Features in GMM-HMM Based Acoustic Modeling for LVCSR</b> .....	104
<i>Zhi-Jie Yan, Qiang Huo, Jian Xu</i>	
<b>Improved Feature Processing for Deep Neural Networks</b> .....	109
<i>Shakti P. Rath, Daniel Povey, Karel Vesely, Jan Cernocky</i>	
<b>Deep vs. Wide: Depth on a Budget for Robust Speech Recognition</b> .....	114
<i>Oriol Vinyals, Nelson Morgan</i>	

#### **ORAL SESSION 5: SPEECH ACOUSTICS**

*Chairs: Kunitoshi Motoki, Jacqueline Vaissiere*

<b>An Early Case of ``VOT''</b> .....	119
<i>Angelika Braun</i>	
<b>Pitch Pattern Variations in Three Regional Varieties of American English</b> .....	123
<i>Robert Allen Fox, Ewa Jacewicz, Jessica Hart</i>	
<b>Fine-Grain Voice Strength Estimation from Vowel Spectral Cues</b> .....	128
<i>Jean-Sylvain Lienard, Claude Barras</i>	
<b>Linking Loudness Increases in Normal and Lombard Speech to Decreasing Vowel Formant Separation</b> .....	133
<i>Elizabeth Godoy, Catherine Mayo, Yannis Stylianou</i>	
<b>Three-Dimensional Rectangular Vocal-Tract Model with Asymmetric Wall Impedances</b> .....	138
<i>Kunitoshi Motoki</i>	
<b>Quasi Closed Phase Analysis for Glottal Inverse Filtering</b> .....	143
<i>Manu Airaksinen, Brad Story, Paavo Alku</i>	

#### **SPECIAL SESSION 1 (A & B): PARALINGUISTIC CHALLENGE**

*Chairs: Anton Batliner, Erlangen Schuller, Bjorn Schuller*

<b>The INTERSPEECH 2013 Computational Paralinguistics Challenge: Social Signals, Conflict, Emotion, Autism</b> .....	148
<i>Bjorn Schuller, Stefan Steidl, Anton Batliner, Alessandro Vinciarelli, Klaus Scherer, Fabien Ringeval, Mohamed Chetouani, Felix Weninger, Florian Eyben, Erik Marchi, Marcello Mortillaro, Hugues Salamin, Anna Polychroniou, Fabio Valente, Samuel Kim</i>	
<b>Non-Linguistic Vocalisation Recognition Based on Hybrid GMM-SVM Approach</b> .....	153
<i>Artur Janicki</i>	
<b>Characteristic Contours of Syllabic-Level Units in Laughter</b> .....	158
<i>Jieun Oh, Eunjoon Cho, Malcolm Slaney</i>	
<b>Detection of Nonverbal Vocalizations Using Gaussian Mixture Models: Looking for Fillers and Laughter in Conversational Speech</b> .....	163
<i>Teun F. Krikke, Khiet P. Truong</i>	
<b>Using Phonetic Patterns for Detecting Social Cues in Natural Conversations</b> .....	168
<i>Johannes Wagner, Florian Lingenfelter, Elisabeth Andre</i>	
<b>Paralinguistic Event Detection from Speech Using Probabilistic Time-Series Smoothing and Masking</b> .....	173
<i>Rahul Gupta, Kartik Audhkhasi, Sungbok Lee, Shrikanth Narayanan</i>	
<b>Detecting Laughter and Filled Pauses Using Syllable-Based Features</b> .....	178
<i>Gouzhen An, David Guy Brizan, Andrew Rosenberg</i>	

<b>Classifying Language-Related Developmental Disorders from Speech Cues: The Promise and the Potential Confounds</b> .....	182
<i>Daniel Bone, Theodora Chaspari, Kartik Audkhasi, James Gibson, Andreas Tsiartas, Maarten Van Segbroeck, Ming Li, Sungbok Lee, Shrikanth Narayanan</i>	
<b>Classification of Developmental Disorders from Speech Signals Using Submodular Feature Selection</b> .....	187
<i>Katrin Kirchhoff, Yuzong Liu, Jeff Bilmes</i>	
<b>Robust and Accurate Features for Detecting and Diagnosing Autism Spectrum Disorders</b> .....	191
<i>Meysam Asgari, Alireza Bayestehtashk, Izhak Shafran</i>	
<b>Suprasegmental Information Modelling for Autism Disorder Spectrum and Specific Language Impairment Classification</b> .....	195
<i>David Martinez, Dayana Ribas, Eduardo Lleida, Alfonso Ortega, Antonio Miguel</i>	
<b>Let Me Finish: Automatic Conflict Detection Using Speaker Overlap</b> .....	200
<i>Felix Grezes, Justin Richards, Andrew Rosenberg</i>	
<b>GMM Based Speaker Variability Compensated System for Interspeech 2013 ComParE Emotion Challenge</b> .....	205
<i>Vidhyasaharan Sethu, Julien Epps, Eliathamby Ambikairajah, Haizhou Li</i>	
<b>Random Subset Feature Selection in Automatic Recognition of Developmental Disorders, Affective States, and Level of Conflict from Speech</b> .....	210
<i>Okko Rasanen, Jouni Pohjalainen</i>	
<b>Ensemble of Machine Learning and Acoustic Segment Model Techniques for Speech Emotion and Autism Spectrum Disorders Recognition</b> .....	215
<i>Hung-yi Lee, Ting-yao Hu, How Jing, Yun-Fan Chang, Yu Tsao, Yu-Cheng Kao, Tsang-Long Pao</i>	
<b>Detecting Autism, Emotions and Social Signals Using AdaBoost</b> .....	220
<i>Gabor Gosztolya, Robert Busa-Fekete, Laszlo Toth</i>	

## **POSTER SESSION 1: PERCEPTION OF PROSODY**

*Chair: Carlos Gussenhoven*

<b>Resistance is Futile --- The Intonation Between Continuation Rise and Calling Contour in German</b> .....	225
<i>Oliver Niebuhr</i>	
<b>The Influence of F0 Contour Continuity on Prominence Perception</b> .....	230
<i>Hansjorg Mixdorff, Oliver Niebuhr</i>	
<b>Native English Listeners' Perceptions of Prosody in L1 and L2 Reading</b> .....	235
<i>Caroline L. Smith, Paul Edmunds</i>	
<b>Naturalness Judgement of L2 Mandarin Chinese --- Does Timing Matter?</b> .....	239
<i>Chiharu Tsurutani, Dean Luo</i>	
<b>Language Background Affects the Strength of the Pitch Bias in a Duration Discrimination Task</b> .....	243
<i>Daniel Aalto, Juraj Simko, Martti Vainio</i>	
<b>Pitch and Lengthening as Cues to Turn Transition in Swedish</b> .....	248
<i>Margaret Zellers</i>	
<b>Perception of Glottalization in Varying Pitch Contexts Across Languages</b> .....	253
<i>Maria Paola Bissiri, Margaret Zellers</i>	
<b>Exemplar-Based Pitch Accent Categorisation Using the Generalized Context Model</b> .....	258
<i>Michael Walsh, Katrin Schweitzer, Nadja Schaffler</i>	
<b>Double Contrast is Signalled by Prenuclear and Nuclear Accent Types Alone, Not by f0-Plateaux</b> .....	263
<i>Bettina Braun, Yuki Asano</i>	
<b>Word Stress Perception in European Portuguese</b> .....	267
<i>Susana Correia, Sonia Frota, Joseph Butler, Marina Vigarito</i>	
<b>Using Generalized Additive Models and Random Forests to Model Prosodic Prominence in German</b> .....	272
<i>Denis Arnold, Petra Wagner, R. Harald Baayen</i>	
<b>Perceiving Speech Rate Differences Between Natural and Time-Scale Modified Utterances</b> .....	277
<i>Hartmut R. Pfitzinger, Hansjorg Mixdorff</i>	

## **POSTER SESSION 2: PROSODY, PHONETICS OF LANGUAGE VARIETIES**

*Chair: Lya Meister*

<b>On the Robustness of Some Acoustic Parameters for Signalling Word Stress Across Styles in Brazilian Portuguese</b> .....	282
<i>Plinio A. Barbosa, Anders Eriksson, Joel Akesson</i>	

<b>Reexamine the Sandhi Rules and the Merging Tones in Hakka Language</b> .....	287
<i>Shao-ren Lyu, Ho-hsien Pan</i>	
<b>A Preliminary Spectral Analysis of Palatal and Velar Stop Bursts in Pitjantjatjara</b> .....	291
<i>Marija Tabain, Richard Beare, Andrew Butcher</i>	
<b>Presentational Focus Realisation in Nalbaria Variety of Assamese</b> .....	296
<i>Shakuntala Mahanta, A.I. Twaha</i>	
<b>On the Relation Between Intonational Phrasing and Pitch Accent Distribution. Evidence from European Portuguese Varieties</b> .....	300
<i>Marisa Cruz, Sonia Frota</i>	
<b>How Are Word-Final Schwas Different in the North and South of France?</b> .....	305
<i>Rena Nemoto, Martine Adda-Decker</i>	
<b>Modeling Postcolonial Language Varieties: Challenges and Lessons Learned from Mozambican Portuguese</b> .....	310
<i>Simone Ashby, Silvia Barbosa, Catarina Silva, Paulino Fumo, Jose Pedro Ferreira</i>	
<b>Prosody of Contrastive Focus in Estonian</b> .....	315
<i>Heete Sahlkai, Mari-Liis Kalvik, Meelis Mihkla</i>	
<b>Exploring the Connection of Acoustic and Distinctive Features</b> .....	320
<i>Thomas Kisler, Uwe D. Reichel</i>	
<b>A Physiological Analysis of the Tense/Lax Vowel Contrast in Two Varieties of German</b> .....	325
<i>Conceicao Cunha, Jonathan Harrington, Phil Hoole</i>	
<b>Production of Estonian Quantity Contrasts by Native Speakers of Finnish</b> .....	330
<i>Einar Meister, Lya Meister</i>	
<b>Aerodynamic and Durational Cues of Phonological Voicing in Whisper</b> .....	335
<i>Yohann Meynadier, Yulia Gaydina</i>	
<b>Information Theoretic Syllable Structure and its Relation to the c-Center Effect</b> .....	340
<i>Uwe D. Reichel</i>	
<b>The Bulgarian Stressed and Unstressed Vowel System. A Corpus Study</b> .....	345
<i>Bistra Andreeva, William Barry, Jacques Koreman</i>	

### **POSTER SESSION 3: SPEECH SYNTHESIS I**

*Chair: Marcela Charfuelan*

<b>Training an Articulatory Synthesizer with Continuous Acoustic Data</b> .....	349
<i>Santitham Prom-on, Peter Birkholz, Yi Xu</i>	
<b>Estimating Speaker-Specific Intonation Patterns Using the Linear Alignment Model</b> .....	354
<i>Geza Kiss, Jan P.H. van Santen</i>	
<b>Factored Maximum Likelihood Kernelized Regression for HMM-Based Singing Voice Synthesis</b> .....	359
<i>June Sig Sung, Doo Hwa Hong, Hyun Woo Koo, Nam Soo Kim</i>	
<b>Improvements to HMM-Based Speech Synthesis Based on Parameter Generation with Rich Context Models</b> .....	364
<i>Shinnosuke Takamichi, Tomoki Toda, Yoshinori Shiga, Sakriani Sakti, Graham Neubig, Satoshi Nakamura</i>	
<b>Voice Conversion in High-Order Eigen Space Using Deep Belief Nets</b> .....	369
<i>Toru Nakashika, Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Arika</i>	
<b>Voice Conversion for Non-Parallel Datasets Using Dynamic Kernel Partial Least Squares Regression</b> .....	373
<i>Hanna Silen, Jani Nurminen, Elina Helander, Moncef Gabbouj</i>	
<b>A Style Control Technique for Singing Voice Synthesis Based on Multiple-Regression HSMM</b> .....	378
<i>Takashi Nose, Misa Kanemoto, Tomoki Koriyama, Takao Kobayashi</i>	
<b>Predicting the Quality of Text-to-Speech Systems from a Large-Scale Feature Set</b> .....	383
<i>Florian Hinterleitner, Christoph R. Norrenbrock, Sebastian Moller, Ulrich Heute</i>	
<b>Speaker-Specific Retraining for Enhanced Compression of Unit Selection Text-to-Speech Databases</b> .....	388
<i>Jani Nurminen, Hanna Silen, Moncef Gabbouj</i>	
<b>Avatar Therapy: An Audio-Visual Dialogue System for Treating Auditory Hallucinations</b> .....	392
<i>Mark Huckvale, Julian Leff, Geoff Williams</i>	
<b>Optimizations and Fitting Procedures for the Liljencrants-Fant Model for Statistical Parametric Speech Synthesis</b> .....	397
<i>Prasanna Kumar Muthukumar, Alan W. Black, H. Timothy Bunnell</i>	
<b>Analysis and Modeling of “Focus” in Context</b> .....	402
<i>Dirk Hovy, Gopala Krishna Anumanchipalli, Alok Parlkar, Caroline Vaughn, Adam Lammert, Eduard Hovy, Alan W. Black</i>	

## **ORAL SESSION 6: PERCEPTION, DIALECTAL DIFFERENCES**

*Chairs: Catherine Best, Sydney Munson, Benjamin Munson*

<b>Production and Perception of Pseudo-V1CV2 Outside the Vowel Triangle: Speech Illusion Effects</b> .....	407
<i>Thi Anh Xuan Tran, Viet Son Nguyen, Eric Castelli, Rene Carre</i>	
<b>Recent Evolution of Non-Standard Consonantal Variants in French Broadcast News</b> .....	412
<i>Maria Candea, Martine Adda-Decker, Lori Lamel</i>	
<b>Architekt or Archtekt? Perception of Devoiced Vowels Produced by Japanese Speakers of German</b> .....	417
<i>Frank Zimmerer, Rei Yasuda, Henning Reetz</i>	
<b>Comparing Vowel Category Response Surfaces Over Age-Varying Maximal Vowel Spaces Within and Across Language Communities</b> .....	421
<i>Andrew R. Plummer, Lucie Menard, Benjamin Munson, Mary E. Beckman</i>	
<b>Perceived Vocal Attractiveness Across Dialects is Similar but not Uniform</b> .....	426
<i>Molly Babel, Grant McGuire</i>	
<b>Mutual Intelligibility of American, Chinese and Dutch-Accented Speakers of English Tested by SUS and SPIN Sentences</b> .....	431
<i>Hongyan Wang, Vincent J. van Heuven</i>	

## **ORAL SESSION 7: SPEECH ENHANCEMENT — SINGLE CHANNEL**

*Chairs: Yifan Gong, Redmond Saruwatari, Hiroshi Saruwatari*

<b>Speech Enhancement Based on Deep Denoising Autoencoder</b> .....	436
<i>Xugang Lu, Yu Tsao, Shigeki Matsuda, Chiori Hori</i>	
<b>Musical Noise Analysis for Bayesian Minimum Mean-Square Error Speech Amplitude Estimators Based on Higher-Order Statistics</b> .....	441
<i>Hiroshi Saruwatari, Suzumi Kanehara, Ryoichi Miyazaki, Kiyohiro Shikano, Kazunobu Kondo</i>	
<b>Non-Negative Matrix Factorization with Linear Constraints for Single-Channel Speech Enhancement</b> .....	446
<i>Nikolay Lyubimov, Mikhail Kotov</i>	
<b>A Single Channel Speech Enhancement Approach by Combining Statistical Criterion and Multi-Frame Sparse Dictionary Learning</b> .....	451
<i>Hung-Wei Tseng, Srikanth Vishnubhotla, Mingyi Hong, Xiangfeng Wang, Jinjun Xiao, Zhi-Quan Luo, Tao Zhang</i>	
<b>Speech Enhancement Using Convolutional Nonnegative Matrix Factorization with Cosparsity Regularization</b> .....	456
<i>Majid Mirbagheri, Yanbo Xu, Sahar Akram, Shihab Shamma</i>	
<b>Joint Stochastic-Deterministic Wiener Filtering with Recursive Bayesian Estimation of Deterministic Speech</b> .....	460
<i>Matthew McCallum, Bernard Guillemin</i>	

## **ORAL SESSION 8: DIALOG MODELING**

*Chairs: Olivier Pietquin, Metz Bellegarda, Jerome Bellegarda*

<b>Automatic Self-Supervised Learning of Associations Between Speech and Text</b> .....	465
<i>Juho Knuutila, Okko Rasanen, Unto K. Laine</i>	
<b>Particle Swarm Optimisation of Spoken Dialogue System Strategies</b> .....	470
<i>Lucie Daubigney, Matthieu Geist, Olivier Pietquin</i>	
<b>Model-Based Bayesian Reinforcement Learning for Dialogue Management</b> .....	475
<i>Pierre Lison</i>	
<b>Evaluating Spoken Dialogue Models Under the Interactive Pattern Recognition Framework</b> .....	480
<i>Fabrizio Ghigi, Maria Ines Torres, Raquel Justo, Jose-Miguel Benedi</i>	
<b>Multi-Layer Mutually Reinforced Random Walk with Hidden Parameters for Improved Multi-Party Meeting Summarization</b> .....	485
<i>Yun-Nung Chen, Florian Metz</i>	
<b>A Recursive Dialogue Game Framework with Optimal Policy Offering Personalized Computer-Assisted Language Learning</b> .....	490
<i>Pei-hao Su, Yow-Bang Wang, Tsung-Hsien Wen, Tien-han Yu, Lin-shan Lee</i>	

## **ORAL SESSION 9: ASR — LEXICAL, PROSODIC AND CROSS/MULTI-LINGUAL**

*Chairs: Kate Knill, Torbjorn Svendsen*

<b>Improving LVCSR with Hidden Conditional Random Fields for Grapheme-to-Phoneme Conversion</b> .....	495
<i>Stefan Hahn, Patrick Lehnen, Simon Wiesler, Ralf Schluter, Hermann Ney</i>	
<b>Context-Dependent Phone Mapping for LVCSR of Under-Resourced Languages</b> .....	500
<i>Van Hai Do, Xiong Xiao, Eng Siong Chng, Haizhou Li</i>	
<b>Improving Grapheme-Based ASR by Probabilistic Lexical Modeling Approach</b> .....	505
<i>Ramya Rasipuram, Mathew Magimai-Doss</i>	
<b>Crosslingual Tandem-SGMM: Exploiting Out-of-Language Data for Acoustic Model and Feature Level Adaptation</b> .....	510
<i>Petr Motlicek, David Imseng, Philip N. Garner</i>	
<b>Multilingual Multilayer Perceptron for Rapid Language Adaptation Between and Across Language Families</b> .....	515
<i>Ngoc Thang Vu, Tanja Schultz</i>	
<b>Modeling Prosodic Sequences with K-Means and Dirichlet Process GMMs</b> .....	520
<i>Andrew Rosenberg</i>	

## **ORAL SESSION 10: PHONETIC CONVERGENCE**

*Chairs: Veronique Delvaux, Jason Shaw*

<b>Convergence of Articulation Rate in Spontaneous Speech</b> .....	525
<i>Antje Schweitzer, Natalie Lewandowski</i>	
<b>Phonetic Convergence in Shadowed Speech: A Comparison of Perceptual and Acoustic Measures</b> .....	530
<i>Jennifer S. Pardo</i>	
<b>Pitch and Duration as a Basis for Entrainment of Overlapped Speech Onsets</b> .....	535
<i>Marcin Włodarczyk, Juraj Simko, Petra Wagner</i>	
<b>Investigating Fine Temporal Dynamics of Prosodic and Lexical Accommodation</b> .....	539
<i>Francesca Bonin, Celine De Looze, Sucheta Ghosh, Emer Gilmartin, Carl Vogel, Anna Polychroniou, Hugues Salamin, Alessandro Vinciarelli, Nick Campbell</i>	
<b>Spontaneous and Explicit Speech Imitation</b> .....	544
<i>Jeesun Kim, Ruben Demirdjian, Chris Davis</i>	
<b>Imitation Interacts with One's Second-Language Phonology But it Does Not Operate Cross-Linguistically</b> .....	548
<i>Vaclav Jonas Podlipsky, Sarka Simackova, Katerina Chladkova</i>	

## **POSTER SESSION 4: SPEECH PRODUCTION, ACQUISITION AND DEVELOPMENT I**

*Chair: Takayuki Arai*

<b>Prosodic Markings of Semantic Predictability in Taiwan Mandarin</b> .....	553
<i>Po-jen Hsieh</i>	
<b>How Did it Work? Historic Phonetic Devices Explained by Coeval Photographs</b> .....	558
<i>Rudiger Hoffmann, Dieter Mehnert, Rolf Dietzel</i>	
<b>Eliciting Speech with Sentence Lists --- A Critical Evaluation with Special Emphasis on Segmental Anchoring</b> .....	563
<i>Lea S. Kohtz, Oliver Niebuhr</i>	
<b>An MRI-Based Acoustic Study of Mandarin Vowels</b> .....	568
<i>Yuguang Wang, Jianwu Dang, Xi Chen, Jianguo Wei, Hongcui Wang, Kiyoshi Honda</i>	
<b>Melody Metrics for Prosodic Typology: Comparing English, French and Chinese</b> .....	572
<i>Daniel Hirst</i>	
<b>Velic Coordination in French Nasals: A Real-Time Magnetic Resonance Imaging Study</b> .....	577
<i>Michael Proctor, Louis Goldstein, Adam Lammert, Dani Byrd, Asterios Toutios, Shrikanth Narayanan</i>	
<b>Learning to Imitate Adult Speech with the KLAIR Virtual Infant</b> .....	582
<i>Mark Huckvale, Amrita Sharma</i>	
<b>Physics-Based Synthesis of Disordered Voices</b> .....	587
<i>Jorge C. Lucero, Jean Schoentgen, Mara Behlau</i>	
<b>Place Assimilation and Articulatory Strategies: The Case of Sibilant Sequences in French as L1 and L2</b> .....	592
<i>Sonia d'Apolito, Barbara Gili Fivela</i>	



<b>Effects of Lexical Class and Lemma Frequency on German Homographs</b> .....	597
<i>Barbara Samlowski, Petra Wagner, Bernd Mobius</i>	
<b>Measuring Laryngealization in Running Speech: Interaction with Contrastive Tones in Yalalag Zapotec</b> .....	602
<i>Leonardo Lancia, Heriberto Avelino, Daniel Voigt</i>	
<b>A Neural Oscillator Model of Speech Timing and Rhythm</b> .....	607
<i>Erin Rusaw</i>	
<b>Observations of Perseverative Coarticulation in Lateral Approximants Using MRI</b> .....	612
<i>Nicole Wong, Maojing Fu, Zhi-Pei Liang, Ryan K. Shosted, Bradley P. Sutton</i>	

## **POSTER SESSION 5: GENERAL TOPICS IN ASR**

*Chair: Tanel Alumae*

<b>Comparing Computation in Gaussian mixture and Neural Network Based Large-Vocabulary Speech Recognition</b> .....	617
<i>Vishwa Gupta, Gilles Boulianne</i>	
<b>Simultaneous Perturbation Stochastic Approximation for Automatic Speech Recognition</b> .....	622
<i>Daniel Stein, Jochen Schwenninger, Michael Stadtschmitzer</i>	
<b>Hardware/Software Codesign for Mobile Speech Recognition</b> .....	627
<i>David Sheffield, Michael Anderson, Yunsup Lee, Kurt Keutzer</i>	
<b>Exploiting the Succeeding Words in Recurrent Neural Network Language Models</b> .....	632
<i>Yangyang Shi, Martha Larson, Pascal Wiggers, Catholijn M. Jonker</i>	
<b>Speech Acoustic Unit Segmentation Using Hierarchical Dirichlet Processes</b> .....	637
<i>Amir Hossein Harati Nejad Torbati, Joseph Picone, Marc Sobel</i>	
<b>Transducer-Based Speech Recognition with Dynamic Language Models</b> .....	642
<i>Munir Georges, Stephan Kanthak, Dietrich Klakow</i>	
<b>A Method for Structure Estimation of Weighted Finite-State Transducers and its Application to Grapheme-to-Phoneme Conversion</b> .....	647
<i>Yotaro Kubo, Takaaki Hori, Atsushi Nakamura</i>	
<b>Combining Forward-Based and Backward-Based Decoders for Improved Speech Recognition Performance</b> .....	652
<i>Denis Jouvet, Dominique Fohr</i>	
<b>iVector-Based Acoustic Data Selection</b> .....	657
<i>Olivier Siohan, Michiel Bacchiani</i>	
<b>Accurate and Compact Large Vocabulary Speech Recognition on Mobile Devices</b> .....	662
<i>Xin Lei, Andrew Senior, Alexander Gruenstein, Jeffrey Sorensen</i>	
<b>Pre-Initialized Composition for Large-Vocabulary Speech Recognition</b> .....	666
<i>Cyril Allauzen, Michael Riley</i>	
<b>Speaker Dependent Activation Keyword Detector Based on GMM-UBM</b> .....	671
<i>Evelyn Kurniawati, Sapna George</i>	
<b>Written-Domain Language Modeling for Automatic Speech Recognition</b> .....	675
<i>Hasim Sak, Yun-hsuan Sung, Françoise Beaufays, Cyril Allauzen</i>	

## **POSTER SESSION 6: VOICE ACTIVITY DETECTION AND SPEECH SEGMENTATION**

*Chair: Ascension Gallardo Antolin*

<b>Detecting Words in Speech Using Linear Separability in a Bag-of-Events Vector Space</b> .....	680
<i>Maarten Versteegh, Louis ten Bosch</i>	
<b>On the Improvement of Multimodal Voice Activity Detection</b> .....	685
<i>Matt Burlick, Dimitrios Dimitriadis, Eric Zavesky</i>	
<b>Using Linguistic Information to Detect Overlapping Speech</b> .....	690
<i>Jurgen T. Geiger, Florian Eyben, Nicholas Evans, Bjorn Schuller, Gerhard Rigoll</i>	
<b>Incremental Acoustic Subspace Learning for Voice Activity Detection Using Harmonicity-Based Features</b> .....	695
<i>Jiaxing Ye, Takumi Kobayashi, Masahiro Murakawa, Tetsuya Higuchi</i>	
<b>Endpoint Detection Using Weighted Finite State Transducer</b> .....	700
<i>Hoon Chung, SungJoo Lee, YunKeun Lee</i>	

<b>A Robust Frontend for VAD: Exploiting Contextual, Discriminative and Spectral Cues of Human Voice</b> .....	704
<i>Maarten Van Segbroeck, Andreas Tsirtas, Shrikanth Narayanan</i>	
<b>All for One: Feature Combination for Highly Channel-Degraded Speech Activity Detection</b> .....	709
<i>Martin Graciarena, Abeer Alwan, Dan Ellis, Horacio Franco, Luciana Ferrer, John H.L. Hansen, Adam Janin, Byung-Suk Lee, Yun Lei, Vikramjit Mitra, Nelson Morgan, Seyed Omid Sadjadi, T.J. Tsai, Nicolas Scheffer, Lee Ngee Tan, Benjamin Williams</i>	
<b>Superposed Speech Localisation Using Frequency Tracking</b> .....	714
<i>Maxime Le Coz, Julien Pinquier, Regine Andre-Obrecht</i>	
<b>Multi-Band Long-Term Signal Variability Features for Robust Voice Activity Detection</b> .....	718
<i>Andreas Tsirtas, Theodora Chaspari, Nossos Katsamanis, Prasanta Kumar Ghosh, Ming Li, Maarten Van Segbroeck, Alexandros Potamianos, Shrikanth Narayanan</i>	
<b>A Low-Complexity Voice Activity Detector for Smart Hearing Protection of Hyperacusic Persons</b> .....	723
<i>Narimene Lezzoum, Ghyslain Gagnon, Jeremie Voix</i>	
<b>Speech Activity Detection on YouTube Using Deep Neural Networks</b> .....	728
<i>Neville Ryant, Mark Liberman, Jiahong Yuan</i>	
<b>Speaker and Noise Independent Voice Activity Detection</b> .....	732
<i>Francois G. Germain, Dennis L. Sun, Gautham J. Mysore</i>	
<b>Confidence-Based Scoring: A Useful Diagnostic Tool for Detection Tasks</b> .....	737
<i>T.J. Tsai, Adam Janin</i>	
<b>Concurrent Processing of Voice Activity Detection and Noise Reduction Using Empirical Mode Decomposition and Modulation Spectrum Analysis</b> .....	742
<i>Yasuaki Kanai, Shota Morita, Masashi Unoki</i>	

## **SHOW & TELL-1: SHOW AND TELL SESSION 1**

*Chairs: Laurence Devillers, Fabrice Lefevre*

<b>The Furhat Social Companion Talking Head</b> .....	747
<i>Samer Al Moubayed, Jonas Beskow, Gabriel Skantze</i>	
<b>Audition: The Most Important Sense for Humanoid Robots?</b> .....	750
<i>Rodolphe Gelin, Gabriele Barbieri</i>	
<b>Ultraspeech-player: Intuitive Visualization of Ultrasound Articulatory Data for Speech Therapy and Pronunciation Training</b> .....	752
<i>Thomas Hueber</i>	
<b>Laughter Modulation: From Speech to Speech-Laugh</b> .....	754
<i>Jieun Oh, Ge Wang</i>	
<b>ReFr: An Open-Source Reranker Framework</b> .....	756
<i>Daniel M. Bikel, Keith B. Hall</i>	
<b>Embedding Speech Recognition to Control Lights</b> .....	759
<i>Alessandro Sosi, Fabio Brugnara, Luca Cristoforetti, Marco Matassoni, Mirco Ravanelli, Maurizio Omologo</i>	
<b>The MUTE Silent Speech Recognition System</b> .....	761
<i>Geoffrey S. Meltzner, James T. Heaton, Yubin Deng</i>	
<b>The Edinburgh Speech Production Facility DoubleTalk Corpus</b> .....	764
<i>James M. Scobbie, Alice Turk, Christian Geng, Simon King, Robin Lickley, Korin Richmond</i>	
<b>Lexee: A Cloud-Based Platform for Building and Deploying Voice-Enabled Mobile Applications</b> .....	767
<i>Dmitry Sityaev, Jonathan Hotz, Vadim Snitkovsky</i>	
<b>Visualizing Articulatory Data with VisArtico</b> .....	770
<i>Slim Ouni</i>	
<b>A Tool to Elicit and Collect Multicultural and Multimodal Laughter</b> .....	773
<i>Mariette Soury, Clement Gossart, Martine Adda-Decker, Laurence Devillers</i>	
<b>Design of a Mobile App for Interspeech Conferences: Towards an Open Tool for the Spoken Language Community</b> .....	775
<i>Robert Schleicher, Tilo Westermann, Jinjin Li, Moritz Lawitschka, Benjamin Mateev, Ralf Reichmuth, Sebastian Moller</i>	

## **ORAL SESSION 11: DISCOURSE, INTONATION, PROSODY**

*Chairs: Plinio Barbosa, Marija Tabain*

<b>The Acoustics of Word Stress in Swedish: A Function of Stress Level, Speaking Style and Word Accent</b> .....	778
<i>Anders Eriksson, Plinio A. Barbosa, Joel Akesson</i>	
<b>Intonational Contrasts Encode Speaker's Certainty in Neutral vs. Incredulity Declarative Questions in French</b> .....	783
<i>Amandine Michelas, Cristel Portes, Maud Champagne-Lavau</i>	

## **VOLUME 2**

<b>Prosodic Changes Pre-Announcing a Syntactic Completion Point in Japanese Utterance</b> .....	788
<i>Yuichi Ishimoto, Mika Enomoto, Hitoshi Iida</i>	
<b>Prosodic Encoding of Declarative, Interrogative and Imperative Sentences in Jaminjung, a Language of Australia</b> .....	793
<i>Candide Simard</i>	
<b>Crosslinguistic Priming in Interactive Reference: Evidence for Conceptual Alignment in Speech Production</b> .....	798
<i>Anne Vullings, Martijn Goudbeek, Emiel Kraemer</i>	
<b>A Cross-Linguistic Study on Turn-Taking and Temporal Alignment in Verbal Interaction</b> .....	803
<i>Spyros Kousidis, David Schlangen, Stavros Skopeteas</i>	

## **ORAL SESSION 12: SOURCE SEPARATION**

*Chairs: Shoji Makino, Emmanuel Vincent*

<b>Discriminative Nonnegative Dictionary Learning Using Cross-Coherence Penalties for Single Channel Source Separation</b> .....	808
<i>Emad M. Grais, Hakan Erdogan</i>	
<b>Monaural Speech Segregation Based on Pitch Track Correction Using an Ensemble Kalman Filter</b> .....	813
<i>Han-Gyu Kim, Gil-Jin Jang, Jeong-Sik Park, Yung-Hwan Oh</i>	
<b>Voice Activity Classification for Automatic Bi-Speaker Adaptive Beamforming in Speech Separation</b> .....	817
<i>Thuy N. Tran, William Cowley, Andre Pollok</i>	
<b>Blind Source Separation Using Spatially Distributed Microphones Based on Microphone-Location Dependent Source Activities</b> .....	822
<i>Keisuke Kinoshita, Mehrez Souden, Tomohiro Nakatani</i>	
<b>Non-Negative Tensor Factorisation of Modulation Spectrograms for Monaural Sound Source Separation</b> .....	827
<i>Tom Barker, Tuomas Virtanen</i>	
<b>Iterative Sinusoidal-Based Partial Phase Reconstruction in Single-Channel Source Separation</b> .....	832
<i>Mario Kaoru Watanabe, Pejman Mowlae</i>	

## **ORAL SESSION 13: PARALINGUISTIC INFORMATION**

*Chairs: Elizabeth Shriberg, Michael Wagner*

<b>Classification of Speech Under Stress by Modeling the Aerodynamics of the Laryngeal Ventricle</b> .....	837
<i>Xiao Yao, Takatoshi Jitsuhiro, Chiyomi Miyajima, Norihide Kitaoka, Kazuya Takeda</i>	
<b>“Sure, I Did the Right Thing”: A System for Sarcasm Detection in Speech</b> .....	842
<i>Rachel Rakov, Andrew Rosenberg</i>	
<b>Investigating Voice Quality as a Speaker-Independent Indicator of Depression and PTSD</b> .....	847
<i>Stefan Scherer, Giota Stratou, Jonathan Gratch, Louis-Philippe Morency</i>	
<b>A Corpus-Based Study of Elderly and Young Speakers of European Portuguese: Acoustic Correlates and Their Impact on Speech Recognition Performance</b> .....	852
<i>Thomas Pellegrini, Annika Hamalainen, Philippe Boula de Mareuil, Michael Tjalve, Isabel Trancoso, Sara Candeias, Miguel Sales Dias, Daniela Braga</i>	
<b>Modeling Spectral Variability for the Classification of Depressed Speech</b> .....	857
<i>Nicholas Cummins, Julien Epps, Vidhyasaharan Sethu, Michael Breakpear, Roland Goecke</i>	

<b>Sentiment Analysis of Online Spoken Reviews</b> .....	862
<i>Veronica Perez-Rosas, Rada Mihalcea</i>	

### **ORAL SESSION 14: ASR — ROBUSTNESS AGAINST NOISE I**

*Chairs: John Hansen, Denis Jouvet*

<b>Using Twin-HMM-Based Audio-Visual Speech Enhancement as a Front-End for Robust Audio-Visual Speech Recognition</b> .....	867
<i>Ahmed Hussien Abdelaziz, Steffen Zeiler, Dorothea Kolossa</i>	
<b>Spectro-Temporal Directional Derivative Features for Automatic Speech Recognition</b> .....	872
<i>James Gibson, Maarten Van Segbroeck, Antonio Ortega, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	
<b>Attribute-Based Histogram Equalization (HEQ) and its Adaptation for Robust Speech Recognition</b> .....	876
<i>Xiong Xiao, Eng Siong Chng, Haizhou Li</i>	
<b>Modified Cepstral Mean Normalization --- Transforming to Utterance Specific Non-Zero Mean</b> .....	881
<i>Vikas Joshi, N. Vishnu Prasad, S. Umesh</i>	
<b>Damped Oscillator Cepstral Coefficients for Robust Speech Recognition</b> .....	886
<i>Vikramjit Mitra, Horacio Franco, Martin Graciarena</i>	
<b>Regularized MVDR Spectrum Estimation-Based Robust Feature Extractors for Speech Recognition</b> .....	891
<i>Md. Jahangir Alam, Patrick Kenny, Douglas O'Shaughnessy</i>	

### **ORAL SESSION 15: NEURAL BASIS OF SPEECH PERCEPTION**

*Chairs: Anne-Lise Giraud*

<b>Optimization of Sigmoidal Rate-Level Function Based on Acoustic Features</b> .....	896
<i>Victor Poblete, Nestor Becerra Yoma, Richard M. Stern</i>	
<b>Composing Auditory ERPs: Cross-Linguistic Comparison of Auditory Change Complex for Japanese Fricative Consonants</b> .....	901
<i>Makiko Sadakata, Loukianos Spyrou, Mizuki Shingai, Kaoru Sekiyama</i>	
<b>How Voicing, Place and Manner of Articulation Differently Modulate Event-Related Potentials Associated with Response Inhibition</b> .....	906
<i>Nathalie Bedoin, Jennifer Krzonowski, Emmanuel Ferragne</i>	
<b>Categorization of Speech in Early Auditory Evoked Responses</b> .....	911
<i>Ludovic Bellier, Michel Mazzuca, Hung Thai-Van, Anne Caclin, Rafael Laboissiere</i>	
<b>Perception and Production of Italian Vowels: An ERP Study</b> .....	916
<i>Anna Dora Manca, Mirko Grimaldi</i>	
<b>Implicit Learning Leads to Familiarity Effects for Intonation but not for Voice</b> .....	921
<i>Ann-Kathrin Grohe, Bettina Braun</i>	

### **SPECIAL SESSION 2: SPOOFING AND COUNTERMEASURES FOR AUTOMATIC SPEAKER VERIFICATION**

*Chairs: Nicholas Evans, Sophia Antipolis, Sebastien Marcel*

<b>Spoofing and Countermeasures for Automatic Speaker Verification</b> .....	925
<i>Nicholas Evans, Tomi Kinnunen, Junichi Yamagishi</i>	
<b>I-Vectors Meet Imitators: On Vulnerability of Speaker Verification Systems Against Voice Mimicry</b> .....	930
<i>Rosa Gonzalez Hautamaki, Tomi Kinnunen, Ville Hautamaki, Timo Leino, Anne-Maria Laukkanen</i>	
<b>Security Evaluation of i-Vector Based Speaker Verification Systems Against Hill-Climbing Attacks</b> .....	935
<i>Marta Gomez-Barrero, Javier Gonzalez-Dominguez, Javier Galbally, Joaquin Gonzalez-Rodriguez</i>	
<b>A New Speaker Verification Spoofing Countermeasure Based on Local Binary Patterns</b> .....	940
<i>Federico Alegre, Ravichander Vipperla, Asmaa Amehraye, Nicholas Evans</i>	
<b>Voice Transformation-Based Spoofing of Text-Dependent Speaker Verification Systems</b> .....	945
<i>Zvi Kons, Hagai Aronowitz</i>	
<b>Vulnerability Evaluation of Speaker Verification Under Voice Conversion Spoofing: The Effect of Text Constraints</b> .....	950
<i>Zhizheng Wu, Anthony Larcher, Kong Aik Lee, Eng Siong Chng, Tomi Kinnunen, Haizhou Li</i>	

## **POSTER SESSION 7: SPEECH PRODUCTION, ACQUISITION AND DEVELOPMENT II**

Chair: Qiang Fang

<b>Timing Differences in Articulation Between Voiced and Voiceless Stop Consonants: An Analysis of Cine-MRI Data</b> .....	955
<i>Masako Fujimoto, Tatsuya Kitamura, Hiroaki Hatano, Ichiro Fujimoto</i>	
<b>Vocal Tract Cross-Distance Estimation from Real-Time MRI Using Region-of-Interest Analysis</b> .....	959
<i>Adam Lammert, Vikram Ramanarayanan, Michael Proctor, Shrikanth Narayanan</i>	
<b>Syllable Nuclei Detection Using Perceptually Significant Features</b> .....	963
<i>Apoorv Reddy Arrabothu, Nivedita Chennupati, B. Yegnanarayana</i>	
<b>Truncation of Pharyngeal Gesture in English Diphthong [aɪ]</b> .....	968
<i>Fang-Ying Hsieh, Louis Goldstein, Dani Byrd, Shrikanth Narayanan</i>	
<b>The Effect of Word Frequency and Lexical Class on Articulatory-Acoustic Coupling</b> .....	973
<i>Zhaojun Yang, Vikram Ramanarayanan, Dani Byrd, Shrikanth Narayanan</i>	
<b>Discrimination Between Fricative and Affricate in Japanese Using Time and Spectral Domain Variables</b> .....	978
<i>Kimiko Yamakawa, Shigeaki Amano</i>	
<b>L2 Syntax Acquisition: The Effect of Oral and Written Computer Assisted Practice</b> .....	982
<i>Polina Drozdova, Catia Cucchiari, Helmer Strik</i>	
<b>The Physiological Use of the Charismatic Voice in Political Speech</b> .....	987
<i>Rosario Signorello, Didier Demolin</i>	
<b>Crosslinguistic Corpus of Hesitation Phenomena: A Corpus for Investigating First and Second Language Speech Performance</b> .....	991
<i>Ralph L. Rose</i>	
<b>Real-Time Control of a 2D Animation Model of the Vocal Tract Using Optopalatography</b> .....	996
<i>Simon Preuss, Christiane Neuschaefer-Rube, Peter Birkholz</i>	
<b>The Influence of Accentuation and Polysyllabicity on Compensatory Shortening in German</b> .....	1001
<i>Jessica Siddins, Jonathan Harrington, Felicitas Kleber, Ulrich Reubold</i>	
<b>An Investigation of Vowel Epenthesis in Chinese Learners' Production of German Consonants</b> .....	1006
<i>Hongwei Ding, Rudiger Hoffmann</i>	
<b>On the Evaluation of Inversion Mapping Performance in the Acoustic Domain</b> .....	1011
<i>Korin Richmond, Zhen-Hua Ling, Junichi Yamagishi, Benigno Uria</i>	

## **POSTER SESSION 8: SPEECH SYNTHESIS II**

Chair: Olivier Rosac

<b>Probabilistic Speech <math>F_0</math> Contour Model Incorporating Statistical Vocabulary Model of Phrase-Accent Command Sequence</b> .....	1016
<i>Tatsuma Ishihara, Hirokazu Kameoka, Kota Yoshizato, Daisuke Saito, Shigeki Sagayama</i>	
<b>Reconstruction of Continuous Voiced Speech from Whispers</b> .....	1021
<i>Ian Vince McLoughlin, Jingjie Li, Yan Song</i>	
<b>Generating Fundamental Frequency Contours for Speech Synthesis in Yoruba</b> .....	1026
<i>Daniel R. van Niekerk, Etienne Barnard</i>	
<b>Real-Time Voice Conversion Using Artificial Neural Networks with Rectified Linear Units</b> .....	1031
<i>Elias Azarov, Maxim Vashkevich, Denis Likhachov, Alexander Petrovsky</i>	
<b>Generation of Fundamental Frequency Contours for Thai Speech Synthesis Using Tone Nucleus Model</b> .....	1036
<i>Oraphan Krityakien, Keikichi Hirose, Nobuaki Minematsu</i>	
<b>Unsupervised Speaker and Expression Factorization for Multi-Speaker Expressive Synthesis of Ebooks</b> .....	1041
<i>Langzhou Chen, Norbert Braunschweiler</i>	
<b>Which Resemblance is Useful to Predict Phrase Boundary Rise Labels for Japanese Expressive Text-to-Speech Synthesis, Numerically-Expressed Stylistic or Distribution-Based Semantic?</b> .....	1046
<i>Hideharu Nakajima, Hideyuki Mizuno, Osamu Yoshioka, Satoshi Takahashi</i>	
<b>A Targets-Based Superpositional Model of Fundamental Frequency Contours Applied to HMM-Based Speech Synthesis</b> .....	1051
<i>Jinju Ni, Yoshinori Shiga, Chiori Hori, Yutaka Kidawara</i>	

<b>An Investigation of Acoustic Features for Singing Voice Conversion Based on Perceptual Age</b> .....	1056
<i>Kazuhiro Kobayashi, Hironori Doi, Tomoki Toda, Tomoyasu Nakano, Masataka Goto, Graham Neubig, Sakriani Sakti, Satoshi Nakamura</i>	
<b>Effect of MPEG Audio Compression on HMM-Based Speech Synthesis</b> .....	1061
<i>Bajibabu Bollepalli, Tuomo Raitio, Paavo Alku</i>	
<b>Evaluation of a Singing Voice Conversion Method Based on Many-to-Many Eigenvoice Conversion</b> .....	1066
<i>Hironori Doi, Tomoki Toda, Tomoyasu Nakano, Masataka Goto, Satoshi Nakamura</i>	
<b>Statistical Nonparametric Speech Synthesis Using Sparse Gaussian Processes</b> .....	1071
<i>Tomoki Koriyama, Takashi Nose, Takao Kobayashi</i>	
<b>Hybrid Nearest-Neighbor/Cluster Adaptive Training for Rapid Speaker Adaptation in Statistical Speech Synthesis Systems</b> .....	1076
<i>Amir Mohammadi, Cenk Demiroglu</i>	
<b>Uniform Concatenative Excitation Model for Synthesising Speech Without Voiced/Unvoiced Classification</b> .....	1081
<i>Joao P. Cabral</i>	

## **POSTER SESSION 9: METADATA, EVALUATION AND RESOURCES**

*Chair: Florian Schiel*

<b>Efficient Speech Transcription Through Respeaking</b> .....	1086
<i>Matthias Sperber, Graham Neubig, Christian Fugen, Satoshi Nakamura, Alex Waibel</i>	
<b>Annotation and Classification of Political Advertisements</b> .....	1091
<i>Samuel Kim, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	
<b>Using Role Play for Collecting Question-Answer Pairs for Dialogue Agents</b> .....	1096
<i>Ryuichiro Higashinaka, Kohji Dohsaka, Hideki Isozaki</i>	
<b>Individual Differences of Emotional Expression in Speaker's Behavioral and Autonomic Responses</b> .....	1100
<i>Yoshiko Arimoto, Kazuo Okanoya</i>	
<b>Development and Validation of the Conversational Agents Scale (CAS)</b> .....	1105
<i>Ina Wechsung, Benjamin Weiss, Christine Kuhnel, Patrick Ehrenbrink, Sebastian Moller</i>	
<b>Motivational Feedback in Crowdsourcing: A Case Study in Speech Transcription</b> .....	1110
<i>G. Riccardi, A. Ghosh, S.A. Chowdhury, Ali Orkan Bayer</i>	
<b>The Sheffield Wargames Corpus</b> .....	1115
<i>Charles Fox, Yulan Liu, Erich Zwyssig, Thomas Hain</i>	
<b>Formalizing Expert Knowledge for Developing Accurate Speech Recognizers</b> .....	1120
<i>Anuj Kumar, Florian Metz, Wenyi Wang, Matthew Kam</i>	
<b>Analysis of Gaze and Speech Patterns in Three-Party Quiz Game Interaction</b> .....	1125
<i>Samer Al Moubayed, Jens Eklund, Joakim Gustafson</i>	
<b>Methodologies for the Evaluation of Speaker Diarization and Automatic Speech Recognition in the Presence of Overlapping Speech</b> .....	1130
<i>Olivier Galibert</i>	
<b>'Houston, We Have a Solution': Using NASA Apollo Program to Advance Speech and Language Processing Technology</b> .....	1134
<i>Abhijeet Sangwan, Lakshmish Kaushik, Chengzhu Yu, John H.L. Hansen, Douglas W. Oard</i>	

## **ORAL SESSION 16: SPEECH TECHNOLOGY FOR SPEECH DISORDERS**

*Chairs: Corinne Fredouille, Avignon and Elmar Noth*

<b>Performance of the MVOCA Silent Speech Interface Across Multiple Speakers</b> .....	1139
<i>Robin Hofe, Jie Bai, Lam A. Cheah, Stephen R. Ell, James M. Gilbert, Roger K. Moore, Phil D. Green</i>	
<b>Automatic Glottal Tracking from High-Speed Digital Images Using a Continuous Normalized Cross Correlation</b> .....	1143
<i>Gustavo Andrade-Miranda, Juan Ignacio Godino-Llorente</i>	
<b>Automatic Evaluation of Parkinson's Speech --- Acoustic, Prosodic and Voice Related Cues</b> .....	1148
<i>Tobias Bocklet, Stefan Steidl, Elmar Noth, Sabine Skodda</i>	
<b>Comparison of Approaches for an Efficient Phonetic Decoding</b> .....	1153
<i>Luiza Orosanu, Denis Jouvet</i>	
<b>Learning Speaker-Specific Pronunciations of Disordered Speech</b> .....	1158
<i>H. Christensen, Phil D. Green, Thomas Hain</i>	

<b>Adapting a Speech into Sign Language Translation System to a New Domain</b> .....	1163
<i>V. Lopez-Ludena, R. San-Segundo, C. Gonzalez-Morcillo, J.C. Lopez, E. Ferreiro</i>	

### **ORAL SESSION 17: SPEECH ANALYSIS II**

*Chairs: Paavo Alku, Helsinki and Abeer Alwan*

<b>Assessing the Intelligibility Impact of Vowel Space Expansion via Clear Speech-Inspired Frequency Warping</b> .....	1168
<i>Elizabeth Godoy, M. Koutsogiannaki, Yannis Stylianou</i>	
<b>Prediction of Intelligibility of Noisy and Time-Frequency Weighted Speech Based on Mutual Information Between Amplitude Envelopes</b> .....	1173
<i>Jesper Jensen, Cees H. Taal</i>	
<b>Frequency-Adaptive Post-Filtering for Intelligibility Enhancement of Narrowband Telephone Speech</b> .....	1178
<i>Emma Jokinen, Marko Takanen, Paavo Alku</i>	
<b>Comparative Investigation of Objective Speech Intelligibility Prediction Measures for Noise-Reduced Signals in Mandarin and Japanese</b> .....	1183
<i>Junfeng Li, Fei Chen, Masato Akagi, Yonghong Yan</i>	
<b>Monitoring the Effects of Temporal Clipping on VoIP Speech Quality</b> .....	1187
<i>Andrew Hines, Jan Skoglund, Anil Kokaram, Naomi Harte</i>	
<b>The Spectral Dynamics of Vowels in Mandarin Chinese</b> .....	1192
<i>Jiahong Yuan</i>	

### **ORAL SESSION 18: DISCRIMINATIVE TRAINING METHODS FOR LANGUAGE MODELING**

*Chairs: Hermann Ney, Murat Saraclar*

<b>CSLM --- A Modular Open-Source Continuous Space Language Modeling Toolkit</b> .....	1197
<i>Holger Schwenk</i>	
<b>Speed Up of Recurrent Neural Network Language Models with Sentence Independent Subsampling Stochastic Gradient Descent</b> .....	1202
<i>Yangyang Shi, Mei-Yuh Hwang, Kaisheng Yao, Martha Larson</i>	
<b>Improving Unsupervised Language Model Adaptation with Discriminative Data Filtering</b> .....	1207
<i>Shuangyu Chang, Michael Levit, Partha Parthasarathy, Benoit Dumoulin</i>	
<b>Lightly Supervised Training for Risk-Based Discriminative Language Models</b> .....	1212
<i>Akio Kobayashi, Takahiro Oku, Yuya Fujita, Shoei Sato</i>	
<b>Investigation of MT-Based ASR Confusion Models for Semi-Supervised Discriminative Language Modeling</b> .....	1217
<i>Erinc Dikici, Emily Prud'hommeaux, Brian Roark, Murat Saraclar</i>	
<b>Unsupervised Discriminative Language Modeling Using Error Rate Estimator</b> .....	1222
<i>Takanobu Oba, Atsunori Ogawa, Takaaki Hori, Hirokazu Masataki, Atsushi Nakamura</i>	

### **ORAL SESSION 19: ASR — ADAPTIVE TRAINING**

*Chairs: Jen-Tzung Chien, Taiwan and Jean-Luc Gauvain*

<b>A Region-Specific Feature-Space Transformation for Speaker Adaptation and Singularity Analysis of Jacobian Matrix</b> .....	1227
<i>Shakti P. Rath, Lukas Burget, Martin Karafiat, Ondrej Glembek, Jan Cernocky</i>	
<b>An Explicit Independence Constraint for Factorised Adaptation in Speech Recognition</b> .....	1232
<i>Y.-Q. Wang, M.J.F. Gales</i>	
<b>Asynchronous Factorisation of Speaker and Background with Feature Transforms in Speech Recognition</b> .....	1237
<i>Oscar Saz, Thomas Hain</i>	
<b>Cluster Adaptive Training with Factorized Decision Trees for Speech Recognition</b> .....	1242
<i>Kai Yu, Hainan Xu</i>	
<b>Rapid and Effective Speaker Adaptation of Convolutional Neural Network Based Models for Speech Recognition</b> .....	1247
<i>Ossama Abdel-Hamid, Hui Jiang</i>	
<b>Text-to-Speech Inspired Duration Modeling for Improved Whole-Word Acoustic Models</b> .....	1252
<i>Keith Kintzley, Aren Jansen, Hynek Hermansky</i>	

## **ORAL SESSION 20: SPEECH ACQUISITION AND DEVELOPMENT**

*Chairs: Fangfang Li, Mark Huckvale*

<b>Duration of Early Vocalisations</b> .....	1257
<i>Adele Gregory, Marija Tabain, Michael Robb</i>	
<b>Acoustic Development of Vowel Production in American English Children</b> .....	1262
<i>Jing Yang, Robert Allen Fox</i>	
<b>The Role of Intrinsic Motivations in Learning Sensorimotor Vocal Mappings: A Developmental Robotics Study</b> .....	1267
<i>Clement Moulin-Frier, Pierre-Yves Oudeyer</i>	
<b>Children's Timing and Repair Strategies for Communication in Adverse Listening Conditions</b> .....	1272
<i>Valerie Hazan, Michele Pettinato</i>	
<b>Speech Planning as an Index of Speech Motor Control Maturity</b> .....	1277
<i>Guillaume Barbier, Pascal Perrier, Lucie Menard, Yohan Payan, Mark K. Tiede, Joseph S. Perkell</i>	
<b>The Relationship Between Gender-Differentiated Productions of /s/ and Gender Role Behaviour in Young Children</b> .....	1282
<i>Melissa Kinsman, Fangfang Li</i>	

## **SPECIAL SESSION 3 (A & B): ARTICULATORY DATA ACQUISITION AND PROCESSING**

*Chairs: Slim Ouni, Nancy and Korin Richmond*

<b>Data-Driven Design of a Sentence List for an Articulatory Speech Corpus</b> .....	1286
<i>Jeffrey Berry, Luciano Fadiga</i>	
<b>Faster 3D Vocal Tract Real-Time MRI Using Constrained Reconstruction</b> .....	1291
<i>Yinghua Zhu, Asterios Toutios, Shrikanth Narayanan, Krishna Nayak</i>	
<b>Relevance-Weighted-Reconstruction of Articulatory Features in Deep-Neural-Network-Based Acoustic-to-Articulatory Mapping</b> .....	1296
<i>Claudia Canevari, Leonardo Badino, Luciano Fadiga, Giorgio Metta</i>	
<b>Word Frequency, Vowel Length and Vowel Quality in Speech Production: An EMA Study of the Importance of Experience</b> .....	1301
<i>Fabian Tomaschek, Martijn Wieling, Denis Arnold, R. Harald Baayen</i>	
<b>Towards a Systematic and Quantitative Analysis of Vocal Tract Data</b> .....	1306
<i>Samuel Silva, Antonio Teixeira, Catarina Oliveira, Paula Martins</i>	
<b>A Two-Step Technique for MRI Audio Enhancement Using Dictionary Learning and Wavelet Packet Analysis</b> .....	1311
<i>Colin Vaz, Vikram Ramanarayanan, Shrikanth Narayanan</i>	
<b>Electromagnetic Articulography with AG500 and AG501</b> .....	1315
<i>Massimo Stella, Antonio Stella, Francesco Sigona, Paolo Bernardini, Mirko Grimaldi, Barbara Gili Fivela</i>	
<b>Development and Implementation of Fiducial Markers for Vocal Tract MRI Imaging and Speech Articulatory Modelling</b> .....	1320
<i>Pierre Badin, Julian Andres Valdes Vargas, Arielle Koncki, Laurent Lamalle, Christophe Savariaux</i>	
<b>Functional Data Analysis of Tongue Articulation in Palatal Vowels: Gothenburg and Malmohus Swedish /i:, y:, u:/</b> .....	1325
<i>Susanne Schotz, Johan Frid, Lars Gustafsson, Anders Lofqvist</i>	
<b>SMASH: A Tool for Articulatory Data Processing and Analysis</b> .....	1330
<i>Jordan R. Green, Jun Wang, David L. Wilson</i>	

## **POSTER SESSION 10: TOPICS IN SPEECH PERCEPTION AND EMOTION**

*Chair: Angelika Braun*

<b>Emotion Recognition of Conversational Affective Speech Using Temporal Course Modeling</b> .....	1335
<i>Jen-Chun Lin, Chung-Hsien Wu, Wen-Li Wei</i>	
<b>The Role of Empathy in the Recognition of Vocal Emotions</b> .....	1340
<i>Rene Altrov, Hille Pajupuu, Jaan Pajupuu</i>	
<b>Electrophysiological Evidence for Benefits of Imitation During the Processing of Spoken Words Embedded in Sentential Contexts</b> .....	1344
<i>Angele Brunelliere, Sophie Dufour</i>	
<b>Compensatory Speech Response to Time-Scale Altered Auditory Feedback</b> .....	1349
<i>Rintaro Ogane, Masaaki Honda</i>	



<b>Bhattacharyya Distance Based Emotional Dissimilarity Measure in Multi-Dimensional Space for Emotion Classification</b> .....	1354
<i>Tin Lay Nwe, Trung Hieu Nguyen, Dilip Kumar Limbu</i>	
<b>On the Enhancement of Dereverberation Algorithms Based on a Perceptual Evaluation Criterion</b> .....	1359
<i>Thiago de M. Prego, Amaro A. de Lima, Sergio L. Netto</i>	
<b>Revisiting Pitch Slope and Height Effects on Perceived Duration</b> .....	1364
<i>Carlos Gussenhoven, Wencui Zhou</i>	
<b>Adaptation to Natural Fast Speech and Time-Compressed Speech in Children</b> .....	1369
<i>Helene Guiraud, Emmanuel Ferragne, Nathalie Bedoin, Veronique Boulenger</i>	
<b>Modeling Durational Incompressibility</b> .....	1374
<i>Andreas Windmann, Juraj Simko, Britta Wrede, Petra Wagner</i>	
<b>Perceived Prosodic Correlates of Smiled Speech in Spontaneous Data</b> .....	1379
<i>Caroline Emond, Lucie Menard, Marty Laforest</i>	
<b>Predicting Speech Quality Based on Interactivity and Delay</b> .....	1383
<i>Alexander Raake, Katrin Schoenenberg, Janto Skowronek, Sebastian Egger</i>	
<b>Perceptual, Acoustic and Electroglottographic Correlates of 3 Aggressive Attitudes in French: A Pilot Study</b> .....	1388
<i>Charlotte Kouklia, Nicolas Audibert</i>	

**POSTER SESSION 11: DISCOURSE AND MACHINE LEARNING, PARALINGUISTIC AND NONLINGUISTIC CUES**

*Chair: Martijn Goudbeek*

<b>Theme Identification in Telephone Service Conversations Using Quaternions of Speech Features</b> .....	1393
<i>Mohamed Morchid, Georges Linares, Marc El-Beze, Renato De Mori</i>	
<b>Detection of Laughter in Children's Speech Using Spectral and Prosodic Acoustic Features</b> .....	1398
<i>Hrishikesh Rao, Jonathan C. Kim, Agata Rozga, Mark A. Clements</i>	
<b>Classification of Cooperative and Competitive Overlaps in Speech Using Cues from the Context, Overlapper, and Overlappee</b> .....	1403
<i>Khiet P. Truong</i>	
<b>Annotation and Detection of Conflict Escalation in Political Debates</b> .....	1408
<i>Samuel Kim, Fabio Valente, Alessandro Vinciarelli</i>	
<b>Machine Learning of Probabilistic Phonological Pronunciation Rules from the Italian CLIPS Corpus</b> .....	1413
<i>Florian Schiel, Mary Stevens, Uwe D. Reichel, Francesco Cutugno</i>	
<b>Human Perception of Alcoholic Intoxication in Speech</b> .....	1418
<i>Barbara Baumeister, Florian Schiel</i>	
<b>Phonetic Manifestation and Influence of Zero Anaphora in Chinese Reading Texts</b> .....	1423
<i>Luying Hou, Yuan Jia, Aijun Li</i>	
<b>Diacritics Restoration for Arabic Dialect Texts</b> .....	1428
<i>S. Harrat, M. Abbas, K. Meftouh, K. Smaili</i>	
<b>Effects of Talk-Spurt Silence Boundary Thresholds on Distribution of Gaps and Overlaps</b> .....	1433
<i>Marcin Wlodarczyk, Petra Wagner</i>	
<b>Final Lengthening in Russian: A Corpus-Based Study</b> .....	1437
<i>Tatiana Kachkovskaia, Nina Volskaya, Pavel Skrelin</i>	
<b>From Segmentation Bootstrapping to Transcription-to-Word Conversion</b> .....	1442
<i>Uwe D. Reichel</i>	
<b>Manual and Automatic Tone Annotation: The Case of an Endangered Language from North Vietnam “Mo Piu”</b> .....	1447
<i>Genevieve Caelen-Haumont, Katarina Bartkova</i>	
<b>Non-Canonical Syntactic Structures in Discourse: Tonality, Tonicity and Tones in English (Semi-)Spontaneous Speech</b> .....	1452
<i>Laetitia Leonarduzzi, Sophie Herment</i>	
<b>Prediction of Strategy and Outcome as Negotiation Unfolds by Using Basic Verbal and Behavioral Features</b> .....	1457
<i>Elnaz Nouri, Sunghyun Park, Stefan Scherer, Jonathan Gratch, Peter Carnevale, Louis-Philippe Morency, David Traum</i>	

## **POSTER SESSION 12: LANGUAGE IDENTIFICATION, SPEAKER DIARIZATION**

Chair: Pietro Laface

<b>Unsupervised Naming of Speakers in Broadcast TV: Using Written Names, Pronounced Names or Both?</b> .....	1461
<i>Johann Poignant, Laurent Besacier, Viet Bac Le, Sophie Rosset, Georges Quenot</i>	
<b>Integer Linear Programming for Speaker Diarization and Cross-Modal Identification in TV Broadcast</b> .....	1466
<i>Herve Bredin, Johann Poignant</i>	
<b>Native Accent Classification via I-Vectors and Speaker Compensation Fusion</b> .....	1471
<i>Andrea DeMarco, Stephen J. Cox</i>	
<b>An Open-Source State-of-the-Art Toolbox for Broadcast News Diarization</b> .....	1476
<i>Mickael Rouvier, Gregor Dupuy, Paul Gay, Elie Khoury, Teva Merlin, Sylvain Meignier</i>	
<b>Audio Event Classification Using Deep Neural Networks</b> .....	1481
<i>Zvi Kons, Orith Toledo-Ronen</i>	
<b>Code-Switching Event Detection Based on Delta-BIC Using Phonetic Eigenvoice Models</b> .....	1486
<i>Wei-Bin Liang, Chung-Hsien Wu, Chun-Shan Hsu</i>	
<b>Automatic Estimation of Dialect Mixing Ratio for Dialect Speech Recognition</b> .....	1491
<i>Naoki Hirayama, Koichiro Yoshino, Katsutoshi Itoyama, Shinsuke Mori, Hiroshi G. Okuno</i>	
<b>The Albayzin 2012 Language Recognition Evaluation</b> .....	1496
<i>Luis Javier Rodriguez-Fuentes, Niko Brummer, Mikel Penagarikano, Amparo Varona, German Bordel, Mireia Diez</i>	
<b>TRAP Language Identification System for RATS Phase II Evaluation</b> .....	1501
<i>Kyu J. Han, Sriram Ganapathy, Ming Li, Mohamed K. Omar, Shrikanth Narayanan</i>	
<b>Improving Language Identification Robustness to Highly Channel-Degraded Speech Through Multiple System Fusion</b> .....	1506
<i>Aaron Lawson, Mitchell McLaren, Yun Lei, Vikramjit Mitra, Nicolas Scheffer, Luciana Ferrer, Martin Graciarena</i>	

## **ORAL SESSION 21: METADATA, EVALUATION AND RESOURCES**

Chairs: Maxine Eskenazi, Ilya Oparin

<b>Annotation Errors Detection in TTS Corpora</b> .....	1510
<i>Jindrich Matousek, Daniel Tihelka</i>	
<b>Technique for Automatic Sentence Level Alignment of Long Speech and Transcripts</b> .....	1515
<i>Imran Ahmed, Sunil Kumar Koppurapu</i>	
<b>Text-to-Speech Alignment of Long Recordings Using Universal Phone Models</b> .....	1519
<i>Sarah Hoffmann, Beat Pfister</i>	
<b>Lightly Supervised Discriminative Training of Grapheme Models for Improved Sentence-Level Alignment of Speech and Text Data</b> .....	1524
<i>Adriana Stan, Peter Bell, Junichi Yamagishi, Simon King</i>	
<b>Automatic Social Role Recognition in Professional Meetings Using Conditional Random Fields</b> .....	1529
<i>Ashtosh Sapru, Herve Bourlard</i>	
<b>Same Same But Different --- An Acoustical Comparison of the Automatic Segmentation of High Quality and Mobile Telephone Speech</b> .....	1534
<i>Christoph Draxler, Hanna S. Feiser</i>	

## **ORAL SESSION 22: SPEECH SYNTHESIS — PROSODY AND EMOTION**

Chairs: Nick Campbell, Emily Mower Provost

<b>Multi-Centroidal Duration Generation Algorithm for HMM-Based TTS</b> .....	1539
<i>Yongguo Kang, Jian Li, Yan Deng, Miaomiao Wang</i>	
<b>Analysis and Synthesis of Shouted Speech</b> .....	1543
<i>Tuomo Raitio, Antti Suni, Jouni Pohjalainen, Manu Airaksinen, Martti Vainio, Paavo Alku</i>	
<b>Robust Estimation of Multiple-Regression HMM Parameters for Dimension-Based Expressive Dialogue Speech Synthesis</b> .....	1548
<i>Tomohiro Nagata, Hiroki Mori, Takashi Nose</i>	
<b>A New Prosody Annotation Protocol for Live Sports Commentaries</b> .....	1553
<i>Sandrine Brognaux, Benjamin Picart, Thomas Drugman</i>	

## VOLUME 3

**Unsupervised Prominence Prediction for Speech Synthesis**..... 1558  
*Mahmoosh Mehrabani, Taniya Mishra, Alistair Conkie*

**Expressive Speech Synthesis in MARY TTS Using Audiobook Data and EmotionML** ..... 1563  
*Marcela Charfuelan, Ingmar Steiner*

### **ORAL SESSION 23: SPOKEN LANGUAGE INFORMATION RETRIEVAL**

*Chairs: Giuseppe Di Fabbrizio, Haizhou Li*

**Using Dialog-Activity Similarity for Spoken Information Retrieval**..... 1568  
*Nigel G. Ward, Steven D. Werner*

**A Hybrid HMM/DNN Approach to Keyword Spotting of Short Words**..... 1573  
*I-Fan Chen, Chin-Hui Lee*

**Leveraging Locality for Topic Identification of Conversational Speech**..... 1578  
*Jonathan Wintrode*

**Person Name Spotting by Combining Acoustic Matching and LDA Topic Models**..... 1583  
*Gregory Senay, Benjamin Bigot, Richard Dufour, Georges Linares, Corinne Fredouille*

**Using Phonological Phrase Segmentation to Improve Automatic Keyword Spotting for the Highly Agglutinating Hungarian Language**..... 1588  
*Gyorgy Szaszak, Andras Beke*

**Leveraging Knowledge Graphs for Web-Scale Unsupervised Semantic Parsing**..... 1593  
*Larry Heck, Dilek Hakkani-Tur, Gokhan Tur*

### **ORAL SESSION 24: SPEAKER RECOGNITION**

*Chairs: Tomi Kinnunen, Joensuu and Nicolas Scheffer*

**Fast and Memory Effective I-Vector Extraction Using a Factorized Sub-Space** ..... 1598  
*Sandro Cumani, Pietro Laface*

**Effective Estimation of a Multi-Session Speaker Model Using Information on Signal Parameters** ..... 1603  
*Konstantin Simonchik, Andrey Shulipa, Timur Pekhovsky*

**Automatic Regularization of Cross-Entropy Cost for Speaker Recognition Fusion**..... 1608  
*Ville Hautamaki, Kong Aik Lee, David A. van Leeuwen, R. Saeidi, Anthony Larcher, Tomi Kinnunen, Taufiq Hasan, Seyed Omid Sadjadi, Gang Liu, Hynek Boril, John H.L. Hansen, Benoit Fauve*

**Speaker Verification Based on Fusion of Acoustic and Articulatory Information**..... 1613  
*Ming Li, Jangwon Kim, Prasanta Kumar Ghosh, Vikram Ramanarayanan, Shrikanth Narayanan*

**The Distribution of Calibrated Likelihood-Ratios in Speaker Recognition**..... 1618  
*David A. van Leeuwen, Niko Brummer*

**Eigenageing Compensation for Speaker Verification**..... 1623  
*Finnian Kelly, Niko Brummer, Naomi Harte*

### **ORAL SESSION 25: MULTIMODAL SPEECH PERCEPTION**

*Chairs: Chris Davis, Sydney and Lucie Menard*

**Effects of Mouth-Only and Whole-Face Displays on Audio-Visual Speech Perception in Noise: Is the Vision of a Talker's Full Face Truly the Most Efficient Solution?** ..... 1628  
*Grozdana Erjavec, Denis Legros*

**Acoustic and Visual Phonetic Features in the McGurk Effect --- An Audiovisual Speech Illusion**..... 1633  
*Kaisa Tiippana, Mikko Tiainen, Lari Vainio, Martti Vainio*

**The Effect of Visual Speech Timing and Form Cues on the Processing of Speech and Nonspeech**..... 1638  
*Chris Davis, Jeusun Kim*

**Effect of Context, Rebinding and Noise, on Audiovisual Speech Fusion**..... 1642  
*Ganesh Attigodu Chandrashekar, Frederic Berthommier, Olha Nahorna, Jean-Luc Schwartz*

**Social Face to Face Communication --- American English Attitudinal Prosody**..... 1647  
*Albert Rilliard, Donna Erickson, Takaaki Shochi, Joao Antonio de Moraes*

**Adaptation of Respiratory Patterns in Collaborative Reading**..... 1652  
*Gerard Bailly, Amelie Rochet-Capellan, Coriandre Vilain*

## **POSTER SESSION 13: SPEECH ANALYSIS**

Chair: Tom Quatier

<b>A Comparative Study of Glottal Open Quotient Estimation Techniques</b> .....	1657
<i>John Kane, Stefan Scherer, Louis-Philippe Morency, Christer Gobl</i>	
<b>Estimation of Multiple-Branch Vocal Tract Models: The Influence of Prior Assumptions</b> .....	1662
<i>Christian H. Kasess, Wolfgang Kreuzer</i>	
<b>Detecting Overlapping Speech with Long Short-Term Memory Recurrent Neural Networks</b> .....	1667
<i>Jurgen T. Geiger, Florian Eyben, Bjorn Schuller, Gerhard Rigoll</i>	
<b>Evaluation of Fundamental Validity in Applying AR-HMM with Automatic Topology Generation to Pathology Voice Analysis</b> .....	1672
<i>Akira Sasou</i>	
<b>Significance of Instants of Significant Excitation for Source Modeling</b> .....	1676
<i>Nagaraj Adiga, S.R.M. Prasanna</i>	
<b>Significance of Variable Height-Bandwidth Group Delay Filters in the Spectral Reconstruction of Speech</b> .....	1681
<i>Devanshu Arya, Anant Raj, Rajesh M. Hegde</i>	
<b>Nonlinear Prediction of Speech Signal Using Volterra-Wiener Series</b> .....	1686
<i>Hemant A. Patil, Tanvina B. Patel</i>	
<b>Evaluation of Speech-Based Protocol for Detection of Early-Stage Dementia</b> .....	1691
<i>Aharon Satt, Alexander Sorin, Orith Toledo-Ronen, Oren Barkan, Ioannis Kompatsiaris, Athina Kokonozi, Magda Tsolaki</i>	
<b>Instantaneous Harmonic Representation of Speech Using Multicomponent Sinusoidal Excitation</b> .....	1696
<i>Elias Azarov, Maxim Vashkevich, Alexander Petrovsky</i>	
<b>A Quantitative Comparison of Glottal Closure Instant Estimation Algorithms on a Large Variety of Singing Sounds</b> .....	1701
<i>Onur Babacan, Thomas Drugman, Nicolas d'Alessandro, Nathalie Henrich, Thierry Dutoit</i>	
<b>Automatic Gender Recognition in Normal and Pathological Speech</b> .....	1706
<i>J.A. Gomez-Garcia, Juan Ignacio Godino-Llorente, G. Castellanos-Dominguez</i>	
<b>Unsupervised Vocal-Tract Length Estimation Through Model-Based Acoustic-to-Articulatory Inversion</b> .....	1711
<i>Shanqing Cai, H. Timothy Bunnell, Rupal Patel</i>	
<b>Model Order Estimation Using Bayesian NMF for Discovering Phone Patterns in Spoken Utterances</b> .....	1716
<i>Sayeh Mirzaei, Hugo Van hamme, Yaser Norouzi</i>	

## **POSTER SESSION 14: ASR — FEATURE EXTRACTION**

Chair: Long Nguyen

<b>Convolutional Deep Rectifier Neural Nets for Phone Recognition</b> .....	1721
<i>Laszlo Toth</i>	
<b>Pitch Synchronous Spectral Analysis for a Pitch Dependent Recognition of Voiced Phonemes - PISAR</b> .....	1726
<i>Hans-Gunter Hirsch</i>	
<b>New Parameters for Automatic Speech Recognition Based on the Mammalian Cochlea Model Using Resonance Analysis</b> .....	1731
<i>Jose Luis Oropeza Rodriguez</i>	
<b>Using an Autoencoder with Deformable Templates to Discover Features for Automated Speech Recognition</b> .....	1736
<i>Navdeep Jaitly, Geoffrey E. Hinton</i>	
<b>Speaking Rate Normalization with Lattice-Based Context-Dependent Phoneme Duration Modeling for Personalized Speech Recognizers on Mobile Devices</b> .....	1740
<i>Ching-Feng Yeh, Hung-yi Lee, Lin-shan Lee</i>	
<b>Subspace Models for Bottleneck Features</b> .....	1745
<i>Jun Qi, Dong Wang, Javier Tejedor</i>	
<b>Bottleneck Features Based on Gammatone Frequency Cepstral Coefficients</b> .....	1750
<i>Jun Qi, Dong Wang, Ji Xu, Javier Tejedor</i>	
<b>Cross-Entropy vs. Squared Error Training: A Theoretical and Experimental Comparison</b> .....	1755
<i>Pavel Golik, Patrick Doetsch, Hermann Ney</i>	

<b>Acoustic Features for Detection of Phonemic Aspiration in Voiced Plosives</b> .....	1760
<i>Vaishali Patil, Preeti Rao</i>	
<b>Estimating Phoneme Class Conditional Probabilities from Raw Speech Signal Using Convolutional Neural Networks</b> .....	1765
<i>Dimitri Palaz, Ronan Collobert, Mathew Magimai-Doss</i>	
<b>Hierarchical Models Based on a Continuous Acoustic Space to Identify Phonological Features</b> .....	1770
<i>Javier Mikel Olaso, Maria Ines Torres</i>	
<b>Locality Sensitive Hashing for Fast Computation of Correlational Manifold Learning Based Feature Space Transformations</b> .....	1775
<i>Vikrant Singh Tomar, Richard C. Rose</i>	
<b>Evaluating Speech Features with the Minimal-Pair ABX Task: Analysis of the Classical MFC/PLP Pipeline</b> .....	1780
<i>Thomas Schatz, Vijayaditya Peddinti, Francis Bach, Aren Jansen, Hynek Hermansky, Emmanuel Dupoux</i>	

## **POSTER SESSION 15: ASR — PRONUNCIATION, PROSODIC AND NEW PARADIGMS**

*Chair: Tanja Schultz*

<b>Knowledge Integration for Improving Performance in LVCSR</b> .....	1785
<i>Chen-Yu Chiang, Sabato Marco Siniscalchi, Sin-Horng Chen, Chin-Hui Lee</i>	
<b>Inter-Speaker Variability in Audio-Visual Classification of Word Prominence</b> .....	1790
<i>Martin Heckmann</i>	
<b>Parameter Clustering for Temporally Varying Weight Regression for Automatic Speech Recognition</b> .....	1795
<i>Shilin Liu, Khe Chai Sim</i>	
<b>Phone Duration Modeling Using Clustering of Rich Contexts</b> .....	1800
<i>Tanel Alumae, Rena Nemoto</i>	
<b>Human Mouth State Detection Using Low Frequency Ultrasound</b> .....	1805
<i>Farzaneh Ahmadi, Mousa Ahmadi, Ian Vince McLoughlin</i>	
<b>Lexical Stress Detection for L2 English Speech Using Deep Belief Networks</b> .....	1810
<i>Kun Li, Xiaojun Qian, Shiyin Kang, Helen Meng</i>	
<b>MLP-HMM Two-Stage Unsupervised Training for Low-Resource Languages on Conversational Telephone Speech Recognition</b> .....	1815
<i>Yanmin Qian, Jia Liu</i>	
<b>Failure Transitions for Joint n-Gram Models and G2P Conversion</b> .....	1820
<i>Josef R. Novak, Nobuaki Minematsu, Keikichi Hirose</i>	
<b>Generative Modeling of Speech <math>F_0</math> Contours</b> .....	1825
<i>Hirokazu Kameoka, Kota Yoshizato, Tatsuma Ishihara, Yasunori Ohishi, Kunio Kashino, Shigeki Sagayama</i>	
<b>G2P Variant Prediction Techniques for ASR and STD</b> .....	1830
<i>Marelle H. Davel, Charl van Heerden, Etienne Barnard</i>	
<b>Rhythm Analysis of Second-Language Speech Through Low-Frequency Auditory Features</b> .....	1835
<i>Jing Zheng, Joseph Tepperman</i>	
<b>Graph-Based Semi-Supervised Learning for Phone and Segment Classification</b> .....	1839
<i>Yuzong Liu, Katrin Kirchhoff</i>	
<b>Selective Use of Gaze Information to Improve ASR Performance in Noisy Environments by Cache-Based Class Language Model Adaptation</b> .....	1843
<i>Ao Shen, Neil Cooke, Martin Russell</i>	
<b>Deep Segmental Neural Networks for Speech Recognition</b> .....	1848
<i>Ossama Abdel-Hamid, Li Deng, Dong Yu, Hui Jiang</i>	
<b>Quantifying Cross-Linguistic Variation in Grapheme-to-Phoneme Mapping</b> .....	1853
<i>Martine Coene, Annemiek Hammer, Wojtek Kowalczyk, Louis ten Bosch, Bart Vaerenberg, Paul J. Govaerts</i>	

## **SHOW & TELL-2: SHOW AND TELL SESSION 2**

*Chairs: Laurence Devillers, Fabrice Lefevre*

<b>The Speech Recognition Virtual Kitchen</b> .....	1857
<i>Florian Metz, Eric Fosler-Lussier, Rebecca Bates</i>	
<b>Multilingual Web Conferencing Using Speech-to-Speech Translation</b> .....	1860
<i>John Chen, Shufei Wen, Vivek Kumar Rangarajan Sridhar, Srinivas Bangalore</i>	
<b>ROCme! Software for the Recording and Management of Speech Corpora</b> .....	1863
<i>Emmanuel Ferragne, Sebastien Flavie, Christian Fressard</i>	

<b>Voice Search in Mobile Applications with the Rootvole Framework</b> .....	1865
<i>Felix Burkhardt</i>	
<b>On-Line Audio Dilation for Human Interaction</b> .....	1868
<i>John S. Novak III, Jason Archer, Valeriy Shafiro, Robert Kenyon, Jason Leigh</i>	
<b>Phase-Aware Single-Channel Speech Enhancement</b> .....	1871
<i>Pejman Mowlae, Mario Kaoru Watanabe, R. Saeidi</i>	
<b>A Free Online Accent and Intonation Dictionary for Teachers and Learners of Japanese</b> .....	1874
<i>Hiroko Hirano, Ibuki Nakamura, Nobuaki Minematsu, Masayuki Suzuki, Chieko Nakagawa, Noriko Nakamura, Yukinori Tagawa, Keikichi Hirose, Hiroya Hashimoto</i>	
<b>Reactive Accent Interpolation Through an Interactive Map Application</b> .....	1876
<i>Maria Astrinaki, Junichi Yamagishi, Simon King, Nicolas d'Alessandro, Thierry Dutoit</i>	
<b>A Non-Experts User Interface for Obtaining Automatic Diagnostic Spelling Evaluations for Learners of the German Writing System</b> .....	1878
<i>Kay Berkling</i>	

### **ORAL SESSION 26: DIALOG SYSTEMS**

*Chairs: Gary Geunbae Lee, Pohang and Sophie Rosset*

<b>Estimation of Interest and Comprehension Level of Audience Through Multi-Modal Behaviors in Poster Conversations</b> .....	1881
<i>Tatsuya Kawahara, Soichiro Hayashi, Katsuya Takanashi</i>	
<b>A New DNN-Based High Quality Pronunciation Evaluation for Computer-Aided Language Learning (CALL)</b> .....	1885
<i>Wenping Hu, Yao Qian, Frank K. Soong</i>	
<b>A Multi-Domain Dialog System to Integrate Heterogeneous Spoken Dialog Systems</b> .....	1890
<i>Joaquín Planells, Lluís-F. Hurtado, Encarna Segarra, Emilio Sanchis</i>	
<b>Development and Evaluation of Spoken Dialog Systems with One or Two Agents</b> .....	1895
<i>Yuki Todo, Ryota Nishimura, Kazumasa Yamamoto, Seiichi Nakagawa</i>	
<b>User Feedback in Human-Robot Interaction: Prosody, Gaze and Timing</b> .....	1900
<i>Gabriel Skantze, Catharine Oertel, Anna Hjalmarsson</i>	
<b>KPCatcher - A Keyphrase Extraction System for Enterprise Videos</b> .....	1905
<i>Yongxin Taylor Xi, Matthias Paulik, Venkata Ramana Gadde, Ananth Sankar</i>	

### **ORAL SESSION 27: SPEECH ANALYSIS III**

*Chairs: Hideki Kawahara, Yannis Stylianou*

<b>Pitch-Gesture Modeling Using Subband Autocorrelation Change Detection</b> .....	1910
<i>Malcolm Slaney, Elizabeth Shriberg, Jui-Ting Huang</i>	
<b>Analysis of Emotional Speech at Subsegmental Level</b> .....	1915
<i>P. Gangamohan, Sudarsana Reddy Kadiri, B. Yegnanarayana</i>	
<b>Periodicity Extraction for Voiced Sounds with Multiple Periodicity</b> .....	1920
<i>Masanori Morise, Hideki Kawahara, Kenji Ozawa</i>	
<b>Modelling and Estimation of the Fundamental Frequency of Speech Using a Hidden Markov Model</b> .....	1925
<i>John H. Taylor, Ben Milner</i>	
<b>Extended Weighted Linear Prediction Using the Autocorrelation Snapshot --- A Robust Speech Analysis Method and its Application to Recognition of Vocal Emotions</b> .....	1930
<i>Jouni Pohjalainen, Paavo Alku</i>	
<b>Improving the Accuracy and the Robustness of Harmonic Model for Pitch Estimation</b> .....	1935
<i>Meysam Asgari, Izhak Shafran</i>	

### **ORAL SESSION 28: ASR — PRONUNCIATION VARIANTS AND MODELING**

*Chairs: Frederic Bechet, Eric Fosler-Lussier*

<b>Discriminative Pronunciation Modeling Based on Minimum Phone Error Training</b> .....	1940
<i>Meixu Song, Qingqing Zhang, Jieli Pan, Yonghong Yan</i>	
<b>Grapheme-to-Phoneme Conversion Based on Adaptive Regularization of Weight Vectors</b> .....	1945
<i>Keigo Kubo, Sakriani Sakti, Graham Neubig, Tomoki Toda, Satoshi Nakamura</i>	
<b>An Efficient Method to Estimate Pronunciation from Multiple Utterances</b> .....	1950
<i>Tofigh Naghibi, Sarah Hoffmann, Beat Pfister</i>	

<b>Category-Based Phoneme-to-Grapheme Transliteration</b> .....	1955
<i>Willem D. Basson, Marelle H. Davel</i>	
<b>Discriminative Training of WFST Factors with Application to Pronunciation Modeling</b> .....	1960
<i>Preethi Jyothi, Eric Fosler-Lussier, Karen Livescu</i>	
<b>Discriminative Training of a Phoneme Confusion Model for a Dynamic Lexicon in ASR</b> .....	1965
<i>Penny Karanasou, Francois Yvon, Thomas Lavergne, Lori Lamel</i>	

### **ORAL SESSION 29: SPEAKER RECOGNITION EVALUATION**

*Chairs: Niko Brummer, Patrick Kenny*

<b>The 2012 NIST Speaker Recognition Evaluation</b> .....	1970
<i>Craig S. Greenberg, Vincent M. Stanford, Alvin F. Martin, Meghana Yadagiri, George R. Doddington, John J. Godfrey, Jaime Hernandez-Cordero</i>	
<b>Likelihood-Ratio Calibration Using Prior-Weighted Proper Scoring Rules</b> .....	1975
<i>Niko Brummer, George R. Doddington</i>	
<b>A Noise-Robust System for NIST 2012 Speaker Recognition Evaluation</b> .....	1980
<i>Luciana Ferrer, Mitchell McLaren, Nicolas Scheffer, Yun Lei, Martin Graciarena, Vikramjit Mitra</i>	
<b>I4U Submission to NIST SRE 2012: A Large-Scale Collaborative Effort for Noise-Robust Speaker Verification</b> .....	1985
<i>R. Saeidi</i>	
<b>Improved Unsupervised NAP Training Dataset Design for Speaker Recognition</b> .....	1990
<i>Hanwu Sun, Bin Ma</i>	
<b>Nuance -- Politecnico di Torino's 2012 NIST Speaker Recognition Evaluation System</b> .....	1995
<i>Daniele Colibro, Claudio Vair, Kevin Farrell, Nir Krause, Gennady Karvitsky, Sandro Cumani, Pietro Laface</i>	

### **ORAL SESSION 30: PHYSIOLOGY AND MODELS OF SPEECH PRODUCTION**

*Chairs: Masaaki Honda, Amelie Rochet Capellan*

<b>A Perceptually and Physiologically Motivated Voice Source Model</b> .....	2000
<i>Gang Chen, Marc Garellek, Jody Kreiman, Bruce R. Gerratt, Abeer Alwan</i>	
<b>Stable Articulatory Tasks and Their Variable Formation: Tamil Retroflex Consonants</b> .....	2005
<i>Caitlin Smith, Michael Proctor, Khalil Iskarous, Louis Goldstein, Shrikanth Narayanan</i>	
<b>Articulatory Settings Facilitate Mechanically Advantageous Motor Control of Vocal Tract Articulators</b> .....	2009
<i>Vikram Ramanarayanan, Adam Lammert, Louis Goldstein, Shrikanth Narayanan</i>	
<b>The Interplay of Linguistic Structure and Breathing in German Spontaneous Speech</b> .....	2013
<i>Amelie Rochet-Capellan, Susanne Fuchs</i>	
<b>Physical Models of the Vocal Tract with a Flapping Tongue for Flap and Liquid Sounds</b> .....	2018
<i>Takayuki Arai</i>	
<b>Articulatory Copy Synthesis from Cine X-Ray Films</b> .....	2023
<i>Yves Laprie, Matthieu Loosvelt, Shinji Maeda, Rudolph Sock, Fabrice Hirsch</i>	

### **SPECIAL SESSION 4: SPEECH SCIENCE IN END-USER APPLICATIONS**

*Chairs: Felix Burkhardt, Berlin and Juergen Schroeter*

<b>Large-Scale Personal Assistant Technology Deployment: The Siri Experience</b> .....	2028
<i>Jerome R. Bellegarda</i>	
<b>Evaluating an Adaptive Dialog System for the Public</b> .....	2033
<i>Benjamin Weiss, Simon Willkomm, Sebastian Moller</i>	
<b>Self-Taught Assistive Vocal Interfaces: An Overview of the ALADIN Project</b> .....	2038
<i>Jort F. Gemmeke, Bart Ons, Netsanet Tessema, Hugo Van hamme, Janneke van de Loo, Guy De Pauw, Walter Daelemans, Jonathan Huyghe, Jan Derboven, Lode Vuegen, Bert Van Den Broeck, Peter Karsmakers, Bart Vanrumste</i>	
<b>Affect Recognition in Real-Life Acoustic Conditions - A New Perspective on Feature Selection</b> .....	2043
<i>Florian Eyben, Felix Weninger, Bjorn Schuller</i>	
<b>A Distributed System for Recognizing Home Automation Commands and Distress Calls in the Italian Language</b> .....	2048
<i>Emanuele Principi, Stefano Squartini, Francesco Piazza, Danilo Fuselli, Maurizio Bonifazi</i>	
<b>Probabilistic Trainable Segmenter for Call Center Audio Using Multiple Features</b> .....	2053
<i>Nina Zinovieva, Xiaodan Zhuang, Pat Peterson, Joe Alwan, Rohit Prasad</i>	

<b>Voice Search in Mobile Applications and the Use of Linked Open Data</b> .....	2058
<i>Felix Burkhardt, Hans Ulrich Nageli</i>	
<b>Evaluation of a Real-Time Voice Order Recognition System from Multiple Audio Channels in a Home</b> .....	2061
<i>Michel Vacher, Benjamin Lecouteux, Dan Istrate, Thierry Joubert, Francois Portet, Mohamed Sehili, Pedro Chahuara</i>	
<b>In-Home Detection of Distress Calls: The Case of Aged Users</b> .....	2064
<i>Frederic Aman, Michel Vacher, Solange Rossato, Francois Portet</i>	
<b>Data Driven Methods for Utterance Semantic Tagging</b> .....	2067
<i>Ding Liu, Anthea Cheung, Anna Margolis, Patrick Redmond, Jun-won Suh, Chao Wang</i>	
<b>The AT&amp;T Speech API: A Study on Practical Challenges for Customized Speech to Text Service</b> .....	2070
<i>E. Gouvea, A. Moreno-Daniel, A. Reddy, R. Chengalvarayan, D. Thomson, A. Ljolje</i>	
<b>In-Vehicle Destination Entry by Voice: Practical Aspects</b> .....	2073
<i>Bart D'hoore, Alfred Wiesen</i>	

## **POSTER SESSION 16: PERCEPTION OF NON NATIVE SOUNDS**

*Chair: Ricardo Gutierrez-Osuna*

<b>Intelligibility at a Multilingual Cocktail Party: Effect of Concurrent Language Knowledge</b> .....	2076
<i>Aurore Gautreau, Michel Hoen, Fanny Meunier</i>	
<b>Regional Accents Affect Speech Intelligibility in a Multitalker Environment</b> .....	2080
<i>Ewa Jacewicz, Robert Allen Fox</i>	
<b>Perception of English Minimal Pairs in Noise by Japanese Listeners: Does Clear Speech for L2 Listeners Help?</b> .....	2085
<i>Shinichi Tokuma, Won Tokuma</i>	
<b>Salento Italian Listeners' Perception of American English Vowels</b> .....	2090
<i>Bianca Sisinni, Paola Escudero, Mirko Grimaldi</i>	
<b>TP 3.1 Software: A Tool for Designing Audio, Visual, and Audiovisual Perceptual Training Tasks and Perception Tests</b> .....	2094
<i>Andreia Schurt Rauber, Anabela Rato, Denise Cristina Kluge, Giane Rodrigues dos Santos</i>	
<b>Effect of Linguistic Masker on the Intelligibility of Mandarin Sentences</b> .....	2098
<i>Fei Chen, Junfeng Li, Lena L.N. Wong, Yonghong Yan</i>	
<b>The Learning and Generalization of Contrasts Consistent or Inconsistent with Native Biases</b> .....	2102
<i>Kyuwon Moon, Meghan Sumner</i>	
<b>L2 English Learners' Recognition of Words Spoken in Familiar versus Unfamiliar English Accents</b> .....	2107
<i>Jia Ying, Jason A. Shaw, Catherine T. Best</i>	
<b>The Effects of Perceptual and/or Productive Training on the Perception and Production of English Vowels /i/and /i:/ by Cantonese ESL Learners</b> .....	2112
<i>Janice Wing Sze Wong</i>	
<b>On the Role of L1 Speech Production in L2 Perception: Evidence from Spanish Learners of French</b> .....	2117
<i>Natalia Kartushina, Ulrich Hans Frauenfelder</i>	
<b>Looking for Lexical Feedback Effects in /t/→/k/ Repairs</b> .....	2122
<i>Pierre Halle, Natalia Kartushina, Juan Segui, Ulrich Hans Frauenfelder</i>	
<b>Recognizing Words Across Regional Accents: The Role of Perceptual Assimilation in Lexical Competition</b> .....	2127
<i>Catherine T. Best, Jason A. Shaw, Elizabeth Clancy</i>	

## **POSTER SESSION 17: SPEECH DISORDERS — DATA AND METHODOLOGY**

*Chair: Heidi Christensen*

<b>Dysarthria Intelligibility Assessment in a Factor Analysis Total Variability Space</b> .....	2132
<i>David Martinez, Phil D. Green, H. Christensen</i>	
<b>Perceptual Interference Between Regional Accent and Voice/Speech Disorders</b> .....	2137
<i>Alain Ghio, Mederic Gasquet-Cyrus, Juliette Roquel, Antoine Giovanni</i>	
<b>Linguistic Disfluency in Narrative Speech: Evidence from Story-Telling in 6-Year Olds</b> .....	2142
<i>Ingrida Balciuniene</i>	
<b>Assessing the Utility of Judgments of Children's Speech Production Made by Untrained Listeners in Uncontrolled Listening Environments</b> .....	2146
<i>Benjamin Munson</i>	



<b>Consonant Distortions in Dysarthria Due to Parkinson's Disease, Amyotrophic Lateral Sclerosis and Cerebellar Ataxia</b> .....	2151
<i>Tanja Kocjancic Antolik, Cecile Fougeron</i>	
<b>Study of Coarticulation and F2 Transitions in French and Italian Adult Stutterers</b> .....	2156
<i>Marine Verdurand, Solange Rossato, Lionel Granjon, Daria Balbo, Claudio Zmarich</i>	
<b>Automatic Tracheoesophageal Voice Typing Using Acoustic Parameters</b> .....	2161
<i>Renee P. Clapham, Corina J. Van As-Brooks, Michiel W.M. Van den Brekel, Frans J.M. Hilgers, Rob J.J.H. Van Son</i>	
<b>Burst-Based Features for the Classification of Pathological Voices</b> .....	2166
<i>Julie Mauclair, Lionel Koenig, Marina Robert, Peggy Gatignol</i>	
<b>Classification of Depression State Based on Articulatory Precision</b> .....	2171
<i>Brian S. Helfer, Thomas F. Quatieri, James R. Williamson, Daryush D. Mehta, Rachele Horwitz, Bea Yu</i>	
<b>Using Text and Acoustic Features to Diagnose Progressive Aphasia and its Subtypes</b> .....	2176
<i>Kathleen C. Fraser, Frank Rudzicz, Elizabeth Rochon</i>	

### **POSTER SESSION 18: SEARCH AND COMPUTATIONAL ISSUES IN LVCSR**

*Chair: Thomas Hain*

<b>Multi-Domain Neural Network Language Model</b> .....	2181
<i>Tanel Alumae</i>	
<b>Improving Lightly Supervised Training for Broadcast Transcription</b> .....	2186
<i>Y. Long, M.J.F. Gales, P. Lanchantin, X. Liu, M.S. Seigel, P.C. Woodland</i>	
<b>Weakly Supervised Parsing with Rules</b> .....	2191
<i>C. Cerisara, A. Lorenzo, P. Kral</i>	
<b>Relative Error Bounds for Statistical Classifiers Based on the <math>f</math>-Divergence</b> .....	2196
<i>Markus Nussbaum-Thom, Eugen Beck, Tamer Alkhouti, Ralf Schluter, Hermann Ney</i>	
<b>Experiments Towards a Better LVCSR System for Tamil</b> .....	2201
<i>Melvin Jose Johnson Premkumar, Ngoc Thang Vu, Tanja Schultz</i>	
<b>A Hybrid Language Model for Open-Vocabulary Thai LVCSR</b> .....	2206
<i>Kwanchiva Thangthai, Ananlada Chotimongkol, Chai Wutiwiewatchai</i>	
<b>Hierarchical Pitman-Yor and Dirichlet Process for Language Model</b> .....	2211
<i>Jen-Tzung Chien, Ying-Lan Chang</i>	
<b>Unsupervised Confidence Calibration Using Examples of Recognized Words and Their Contexts</b> .....	2216
<i>Taichi Asami, Satoshi Kobashikawa, Hirokazu Masataki, Osamu Yoshioka, Satoshi Takahashi</i>	
<b>Multilingual Hierarchical MRATA Features for ASR</b> .....	2221
<i>Zoltan Tuske, Ralf Schluter, Hermann Ney</i>	
<b>Heuristic Selection of Training Sentences from Historical TV Guide for Semi-Supervised LM Adaptation</b> .....	2226
<i>Harry M. Chang</i>	
<b>Combination of Random Indexing Based Language Model and N-Gram Language Model for Speech Recognition</b> .....	2231
<i>Dominique Fohr, Odile Mella</i>	
<b>Improving Low-Resource CD-DNN-HMM Using Dropout and Multilingual DNN Training</b> .....	2236
<i>Yajie Miao, Florian Metze</i>	
<b>Finding Recurrent Out-of-Vocabulary Words</b> .....	2241
<i>Long Qin, Alexander Rudnicky</i>	
<b>Using Conversational Word Bursts in Spoken Term Detection</b> .....	2246
<i>Justin Chiu, Alexander Rudnicky</i>	

### **ORAL SESSION 31: SPEECH AND HEARING DISORDERS**

*Chairs: Suzanne Boyce, Valerie Hazan*

<b>Brain Activations in Speech Recovery Process After Intra-Oral Surgery: An fMRI Study</b> .....	2251
<i>Audrey Acher, Marc Sato, Laurent Lamalle, Coriandre Vilain, Arnaud Attye, Alexandre Krainik, Georges Bettega, Christian Adrien Righini, Brice Carlot, Muriel Brix, Pascal Perrier</i>	
<b>Acoustic and Perceptual Analysis of Vocal Tremor</b> .....	2256
<i>Christophe Mertens, Jean Schoentgen, Francis Grenez, Sabine Skodda</i>	

<b>Lexical Tone Perception in Thai Normal-Hearing Adults and Those Using Hearing Aids: A Case Study</b> .....	2261
<i>C. Tantibundhit, C. Onsuwan, N. Klangpornkun, P. Phienphanich, T. Saimai, N. Saimai, P. Pitathawatchai, Chai Wutiwivatchai</i>	
<b>Evaluation of a Bone-Conducted Ultrasonic Hearing Aid in Vocal Emotion Transmission</b> .....	2266
<i>Takayuki Kagomiya, Seiji Nakagawa</i>	
<b>Processing of /h/ and /u/ in Italian Cochlear-Implant Children: A Behavioral and Neurophysiologic Study</b> .....	2271
<i>Luigia Garrapa, Davide Bottari, Mirko Grimaldi, Francesco Pavani, Andrea Calabrese, Michele De Benedetto, Silvano Vitale</i>	
<b>Predicting the Bilateral Advantage in Cochlear Implantees Using a Non-Intrusive Speech Intelligibility Measure</b> .....	2276
<i>Stefano Cosentino, Tiago H. Falk, David McAlpine</i>	

### **ORAL SESSION 32: SPEECH AND AUDIO SEGMENTATION**

*Chairs: Regine Andre-Obrecht, Richard Stern*

<b>A Blind Segmentation Approach to Acoustic Event Detection Based on I-Vector</b> .....	2281
<i>Zhen Huang, You-Chi Cheng, Kehuang Li, Ville Hautamaki, Chin-Hui Lee</i>	
<b>A Dynamic Programming Framework for Neural Network-Based Automatic Speech Segmentation</b> .....	2286
<i>Van Zyl van Vuuren, Louis ten Bosch, Thomas Niesler</i>	
<b>Acoustic Segmentation of Speech Using Zero Time Liftering (ZTL)</b> .....	2291
<i>RaviShankar Prasad, B. Yegnanarayana</i>	
<b>Unsupervised Mining of Acoustic Subword Units with Segment-Level Gaussian Posteriorgrams</b> .....	2296
<i>Haipeng Wang, Tan Lee, Cheung-Chi Leung, Bin Ma, Haizhou Li</i>	
<b>Combination of Auditory Attention Features with Phone Posteriors for Better Automatic Phoneme Segmentation</b> .....	2301
<i>Ozlem Kalinli</i>	
<b>Automatic Phonetic Segmentation Using Boundary Models</b> .....	2305
<i>Jiahong Yuan, Neville Ryant, Mark Liberman, Andreas Stolcke, Vikramjit Mitra, Wen Wang</i>	

### **ORAL SESSION 33: SPEECH SYNTHESIS — VARIOUS TOPICS**

*Chairs: Keiichi Tokuda, Petra Wagner*

<b>HMM-Based TTS for Hanoi Vietnamese: Issues in Design and Evaluation</b> .....	2310
<i>Thi Thu Trang Nguyen, Christophe D'Alessandro, Albert Rilliard, Do Dat Tran</i>	

## **VOLUME 4**

<b>HMM-Based Synthesis of Creaky Voice</b> .....	2315
<i>Tuomo Raitio, John Kane, Thomas Drugman, Christer Gobl</i>	
<b>Integrating Conditional Random Fields and Joint Multi-Gram Model with Syllabic Features for Grapheme-to-Phone Conversion</b> .....	2320
<i>Xiaoxuan Wang, Khe Chai Sim</i>	
<b>Structure Learning in Hidden Conditional Random Fields for Grapheme-to-Phoneme Conversion</b> .....	2325
<i>Patrick Lehnen, Alexandre Allauzen, Thomas Lavergne, Francois Yvon, Stefan Hahn, Hermann Ney</i>	
<b>TUNDRA: A Multilingual Corpus of Found Data for TTS Research Created with Light Supervision</b> .....	2330
<i>Adriana Stan, O. Watts, Y. Mamiya, M. Giurgiu, Robert A.J. Clark, Junichi Yamagishi, Simon King</i>	
<b>Minimum Mean Squared Error Based Warped Complex Cepstrum Analysis for Statistical Parametric Speech Synthesis</b> .....	2335
<i>Ranniere Maia, M.J.F. Gales, Yannis Stylianou, Masami Akamine</i>	

### **ORAL SESSION 34: ASR — DISCRIMINATIVE TRAINING**

*Chairs: Karen Livescu, Phil Woodland*

<b>Augmented Conditional Random Fields Modeling Based on Discriminatively Trained Features</b> .....	2340
<i>Yasser Hifny</i>	

<b>Sequence-Discriminative Training of Deep Neural Networks</b> .....	2344
<i>Karel Vesely, Arnab Ghoshal, Lukas Burget, Daniel Povey</i>	
<b>Discriminatively Trained Sparse Inverse Covariance Matrices for Low Resource Acoustic Modeling</b> .....	2349
<i>Weibin Zhang, Pascale Fung</i>	
<b>Discriminative Training of Acoustic Models for System Combination</b> .....	2354
<i>Yuuki Tachioka, Shinji Watanabe</i>	
<b>Semi-Supervised GMM and DNN Acoustic Model Training with Multi-System Combination and Confidence Re-Calibration</b> .....	2359
<i>Yan Huang, Dong Yu, Yifan Gong, Chaojun Liu</i>	
<b>Restructuring of Deep Neural Network Acoustic Models with Singular Value Decomposition</b> .....	2364
<i>Jian Xue, Jinyu Li, Yifan Gong</i>	

### **ORAL SESSION 35: L2 ACQUISITION, MULTILINGUALISM**

*Chairs: Kathy Huet, Caroline Smith*

<b>Large-Scale Characterization of Mandarin Pronunciation Errors Made by Native Speakers of European Languages</b> .....	2369
<i>Nancy F. Chen, Vivaek Shivakumar, Mahesh Harikumar, Bin Ma, Haizhou Li</i>	
<b>Production Training in Second Language Acquisition: A Comparison Between Objective Measures and Subjective Judgments</b> .....	2374
<i>Veronique Delvaux, Kathy Huet, Myriam Piccaluga, Bernard Harmegnies</i>	
<b>The Production and Perception of Voice Onset Time in English-Speaking Children Enrolled in a French Immersion Program</b> .....	2379
<i>Nicole Netelenbos, Fangfang Li</i>	
<b>Pronunciation Errors by Spanish Learners of Dutch: A Data-Driven Study for ASR-Based Pronunciation Training</b> .....	2384
<i>Pepi Burgos, Catia Cucchiari, Roeland van Hout, Helmer Strik</i>	
<b>Realisation of Tonal Alignment in the English of Japanese-English Late Bilinguals</b> .....	2389
<i>Calbert Graham, Brechtje Post</i>	
<b>The Influence of Language and Speech Task Upon Creaky Voice Use Among Six Young American Women Learning French</b> .....	2394
<i>Agathe Benoist-lucy, Claire Pillot-Loiseau</i>	

### **SPECIAL SESSION 5 (A & B): CHILD COMPUTER INTERACTION**

*Chairs: Keelan Evanini, Johan Schalkwyk*

<b>Acoustic-Prosodic, Turn-Taking, and Language Cues in Child-Psychologist Interactions for Varying Social Demand</b> .....	2399
<i>Daniel Bone, Chi-Chun Lee, Theodora Chaspari, Matthew P. Black, Marian E. Williams, Sungbok Lee, Pat Levitt, Shrikanth Narayanan</i>	
<b>A Preliminary Study of Child Vocalization on a Parallel Corpus of US and Shanghainese Toddlers</b> .....	2404
<i>Hynek Boril, Qian Zhang, Pongtep Angkittrakul, John H.L. Hansen, Dongxin Xu, Jill Gilkerson, Jeffrey A. Richards</i>	
<b>A Survey About Databases of Children's Speech</b> .....	2409
<i>Felix Claus, Hamurabi Gamboa Rosales, Rico Petrick, Horst-Udo Hain, Rüdiger Hoffmann</i>	
<b>Affective Evaluation of Multimodal Dialogue Games for Preschoolers Using Physiological Signals</b> .....	2414
<i>Vassiliki Kouloumenta, Manolis Perakakis, Alexandros Potamianos</i>	
<b>Amplitude Modulation Features for Emotion Recognition from Speech</b> .....	2419
<i>Md. Jahangir Alam, Yazid Attabi, Pierre Dumouchel, Patrick Kenny, Douglas O'Shaughnessy</i>	
<b>Analyzing Eye-Voice Coordination in Rapid Automated Naming</b> .....	2424
<i>Daniel Bone, Chi-Chun Lee, Vikram Ramanarayanan, Shrikanth Narayanan, Renske S. Hoedemaker, Peter C. Gordon</i>	
<b>Analyzing the Structure of Parent-Moderated Narratives from Children with ASD Using an Entity-Based Approach</b> .....	2429
<i>Theodora Chaspari, Emily Mower Provost, Shrikanth Narayanan</i>	
<b>Automated Speech Scoring for Non-Native Middle School Students with Multiple Task Types</b> .....	2434
<i>Keelan Evanini, Xinhao Wang</i>	
<b>Identification of Gender from Children's Speech by Computers and Humans</b> .....	2439
<i>Saeid Safavi, Peter Jancovic, Martin Russell, Michael Carey</i>	

<b>On Why Japanese /r/ Sounds are Difficult for Children to Acquire</b> .....	2444
<i>Takayuki Arai</i>	

## **POSTER SESSION 19: SPEAKER RECOGNITION**

*Chair: Craig Greenberg*

<b>Anchor and UBM-Based Multi-Class MLLR M-Vector System for Speaker Verification</b> .....	2449
<i>A.K. Sarkar, Claude Barras</i>	
<b>Ensemble Approach in Speaker Verification</b> .....	2454
<i>Leibny Paola Garcia Perera, Bhiksha Raj, Juan Arturo Nolasco-Flores</i>	
<b>Sequential Model Adaptation for Speaker Verification</b> .....	2459
<i>Jun Wang, Dong Wang, Xiaojun Wu, Thomas Fang Zheng, Javier Tejedor</i>	
<b>Improving Short Utterance Based I-Vector Speaker Recognition Using Source and Utterance-Duration Normalization Techniques</b> .....	2464
<i>A. Kanagasundaram, D. Dean, Javier Gonzalez-Dominguez, S. Sridharan, D. Ramos, Joaquín Gonzalez-Rodriguez</i>	
<b>On Leveraging Conversational Data for Building a Text Dependent Speaker Verification System</b> .....	2469
<i>Hagai Aronowitz, Oren Barkan</i>	
<b>THU-EE System Fusion for the NIST 2012 Speaker Recognition Evaluation</b> .....	2473
<i>Wei-Qiang Zhang, Zhiyi Li, Weiwei Liu, Jia Liu</i>	
<b>Subspace-Constrained Supervector PLDA for Speaker Verification</b> .....	2478
<i>Daniel Garcia-Romero, Alan McCree</i>	
<b>Augmenting Short-Term Cepstral Features with Long-Term Discriminative Features for Speaker Verification of Telephone Data</b> .....	2483
<i>Cong-Thanh Do, Claude Barras, Viet Bac Le, A.K. Sarkar</i>	
<b>Using Group Delay Functions from All-Pole Models for Speaker Recognition</b> .....	2488
<i>Padmanabhan Rajan, Tomi Kinnunen, Cemal Hanilci, Jouni Pohjalainen, Paavo Alku</i>	
<b>Secure Binary Embeddings of Front-End Factor Analysis for Privacy Preserving Speaker Verification</b> .....	2493
<i>Jose Portelo, Alberto Abad, Bhiksha Raj, Isabel Trancoso</i>	
<b>On Von-Mises Fisher Mixture Model in Text-Independent Speaker Identification</b> .....	2498
<i>Jalil Taghia, Zhanyu Ma, Arne Leijon</i>	
<b>Using Phone Log-Likelihood Ratios as Features for Speaker Recognition</b> .....	2503
<i>Mireia Diez, Amparo Varona, Mikel Penagarikano, Luis Javier Rodriguez-Fuentes, German Bordel</i>	
<b>Handling Recordings Acquired Simultaneously Over Multiple Channels with PLDA</b> .....	2508
<i>Jesus Villalba, Mireia Diez, Amparo Varona, Eduardo Lleida</i>	
<b>Bayesian Distance Metric Learning on i-Vector for Speaker Verification</b> .....	2513
<i>Xiao Fang, Najim Dehak, James Glass</i>	
<b>Merging Human and Automatic System Decisions to Improve Speaker Recognition Performance</b> .....	2518
<i>Rosa Gonzalez Hautamaki, Ville Hautamaki, Padmanabhan Rajan, Tomi Kinnunen</i>	

## **POSTER SESSION 20: DIALOG SYSTEMS AND APPLICATIONS I**

*Chair: Martine Garnier-Rizet*

<b>Recurrent Neural Networks for Language Understanding</b> .....	2523
<i>Kaisheng Yao, Geoffrey Zweig, Mei-Yuh Hwang, Yangyang Shi, Dong Yu</i>	
<b>A Study on LVCSR and Keyword Search for Tagalog</b> .....	2528
<i>Korbinian Riedhammer, Van Hai Do, James Hieronymus</i>	
<b>Characterising Depressed Speech for Classification</b> .....	2533
<i>Sharifa Alghowinem, Roland Goecke, Michael Wagner, Julien Epps, Gordon Parker, Michael Breakspear</i>	
<b>Combining Acoustic Name Spotting and Continuous Context Models to improve Spoken Person Name Recognition in Speech</b> .....	2538
<i>Benjamin Bigot, Gregory Senay, Georges Linares, Corinne Fredouille, Richard Dufour</i>	
<b>A Resource-Dependent Approach to Word Modeling for Keyword Spotting</b> .....	2543
<i>I-Fan Chen, Chin-Hui Lee</i>	
<b>Markers of Confidence and Correctness in Spoken Medical Narratives</b> .....	2548
<i>Kathryn Womack, Cecilia Ovesdotter Alm, Cara Calvelli, Jeff B. Pelz, Pengcheng Shi, Anne Haake</i>	

<b>Development of a Web Framework for Teaching and Learning Japanese Prosody: OJAD (Online Japanese Accent Dictionary)</b> .....	2553
<i>Ibuki Nakamura, Nobuaki Minematsu, Masayuki Suzuki, Hiroko Hirano, Chieko Nakagawa, Noriko Nakamura, Yukinori Tagawa, Keikichi Hirose, Hiroya Hashimoto</i>	
<b>Addressee Detection for Dialog Systems Using Temporal and Spectral Dimensions of Speaking Style</b> .....	2558
<i>Elizabeth Shriberg, Andreas Stolcke, Suman Ravuri</i>	
<b>Analysis of Factors Involved in the Choice of Rising or Non-Rising Intonation in Question Utterances Appearing in Conversational Speech</b> .....	2563
<i>Hiroaki Hatano, Miyako Kiso, Carlos T. Ishi</i>	
<b>IsNL? A Discriminative Approach to Detect Natural Language Like Queries for Conversational Understanding</b> .....	2568
<i>Asli Celikyilmaz, Gokhan Tur, Dilek Hakkani-Tur</i>	
<b>Automatic Accent Quantification of Indian Speakers of English</b> .....	2573
<i>Jian Cheng, Nikhil Bojja, Xin Chen</i>	
<b>Semantic Parsing Using Word Confusion Networks with Conditional Random Fields</b> .....	2578
<i>Gokhan Tur, Anoop Deoras, Dilek Hakkani-Tur</i>	
<b>Timing Responses to Questions in Dialogue</b> .....	2583
<i>Sofia Strombergsson, Anna Hjalmarsson, Jens Edlund, David House</i>	
<b>BUT BABEL System for Spontaneous Cantonese</b> .....	2588
<i>Martin Karafiat, Frantisek Grezl, Mirko Hannemann, Karel Vesely, Jan Cernocky</i>	
<b>Semi-Supervised Manifold Learning Approaches for Spoken Term Verification</b> .....	2593
<i>Atta Norouzi, Richard C. Rose, Aren Jansen</i>	
<b>Language Modeling for Mixed Language Speech Recognition Using Weighted Phrase Extraction</b> .....	2598
<i>Ying Li, Pascale Fung</i>	

## **POSTER SESSION 21: SPOKEN MACHINE TRANSLATION AND SPEECH NATURAL LANGUAGE PROCESSING**

*Chair: Chiori Hori*

<b>Quality Assessment of Asymmetric Multiparty Telephone Conferences: A Systematic Method from Technical Degradations to Perceived Impairments</b> .....	2603
<i>Janto Skowronek, Julian Herlinghaus, Alexander Raake</i>	
<b>User Activity Estimation Method Based on Probabilistic Generative Model of Acoustic Event Sequence with User Activity and its Subordinate Categories</b> .....	2608
<i>Keisuke Imoto, Suehiro Shimauchi, Hisashi Uematsu, Hitoshi Ohmuro</i>	
<b>Generalizing Continuous-Space Translation of Paralinguistic Information</b> .....	2613
<i>Takatomo Kano, Shinnosuke Takamichi, Sakriani Sakti, Graham Neubig, Tomoki Toda, Satoshi Nakamura</i>	
<b>An Empirical Comparison of Joint Optimization Techniques for Speech Translation</b> .....	2618
<i>Masaya Ohgushi, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura</i>	
<b>A Sequential Repetition Model for Improved Disfluency Detection</b> .....	2623
<i>Mari Ostendorf, Sangyun Hahn</i>	
<b>Disfluency Detection Based on Prosodic Features for University Lectures</b> .....	2628
<i>Henrique Medeiros, Helena Moniz, Fernando Batista, Isabel Trancoso, Luis Nunes</i>	
<b>What's the Difference? Comparing Humans and Machines on the Aurora 2 Speech Recognition Task</b> .....	2633
<i>Bernd T. Meyer</i>	
<b>Calibration of Distance Measures for Unsupervised Query-by-Example</b> .....	2638
<i>Michele Gubian, Lou Boves, Maarten Versteegh</i>	
<b>Indexing Multimedia Documents with Acoustic Concept Recognition Lattices</b> .....	2643
<i>Diego Castan, Murat Akbacak</i>	
<b>MINT.tools: Tools and Adaptors Supporting Acquisition, Annotation and Analysis of Multimodal Corpora</b> .....	2648
<i>Spyros Kousidis, Thies Pfeiffer, David Schlangen</i>	

## **SHOW & TELL-3: SHOW AND TELL SESSION 3**

*Chairs: Laurence Devillers, Fabrice Lefevre*

<b>Simple<sup>4</sup>All</b> .....	2653
<i>Robert A.J. Clark</i>	

<b>On-Line Learning of Lexical Items and Grammatical Constructions via Speech, Gaze and Action-Based Human-Robot Interaction</b> .....	2656
<i>Gregoire Pointeau, Maxime Petit, Xavier Hinaut, Guillaume Gibert, Peter Ford Dominey</i>	
<b>Development of a Pronunciation Training System Based on Auditory-Visual Elements</b> .....	2659
<i>Haruko Miyakoda</i>	
<b>Real-Time and Non-Real-Time Voice Conversion Systems with Web Interfaces</b> .....	2661
<i>Elias Azarov, Maxim Vashkevich, Denis Likhachov, Alexander Petrovsky</i>	
<b>Application of the NAO Humanoid Robot in the Treatment of Bone Marrow-Transplanted Children (Demo)</b> .....	2663
<i>E. Csala, G. Nemeth, Cs. Zainko</i>	
<b>Photo-Realistic Expressive Text to Talking Head Synthesis</b> .....	2666
<i>Vincent Wan, Robert Anderson, Art Blokland, Norbert Braunschweiler, Langzhou Chen, BalaKrishna Kolluru, Javier Latorre, Ranniery Maia, Bjorn Stenger, Kayoko Yanagisawa, Yannis Stylianou, Masami Akamine, M.J.F. Gales, Roberto Cipolla</i>	
<b>Demonstration of LAPSyD: Lyon-Albuquerque Phonological Systems Database</b> .....	2669
<i>Ian Maddieson, Sebastien Flavier, Egidio Marsico, Francois Pellegrino</i>	
<b>SpeechMark Acoustic Landmark Tool: Application to Voice Pathology</b> .....	2671
<i>Suzanne Boyce, Marisha Speights, Keiko Ishikawa, Joel MacAuslan</i>	
<b>MODIS: An Audio Motif Discovery Software</b> .....	2674
<i>Laurence Catanese, Nathan Souvira-Labastie, Bingqing Qu, Sebastien Campion, Guillaume Gravier, Emmanuel Vincent, Frederic Bimbot</i>	

### **ORAL SESSION 36: LANGUAGE MODEL ADAPTATION**

*Chairs: Michiel Bacchiani, Richard Schwartz*

<b>Fitting Long-Range Information Using Interpolated Distanced N-Grams and Cache Models into a Latent Dirichlet Language Model for Speech Recognition</b> .....	2677
<i>Md. Akmal Haidar, Douglas O'Shaughnessy</i>	
<b>Incorporating Proximity Information for Relevance Language Modeling in Speech Recognition</b> .....	2682
<i>Yi-Wen Chen, Bo-Han Hao, Kuan-Yu Chen, Berlin Chen</i>	
<b>Instance-Based On-Line Language Model Adaptation</b> .....	2687
<i>Ali Orkan Bayer, G. Riccardi</i>	
<b>Unsupervised Topic Adaptation for Morph-Based Speech Recognition</b> .....	2692
<i>Andre Mansikkaniemi, Mikko Kurimo</i>	
<b>Unsupervised Language Model Adaptation for Automatic Speech Recognition of Broadcast News Using Web 2.0</b> .....	2697
<i>Tim Schlippe, Lukasz Gren, Ngoc Thang Vu, Tanja Schultz</i>	
<b>Recurrent Neural Network Based Language Model Personalization by Social Network Crowdsourcing</b> .....	2702
<i>Tsung-Hsien Wen, Aaron Heide, Hung-yi Lee, Yu Tsao, Lin-shan Lee</i>	

### **ORAL SESSION 37: SPOKEN LANGUAGE SUMMARIZATION AND UNDERSTANDING**

*Chairs: Dilek Hakkani-Tur, Jingjing Liu*

<b>Language-Independent Call Routing Using the Large Margin Estimation Principle</b> .....	2707
<i>Moataz El Ayadi, Mohamed Afify</i>	
<b>Deep Belief Network Based Semantic Taggers for Spoken Language Understanding</b> .....	2712
<i>Anoop Deoras, Ruhi Sarikaya</i>	
<b>Error-Corrective Discriminative Joint Decoding of Automatic Spoken Language Transcription and Understanding</b> .....	2717
<i>Bassam Jabaian, Fabrice Lefevre</i>	
<b>Detecting Summarization Hot Spots in Meetings Using Group Level Involvement and Turn-Taking Features</b> .....	2722
<i>Catherine Lai, Jean Carletta, Steve Renals</i>	
<b>Supervised Spoken Document Summarization Based on Structured Support Vector Machine with Utterance Clusters as Hidden Variables</b> .....	2727
<i>Sz-Rung Shiang, Hung-yi Lee, Lin-shan Lee</i>	
<b>Web Data Harvesting for Speech Understanding Grammar Induction</b> .....	2732
<i>Ioannis Klasinas, Alexandros Potamianos, Elias Iosif, Spiros Georgiladakis, Gianluca Marneli</i>	

## **ORAL SESSION 38: SPEECH SYNTHESIS — MULTIMODAL AND ARTICULATORY SYNTHESIS**

*Chairs: Jonas Beskow, Jianwu Dang*

<b>Articulatory Synthesis of French Connected Speech from EMA Data</b> .....	2737
<i>Asterios Toutios, Shrikanth Narayanan</i>	
<b>A New Language Independent, Photo-Realistic Talking Head Driven by Voice Only</b> .....	2742
<i>Xinjian Zhang, Lijuan Wang, Gang Li, Frank Seide, Frank K. Soong</i>	
<b>Binocular Photometric Stereo Acquisition and Reconstruction for 3D Talking Head Applications</b> .....	2747
<i>Chaoyang Wang, Lijuan Wang, Yasuyuki Matsushita, Bojun Huang, Magnetron Chen, Frank K. Soong</i>	
<b>Speaker Adaptation of an Acoustic-Articulatory Inversion Model Using Cascaded Gaussian Mixture Regressions</b> .....	2752
<i>Thomas Hueber, Gerard Bailly, Pierre Badin, Frederic Elisei</i>	
<b>Articulatory Features for Speech-Driven Head Motion Synthesis</b> .....	2757
<i>Atef Ben-Youssef, Hiroshi Shimodaira, David Adam Braude</i>	
<b>Template-Warping Based Speech Driven Head Motion Synthesis</b> .....	2762
<i>David Adam Braude, Hiroshi Shimodaira, Atef Ben-Youssef</i>	

## **ORAL SESSION 39: SPEAKER DIARIZATION AND RECOGNITION**

*Chairs: Sylvain Meignier, Jason Pelecanos*

<b>ALIZE 3.0 - Open Source Toolkit for State-of-the-Art Speaker Recognition</b> .....	2767
<i>Anthony Larcher, Jean-Francois Bonastre, Benoit Fauve, Kong Aik Lee, Christophe Levy, Haizhou Li, John S.D. Mason, Jean-Yves Parfait</i>	
<b>New Cosine Similarity Scorings to Implement Gender-Independent Speaker Verification</b> .....	2772
<i>Mohammed Senoussaoui, Patrick Kenny, Pierre Dumouchel, Najim Dehak</i>	
<b>Improving Speaker Identification in TV-Shows Using Person Name Detection in Overlaid Text and Speech</b> .....	2777
<i>Delphine Charlet, Corinne Fredouille, Geraldine Damnati, Gregory Senay</i>	
<b>Exploring Methods of Improving Speaker Accuracy for Speaker Diarization</b> .....	2782
<i>Mary Tai Knox, Nikki Mirghafori, Gerald Friedland</i>	
<b>Combining Deep Speaker Specific Representations with GMM-SVM for Speaker Verification</b> .....	2787
<i>Ryan Price, Sangeeta Biswas, Koichi Shinoda</i>	
<b>Using Spectral Moments as a Speaker Specific Feature in Nasals and Fricatives</b> .....	2792
<i>Carola Schindler, Christoph Draxler</i>	

## **ORAL SESSION 40: MODELS OF SPEECH PERCEPTION**

*Chairs: Louis Ten Bosch, Nijmegen and Odile Bagou*

<b>A Computational Model of Perceptuo-Motor Processing in Speech Perception: Learning to Imitate and Categorize Synthetic CV Syllables</b> .....	2796
<i>Raphael Laurent, Jean-Luc Schwartz, Pierre Bessiere, Julien Diard</i>	
<b>Talker-Specific Perceptual Processing: Influences on Internal Category Structure</b> .....	2801
<i>Rachel M. Theodore</i>	
<b>Elicitation and Analysis of a Corpus of Robust Noise-Induced Word Misperceptions in Spanish</b> .....	2806
<i>Maria Luisa Garcia Lecumberri, Attila Mate Toth, Yan Tang, Martin Cooke</i>	
<b>Vocabulary Structure and Spoken-Word Recognition: Evidence from French Reveals the Source of Embedding Asymmetry</b> .....	2811
<i>Anne Cutler, Laurence Bruggeman</i>	
<b>How do Multiple Sublexical Cues Converge in Lexical Segmentation? An Artificial Language Learning Study</b> .....	2816
<i>Odile Bagou, Ulrich Hans Frauenfelder</i>	
<b>Towards an End-to-End Computational Model of Speech Comprehension: Simulating a Lexical Decision Task</b> .....	2821
<i>Louis ten Bosch, Lou Boves, Mirjam Ernestus</i>	

## **POSTER SESSION 22: PARALINGUISTIC INFORMATION**

*Chair: Julia Hirschberg*

<b>Demographic Recommendation by Means of Group Profile Elicitation Using Speaker Age and Gender Recognition</b> .....	2826
<i>Sven Ewan Shepstone, Zheng-Hua Tan, Soren Holdt Jensen</i>	
<b>Affective Classification of Generic Audio Clips Using Regression Models</b> .....	2831
<i>Nikolaos Malandrakis, Shiva Sundaram, Alexandros Potamianos</i>	
<b>A Preliminary Study of Cross-Lingual Emotion Recognition from Speech: Automatic Classification versus Human Perception</b> .....	2836
<i>Je Hun Jeon, Duc Le, Rui Xia, Yang Liu</i>	
<b>Active Learning for Dimensional Speech Emotion Recognition</b> .....	2840
<i>Wenjing Han, Haifeng Li, Huabin Ruan, Lin Ma, Jiayin Sun, Bjorn Schuller</i>	
<b>Auditory Detectability of Vocal Ageing and its Effect on Forensic Automatic Speaker Recognition</b> .....	2845
<i>Finnian Kelly, Naomi Harte</i>	
<b>Comparative Study of Speaker Personality Traits Recognition in Conversational and Broadcast News Speech</b> .....	2850
<i>Firoj Alam, G. Riccardi</i>	
<b>Active Learning by Label Uncertainty for Acoustic Emotion Recognition</b> .....	2855
<i>Zixing Zhang, Jun Deng, Erik Marchi, Bjorn Schuller</i>	
<b>Modeling Therapist Empathy and Vocal Entrainment in Drug Addiction Counseling</b> .....	2860
<i>Bo Xiao, Panayiotis G. Georgiou, Zac E. Imel, David C. Atkins, Shrikanth Narayanan</i>	
<b>Estimating Callers' Levels of Knowledge in Call Center Dialogues</b> .....	2865
<i>Chiaki Miyazaki, Ryuichiro Higashinaka, Toshiro Makino, Yoshihiro Matsuo</i>	
<b>Energy and F0 Contour Modeling with Functional Data Analysis for Emotional Speech Detection</b> .....	2870
<i>Juan Pablo Arias, Carlos Busso, Nestor Becerra Yoma</i>	
<b>Incremental Emotion Recognition</b> .....	2875
<i>Taniya Mishra, Dimitrios Dimitriadis</i>	
<b>Comparison of Spectrum Estimators in Speaker Verification: Mismatch Conditions Induced by Vocal Effort</b> .....	2880
<i>Cemal Hanilci, Tomi Kinnunen, Padmanabhan Rajan, Jouni Pohjalainen, Paavo Alku, Figen Ertas</i>	
<b>Using Denoising Autoencoder for Emotion Recognition</b> .....	2885
<i>Rui Xia, Yang Liu</i>	

## **POSTER SESSION 23: SPEECH AND AUDIO SIGNAL PROCESSING**

*Chair: Piero Cosi*

<b>A Phase-Modified Approach for TDE-Based Acoustic Localization</b> .....	2889
<i>Georgios Athanasopoulos, Werner Verhelst</i>	
<b>Interference Robust DOA Estimation of Human Speech by Exploiting Historical Information and Temporal Correlation</b> .....	2894
<i>Wei Xue, Shan Liang, Wenju Liu</i>	
<b>Identifying New Bird Species from Differences in Birdsong</b> .....	2899
<i>Naomi Harte, Sadhbh Murphy, David J. Kelly, Nicola M. Marples</i>	
<b>Controlling "Shout" Expression in a Japanese POP Singing Performance: Analysis and Suppression Study</b> .....	2904
<i>Yuri Nishigaki, Ken-Ichi Sakakibara, Masanori Morise, Ryuichi Nisimura, Toshio Irino, Hideki Kawahara</i>	
<b>Dimensionality Analysis of Singing Speech Based on Locality Preserving Projections</b> .....	2909
<i>Mahnoosh Mehrabani, John H.L. Hansen</i>	
<b>Audio Classification Using Dominant Spatial Patterns in Time-Frequency Space</b> .....	2914
<i>Md. Khademul Islam Molla, Keikichi Hirose</i>	
<b>Spectro-Temporal Modulation Based Singing Detection Combined with Pitch-Based Grouping for Singing Voice Separation</b> .....	2919
<i>Tse-En Lin, Chung-Chien Hsu, Yi-Cheng Chen, Jian-Hueng Chen, Tai-Shih Chi</i>	
<b>NMF-Based Temporal Feature Integration for Acoustic Event Classification</b> .....	2923
<i>Jimmy Ludena-Choez, Ascension Gallardo-Antolin</i>	
<b>Robust Audio-Codebooks for Large-Scale Event Detection in Consumer Videos</b> .....	2928
<i>Shourabh Rawat, Peter F. Schulam, Susanne Burger, Duo Ding, Yipei Wang, Florian Metze</i>	



<b>Person Identification Using Biometric Markers from Footsteps Sound</b> .....	2933
<i>M. Umair Bin Altaf, Taras Butko, Bing-Hwang Juang</i>	
<b>Learning Binaural Spectrogram Features for Azimuthal Speaker Localization</b> .....	2938
<i>Wiktor Mlynarski</i>	
<b>An Unsupervised Bayesian Classifier for Multiple Speaker Detection and Localization</b> .....	2942
<i>Youssef Oualil, Friedrich Faubel, Dietrich Klakow</i>	
<b>Joint Recognition and Direction-of-Arrival Estimation of Simultaneous Meeting-Room Acoustic Events</b> .....	2947
<i>Rupayan Chakraborty, Climent Nadeu</i>	
<b>Audio Self Organized Units for High-Level Event Detection</b> .....	2952
<i>Xiaodan Zhuang, Shuang Wu, Pradeep Natarajan, Rohit Prasad, Prem Natarajan</i>	

### **POSTER SESSION 24: ASR — ROBUSTNESS AGAINST NOISE III**

*Chair: Nelson Morgan*

<b>Distribution-Based Feature Normalization for Robust Speech Recognition Leveraging Context and Dynamics Cues</b> .....	2957
<i>Yu-Chen Kao, Berlin Chen</i>	
<b>An Investigation of Temporally Varying Weight Regression for Noise Robust Speech Recognition</b> .....	2962
<i>Shilin Liu, Khe Chai Sim</i>	
<b>Feature Space Generalized Variable Parameter HMMs for Noise Robust Recognition</b> .....	2967
<i>Yang Li, Xunying Liu, Lan Wang</i>	
<b>Bidirectional Truncated Recurrent Neural Networks for Efficient Speech Denoising</b> .....	2972
<i>Philemon Brakel, Dirk Stroobandt, Benjamin Schrauwen</i>	
<b>Multi-Stream Recognition of Noisy Speech with Performance Monitoring</b> .....	2977
<i>Ehsan Variani, Feipeng Li, Hynek Hermansky</i>	
<b>Model-Based Noise Suppression Using Unsupervised Estimation of Hidden Markov Model for Non-Stationary Noise</b> .....	2981
<i>Masakiyo Fujimoto, Tomohiro Nakatani</i>	
<b>Joint Noise Cancellation and Dereverberation Using Multi-Channel Linearly Constrained Minimum Variance Filter</b> .....	2986
<i>Karan Nathwani, Rajesh M. Hegde</i>	
<b>Is Speech Enhancement Pre-Processing Still Relevant When Using Deep Neural Networks for Acoustic Modeling?</b> .....	2991
<i>Marc Delcroix, Yotaro Kubo, Tomohiro Nakatani, Atsushi Nakamura</i>	
<b>Histogram Equalization of Real and Imaginary Modulation Spectra for Noise-Robust Speech Recognition</b> .....	2996
<i>Hsin-Ju Hsieh, Berlin Chen, Jieh-weih Hung</i>	
<b>An Investigation of Spectral Restoration Algorithms for Deep Neural Networks Based Noise Robust Speech Recognition</b> .....	3001
<i>Bo Li, Yu Tsao, Khe Chai Sim</i>	
<b>Bounded Conditional Mean Imputation with an Approximate Posterior</b> .....	3006
<i>Ulpu Remes</i>	
<b>Mixtures of Bayesian Joint Factor Analyzers for Noise Robust Automatic Speech Recognition</b> .....	3011
<i>Xiaodong Cui, Vaibhava Goel, Brian Kingsbury</i>	
<b>Robust Speech Enhancement Techniques for ASR in Non-Stationary Noise and Dynamic Environments</b> .....	3016
<i>Gang Liu, Dimitrios Dimitriadis, Enrico Bocchieri</i>	

### **ORAL SESSION 41: LINGUISTIC SYSTEMS, PHONETICS-PHONOLOGY INTERFACE**

*Chairs: Ioana Chitoran, Henning Reetz*

<b>LAPSyD: Lyon-Albuquerque Phonological Systems Database</b> .....	3021
<i>Ian Maddieson, Sebastien Flavier, Egidio Marsico, Christophe Coupe, Francois Pellegrino</i>	
<b>The Duration Compensation Issue Revisited</b> .....	3026
<i>Plinio A. Barbosa</i>	
<b>Cross-Language Comparison of Functional Load for Vowels, Consonants, and Tones</b> .....	3031
<i>Yoon Mi Oh, Francois Pellegrino, Christophe Coupe, Egidio Marsico</i>	

<b>Notes on So-Called Inter-Speaker Difference in Spontaneous Speech: The Case of Japanese Voiced Obstruent</b> .....	3036
<i>Kikuo Maekawa</i>	
<b>The Role of the Pharynx and Tongue in Enhancement of Vowel Nasalization: A Real-Time MRI Investigation of French Nasal Vowels</b> .....	3041
<i>Christopher Carignan, Ryan K. Shosted, Maojing Fu, Zhi-Pei Liang, Bradley P. Sutton</i>	
<b>Assimilation of Word-Final Nasals to Following Word-Initial Place of Articulation in UK English</b> .....	3046
<i>Margaret E.L. Renwick, Ladan Baghai-Ravary, Rosalind Temple, John S. Coleman</i>	

#### **ORAL SESSION 42: SPEECH SYNTHESIS — VOICE CONVERSION**

*Chairs: Yao Qian, Tomoki Toda*

<b>Joint Spectral Distribution Modeling Using Restricted Boltzmann Machines for Voice Conversion</b> .....	3051
<i>Ling-Hui Chen, Zhen-Hua Ling, Yan Song, Li-Rong Dai</i>	
<b>Exemplar-Based Unit Selection for Voice Conversion Utilizing Temporal Information</b> .....	3056
<i>Zhizheng Wu, Tuomas Virtanen, Tomi Kinnunen, Eng Siong Chng, Haizhou Li</i>	

### **VOLUME 5**

<b>Alleviating the Over-Smoothing Problem in GMM-Based Voice Conversion with Discriminative Training</b> .....	3061
<i>Hsin-Te Hwang, Yu Tsao, Hsin-Min Wang, Yih-Ru Wang, Sin-Horng Chen</i>	
<b>A Hybrid Approach to Electrolaryngeal Speech Enhancement Based on Spectral Subtraction and Statistical Voice Conversion</b> .....	3066
<i>Kou Tanaka, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura</i>	
<b>A Digital Signal Processor Implementation of Silent/Electrolaryngeal Speech Enhancement based on Real-Time Statistical Voice Conversion</b> .....	3071
<i>Takuto Moriguchi, Tomoki Toda, Motoaki Sano, Hiroshi Sato, Graham Neubig, Sakriani Sakti, Satoshi Nakamura</i>	
<b>Foreign Accent Conversion Through Voice Morphing</b> .....	3076
<i>Sandesh Aryal, Daniel Felps, Ricardo Gutierrez-Osuna</i>	

#### **ORAL SESSION 43: LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION SYSTEMS**

*Chairs: Steve Renals, Dirk Van Compernelle*

<b>Empirical Link Between Hypothesis Diversity and Fusion Performance in an Ensemble of Automatic Speech Recognition Systems</b> .....	3081
<i>Kartik Audhkhasi, Andreas M. Zavou, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	
<b>A Lecture Transcription System Combining Neural Network Acoustic and Language Models</b> .....	3086
<i>Peter Bell, Hitoshi Yamamoto, Pawel Swietojanski, Youzheng Wu, Fergus McInnes, Chiori Hori, Steve Renals</i>	
<b>Neural Network Acoustic Models for the DARPA RATS Program</b> .....	3091
<i>Hagen Soltau, Hong-Kwang Kuo, Lidia Mangu, George Saon, Tomas Beran</i>	
<b>Improved Models for Automatic Punctuation Prediction for Spoken and Written Text</b> .....	3096
<i>Nicola Ueffing, Maximilian Bisani, Paul Vozila</i>	
<b>Some Issues Affecting the Transcription of Hungarian Broadcast Audio</b> .....	3101
<i>Anindya Roy, Lori Lamel, Thiago Fraga-Silva, Jean-Luc Gauvain, Ilya Oparin</i>	
<b>Development of the RWTH Transcription System for Slovenian</b> .....	3106
<i>Pavel Golik, Zoltan Tuske, Ralf Schluter, Hermann Ney</i>	

#### **ORAL SESSION 44: ROBUST SPEAKER RECOGNITION**

*Chairs: Najim Dehak, Lukas Burget*

<b>Noise Robust Speaker Verification with Delta Cepstrum Normalization</b> .....	3111
<i>Naoyuki Kanda, Ryu Takeda, Yasunari Obuchi</i>	
<b>R-Norm: Improving Inter-Speaker Variability Modelling at the Score Level via Regression Score Normalisation</b> .....	3116
<i>David Vandyke, Michael Wagner, Roland Goecke</i>	

<b>Frequency Warping and Robust Speaker Verification: A Comparison of Alternative Mel-Scale Representations</b> .....	3121
<i>Tomi Kinnunen, Md. Jahangir Alam, Pavel Matejka, Patrick Kenny, Jan Cernocky, Douglas O'Shaughnessy</i>	
<b>Acoustic Factor Analysis Based Universal Background Model for Robust Speaker Verification in Noise</b> .....	3126
<i>Taufiq Hasan, John H.L. Hansen</i>	
<b>A New Bayesian Network to Assess the Reliability of Speaker Verification Decisions</b> .....	3131
<i>Jesus Villalba, Eduardo Lleida, Alfonso Ortega, Antonio Miguel</i>	
<b>The IBM RATS Phase II Speaker Recognition System: Overview and Analysis</b> .....	3136
<i>Weizhong Zhu, Sibel Yaman, Jason Pelecanos</i>	

### **ORAL SESSION 45: ACOUSTIC AND ARTICULATORY CUES IN SPEECH PERCEPTION**

*Chairs: Mary Beckman, David House*

<b>Vowel Identity Conditions the Time Course of Tone Recognition</b> .....	3141
<i>Jason A. Shaw, Michael D. Tyler, Benjawan Kasisopa, Yuan Ma, Michael Proctor, Chong Han, Donald Derrick, Denis Burnham</i>	
<b>Changes in the Role of Intensity as a Cue for Fricative Categorisation</b> .....	3146
<i>Odette Scharenborg, Esther Janse</i>	
<b>Weighting of Acoustic Cues Shifts to Frication Duration in Identification of Fricatives/Affricates When Auditory Properties Are Degraded Due to Aging</b> .....	3151
<i>Keiichi Yasu, Takayuki Arai, Kei Kobayashi, Mitsuko Shindo</i>	
<b>Duration as a Secondary Cue for Perception of Voicing and Tone in Shanghai Chinese</b> .....	3156
<i>Jiayin Gao, Pierre Halle</i>	
<b>Development of Central Auditory Processes and Their Links with Language Skills in Typically Developing Children</b> .....	3161
<i>Marie Deckerle, Fanny Meunier, Marie-Ange N'Guyen, Estelle Gillet-Perret, Delphine Lassus-Sangosse, Sophie Donnadiou</i>	
<b>Show Me What You Listen to! Auditory Classification Images Can Reveal the Processing of Fine Acoustic Cues During Speech Categorization</b> .....	3166
<i>Leo Varnet, Kenneth Knoblauch, Fanny Meunier, Michel Hoen</i>	

### **POSTER SESSION 25: SPEECH PRODUCTION — DATA AND MODELS**

*Chair: Claire Pillot-Loiseau*

<b>The Organ Stop "Vox Humana" as a Model for a Vowel Synthesiser</b> .....	3171
<i>Fabian Brackhane, Jurgen Trouvain</i>	
<b>Information Theoretic Acoustic Feature Selection for Acoustic-to-Articulatory Inversion</b> .....	3176
<i>Prasanta Kumar Ghosh, Shrikanth Narayanan</i>	
<b>Formant Contours in Czech Vowels: Speaker-Discriminating Potential</b> .....	3181
<i>Dita Fejlova, David Lukes, Radek Skarnitzl</i>	
<b>An Anisotropic Diffusion Filter Based on Multidirectional Separability</b> .....	3186
<i>Shen Liu, Jianguo Wei, Xin Wang, Wenhuan Lu, Qiang Fang, Jianwu Dang</i>	
<b>The Phonological Voicing Contrast in Czech: An EPG Study of Phonated and Whispered Fricatives</b> .....	3190
<i>Radek Skarnitzl, Pavel Sturm, Pavel Machac</i>	
<b>Vowel and Prosodic Factor Dependent Variations of Vocal-Tract Length</b> .....	3195
<i>Shinji Maeda, Yves Laprie</i>	
<b>Word Identification Using Phonetic Features: Towards a Method to Support Multivariate fMRI Speech Decoding</b> .....	3200
<i>Tijl Grootswagers, Karen Dijkstra, Louis ten Bosch, Alex Brandmeyer, Makiko Sadakata</i>	
<b>Analysis of Breathily, Modal and Pressed Phonation Based on Low Frequency Spectral Density</b> .....	3205
<i>Dhananjaya Gowda, Mikko Kurimo</i>	
<b>Is the Vowel Length Contrast in Japanese Exaggerated in Infant-Directed Speech?</b> .....	3210
<i>Keiichi Tajima, Kuniyoshi Tanaka, Andrew Martin, Reiko Mazuka</i>	
<b>Investigating the Relationship Between Glottal Area Waveform Shape and Harmonic Magnitudes Through Computational Modeling and Laryngeal High-Speed Videendoscopy</b> .....	3215
<i>Gang Chen, Robin A. Samlan, Jody Kreiman, Abeer Alwan</i>	
<b>Formant Frequency Tracking Using Gaussian Mixtures with Maximum <i>a posteriori</i> Adaptation</b> .....	3220
<i>Jonathan C. Kim, Hrishikesh Rao, Mark A. Clements</i>	

<b>Devoicing of Vowels in German, a Comparison of Japanese and German Speakers</b> .....	3225
<i>Rei Yasuda, Frank Zimmerer</i>	
<b>Identifying Consonantal Tasks via Measures of Tongue Shaping: A Real-Time MRI Investigation of the Production of Vocalized Syllabic /l/ in American English</b> .....	3229
<i>Caitlin Smith, Adam Lammert</i>	

## **POSTER SESSION 26: SPEECH ENHANCEMENT**

*Chair: Maurizio Omologo*

<b>A Speech Enhancement Method by Coupling Speech Detection and Spectral Amplitude Estimation</b> .....	3233
<i>Feng Deng, Chang-chun Bao, Feng Bao</i>	
<b>Late Reverberation Suppression Using MMSE Modulation Spectral Estimation</b> .....	3238
<i>Chenxi Zheng, Wai-Yip Chan</i>	
<b>A New Statistical Excitation Mapping for Enhancement of Throat Microphone Recordings</b> .....	3243
<i>M.A. Tugtekin Turan, Engin Erzin</i>	
<b>Classification Based Binaural Dereverberation</b> .....	3248
<i>Nicoleta Roman, Michael I. Mandel</i>	
<b>Target-to-Non-Target Directional Ratio Estimation Based on Dual-Microphone Phase Differences for Target-Directional Speech Enhancement</b> .....	3253
<i>Seon Man Kim, Hong Kook Kim</i>	
<b>Speech Spectrum Restoration Based on Conditional Restricted Boltzmann Machine</b> .....	3258
<i>Xugang Lu, Shigeki Matsuda, Chiori Hori</i>	
<b>Speaker Separation Using Visual Speech Features and Single-Channel Audio</b> .....	3263
<i>Faheem Khan, Ben Milner</i>	
<b>Spectral Modulation Sensitivity Based Perceptual Acoustic Echo Cancellation</b> .....	3268
<i>Wei-Lun Chuang, Kah-Meng Cheong, Chung-Chien Hsu, Tai-Shih Chi</i>	
<b>Speech Enhancement Using Compressed Sensing</b> .....	3273
<i>Vinayak Abrol, Pulkit Sharma, Anil Kumar Sao</i>	
<b>Spectro-Temporal Post-Enhancement Using MMSE Estimation in NMF Based Single-Channel Source Separation</b> .....	3278
<i>Emad M. Grais, Hakan Erdogan</i>	
<b>A Pitch-Based Spectral Enhancement Technique for Robust Speech Processing</b> .....	3283
<i>Kantapon Kaewtip, Lee Ngee Tan, Abeer Alwan</i>	
<b>Stochastic-Deterministic Signal Modelling for the Tracking of Pitch in Noise and Speech Mixtures Using Factorial HMMs</b> .....	3288
<i>Matthew McCallum, Bernard Guillemin</i>	
<b>Restoration of Clipped Signals with Application to Speech Recognition</b> .....	3293
<i>Shay Maymon, Etienne Marcheret, Vaibhava Goel</i>	
<b>On the Robustness of Distributed EM Based BSS in Asynchronous Distributed Microphone Array Scenarios</b> .....	3297
<i>Yasufumi Uezu, Keisuke Kinoshita, Mehrez Souden, Tomohiro Nakatani</i>	

## **POSTER SESSION 27: ASR — ACOUSTIC MODELING**

*Chair: Arnab Ghoshal*

<b>Infinite Support Vector Machines in Speech Recognition</b> .....	3302
<i>Jingzhou Yang, Rogier C. van Dalen, M.J.F. Gales</i>	
<b>An On-Line Incremental Speaker Adaptation Technique for Audio Stream Transcription</b> .....	3307
<i>Diego Giuliani, Fabio Brugnara</i>	
<b>Accent- and Speaker-Specific Polyphone Decision Trees for Non-Native Speech Recognition</b> .....	3312
<i>Dominic Telaar, Mark C. Fuhs</i>	
<b>Investigations on Hessian-Free Optimization for Cross-Entropy Training of Deep Neural Networks</b> .....	3316
<i>Simon Wiesler, Jinyu Li, Jian Xue</i>	
<b>Cross-Lingual Acoustic Model Adaptation Based on Transfer Vector Field Smoothing with MAP</b> .....	3321
<i>Masahiro Saiko, Shigeki Matsuda, Ken Hanazawa, Ryosuke Isotani, Chiori Hori</i>	
<b>N-Best Rescoring by Phoneme Classifiers Using Subclass AdaBoost Algorithm</b> .....	3326
<i>Hiroshi Fujimura, Yusuke Shinohara, Takashi Masuko</i>	
<b>Stream Selection and Integration in Multistream ASR Using GMM-Based Performance Monitoring</b> .....	3331
<i>Tetsuji Ogawa, Feipeng Li, Hynek Hermansky</i>	

<b>VTLN Based on the Linear Interpolation of Contiguous Mel Filter-Bank Energies</b> .....	3336
<i>Nestor Becerra Yoma, Claudio Garreton, Fernando Huenupan, Ignacio Catalan, Jorge Wuth</i>	
<b>Context-Dependent Modeling and Speaker Normalization Applied to Reservoir-Based Phone Recognition</b> .....	3341
<i>Fabian Triefenbach, Azarakhsh Jalalvand, Kris Demuynck, Jean-Pierre Martens</i>	
<b>Interpolation of Acoustic Models for Speech Recognition</b> .....	3346
<i>Thiago Fraga-Silva, Jean-Luc Gauvain, Lori Lamel</i>	
<b>Training Log-Linear Acoustic Models in Higher-Order Polynomial Feature Space for Speech Recognition</b> .....	3351
<i>M. Tahir, H. Huang, Ralf Schluter, Hermann Ney, Louis ten Bosch, Bert Cranen, Lou Boves</i>	
<b>Comparison of Spectral Analysis Methods for Automatic Speech Recognition</b> .....	3355
<i>Venkata Neelima Parinam, Chandra Vootkuri, Stephen A. Zahorian</i>	
<b>Synthetic Speaker Models Using VTLN to Improve the Performance of Children in Mismatched Speaker Conditions for ASR</b> .....	3360
<i>D.R. Sanand, T. Svendsen</i>	
<b>Exploring Convolutional Neural Network Structures and Optimization Techniques for Speech Recognition</b> .....	3365
<i>Ossama Abdel-Hamid, Li Deng, Dong Yu</i>	

### **SPECIAL EVENT: ESCA/ISCA ANNIVERSARY**

*Chair: Joseph Mariani*

<b>Rediscovering 25 Years of Discoveries in Spoken Language Processing: A Preliminary ISCA Archive Analysis (ABSTRACT)</b> .....	3370
<i>J. Mariani, P. Paroubek, G. Francopoulo, M. Delaborde</i>	

### **ORAL SESSION 46: LANGUAGE MODELING FOR CONVERSATIONAL SPEECH**

*Chairs: Katrin Kirchhoff, Giuseppe Riccardi*

<b>Feature-Rich Sub-Lexical Language Models Using a Maximum Entropy Approach for German LVCSR</b> .....	3371
<i>M. Ali Basha Shaik, Amr El-Desoky Mousa, Ralf Schluter, Hermann Ney</i>	
<b>Morpheme Level Hierarchical Pitman-Yor Class-Based Language Models for LVCSR of Morphologically Rich Languages</b> .....	3376
<i>Amr El-Desoky Mousa, M. Ali Basha Shaik, Ralf Schluter, Hermann Ney</i>	
<b>Discriminatively Trained Dependency Language Modeling for Conversational Speech Recognition</b> .....	3381
<i>Benjamin Lambert, Bhiksha Raj, Rita Singh</i>	
<b>Prefix Tree Based N-best list Re-scoring for Recurrent Neural Network Language Model Used in Speech Recognition System</b> .....	3386
<i>Yujing Si, Qingqing Zhang, Ta Li, Jielin Pan, Yonghong Yan</i>	
<b>Cross-Domain Paraphrasing for Improving Language Modelling Using Out-of-Domain Data</b> .....	3391
<i>X. Liu, M.J.F. Gales, P.C. Woodland</i>	
<b>Viterbi Decoding for Latent Words Language Models Using Gibbs Sampling</b> .....	3396
<i>Ryo Masumura, Hirokazu Masataki, Takanobu Oba, Osamu Yoshioka, Satoshi Takahashi</i>	

### **ORAL SESSION 47: SPEECH ENHANCEMENT AND CODING**

*Chairs: Gernot Kubin, Isabel Trancoso*

<b>Computationally Efficient Objective Function for Algebraic Codebook Optimization in ACELP</b> .....	3401
<i>Tom Backstrom</i>	
<b>Speech Quality Prediction for Artificial Bandwidth Extension Algorithms</b> .....	3406
<i>Sebastian Moller, Emilia Kelaïdi, Friedemann Koster, Nicolas Cote, Patrick Bauer, Tim Fingscheidt, Thomas Schlien, Hannu Pulakka, Paavo Alku</i>	
<b>Speech Enhancement with Weighted Denoising Auto-Encoder</b> .....	3411
<i>Bing-yin Xia, Chang-chun Bao</i>	
<b>Syllable-Based Pitch Encoding for Low Bit Rate Speech Coding with Recognition/Synthesis Architecture</b> .....	3416
<i>Milos Cernak, Xingyu Na, Philip N. Garner</i>	

<b>Artificial Bandwidth Extension Based on Regularized Piecewise Linear Mapping with Discriminative Region Weighting and Long-Span Features</b> .....	3420
<i>Nguyen Duc Duy, Masayuki Suzuki, Nobuaki Minematsu, Keikichi Hirose</i>	
<b>Enhanced Muting Method in Packet Loss Concealment of ITU-T G.722 Employing Optimized Sigmoid Function</b> .....	3425
<i>Bong-Ki Lee, Chungsoo Lim, Jihwan Park, Joon-Hyuk Chang</i>	

## **ORAL SESSION 48: SPOKEN MACHINE TRANSLATION AND SPEECH NATURAL LANGUAGE PROCESSING**

Chairs: Satoshi Nakamura, Monika Wozczynna

<b>Automatic Human Utility Evaluation of ASR Systems: Does WER Really Predict Performance?</b> .....	3430
<i>Benoit Favre, Kyla Cheung, Siavash Kazemian, Adam Lee, Yang Liu, Cosmin Munteanu, Ani Nenkova, Dennis Ochei, Gerald Penn, Stephen Tratz, Clare Voss, Frauke Zeller</i>	
<b>Corpus Analysis of Simultaneous Interpretation Data for Improving Real Time Speech Translation</b> .....	3435
<i>Vivek Kumar Rangarajan Sridhar, John Chen, Srinivas Bangalore</i>	
<b>A Real-World System for Simultaneous Translation of German Lectures</b> .....	3440
<i>Eunah Cho, Christian Fugen, Teresa Hermann, Kevin Kilgour, Mohammed Mediani, Christian Mohr, Jan Niehues, Kay Rottmann, Christian Saam, Sebastian Stuker, Alex Waibel</i>	
<b>Freestyle: A Challenge-Response System for Hip Hop Lyrics via Unsupervised Induction of Stochastic Transduction Grammars</b> .....	3445
<i>Dekai Wu, Karteek Addanki, Markus Saers</i>	
<b>Toward Transfer of Acoustic Cues of Emphasis Across Languages</b> .....	3450
<i>Andreas Tsiartas, Panayiotis G. Georgiou, Shrikanth Narayanan</i>	
<b>Simple, Lexicalized Choice of Translation Timing for Simultaneous Speech Translation</b> .....	3454
<i>Tomoki Fujita, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura</i>	

## **ORAL SESSION 49: ASR — ROBUSTNESS AGAINST NOISE II**

Chairs: Li Deng, George Saon

<b>Noise Adaptive Training for Subspace Gaussian Mixture Models</b> .....	3459
<i>Liang Lu, Arnab Ghoshal, Steve Renals</i>	
<b>The IBM Speech Activity Detection System for the DARPA RATS Program</b> .....	3464
<i>George Saon, Samuel Thomas, Hagen Soltau, Sriram Ganapathy, Brian Kingsbury</i>	
<b>Conditional Emission Densities for Combining Speech Enhancement and Recognition Systems</b> .....	3469
<i>Armin Sehr, Takuya Yoshioka, Marc Delcroix, Keisuke Kinoshita, Tomohiro Nakatani, Roland Maas, Walter Kellermann</i>	
<b>Channel Selection Using N-Best Hypothesis for Multi-Microphone ASR</b> .....	3474
<i>Martin Wolf, Climent Nadeu</i>	
<b>Reverberant Speech Recognition Based on Denoising Autoencoder</b> .....	3479
<i>Takaaki Ishii, Hiroki Komiyama, Takahiro Shinozaki, Yasuo Horiuchi, Shingo Kuroiwa</i>	
<b>Adaptive Stereo-Based Stochastic Mapping</b> .....	3484
<i>Shay Maymon, Pierre Dognin, Xiaodong Cui, Vaibhava Goel</i>	

## **ORAL SESSION 50: ARTICULATORY AND ACOUSTIC CUES OF SPEECH PROSODY**

Chairs: Aijun Li, Oliver Niebuhr

<b>The Interplay of Intonation and Complex Lexical Tones: How Speaker Attitudes Affect the Realization of Glottalization on Vietnamese Sentence-Final Particles</b> .....	3489
<i>Thi-Lan Nguyen, Alexis Michaud, Do Dat Tran, Dang-Khoa Mac</i>	
<b>The Voice Prominence Hypothesis: The Interplay of F0 and Voice Source Features in Accentuation</b> .....	3494
<i>Ailbhe Ni Chasaide, Irena Yanushevskaya, John Kane, Christer Gobl</i>	
<b>Mora-Based Pre-Low Raising in Japanese Pitch Accent</b> .....	3499
<i>Albert Lee, Yi Xu, Santitham Prom-on</i>	
<b>Prosodic Cues of Sarcastic Speech in French: Slower, Higher, Wider</b> .....	3504
<i>Helene Loevenbruck, Mohamed Ameer Ben Jannet, Mariapaola D'Imperio, Mathilde Spini, Maud Champagne-Lavau</i>	
<b>Correlates of Contrastive Focus in Congenitally Blind Adults and Sighted Adults</b> .....	3509
<i>Lucie Menard, Annie Leclerc, Mark K. Tiede, Amelie Premont, Christine Turgeon, Pamela Trudeau-Fisette, Dominique Cote</i>	

<b>Is Protrusion of French Rounded Vowels Affected by Prosodic Positions?</b> .....	3514
<i>Laurianne Georgeton, Nicolas Audibert</i>	

### **SPECIAL SESSION 6: INTELLIGIBILITY-ENHANCING SPEECH MODIFICATIONS**

*Chairs: Martin Cooke, Catherine Mayo*

<b>Intelligibility-Enhancing Speech Modifications: The Hurricane Challenge</b> .....	3519
<i>Martin Cooke, Catherine Mayo, Cassia Valentini-Botinhao</i>	
<b>Statistical Synthesizer with Embedded Prosodic and Spectral Modifications to Generate Highly Intelligible Speech in Noise</b> .....	3524
<i>D. Erro, T.C. Zorila, Yannis Stylianou, E. Navas, I. Hernaez</i>	
<b>Lombard Modified Text-to-Speech Synthesis for Improved Intelligibility: Submission for the Hurricane Challenge 2013</b> .....	3529
<i>Antti Suni, Reima Karhila, Tuomo Raitio, Mikko Kurimo, Martti Vainio, Paavo Alku</i>	
<b>Combining Perceptually-Motivated Spectral Shaping with Loudness and Duration Modification for Intelligibility Enhancement of HMM-Based Synthetic Speech in Noise</b> .....	3534
<i>Cassia Valentini-Botinhao, Junichi Yamagishi, Simon King, Yannis Stylianou</i>	
<b>Increasing Speech Intelligibility via Spectral Shaping with Frequency Warping and Dynamic Range Compression Plus Transient Enhancement</b> .....	3539
<i>Elizabeth Godoy, Yannis Stylianou</i>	
<b>Improving Speech Intelligibility in Noise by SII-Dependent Preprocessing Using Frequency-Dependent Amplification and Dynamic Range Compression</b> .....	3544
<i>Henning Schepker, Jan Rannies, Simon Doclo</i>	
<b>SII-Based Speech Preprocessing for Intelligibility Improvement in Noise</b> .....	3549
<i>Cees H. Taal, Jesper Jensen</i>	
<b>Rephrasing-Based Speech Intelligibility Enhancement</b> .....	3554
<i>Mengqiu Zhang, Petko N. Petkov, W. Bastiaan Kleijn</i>	
<b>Information-Preserving Temporal Reallocation of Speech in the Presence of Fluctuating Maskers</b> .....	3559
<i>Vincent Aubanel, Martin Cooke</i>	
<b>Preservation of Speech Spectral Dynamics Enhances Intelligibility</b> .....	3564
<i>Petko N. Petkov, W. Bastiaan Kleijn</i>	
<b>An Overview of the VUB Entry for the 2013 Hurricane Challenge</b> .....	3569
<i>Henk Brouckxon, Werner Verhelst</i>	
<b>Improvement of Speech Intelligibility by Reallocation of Spectral Energy</b> .....	3572
<i>Reiko Takou, Nobumasa Seiyama, Atsushi Imai</i>	

### **POSTER SESSION 28: SPEECH TECHNOLOGY FOR SPEECH AND HEARING DISORDERS**

*Chair: Alain Ghio*

<b>Language-Universal Speech Audiometry with Automated Scoring</b> .....	3575
<i>Bart Vaerenberg, Louis ten Bosch, Wojtek Kowalczyk, Martine Coene, Herwig De Smet, Paul J. Govaerts</i>	
<b>Balancing Word Lists in Speech Audiometry Through Large Spoken Language Corpora</b> .....	3580
<i>Annemiek Hammer, Bart Vaerenberg, Wojtek Kowalczyk, Louis ten Bosch, Martine Coene, Paul J. Govaerts</i>	
<b>Developing an Information System for Deaf</b> .....	3584
<i>V. Lopez-Ludena, R. San-Segundo, J. Ferreiros, J.M. Pardo, E. Ferreira</i>	
<b>Dysarthric Speech Recognition Using Dysarthria-Severity-Dependent and Speaker-Adaptive Models</b> .....	3589
<i>Myung Jong Kim, Joohong Yoo, Hoirin Kim</i>	
<b>Voice Pathology Detection and Classification Using MPEG-7 Audio Low-Level Features</b> .....	3594
<i>Ghulam Muhammad, Moutasem Melhem</i>	
<b>Empirical Mode Decomposition-Based Spectral Acoustic Cues for Disordered Voices Analysis</b> .....	3599
<i>Abdellah Kacha, Francis Grenez, Jean Schoentgen</i>	
<b>Exemplar-Based Individuality-Preserving Voice Conversion for Articulation Disorders in Noisy Environments</b> .....	3604
<i>Ryo Aihara, Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Arika</i>	
<b>Combining In-Domain and Out-of-Domain Speech Data for Automatic Recognition of Disordered Speech</b> .....	3609
<i>H. Christensen, M.B. Aniol, Peter Bell, Phil D. Green, Thomas Hain, Simon King, Pawel Swietojanski</i>	

<b>Effects of Envelope Filter Cutoff Frequency on the Intelligibility of Mandarin Noise-Vocoded Speech in Babble Noise: Implications for Cochlear Implants</b> .....	3613
<i>Guangting Mai, James W. Minett, William S.-Y. Wang</i>	

## **POSTER SESSION 29: ROBUST SPEAKER RECOGNITION**

*Chair: Douglas O'shaughnessy*

<b>Multi-Session PLDA Scoring of I-Vector for Partially Open-Set Speaker Detection</b> .....	3618
<i>Kong Aik Lee, Anthony Larcher, Chang Huai You, Bin Ma, Haizhou Li</i>	
<b>Impact of Noise Reduction and Spectrum Estimation on Noise Robust Speaker Identification</b> .....	3623
<i>Keith W. Godin, Seyed Omid Sadjadi, John H.L. Hansen</i>	
<b>Improvement of Distant-Talking Speaker Identification Using Bottleneck Features of DNN</b> .....	3628
<i>Takanori Yamada, Longbiao Wang, Atsuhiko Kai</i>	
<b>Geometric Contamination for GMM/UBM Speaker Verification in Reverberant Environments</b> .....	3632
<i>Alessio Brutti, Maurizio Omologo</i>	
<b>Towards a More Efficient SVM Supervector Speaker Verification System Using Gaussian Reduction and a Tree-Structured Hash</b> .....	3637
<i>Richard D. McClanahan, Phillip L. De Leon</i>	
<b>Improving the PLDA Based Speaker Verification in Limited Microphone Data Conditions</b> .....	3641
<i>A. Kanagasundaram, D. Dean, Javier Gonzalez-Dominguez, S. Sridharan, D. Ramos, Joaquin Gonzalez-Rodriguez</i>	
<b>The I3A Speaker Recognition System for NIST SRE12: Post-Evaluation Analysis</b> .....	3646
<i>Jesus Villalba, Eduardo Lleida, Alfonso Ortega, Antonio Miguel</i>	
<b>Text-Dependent Speaker Recognition Using PLDA with Uncertainty Propagation</b> .....	3651
<i>T. Stafylakis, Patrick Kenny, P. Ouellet, J. Perez, M. Kockmann, Pierre Dumouchel</i>	
<b>Robust Speaker Recognition Using Spectro-Temporal Autoregressive Models</b> .....	3656
<i>Sri Harish Mallidi, Sriram Ganapathy, Hynek Hermansky</i>	
<b>Effect of Multicondition Training on i-Vector PLDA Configurations for Speaker Recognition</b> .....	3661
<i>Padmanabhan Rajan, Tomi Kinnunen, Ville Hautamaki</i>	
<b>Improving Robustness to Compressed Speech in Speaker Recognition</b> .....	3665
<i>Mitchell McLaren, Victor Abrash, Martin Graciarena, Yun Lei, Jan Pesan</i>	
<b>Modulation Features for Noise Robust Speaker Identification</b> .....	3670
<i>Vikramjit Mitra, Mitchell McLaren, Horacio Franco, Martin Graciarena, Nicolas Scheffer</i>	
<b>Minimax i-Vector Extractor for Short Duration Speaker Verification</b> .....	3675
<i>Ville Hautamaki, You-Chi Cheng, Padmanabhan Rajan, Chin-Hui Lee</i>	
<b>Standoff Speaker Recognition: Effects of Recording Distance Mismatch on Speaker Recognition System Performance</b> .....	3680
<i>Mike Fowler, Mark McCurry, Jonathan Bramsen, Kehinde Dunsin, Jeremiah Remus</i>	

## **POSTER SESSION 30: DIALOG SYSTEMS AND APPLICATIONS II**

*Chair: Helen Meng*

<b>Correlates to Intelligibility in Deviant Child Speech --- Comparing Clinical Evaluations to Audience Response System-Based Evaluations by Untrained Listeners</b> .....	3684
<i>Sofia Strombergsson, Christina Taanander</i>	
<b>Using Linguistic Analysis to Characterize Conceptual Units of Thought in Spoken Medical Narratives</b> .....	3689
<i>Kathryn Womack, Cecilia Ovesdotter Alm, Cara Calvelli, Jeff B. Pelz, Pengcheng Shi, Anne Haake</i>	
<b>Interacting with Robots via Speech and Gestures, an Integrated Architecture</b> .....	3694
<i>Francesco Cutugno, Alberto Finzi, Michelangelo Fiore, Enrico Leone, Silvia Rossi</i>	
<b>Incorporating Named Entity Recognition into the Speech Transcription Process</b> .....	3699
<i>Mohamed Hatmi, Christine Jacquin, Emmanuel Morin, Sylvain Meignier</i>	
<b>DTW-Distance-Ordered Spoken Term Detection</b> .....	3704
<i>Teppey Ohno, Tomoyosi Akiba</i>	
<b>Refining Sentence Similarity with Discourse Information in Dialog System</b> .....	3709
<i>Sangkeun Jung, Seung-Hoon Na</i>	
<b>Two-Step Correction of Speech Recognition Errors Based on N-Gram and Long Contextual Information</b> .....	3714
<i>Ryohei Nakatani, Tetsuya Takiguchi, Yasuo Ariki</i>	
<b>Inferring Actor Communities from Videos</b> .....	3718
<i>Sumit Negi, Rammath Balasubramanyan, Santanu Chaudhury</i>	



<b>Multiple Topic Identification in Telephone Conversations</b> .....	3723
<i>Xavier Bost, Marc El-Beze, Renato De Mori</i>	
<b>Variable-Span Out-of-Vocabulary Named Entity Detection</b> .....	3728
<i>Wei Chen, Sankaranarayanan Ananthakrishnan, Rohit Prasad, Prem Natarajan</i>	
<b>On the Feasibility of Using Pupil Diameter to Estimate Cognitive Load Changes for In-Vehicle Spoken Dialogues</b> .....	3733
<i>Andrew L. Kun, Oskar Palinko, Zeljko Medenica, Peter A. Heeman</i>	
<b>Investigation of Recurrent-Neural-Network Architectures and Learning Methods for Spoken Language Understanding</b> .....	3738
<i>Gregoire Mesnil, Xiaodong He, Li Deng, Yoshua Bengio</i>	
<b>Paraphrase Features to Improve Natural Language Understanding</b> .....	3743
<i>Xiaohu Liu, Ruhi Sarikaya, Chris Brockett, Chris Quirk, William B. Dolan</i>	
<b>A Weakly-Supervised Approach for Discovering New User Intents from Search Query Logs</b> .....	3747
<i>Dilek Hakkani-Tur, Asli Celikyilmaz, Larry Heck, Gokhan Tur</i>	
<b>Exploiting Shared Information for Multi-Intent Natural Language Sentence Classification</b> .....	3752
<i>Puyang Xu, Ruhi Sarikaya</i>	
<b>Author Index</b>	