

# **2009 IEEE Workshop on Automatic Speech Recognition & Understanding**

**(ASRU 2009)**

**Merano, Italy  
13-17 December 2009**



**IEEE Catalog Number: CFP09SRW-PRT  
ISBN: 978-1-4244-5478-5**

---

---

## Keynote Papers

---

- |    |   |
|----|---|
| 1  | <b>Generalization Problem in ASR Acoustic Model Training and Adaptation</b><br><i>(Sadaoki Furui)</i>           |
| 11 | <b>It's Not You, It's Me: Automatically Extracting Social Meaning from Speed Dates</b><br><i>(Dan Jurafsky)</i> |

---

---

## Invited Papers

---

- 12      **Toward Machine Translation with Statistics and Syntax and Semantics**  
*(Dekai Wu)*
- 22      **Audio-Visual Automatic Speech Recognition and Related Bimodal Speech Technologies:  
A Review of the State-of-the-Art and Open Problems**  
*(Gerasimos Potamianos)*
- 23      **Trends and Challenges in Language Modeling for Speech Recognition and Machine  
Translation**  
*(Holger Schwenk)*
- 24      **Manipulation of Consonants in Natural Speech**  
*(Jont Allen, Feipeng Li)*
- 25      **Spoken Dialogue Systems: Challenges, and Opportunities for Research**  
*(Jason D. Williams)*
- 26      **Voice-Based Information Retrieval — How Far are We from the Text-Based Information  
Retrieval?**  
*(Lin-shan Lee, Yi-cheng Pan)*
- 44      **Acoustic Modelling for Speech Recognition: Hidden Markov Models and Beyond?**  
*(M.J.F. Gales)*
- 45      **Online Discriminative Learning: Theory and Applications**  
*(Nicolò Cesa-Bianchi)*
- 46      **New Perspectives on Spoken Language Understanding: Does Machine Need to Fully  
Understand Speech?**  
*(Tatsuya Kawahara)*
- 51      **Rapid Language Adaptation Tools for Multilingual Speech Processing**  
*(Tanja Schultz)*

---

---

## Acoustic Features

---

- 52      **Investigations on Features for Log-Linear Acoustic Models in Continuous Speech Recognition**  
*(S. Wiesler, M. Nußbaum-Thom, Georg Heigold, Ralf Schlüter, Hermann Ney)*
- 58      **Leveraging Speech Production Knowledge for Improved Speech Recognition**  
*(Abhijeet Sangwan, John H.L. Hansen)*
- 64      **Optimal Quantization and Bit Allocation for Compressing Large Discriminative Feature Space Transforms**  
*(Etienne Marcheret, Vaibhava Goel, Peder A. Olsen)*
- 70      **Using Temporal Information for Improving Articulatory-Acoustic Feature Classification**  
*(Barbara Schuppler, Joost van Doremalen, Odette Scharenborg, Bert Cranen, Lou Boves)*
- 76      **Log-Linear Framework for Linear Feature Transformations in Speech Recognition**  
*(Muhammad Ali Tahir, Georg Heigold, Christian Plahl, Ralf Schlüter, Hermann Ney)*
- 82      **Multi-View Learning of Acoustic Features for Speaker Recognition**  
*(Karen Livescu, Mark Stoehr)*
- 87      **Large-Margin Feature Adaptation for Automatic Speech Recognition**  
*(Chih-Chieh Cheng, Fei Sha, Lawrence K. Saul)*
- 93      **Articulatory Feature Detection with Support Vector Machines for Integration into ASR and Phone Recognition**  
*(Upendra V. Chaudhari, Michael Picheny)*
- 99      **Transition Features for CRF-Based Speech Recognition and Boundary Detection**  
*(Spiros Dimopoulos, Eric Fosler-Lussier, Chin-Hui Lee, Alexandros Potamianos)*
- 103     **Short-Time Instantaneous Frequency and Bandwidth Features for Speech Recognition**  
*(Pirros Tsiakoulis, Alexandros Potamianos, Dimitrios Dimitriadis)*

---

---

## Acoustic Modeling

---

- 107      **Hidden Conditional Random Fields for Phone Recognition**  
*(Yun-Hsuan Sung, Dan Jurafsky)*
- 113      **Diagonal Priors for Full Covariance Speech Recognition**  
*(Peter Bell, Simon King)*
- 118      **A Study on Hidden Structural Model and its Application to Labeling Sequences**  
*(Yu Qiao, Masayuki Suzuki, Nobuaki Minematsu)*
- 124      **A Hierarchical Structure for Modeling Inter and Intra Phonetic Information for Phoneme Recognition**  
*(Daniel Vázquez, Guillermo Aradilla, Rainer Gruhn, Wolfgang Minker)*
- 130      **Automatic Selection of Recognition Errors by Respeaking the Intended Text**  
*(Keith Vertanen, Per Ola Kristensson)*
- 136      **Improving Online Incremental Speaker Adaptation with Eigen Feature Space MLLR**  
*(Xiaodong Cui, Jian Xue, Bowen Zhou)*
- 141      **Kernel Metric Learning for Phonetic Classification**  
*(Jui-Ting Huang, Xi Zhou, Mark Hasegawa-Johnson, Thomas Huang)*
- 146      **Three-Layer Optimizations for Fast GMM Computations on GPU-Like Parallel Processors**  
*(Kshitij Gupta, John D. Owens)*
- 152      **A Segmental CRF Approach to Large Vocabulary Continuous Speech Recognition**  
*(Geoffrey Zweig, Patrick Nguyen)*
- 158      **Generalized Likelihood Ratio Discriminant Analysis**  
*(Hung-Shin Lee, Berlin Chen)*

---

---

## Robustness in ASR I

---

- 164      **Temporal Envelope Subtraction for Robust Speech Recognition Using Modulation Spectrum**  
*(Sriram Ganapathy, Samuel Thomas, Hynek Hermansky)*
- 170      **Discriminative Adaptive Training with VTS and JUD**  
*(F. Flego, M.J.F. Gales)*
- 176      **Hierarchical Variational Loopy Belief Propagation for Multi-Talker Speech Recognition**  
*(Steven J. Rennie, John R. Hershey, Peder A. Olsen)*
- 182      **SNR Features for Automatic Speech Recognition**  
*(Philip N. Garner)*
- 188      **Power Function-Based Power Distribution Normalization Algorithm for Robust Speech Recognition**  
*(Chanwoo Kim, Richard M. Stern)*
- 194      **Mask Estimation Employing Posterior-Based Representative Mean for Missing-Feature Speech Recognition with Time-Varying Background Noise**  
*(Wooil Kim, John H.L. Hansen)*
- 199      **Noise Robust Model Adaptation Using Linear Spline Interpolation**  
*(Kaustubh Kalgaonkar, Michael L. Seltzer, Alex Acero)*
- 205      **Support Vector Machines for Noise Robust ASR**  
*(M.J.F. Gales, A. Ragni, H. AlDamarki, C. Gautier)*
- 211      **Generalized Cyclic Transformations in Speaker-Independent Speech Recognition**  
*(Florian Müller, Eugene Belilovsky, Alfred Mertins)*
- 216      **MLLR/MAP Adaptation Using Pronunciation Variation for Non-Native Speech Recognition**  
*(Yoo Rhee Oh, Hong Kook Kim)*

---

---

## Robustness in ASR II

---

- 222     **Improving Joint Uncertainty Decoding Performance by Predictive Methods for Noise Robust Speech Recognition**  
*(Haitian Xu, M.J.F. Gales, K.K. Chin)*
- 228     **Improved Decision Trees for Multi-Stream HMM-Based Audio-Visual Continuous Speech Recognition**  
*(Jing Huang, Karthik Visweswariah)*
- 232     **Extended Minimum Classification Error Training in Voice Activity Detection**  
*(Takayuki Arakawa, Haitham Al-Hassanieh, Masanori Tsujikawa, Ryosuke Isotani)*
- 237     **An Improved Parallel Model Combination Method for Noisy Speech Recognition**  
*(Hadi Veisi, Hossein Sameti)*
- 243     **Robust Speech Recognition Using a Small Power Boosting Algorithm**  
*(Chanwoo Kim, Kshitiz Kumar, Richard M. Stern)*
- 249     **Robust Distributed Speech Recognition Using Two-Stage Filtered Minima Controlled Recursive Averaging**  
*(Negar Ghourchian, Sid-Ahmed Selouani, Douglas O'Shaughnessy)*
- 255     **A Study on Hidden Markov Model's Generalization Capability for Speech Recognition**  
*(Xiong Xiao, Jinyu Li, Eng Siong Chng, Haizhou Li, Chin-Hui Lee)*
- 261     **Sub-Band Modulation Spectrum Compensation for Robust Speech Recognition**  
*(Wen-hsiang Tu, Sheng-Yuan Huang, Jieh-weih Hung)*
- 266     **An Improved Perceptual Speech Enhancement Technique Employing a Psychoacoustically Motivated Weighting Factor**  
*(Md. Jahangir Alam, Sid-Ahmed Selouani, Douglas O'Shaughnessy)*
- 271     **MAP Estimation of Online Mapping Parameters in Ensemble Speaker and Speaking Environment Modeling**  
*(Yu Tsao, Shigeki Matsuda, Satoshi Nakamura, Chin-Hui Lee)*

---

---

## Language Modeling and Search

---

- 276     **Dynamic Network Decoding Revisited**  
          (*Hagen Soltau, George Saon*)
- 282     **Iterative Decoding: A Novel Re-Scoring Framework for Confusion Networks**  
          (*Anoop Deoras, Frederick Jelinek*)
- 287     **Island-Driven Search Using Broad Phonetic Classes**  
          (*Tara N. Sainath*)
- 293     **A Multiplatform Speech Recognition Decoder Based on Weighted Finite-State Transducers**  
          (*Emilian Stoimenov, Tanja Schultz*)
- 299     **Scaling Shrinkage-Based Language Models**  
          (*Stanley F. Chen, Lidia Mangu, Bhuvana Ramabhadran, Ruhi Sarikaya, Abhinav Sethy*)
- 305     **Discriminative Training of n-Gram Language Models for Speech Recognition via Linear Programming**  
          (*Vladimir Magdin, Hui Jiang*)
- 311     **Constrained Discriminative Training of N-Gram Language Models**  
          (*Ariya Rastrow, Abhinav Sethy, Bhuvana Ramabhadran*)
- 317     **Self-Supervised Discriminative Training of Statistical Language Models**  
          (*Puyang Xu, Damianos Karakos, Sanjeev Khudanpur*)
- 323     **Towards the Use of Inferred Cognitive States in Language Modeling**  
          (*Nigel G. Ward, Alejandro Vega*)



---

---

## Automatic Speech Recognition Systems

---

- 327     **Syntactic Features for Arabic Speech Recognition**  
          (*Hong-Kwang Jeff Kuo, Lidia Mangu, Ahmad Emami, Imed Zitouni, Young-Suk Lee*)
- 333     **Vietnamese Large Vocabulary Continuous Speech Recognition**  
          (*Ngoc Thang Vu, Tanja Schultz*)
- 339     **The ESAT 2008 System for N-Best Dutch Speech Recognition Benchmark**  
          (*Kris Demuynck, Antti Puurula, Dirk Van Compernelle, Patrick Wambacq*)
- 345     **On Speeding Phoneme Recognition in a Hierarchical MLP Structure**  
          (*Daniel Vásquez, Guillermo Aradilla, Rainer Gruhn, Wolfgang Minker*)
- 349     **Robust Vocabulary Independent Keyword Spotting with Graphical Models**  
          (*Martin Wöllmer, Florian Eyben, Björn Schuller, Gerhard Rigoll*)
- 354     **Integrating Morphology into Automatic Speech Recognition**  
          (*Haşim Sak, Murat Saraçlar, Tunga Güngör*)
- 359     **An Exploration of Large Vocabulary Tools for Small Vocabulary Phonetic Recognition**  
          (*Tara N. Sainath, Bhuvana Ramabhadran, Michael Picheny*)
- 365     **MLP Based Hierarchical System for Task Adaptation in ASR**  
          (*Joel Pinto, Mathew Magimai-Doss, Hervé Bourlard*)
- 371     **Detection of OOV Words by Combining Acoustic Confidence Measures with Linguistic Features**  
          (*Frederik Stouten, Dominique Fohr, Irina Illina*)
- 376     **From Speech to Letters — Using a Novel Neural Network Architecture for Grapheme Based ASR**  
          (*Florian Eyben, Martin Wöllmer, Björn Schuller, Alex Graves*)

---

---

## Spoken Document Retrieval and Summarization

---

- 381      **Graph-Based Submodular Selection for Extractive Summarization**  
*(Hui Lin, Jeff Bilmes, Shasha Xie)*
- 387      **Integrating Prosodic Features in Extractive Meeting Summarization**  
*(Shasha Xie, Dilek Hakkani-Tür, Benoit Favre, Yang Liu)*
- 392      **Extractive Speech Summarization by Active Learning**  
*(Justin Jian Zhang, Ricky Ho Yin Chan, Pascale Fung)*
- 398      **Unsupervised Spoken Keyword Spotting via Segmental DTW on Gaussian Posteriorgrams**  
*(Yaodong Zhang, James R. Glass)*
- 404      **Query-by-Example Spoken Term Detection for OOV Terms**  
*(Carolina Parada, Abhinav Sethy, Bhuvana Ramabhadran)*
- 410      **Spoken Term Detection from Bilingual Spontaneous Speech Using Code-Switched Lattice-Based Structures for Words and Subword Units**  
*(Hung-Yi Lee, Yueh-Lien Tang, Hao Tang, Lin-shan Lee)*
- 416      **Improved Vocabulary Independent Search with Approximate Match Based on Conditional Random Fields**  
*(Upendra V. Chaudhari, Michael Picheny)*
- 421      **Query-By-Example Spoken Term Detection Using Phonetic Posteriorgram Templates**  
*(Timothy J. Hazen, Wade Shen, Christopher White)*
- 427      **Topic-Based Speaker Recognition for German Parliamentary Speeches**  
*(Doris Baum)*
- 432      **Robust Speaker Diarization for Short Speech Recordings**  
*(David Imseng, Gerald Friedland)*

---

---

## Spoken/Multimodal Language Understanding and Dialog Systems

---

- 438     **Ontology-Based Grounding of Spoken Language Understanding**  
          (*Silvia Quarteroni, Marco Dinarelli, Giuseppe Riccardi*)
- 444     **Active Learning for Rule-Based and Corpus-Based Spoken Language Understanding Models**  
          (*Pierre Gotab, Frédéric Béchet, Geraldine Damnati*)
- 450     **Representing the Reinforcement Learning State in a Negotiation Dialogue**  
          (*Peter A. Heeman*)
- 456     **Back-Off Action Selection in Summary Space-Based POMDP Dialogue Systems**  
          (*M. Gašić, F. Lefèvre, F. Jurčiček, S. Keizer, F. Mairesse, B. Thomson, K. Yu, S. Young*)
- 462     **Response Timing Generation and Response Type Selection for a Spontaneous Spoken Dialog System**  
          (*Ryota Nishimura, Seiichi Nakagawa*)
- 468     **Garbage Modeling with Decoys for a Sequential Recognition Scenario**  
          (*Michael Levit, Shuangyu Chang, Bruce Buntschuh*)
- 474     **Correlation-Based Query Relaxation for Example-Based Dialog Modeling**  
          (*Cheongjae Lee, Sungjin Lee, Sangkeun Jung, Kyungduk Kim, Donghyeon Lee, Gary Geunbae Lee*)
- 479     **The Exploration/Exploitation Trade-Off in Reinforcement Learning for Dialogue Management**  
          (*Sebastian Varges, Giuseppe Riccardi, Silvia Quarteroni, Alexei V. Ivanov*)
- 485     **Any Questions? Automatic Question Detection in Meetings**  
          (*Kofi Boakye, Benoit Favre, Dilek Hakkani-Tür*)
- 490     **Weighted Finite State Transducer Based Statistical Dialog Management**  
          (*Chiori Hori, Kiyonori Ohtake, Teruhisa Misu, Hideki Kashioka, Satoshi Nakamura*)

---

---

## Multilingual Speech Processing and Machine Translation

---

- 496      **Automatic Translation from Parallel Speech: Simultaneous Interpretation as MT Training Data**  
*(Matthias Paulik, Alex Waibel)*
- 502      **Reinforcing Language Model for Speech Translation with Auxiliary Data**  
*(Jia Cui, Yonggang Deng, Bowen Zhou)*
- 507      **The Asian Network-Based Speech-to-Speech Translation System**  
*(Sakriani Sakti, Noriyuki Kimura, Michael Paul, Chiori Hori, Eiichiro Sumita, Satoshi Nakamura, Jun Park, Chai Wutiwiwatchai, Bo Xu, Hammam Riza, Karunesh Arora, Chi Mai Luong, Haizhou Li)*
- 513      **Towards Integrated Machine Translation Using Structural Alignment from Syntax-Augmented Synchronous Parsing**  
*(Bing Xiang, Bowen Zhou, Martin Čmejrek)*
- 519      **Phone-to-Word Decoding Through Statistical Machine Translation and Complementary System Combination**  
*(D. Falavigna, M. Gerosa, R. Gretter, D. Giuliani)*
- 525      **Pronunciation Modeling for Dialectal Arabic Speech Recognition**  
*(Hassan Al-Haj, Roger Hsiao, Ian Lane, Alan W. Black, Alex Waibel)*

---

---

## Topics in Speech Recognition

---

- 529 **Speaker De-Identification via Voice Transformation**  
(*Qin Jin, Arthur R. Toth, Tanja Schultz, Alan W. Black*)
- 534 **Multilingual Speaker Age Recognition: Regression Analyses on the Lwazi Corpus**  
(*Michael Feld, Etienne Barnard, Charl van Heerden, Christian Müller*)
- 540 **Comparing Automatic Rich Transcription for Portuguese, Spanish and English Broadcast News**  
(*Fernando Batista, Isabel Trancoso, Nuno J. Mamede*)
- 546 **Discriminative Product-of-Expert Acoustic Mapping for Cross-Lingual Phone Recognition**  
(*Khe Chai Sim*)
- 552 **Acoustic Emotion Recognition: A Benchmark Comparison of Performances**  
(*Björn Schuller, Bogdan Vlasenko, Florian Eyben, Gerhard Rigoll, Andreas Wendemuth*)
- 558 **Local and Global Models for Spontaneous Speech Segment Detection and Characterization**  
(*Richard Dufour, Yannick Estève, Paul Deléglise, Frédéric Béchet*)
- 562 **Lexicon Adaptation for Subword Speech Recognition**  
(*Timo Mertens, Daniel Schneider, Arild Brandrud Næss, Torbjørn Svendsen*)
- 568 **Lattice-Based Lexical Cues for Word Fragment Detection in Conversational Speech**  
(*Kartik Audhkhasi, Panayiotis Georgiou, Shrikanth Narayanan*)
- 574 **Sub-structure-Based Estimation of Pronunciation Proficiency and Classification of Learners**  
(*Masayuki Suzuki, Nobuaki Minematsu, Dean Luo, Keikichi Hirose*)
- 580 **Automatic Detection of Vowel Pronunciation Errors Using Multiple Information Sources**  
(*Joost van Doremalen, Catia Cucchiari, Helmer Strik*)
- 586 **Automatic Punctuation Generation for Speech**  
(*Wenzhu Shen, Roger Peng Yu, Frank Seide, Ji Wu*)