

15th Annual Conference of the International Speech Communication Association (INTERSPEECH 2014)

Celebrating the Diversity of Spoken Languages

**Singapore
14 – 18 September 2014**

Volume 1 of 4

Editors:

**H. Li
P. Ching**

ISBN: 978-1-63439-435-2

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© (2014) 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

Additional copies of this publication are available from:

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

Table of Contents and Access to Abstracts

Multi-Lingual ASR

[Cutler, Anne / Zhang, Yu / Chuangsuwanich, Ekapol / Glass, James R.: "Language ID-based training of multilingual stacked bottleneck features", 1-5.](#)

[Do, Van Hai / Xiao, Xiong / Chng, Eng Siong / Li, Haizhou: "Kernel density-based acoustic model with cross-lingual bottleneck features for resource limited LVCSR", 6-10.](#)

[Vu, Ngoc Thang / Wang, Yuanfan / Klose, Marten / Mihaylova, Zlatka / Schultz, Tanja: "Improving ASR performance on non-native speech using multilingual and crosslingual information", 11-15.](#)

[Knill, Kate M. / Gales, Mark J. F. / Ragni, Anton / Rath, Shakti P.: "Language independent and unsupervised acoustic models for speech recognition and keyword spotting", 16-20.](#)

[Bell, Peter / Driesen, Joris / Renals, Steve: "Cross-lingual adaptation with multi-task adaptive networks", 21-25.](#)

[Razavi, Marzieh / Doss, Mathew Magimai: "On recognition of non-native speech using probabilistic lexical model", 26-30.](#)

Prosody Processing

[Tanaka, Kou / Toda, Tomoki / Neubig, Graham / Sakti, Sakriani / Nakamura, Satoshi: "Direct \$F_0\$ control of an electrolarynx based on statistical excitation feature prediction and its evaluation through simulation", 31-35.](#)

[Niekerk, Daniel R. van / Barnard, Etienne: "A target approximation intonation model for yorùbá TTS", 36-40.](#)

[Vadapalli, Anandaswarup / Prahallad, Kishore: "Learning continuous-valued word representations for phrase break prediction", 41-45.](#)

[Che, Hao / Tao, Jianhua / Li, Ya: "Improving Mandarin prosodic boundary prediction with rich syntactic features", 46-50.](#)

[Dall, Rasmus / Tomalin, Marcus / Wester, Mirjam / Byrne, William / King, Simon: "Investigating automatic & human filled pause insertion for speech synthesis", 51-55.](#)

[Dall, Rasmus / Wester, Mirjam / Corley, Martin: "The effect of filled pauses and speaking rate on speech comprehension in natural, vocoded and synthetic speech", 56-60.](#)

Speaker Recognition — Applications

[Khoury, Elie / Kinnunen, Tomi / Sizov, Aleksandr / Wu, Zhizheng / Marcel, Sébastien: "Introducing i-vectors for joint anti-spoofing and speaker verification", 61-65.](#)

[Leary, Ryan / Andrews, Walter: "Random projections for large-scale speaker search", 66-70.](#)

[Fredouille, Corinne / Charlet, Delphine: "Analysis of i-vector framework for speaker identification in TV-shows", 71-75.](#)

[Laurent, Antoine / Camelin, Nathalie / Raymond, Christian: "Boosting bonsai trees for efficient features combination: application to speaker role identification", 76-80.](#)

[Raimond, Yves / Nixon, Thomas: "Identifying contributors in the BBC world service archive", 81-85.](#)

[Kelly, Finnian / Saeidi, Rahim / Harte, Naomi / Leeuwen, David A. van: "Effect of long-term ageing on i-vector speaker verification", 86-90.](#)

Phonetics and Phonology 1, 2

[Versteegh, Maarten / Seidl, Amanda / Cristia, Alejandrina: "Acoustic correlates of phonological status", 91-95.](#)

[Airaksinen, Manu / Alku, Paavo: "Parameterization of the glottal source with the phase plane plot", 96-100.](#)

[Rose, Phil: "Transcribing tone — a likelihood-based quantitative evaluation of chao's tone letters", 101-105.](#)

[Hamzah, Diyana / German, James Sneed: "Intonational phonology and prosodic hierarchy in malay", 106-110.](#)

[Reichel, Uwe D. / Mády, Katalin: "Comparing parameterizations of pitch register and its discontinuities at prosodic boundaries for Hungarian", 111-115.](#)

[Christodoulides, George / Avanzi, Mathieu: "An evaluation of machine learning methods for prominence detection in French", 116-119.](#)

[Chen, Gang / Park, Soo Jin / Kreiman, Jody / Alwan, Abeer: "Investigating the effect of F0 and vocal intensity on harmonic magnitudes: data from high-speed laryngeal videoendoscopy", 1668-1672.](#)

[Delais-Roussarie, Elisabeth / Lolive, Damien / Yoo, Hiyon / Barbot, Nelly / Rosec, Olivier: "Adapting prosodic chunking algorithm and synthesis system to specific style: the case of dictation", 1673-1677.](#)

[Sung, Jae-Hyun: "The articulation of lexical and post-lexical palatalization in Korean", 1678-1682.](#)

[Archangeli, Diana / Johnston, Samuel / Sung, Jae-Hyun / Fisher, Muriel / Hammond, Michael / Carnie, Andrew: "Articulation and neutralization: a preliminary study of lenition in scottish gaelic", 1683-1687.](#)

[Amino, Kanae / Makinae, Hisanori / Kitamura, Tatsuya: "Nasality in speech and its contribution to speaker individuality", 1688-1692.](#)

[Brown, Jason / Matene, Eden: "Is speech rhythm an intrinsic property of language?", 1693-1697.](#)

[Jackschina, Anke / Schuppler, Barbara / Muhr, Rudolf: "Where /ar/ the /r/s in standard austrian German?", 1698-1702.](#)

[Hu, Fang / Zhang, Minghui: "Diphthongized vowels in the yi county hui Chinese dialect", 1703-1707.](#)

[Dellwo, Volker / Mok, Peggy / Jenny, Mathias: "Rhythmic variability between some asian languages: results from an automatic analysis of temporal characteristics", 1708-1711.](#)

[Braun, Angelika / Decker, Daniela: "Listener estimation of speaker age based on whispered speech", 1712-1716.](#)

[Kasisopa, Benjawan / Attina, Virginie / Burnham, Denis: "The Lombard effect with Thai lexical tones: an acoustic analysis of articulatory modifications in noise", 1717-1721.](#)

Open Domain Situated Conversational Interaction (Special Session)

[Pappu, Aasish / Rudnicky, Alexander I.: "Learning situated knowledge bases through dialog", 120-124.](#)

[Misu, Teruhisa: "Crowdsourcing for situated dialog systems in a moving car", 125-129.](#)

[Higashinaka, Ryuichiro / Meguro, Toyomi / Imamura, Kenji / Sugiyama, Hiroaki / Makino, Toshiro / Matsuo, Yoshihiro: "Evaluating coherence in open domain conversational systems", 130-134.](#)

[Bechet, Frederic / Nasr, Alexis / Favre, Benoit: "Adapting dependency parsing to spontaneous speech for open domain spoken language understanding", 135-139.](#)

[Gašić, M. / Kim, Dongho / Tsiakoulis, Pirros / Breslin, Catherine / Henderson, Matthew / Szummer, M. / Thomson, B. / Young, Steve: "Incremental on-line adaptation of POMDP-based dialogue managers to extended domains", 140-144.](#)

[Robichaud, Jean-Philippe / Crook, Paul A. / Xu, Puyang / Khan, Omar Zia / Sarikaya, Ruhi: "Hypotheses ranking for robust domain classification and tracking in dialogue systems", 145-149.](#)

Speech Production: Models and Acoustics

[Ramanarayanan, Vikram / Goldstein, Louis / Narayanan, Shrikanth S.: "Motor control primitives arising from a learned dynamical systems model of speech articulation", 150-154.](#)

[Yeh, Chia-Hsin / Wang, Chiung-Yao / Tu, Jung-Yueh: "Nonword repetition of taiwanese disyllabic tonal sequences in adults with language attrition", 155-158.](#)

[Windmann, Andreas / Šimko, Juraj / Wagner, Petra: "A unified account of prominence effects in an optimization-based model of speech timing", 159-163.](#)

[Kim, Jangwon / Lee, Sungbok / Narayanan, Shrikanth S.: "Estimation of the movement trajectories of non-crucial articulators based on the detection of crucial moments and physiological constraints", 164-168.](#)

[Sudhakar, Prasad / Ghosh, Prasanta Kumar: "Sparse smoothing of articulatory features from Gaussian mixture model based acoustic-to-articulatory inversion: benefit to speech recognition", 169-173.](#)

[Wang, Jun / Katz, William / Campbell, Thomas F.: "Contribution of tongue lateral to consonant production", 174-178.](#)

[Liu, Min / Shi, Shuju / Zhang, Jinsong: "A preliminary study on acoustic correlates of tone2+tone2 disyllabic word stress in Mandarin", 179-183.](#)

[Abuoudeh, Mohammad / Crouzet, Olivier: "Vowel length impact on locus equation parameters: an investigation on Jordanian Arabic", 184-188.](#)

[Roberts, Philip J. / Reetz, Henning / Lahiri, Aditi: "Corpus-testing a fricative discriminator; or, just how invariant is this invariant?", 189-192.](#)

[Bush, Brian O. / Kain, Alexander: "Modeling coarticulation in continuous speech", 193-197.](#)

[Daoudi, Khalid / Bertrac, Blaise: "On classification between normal and pathological voices using the MEEI-kayPENTAX database: issues and consequences", 198-202.](#)

[Bukmaier, Véronique / Harrington, Jonathan / Reubold, Ulrich / Kleber, Felicitas: "Synchronic variation in the articulation and the acoustics of the Polish three-way place distinction in sibilants and its implications for diachronic change", 203-207.](#)

Extraction of Para-Linguistic Information

[Gupta, Rahul / Georgiou, Panayiotis G. / Atkins, David C. / Narayanan, Shrikanth S.: "Predicting client's inclination towards target behavior change in motivational interviewing and investigating the role of laughter", 208-212.](#)

[Xiao, Bo / Bone, Daniel / Segbroeck, Maarten Van / Imel, Zac E. / Atkins, David C. / Georgiou, Panayiotis G. / Narayanan, Shrikanth S.: "Modeling therapist empathy through prosody in drug addiction counseling", 213-217.](#)

[Bone, Daniel / Lee, Chi-Chun / Potamianos, Alexandros / Narayanan, Shrikanth S.: "An investigation of vocal arousal dynamics in child-psychologist interactions using synchrony measures and a conversation-based model", 218-222.](#)

[Han, Kun / Yu, Dong / Tashev, Ivan: "Speech emotion recognition using deep neural network and extreme learning machine", 223-227.](#)

[Truong, Khiet P. / Westerhof, Gerben J. / Jong, Franciska de / Heylen, Dirk: "An annotation scheme for sighs in spontaneous dialogue", 228-232.](#)

[He, Lei / Dellwo, Volker: "Speaker idiosyncratic variability of intensity across syllables", 233-237.](#)

[Mariooryad, Soroosh / Lotfian, Reza / Busso, Carlos: "Building a naturalistic emotional speech corpus by retrieving expressive behaviors from existing speech corpora", 238-242.](#)

[Safavi, Saeid / Russell, Martin / Jančovič, Peter: "Identification of age-group from children's speech by computers and humans", 243-247.](#)

Spoken Language Understanding

[Morchid, Mohamed / Dufour, Richard / Bouallegue, Mohamed / Linarès, Georges / Mori, Renato De: "Theme identification in human-human conversations with features from specific speaker type hidden spaces", 248-252.](#)

[Marin, Alex / Holenstein, Roman / Sarikaya, Ruhi / Ostendorf, Mari: "Learning phrase patterns for text classification using a knowledge graph and unlabeled data", 253-257.](#)

[Xu, Puyang / Sarikaya, Ruhi: "Targeted feature dropout for robust slot filling in natural language understanding", 258-262.](#)

[Shiang, Sz-Rung / Lee, Hung-yi / Lee, Lin-shan: "Spoken question answering using tree-structured conditional random fields and two-layer random walk", 263-267.](#)

[Sarikaya, Ruhi / Celikyilmaz, Asli / Deoras, Anoop / Jeong, Minwoo: "Shrinkage based features for slot tagging with conditional random fields", 268-272.](#)

[Shi, Yangyang / Pan, Yi-Cheng / Hwang, Mei-Yuh: "Cluster based Chinese abbreviation modeling", 273-277.](#)

[Zhang, Xiantao / Li, Dongchen / Wu, Xihong: "Parsing named entity as syntactic structure", 278-282.](#)

[Tur, Gokhan / Deoras, Anoop / Hakkani-Tür, Dilek: "Detecting out-of-domain utterances addressed to a virtual personal assistant", 283-287.](#)

[Georgiladakis, Spiros / Unger, Christina / Iosif, Elias / Walter, Sebastian / Cimiano, Philipp / Petrakis, Euripides / Potamianos, Alexandros: "Fusion of knowledge-based and data-driven approaches to grammar induction", 288-292.](#)

[Katerenchuk, Denys / Rosenberg, Andrew: "Improving named entity recognition with prosodic features", 293-297.](#)

[Ravuri, Suman V. / Stolcke, Andreas: "Neural network models for lexical addressee detection", 298-302.](#)

[Freeman, Valerie / Chan, Julian / Levow, Gina-Anne / Wright, Richard / Ostendorf, Mari / Zayats, Victoria: "Manipulating stance and involvement using collaborative tasks: an exploratory comparison", 303-307.](#)

Spoken Dialogue Systems

[Ghigi, Fabrizio / Eskenazi, Maxine / Torres, M. Ines / Lee, Sungjin: "Incremental dialog processing in a task-oriented dialog", 308-312.](#)

[Hotta, Naoki / Komatani, Kazunori / Sato, Satoshi / Nakano, Mikio: "Detecting incorrectly-segmented utterances for posteriori restoration of turn-taking and ASR results", 313-317.](#)

[Hassan, Hany / Schwartz, Lee / Hakkani-Tür, Dilek / Tur, Gokhan: "Segmentation and disfluency removal for conversational speech translation", 318-322.](#)

[Watanabe, Shinji / Hershey, John R. / Marks, Tim K. / Fujii, Youichi / Koji, Yusuke: "Cost-level integration of statistical and rule-based dialog managers", 323-327.](#)

[Kim, Dongho / Breslin, Catherine / Tsiakoulis, Pirros / Gašić, M. / Henderson, Matthew / Young, Steve: "Inverse reinforcement learning for micro-turn management", 328-332.](#)

[Kane, John / Yanushevskaya, Irena / Looze, Céline de / Vaughan, Brian / Chasaide, Ailbhe Ní: "Analysing the prosodic characteristics of speech-chunks preceding silences in task-based interactions", 333-337.](#)

DNN Architectures and Robust Recognition

[Sak, Haşim / Senior, Andrew / Beaufays, Françoise: "Long short-term memory recurrent neural network architectures for large scale acoustic modeling", 338-342.](#)

[Saon, George / Soltau, Hagen / Emami, Ahmad / Picheny, Michael: "Unfolded recurrent neural networks for speech recognition", 343-347.](#)

[Tomar, Vikrant Singh / Rose, Richard C.: "Manifold regularized deep neural networks", 348-352.](#)

[Li, Bo / Sim, Khe Chai: "Modeling long temporal contexts for robust DNN-based speech recognition", 353-357.](#)

[Li, Feipeng / Nidadavolu, Phani S. / Hermansky, Hynek: "A long, deep and wide artificial neural net for robust speech recognition in unknown noise", 358-362.](#)

[Seps, Ladislav / Malek, Jiri / Cerva, Petr / Nouza, Jan: "Investigation of deep neural networks for robust recognition of nonlinearly distorted speech", 363-367.](#)

Speaker Recognition — Evaluation and Forensics

[Bansé, Désiré / Doddington, George R. / Garcia-Romero, Daniel / Godfrey, John J. / Greenberg, Craig S. / Martin, Alvin F. / McCree, Alan / Przybocki, Mark / Reynolds, Douglas A.: "Summary and initial results of the 2013-2014 speaker recognition i-vector machine learning challenge", 368-372.](#)

[Leeuwen, David A. van / Brümmer, Niko: "Constrained speaker linking", 373-377.](#)

[Novoselov, Sergey / Pekhovsky, Timur / Simonchik, Konstantin / Shulipa, Andrey: "RBM-PLDA subsystem for the NIST i-vector challenge", 378-382.](#)

[Shum, Stephen H. / Dehak, Najim / Glass, James R.: "Limited labels for unlimited data: active learning for speaker recognition", 383-387.](#)

[Brümmer, Niko / Swart, Albert: "Bayesian calibration for forensic evidence reporting", 388-392.](#)

[Ishihara, Shunichi: "Replicate mismatch between test/background and development databases: the impact on the performance of likelihood ratio-based forensic voice comparison", 393-397.](#)

Speech Production I, II

[Airaksinen, Manu / Bäckström, Tom / Alku, Paavo: "Automatic estimation of the lip radiation effect in glottal inverse filtering", 398-402.](#)

[Rosa, Marcelo de Oliveira: "Simulation of 3d larynges with asymmetric distribution of viscoelastic properties in their vocal folds", 403-407.](#)

[Takemoto, Hironori / Mokhtari, Parham / Kitamura, Tatsuya: "Comparison of vocal tract transfer functions calculated using one-dimensional and three-dimensional acoustic simulation methods", 408-412.](#)

[Kim, Jangwon / Erickson, Donna / Lee, Sungbok / Narayanan, Shrikanth S.: "A study of invariant properties and variation patterns in the converter/distributor model for emotional speech", 413-417.](#)

[Hewer, Alexander / Steiner, Ingmar / Wuhrer, Stefanie: "A hybrid approach to 3d tongue modeling from vocal tract MRI using unsupervised image segmentation and mesh deformation", 418-421.](#)

[Kaburagi, Tokihiko: "Estimation of vocal-tract shape from speech spectrum and speech resynthesis based on a generative model", 422-426.](#)

[Benítez, Andrés / Ramanarayanan, Vikram / Goldstein, Louis / Narayanan, Shrikanth S.: "A real-time MRI study of articulatory setting in second language speech", 701-705.](#)

[Arai, Takayuki: "Retroflex and bunched English /r/ with physical models of the human vocal tract", 706-710.](#)

[Rong, Panying / Yunusova, Yana / Berry, James D. / Zinman, Lorne / Green, Jordan R.: "Parameterization of articulatory pattern in speakers with ALS", 711-715.](#)

[Sujith P., Sujith P. / Ghosh, Prasanta Kumar: "Missing samples estimation in electromagnetic articulography data using equality constrained kalman smoother", 716-720.](#)

[Ji, An / Johnson, Michael T. / Berry, Jeff: "Palate-referenced articulatory features for acoustic-to-articulator inversion", 721-725.](#)

[Uchida, Hidetsugu / Wakamiya, Kohei / Kaburagi, Tokihiko: "A study on the improvement of measurement accuracy of the three-dimensional electromagnetic articulography", 726-730.](#)

INTERSPEECH 2014 Computational Paralinguistics Challenge (ComParE)

[Schuller, Björn / Steidl, Stefan / Batliner, Anton / Epps, Julien / Eyben, Florian / Ringeval, Fabien / Marchi, Erik / Zhang, Yue: "The INTERSPEECH 2014 computational paralinguistics challenge: cognitive & physical load", 427-431.](#)

[Pohjalainen, Jouni / Alku, Paavo: "Filtering and subspace selection for spectral features in detecting speech under physical stress", 432-436.](#)

[Li, Ming: "Automatic recognition of speaker physical load using posterior probability based features from acoustic and phonetic tokens", 437-441.](#)

[Kaya, Heysem / Özkaptan, Tuğçe / Salah, Albert Ali / Gürgen, Sadık Fikret: "Canonical correlation analysis and local fisher discriminant analysis based multi-view acoustic feature reduction for physical load prediction", 442-446.](#)

[Jing, How / Hu, Ting-Yao / Lee, Hung-Shin / Chen, Wei-Chen / Lee, Chi-Chun / Tsao, Yu / Wang, Hsin-Min: "Ensemble of machine learning algorithms for cognitive and physical speaker load detection", 447-451.](#)

[Gosztolya, Gábor / Grósz, Tamás / Busa-Fekete, Róbert / Tóth, László: "Detecting the intensity of cognitive and physical load using AdaBoost and deep rectifier neural networks", 452-456.](#)

[Montacié, Claude / Caraty, Marie-José: "High-level speech event analysis for cognitive load classification", 731-735.](#)

[Nwe, Tin Lay / Nguyen, Trung Hieu / Ma, Bin: "On the use of Bhattacharyya based GMM distance and neural net features for identification of cognitive load levels", 736-740.](#)

[Huckvale, Mark: "Prediction of cognitive load from speech with the VOQAL voice quality toolbox for the interspeech 2014 computational paralinguistics challenge", 741-745.](#)

[Kua, Jia Min Karen / Sethu, Vidhyasaharan / Le, Phu / Ambikairajah, Eliathamby: "The UNSW submission to INTERSPEECH 2014 compare cognitive load challenge", 746-750.](#)

[Segbroeck, Maarten Van / Travadi, Ruchir / Vaz, Colin / Kim, Jangwon / Black, Matthew P. / Potamianos, Alexandros / Narayanan, Shrikanth S.: "Classification of cognitive load from speech using an i-vector framework", 751-755.](#)

Hearing and Perception

[Iyer, Nandini / Thompson, Eric / Simpson, Brian / Romigh, Griffin: "Revisiting the right-ear advantage for speech: implications for speech displays", 457-461.](#)

[Bosch, L. ten / Ernestus, Miriam / Boves, Lou: "Comparing reaction time sequences from human participants and computational models", 462-466.](#)

[Andrei, Valentin / Cucu, Horia / Buzo, Andi / Burileanu, Corneliu: "Detecting the number of competing speakers — human selective hearing versus spectrogram distance based estimator", 467-470.](#)

[Li, Guo / Peng, Gang: "The influence of sensory memory and attention on the context effect in talker normalization", 471-475.](#)

[Lin, Payton / Chen, Fei / Wang, Syu Siang / Lai, Ying-Hui / Tsao, Yu: "Automatic speech recognition with primarily temporal envelope information", 476-480.](#)

[Lai, Ying-Hui / Chen, Fei / Tsao, Yu: "An adaptive envelope compression strategy for speech processing in cochlear implants", 481-484.](#)

[Helfer, Brian S. / Quatieri, Thomas F. / Williamson, James R. / Keyes, Laurel / Evans, Benjamin / Greene, W. Nicholas / Vian, Trina / Lacirignola, Joseph / Shenk, Trey / Talavage, Thomas / Palmer, Jeff / Heaton, Kristin: "Articulatory dynamics and coordination in classifying cognitive change with preclinical mTBI", 485-489.](#)

[Jinbo, Nozomi / Takamichi, Shinnosuke / Toda, Tomoki / Neubig, Graham / Sakti, Sakriani / Nakamura, Satoshi: "A hearing impairment simulation method using audiogram-based approximation of auditory characteristics", 490-494.](#)

[Wang, Dongmei / Kates, James M. / Hansen, John H. L.: "Investigation of the relative perceptual importance of temporal envelope and temporal fine structure between tonal and non-tonal languages", 495-498.](#)

[Fogerty, Daniel / Chen, Fei: "Vowel spectral contributions to English and Mandarin sentence intelligibility", 499-503.](#)

[Mittal, Vinay Kumar / Yegnanarayana, B.: "Significance of aperiodicity in the pitch perception of expressive voices", 504-508.](#)

Cross-Linguistic Studies

[Wester, Mirjam / Lecumberri, María Luisa García / Cooke, Martin: "DIAPIX-FL: a symmetric corpus of problem-solving dialogues in first and second languages", 509-513.](#)

[Coupé, Christophe / Oh, Yoon Mi / Pellegrino, François / Marsico, Egidio: "Cross-linguistic investigations of oral and silent reading", 514-518.](#)

[Coumans, Juul / Hout, Roeland van / Scharenborg, Odette: "Non-native word recognition in noise: the role of word-initial and word-final information", 519-523.](#)

[Wong, Janice Wing Sze: "The effects of high and low variability phonetic training on the perception and production of English vowels /e/-/æ/ by Cantonese ESL learners with high and low L2 proficiency levels", 524-528.](#)

[Burgos, Pepi / Jani, Mátyás / Cucchiaroni, Catia / Hout, Roeland van / Strik, Helmer: "Dutch vowel production by Spanish learners: duration and spectral features", 529-533.](#)

[Lengeris, Angelos / Nicolaidis, Katerina: "English consonant confusions by Greek listeners in quiet and noise and the role of phonological short-term memory", 534-538.](#)

[Detey, Sylvain / Racine, Isabelle / Eychenne, Julien / Kawaguchi, Yuji: "Corpus-based L2 phonological data and semi-automatic perceptual analysis: the case of nasal vowels produced by beginner Japanese learners of French", 539-543.](#)

[Pintér, Gábor / Mizuguchi, Shinobu / Tateishi, Koichi: "Perception of prosodic prominence and boundaries by L1 and L2 speakers of English", 544-547.](#)

[Kalathottukaren, Rose Thomas / Purdy, Suzanne C. / Ballard, Elaine: "Prosody perception, reading accuracy, nonliteral language comprehension, and music and tonal pitch discrimination in school aged children", 548-552.](#)

[Drozdova, Polina / Hout, Roeland van / Scharenborg, Odette: "Phoneme category retuning in a non-native language", 553-557.](#)

[Chiou, Bo-Chang / Chen, Chia-Ping: "Speech emotion recognition with cross-lingual databases", 558-561.](#)

Speaker Diarization

[Inoue, Koji / Wakabayashi, Yukoh / Yoshimoto, Hiromasa / Kawahara, Tatsuya: "Speaker diarization using eye-gaze information in multi-party conversations", 562-566.](#)

[Huang, Che-Wei / Xiao, Bo / Georgiou, Panayiotis G. / Narayanan, Shrikanth S.: "Unsupervised speaker diarization using riemannian manifold clustering", 567-571.](#)

[Delgado, Héctor / Fredouille, Corinne / Serrano, Javier: "Towards a complete binary key system for the speaker diarization task", 572-576.](#)

[Ghaemmaghami, Houman / Dean, David / Sridharan, Sridha: "An iterative speaker re-diarization scheme for improving speaker-based entity extraction in multimedia archives", 577-581.](#)

[Gebre, Binyam Gebrekidan / Wittenburg, Peter / Drude, Sebastian / Huijbregts, Marijn / Heskes, Tom: "Speaker diarization using gesture and speech", 582-586.](#)

[Dupuy, Grégor / Meignier, Sylvain / Estève, Yannick: "Is incremental cross-show speaker diarization efficient for processing large volumes of data?", 587-591.](#)

[Dighe, Pranay / Ferràs, Marc / Boulard, Hervé: "Detecting and labeling speakers on overlapping speech using vector taylor series", 592-596.](#)

[Yella, Sree Harsha / Motlicek, Petr / Boulard, Hervé: "Phoneme background model for information bottleneck based speaker diarization", 597-601.](#)

[Ferràs, Marc / Masneri, Stefano / Schreer, Oliver / Boulard, Hervé: "Diarizing large corpora using multi-modal speaker linking", 602-606.](#)

[Bechet, Frederic / Bendris, Meriem / Charlet, Delphine / Damnati, Géraldine / Favre, Benoit / Rouvier, Mickael / Auguste, Remi / Bigot, Benjamin / Dufour, Richard / Fredouille, Corinne / Linarès, Georges / Martinet, Jean / Senay, Gregory / Tirilly, Pierre: "Multimodal understanding for person recognition in video broadcasts", 607-611.](#)

Robust ASR 1, 2

[Gibson, James / Segbroeck, Maarten Van / Narayanan, Shrikanth S.: "Comparing time-frequency representations for directional derivative features", 612-615.](#)

[Du, Jun / Wang, Qing / Gao, Tian / Xu, Yong / Dai, Li-Rong / Lee, Chin-Hui: "Robust speech recognition with speech enhanced deep neural networks", 616-620.](#)

[Vincent, Emmanuel / Gkiokas, Aggelos / Schnitzer, Dominik / Flexer, Arthur: "An investigation of likelihood normalization for robust ASR", 621-625.](#)

[Spille, Constantin / Meyer, Bernd T.: "Identifying the human-machine differences in complex binaural scenes: what can be learned from our auditory system", 626-630.](#)

[Geiger, Jürgen T. / Zhang, Zixing / Weninger, Felix / Schuller, Björn / Rigoll, Gerhard: "Robust speech recognition using long short-term memory recurrent neural networks for hybrid acoustic modelling", 631-635.](#)

[Liu, Shilin / Sim, Khe Chai: "Joint adaptation and adaptive training of TVWR for robust automatic speech recognition", 636-640.](#)

[Park, Hyung-Min / Maciejewski, Matthew / Kim, Chanwoo / Stern, Richard M.: "Robust speech recognition in reverberant environments using subband-based steady-state monaural and binaural suppression", 2715-2718.](#)

[Zhao, Rui / Li, Jinyu / Gong, Yifan: "Variable-component deep neural network for robust speech recognition", 2719-2723.](#)

[Kao, Yu-Chen / Wang, Yi-Ting / Chen, Berlin: "Effective modulation spectrum factorization for robust speech recognition", 2724-2728.](#)

[Ravuri, Suman V.: "Hybrid MLP/structured-SVM tandem systems for large vocabulary and robust ASR", 2729-2733.](#)

[Kim, Chanwoo / Chin, Kean K. / Bacchiani, Michiel / Stern, Richard M.: "Robust speech recognition using temporal masking and thresholding algorithm", 2734-2738.](#)

[Xie, Xurong / Su, Rongfeng / Liu, Xunying / Wang, Lan: "Deep neural network bottleneck features for generalized variable parameter HMMs", 2739-2743.](#)

[Bu, Suliang / Qian, Yanmin / Yu, Kai: "A novel dynamic parameters calculation approach for model compensation", 2744-2748.](#)

[Hashimoto, Naoaki / Nakano, Shoichi / Yamamoto, Kazumasa / Nakagawa, Seiichi: "Speech recognition based on Itakura-Saito divergence and dynamics/sparseness constraints from mixed sound of speech and music by non-negative matrix factorization", 2749-2753.](#)

[Chung, Yong-Joo: "Noise robust speech recognition based on noise-adapted HMMs using speech feature compensation", 2754-2758.](#)

[Alam, M. J. / Kenny, Patrick / Dumouchel, Pierre / O'Shaughnessy, Douglas: "Noise spectrum estimation using Gaussian mixture model-based speech presence probability for robust speech recognition", 2759-2763.](#)

Implementation of Language Model Algorithms

[Chen, X. / Wang, Y. / Liu, X. / Gales, Mark J. F. / Woodland, Philip C.: "Efficient GPU-based training of recurrent neural network language models using spliced sentence bunch", 641-645.](#)

[Nolden, David / Schlüter, Ralf / Ney, Hermann: "Word pair approximation for more efficient decoding with high-order language models", 646-650.](#)

[Adel, Heike / Kirchhoff, Katrin / Vu, Ngoc Thang / Telaar, Dominic / Schultz, Tanja: "Comparing approaches to convert recurrent neural networks into backoff language models for efficient decoding", 651-655.](#)

[Nolden, David / Soltau, Hagen / Povey, Daniel / Ghahremani, Pegah / Mangu, Lidia / Ney, Hermann: "Removing redundancy from lattices", 656-660.](#)

[Sundermeyer, Martin / Tüske, Zoltán / Schlüter, Ralf / Ney, Hermann: "Lattice decoding and rescoring with long-Span neural network language models", 661-665.](#)

[Levit, Michael / Parthasarathy, Sarangarajan / Chang, Shuangyu / Stolcke, Andreas / Dumoulin, Benoît: "Word-phrase-entity language models: getting more mileage out of n-grams", 666-670.](#)

Speaker Recognition — Noise and Channel Robustness

[Sarkar, Sourjya / Rao, K. Sreenivasa: "A novel boosting algorithm for improved i-vector based speaker verification in noisy environments", 671-675.](#)

[Campbell, W. M.: "Using deep belief networks for vector-based speaker recognition", 676-680.](#)

[Lei, Yun / Ferrer, Luciana / McLaren, Mitchell / Scheffer, Nicolas: "A deep neural network speaker verification system targeting microphone speech", 681-685.](#)

[McLaren, Mitchell / Lei, Yun / Scheffer, Nicolas / Ferrer, Luciana: "Application of convolutional neural networks to speaker recognition in noisy conditions", 686-690.](#)

[Pelecanos, Jason / Zhu, Weizhong / Yaman, Sibel: "SVM based speaker recognition: harnessing trials with multiple enrollment sessions", 691-695.](#)

[Gallardo, Laura Fernández / Wagner, Michael / Möller, Sebastian: "I-vector speaker verification based on phonetic information under transmission channel effects", 696-700.](#)

Speech Synthesis I-III

[Zang, Xiao / Wu, Zhiyong / Meng, Helen / Jia, Jia / Cai, Lianhong: "Using conditional random fields to predict focus word pair in spontaneous spoken English", 756-760.](#)

[Sproat, Richard / Hall, Keith: "Applications of maximum entropy rankers to problems in spoken language processing", 761-764.](#)

[Gonzalvo, Xavi / Podsiadło, Monika: "Text-to-speech with cross-lingual neural network-based grapheme-to-phoneme models", 765-769.](#)

[Nagahama, Daiki / Nose, Takashi / Koriyama, Tomoki / Kobayashi, Takao: "Transform mapping using shared decision tree context clustering for HMM-based cross-lingual speech synthesis", 770-774.](#)

[Ramani, B. / Jeeva, M. P. Actlin / Vijayalakshmi, P. / Nagarajan, T.: "Cross-lingual voice conversion-based polyglot speech synthesizer for indian languages", 775-779.](#)

[Hu, Qiong / Stylianou, Yannis / Maia, Ranniery / Richmond, Korin / Yamagishi, Junichi / Latorre, Javier: "An investigation of the application of dynamic sinusoidal models to statistical parametric speech synthesis", 780-784.](#)

[Patil, Hemant A. / Patel, Tanvina B.: "Chaotic mixed excitation source for speech synthesis", 785-789.](#)

[Sorin, Alexander / Shechtman, Slava / Pollet, Vincent: "Refined inter-segment joining in multi-form speech synthesis", 790-794.](#)

[Zhang, Ran / Wen, Zhengqi / Tao, Jianhua / Li, Ya / Liu, Bing / Lou, Xiaoyan: "A hierarchical viterbi algorithm for Mandarin hybrid speech synthesis system", 795-799.](#)

[Fabre, Diandra / Hueber, Thomas / Badin, Pierre: "Automatic animation of an articulatory tongue model from ultrasound images using Gaussian mixture regression", 2293-2297.](#)

[Tobing, Patrick Lumban / Toda, Tomoki / Neubig, Graham / Sakti, Sakriani / Nakamura, Satoshi / Purwarianti, Ayu: "Articulatory controllable speech modification based on statistical feature mapping with Gaussian mixture models", 2298-2302.](#)

[Ding, Chuang / Zhu, Pengcheng / Xie, Lei / Jiang, Dongmei / Fu, Zhong-Hua: "Speech-driven head motion synthesis using neural networks", 2303-2307.](#)

[Song, Peng / Jin, Yun / Zheng, Wenming / Zhao, Li: "Text-independent voice conversion using speaker model alignment method from non-parallel speech", 2308-2312.](#)

[Chen, Ling-Hui / Ling, Zhen-Hua / Dai, Li-Rong: "Voice conversion using generative trained deep neural networks with multiple frame spectral envelopes", 2313-2317.](#)

[Sanchez, Gerard / Silen, Hanna / Nurminen, Jani / Gabbouj, Moncef: "Hierarchical modeling of F0 contours for voice conversion", 2318-2321.](#)

[Kadowaki, Kento / Ishihara, Tatsuma / Hojo, Nobukatsu / Kameoka, Hirokazu: "Speech prosody generation for text-to-speech synthesis based on generative model of F₀ contours", 2322-2326.](#)

[Chen, Xiayu / Zhang, Yang / Hasegawa-Johnson, Mark: "An iterative approach to decision tree training for context dependent speech synthesis", 2327-2331.](#)

[Nguyen, Thi Thu Trang / Rilliard, Albert / Tran, Do Dat / d'Alessandro, Christophe: "Prosodic phrasing modeling for vietnamese TTS using syntactic information", 2332-2336.](#)

[Koriyama, Tomoki / Suzuki, Hiroshi / Nose, Takashi / Shinozaki, Takahiro / Kobayashi, Takao: "Accent type and phrase boundary estimation using acoustic and language models for automatic prosodic labeling", 2337-2341.](#)

[Fang, Qiang / Wei, Jianguo / Hu, Fang: "Reconstruction of mistracked articulatory trajectories", 2342-2345.](#)

[Chen, Langzhou / Braunschweiler, Norbert: "Enabling controllability for continuous expression space", 2912-2916.](#)

[Nose, Takashi / Ito, Akinori: "Analysis of spectral enhancement using global variance in HMM-based speech synthesis", 2917-2921.](#)

[Valentini-Botinhao, Cassia / Toman, Markus / Pucher, Michael / Schabus, Dietmar / Yamagishi, Junichi: "Intelligibility analysis of fast synthesized speech", 2922-2926.](#)

[López-Peláez, Susana Palmaz / Clark, Robert A. J.: "Speech synthesis reactive to dynamic noise environmental conditions", 2927-2931.](#)

[Baumann, Timo: "Partial representations improve the prosody of incremental speech synthesis", 2932-2936.](#)

[Tsiakoulis, Pirros / Breslin, Catherine / Gašić, M. / Henderson, Matthew / Kim, Dongho / Young, Steve: "Dialogue context sensitive speech synthesis using factorized decision trees", 2937-2941.](#)

[Wang, Xin / Ling, Zhen-Hua / Dai, Li-Rong: "Concept-to-speech generation by integrating syntagmatic features into HMM-based speech synthesis", 2942-2946.](#)

[Gowda, Dhananjaya / Kallasjoki, Heikki / Karhila, Reima / Contan, Cristian / Palomäki, Kalle / Giurgiu, Mircea / Kurimo, Mikko: "On the role of missing data imputation and NMF feature enhancement in building synthetic voices using reverberant speech", 2947-2951.](#)

[Do, C. -T. / Evrard, M. / Leman, A. / d'Alessandro, Christophe / Rilliard, Albert / Crebouw, J. -L.: "Objective evaluation of HMM-based speech synthesis system using kullback-leibler divergence", 2952-2956.](#)

[Latorre, Javier / Yanagisawa, Kayoko / Wan, Vincent / Kolluru, BalaKrishna / Gales, Mark J. F.: "Speech intonation for TTS: study on evaluation methodology", 2957-2961.](#)

Multi-Lingual Cross-Lingual and Low-Resource ASR

[Miao, Yajie / Metze, Florian: "Improving language-universal feature extraction with deep maxout and convolutional neural networks", 800-804.](#)

[Fernandez, Raul / Cui, Jia / Rosenberg, Andrew / Ramabhadran, Bhuvana / Cui, Xiaodong: "Exploiting vocal-source features to improve ASR accuracy for low-resource languages", 805-809.](#)

[Ragni, Anton / Knill, Kate M. / Rath, Shakti P. / Gales, Mark J. F.: "Data augmentation for low resource languages", 810-814.](#)

[Jouvet, Denis / Fohr, Dominique: "About combining forward and backward-based decoders for selecting data for unsupervised training of acoustic models", 815-819.](#)

[Grézl, František / Karafiát, Martin: "Combination of multilingual and semi-supervised training for under-resourced languages", 820-824.](#)

[Vu, Ngoc Thang / Weiner, Jochen / Schultz, Tanja: "Investigating the learning effect of multilingual bottle-neck features for ASR", 825-829.](#)

[Miao, Yajie / Zhang, Hao / Metze, Florian: "Distributed learning of multilingual DNN feature extractors using GPUs", 830-834.](#)

[Rath, Shakti P. / Knill, Kate M. / Ragni, Anton / Gales, Mark J. F.: "Combining tandem and hybrid systems for improved speech recognition and keyword spotting on low resource languages", 835-839.](#)

[Cui, Jia / Ramabhadran, Bhuvana / Cui, Xiaodong / Rosenberg, Andrew / Kingsbury, Brian / Sethy, Abhinav: "Recent improvements in neural network acoustic modeling for LVCSR in low resource languages", 840-844.](#)

[Huang, Yan / Slaney, Malcolm / Seltzer, Michael L. / Gong, Yifan: "Towards better performance with heterogeneous training data in acoustic modeling using deep neural networks", 845-849.](#)

Speech Estimation and Sound Source Separation

[Higuchi, Takuya / Takeda, Hirofumi / Nakamura, Tomohiko / Kameoka, Hirokazu: "A unified approach for underdetermined blind signal separation and source activity detection by multichannel factorial hidden Markov models", 850-854.](#)

[Vaz, Colin / Dimitriadis, Dimitrios / Narayanan, Shrikanth S.: "Enhancing audio source separability using spectro-temporal regularization with NMF", 855-859.](#)

[Mirzaei, Sayeh / Van hamme, Hugo / Norouzi, Yaser: "Blind speech source localization, counting and separation for 2-channel convolutive mixtures in a reverberant environment", 860-864.](#)

[Weninger, Felix / Roux, Jonathan Le / Hershey, John R. / Watanabe, Shinji: "Discriminative NMF and its application to single-channel source separation", 865-869.](#)

[Kawahara, Hideki / Kitamura, Tatsuya / Takemoto, Hironori / Nisimura, Ryuichi / Irino, Toshio: "Vocal tract length estimation based on vowels using a database consisting of 385 speakers and a database with MRI-based vocal tract shape information", 870-874.](#)

[Wang, Haipeng / Lee, Tan / Leung, Cheung-Chi / Ma, Bin / Li, Haizhou: "A graph-based Gaussian component clustering approach to unsupervised acoustic modeling", 875-879.](#)

[Ziaei, Ali / Sangwan, Abhijeet / Hansen, John H. L.: "A speech system for estimating daily word counts", 880-884.](#)

[Lu, Xugang / Tsao, Yu / Matsuda, Shigeki / Hori, Chiori: "Ensemble modeling of denoising autoencoder for speech spectrum restoration", 885-889.](#)

Feature Extraction and Modeling for ASR 1, 2

[Tüske, Zoltán / Golik, Pavel / Schlüter, Ralf / Ney, Hermann: "Acoustic modeling with deep neural networks using raw time signal for LVCSR", 890-894.](#)

[Mitra, Vikramjit / Wang, Wen / Franco, Horacio / Lei, Yun / Bartels, Chris / Graciarena, Martin: "Evaluating robust features on deep neural networks for speech recognition in noisy and channel mismatched conditions", 895-899.](#)

[Sainath, Tara N. / Peddinti, Vijayaditya / Kingsbury, Brian / Fousek, Petr / Ramabhadran, Bhuvana / Nahamoo, David: "Deep scattering spectra with deep neural networks for LVCSR tasks", 900-904.](#)

[Chang, Shuo-Yiin / Morgan, Nelson: "Robust CNN-based speech recognition with Gabor filter kernels", 905-909.](#)

[Lu, Liang / Renals, Steve: "Probabilistic linear discriminant analysis with bottleneck features for speech recognition", 910-914.](#)

[Schatz, Thomas / Peddinti, Vijayaditya / Cao, Xuan-Nga / Bach, Francis / Hermansky, Hynek / Dupoux, Emmanuel: "Evaluating speech features with the minimal-pair ABX task \(II\): resistance to noise", 915-919.](#)

[Geiger, Jürgen T. / Gemmeke, Jort F. / Schuller, Björn / Rigoll, Gerhard: "Investigating NMF speech enhancement for neural network based acoustic models", 2405-2409.](#)

[Lilley, Jason / Mahshie, James / Bunnell, H. Timothy: "Automatic speech feature classification for children with cochlear implants", 2410-2414.](#)

[Tachioka, Yuuki / Watanabe, Shinji / Roux, Jonathan Le / Hershey, John R.: "Sequential maximum mutual information linear discriminant analysis for speech recognition", 2415-2419.](#)

[Ghaffarzadegan, Shabnam / Bořil, Hynek / Hansen, John H. L.: "Model and feature based compensation for whispered speech recognition", 2420-2424.](#)

[Moghimi, Amir R. / Raj, Bhiksha / Stern, Richard M.: "Post-masking: a hybrid approach to array processing for speech recognition", 2425-2429.](#)

[de-la-Calle-Silos, F. / Valverde-Albacete, F. J. / Gallardo-Antolín, A. / Peláez-Moreno, C.: "ASR feature extraction with morphologically-filtered power-normalized cochleograms", 2430-2434.](#)

[Martinez, Angel Mario Castro / Moritz, Niko / Meyer, Bernd T.: "Should deep neural nets have ears? the role of auditory features in deep learning approaches", 2435-2439.](#)

[Fox, Charles / Hain, Thomas: "Extending Limbeam with discrimination and coarse gradients", 2440-2444.](#)

[Mukherjee, Sankar / Mandal, Shyamal Kumar Das: "Generation of F0 contour using deep boltzmann machine and twin Gaussian process hybrid model for bengali language", 2445-2449.](#)

[Morales-Cordovilla, Juan A. / Pessentheiner, Hannes / Hagmüller, Martin / Kubin, Gernot: "Room localization for distant speech recognition", 2450-2453.](#)

[Bahaadini, Sara / Asaei, Afsaneh / Imseng, David / Boulard, Hervé: "Posterior-based sparse representation for automatic speech recognition", 2454-2458.](#)

Speech Analysis I, II

[Tabain, Marija / Butcher, Andrew / Breen, Gavan / Beare, Richard: "Lateral formants in three central australian languages", 920-924.](#)

[Khasanova, Alina / Cole, Jennifer / Hasegawa-Johnson, Mark: "Detecting articulatory compensation in acoustic data through linear regression modeling", 925-929.](#)

[Guo, Jinxi / Liu, Angli / Arsikere, Harish / Alwan, Abeer / Lulich, Steven M.: "The relationship between the second subglottal resonance and vowel class, standing height, trunk length, and F0 variation for Mandarin speakers", 930-934.](#)

[Meenakshi, Nisha / Yarra, Chiranjeevi / Yamini, B. K. / Ghosh, Prasanta Kumar: "Comparison of speech quality with and without sensors in electromagnetic articulograph AG 501 recording", 935-939.](#)

[Albuquerque, Luciana / Oliveira, Catarina / Teixeira, António / Sa-Couto, Pedro / Freitas, João / Dias, Miquel Sales: "Impact of age in the production of European Portuguese vowels", 940-944.](#)

[Yu, Chengzhu / Hansen, John H. L. / Oard, Douglas W.: "'houston, we have a solution!': a case study of the analysis of astronaut speech during NASA apollo 11 for long-term speaker modeling", 945-948.](#)

[Luan, Yi / Wright, Richard / Ostendorf, Mari / Levow, Gina-Anne: "Relating automatic vowel space estimates to talker intelligibility", 2238-2242.](#)

[Kawahara, Hideki / Morise, Masanori / Toda, Tomoki / Banno, Hideki / Nisimura, Ryuichi / Irino, Toshio: "Excitation source analysis for high-quality speech manipulation systems based on an interference-free representation of group delay with minimum phase response compensation", 2243-2247.](#)

[Pedersen, Christian Fischer / Bäckström, Tom: "Sparse time-frequency representation of speech by the vandermonde transform", 2248-2252.](#)

[Nandwana, Mahesh Kumar / Hansen, John H. L.: "Analysis and identification of human scream: implications for speaker recognition", 2253-2257.](#)

[Wang, Dongmei / Loizou, Philipos C. / Hansen, John H. L.: "F0 estimation in noisy speech based on long-term harmonic feature analysis combined with neural network classification", 2258-2262.](#)

[Slaney, Malcolm / Seltzer, Michael L.: "The influence of pitch and noise on the discriminability of filterbank features", 2263-2267.](#)

Speech Technologies and Applications

[Harwath, David / Gruenstein, Alexander / McGraw, Ian: "Choosing useful word alternates for automatic speech recognition correction interfaces", 949-953.](#)

[Chen, X. / Gales, Mark J. F. / Knill, Kate M. / Breslin, Catherine / Chen, Langzhou / Chin, K. K. / Wan, Vincent: "An initial investigation of long-term adaptation for meeting transcription", 954-958.](#)

[Ng, Tim / Hsiao, Roger / Zhang, Le / Karakos, Damianos / Mallidi, Sri Harish / Karafiát, Martin / Veselý, Karel / Szőke, Igor / Zhang, Bing / Nguyen, Long / Schwartz, Richard: "Progress in the BBN keyword search system for the DARPA RATS program", 959-963.](#)

[Nouza, Jan / Cerva, Petr / Zdansky, Jindrich / Blavka, Karel / Bohac, Marek / Silovsky, Jan / Chaloupka, Josef / Kucharova, Michaela / Seps, Ladislav / Malek, Jiri / Rott, Michal: "Speech-to-text technology to transcribe and disclose 100,000+ hours of bilingual documents from historical Czech and Czechoslovak radio archive", 964-968.](#)

[Yılmaz, Emre / Pelemans, Joris / Van hamme, Hugo: "Automatic assessment of children's reading with the FLaVoR decoding using a phone confusion model", 969-972.](#)

[Shaik, M. Ali Basha / Tüske, Zoltán / Tahir, M. Ali / Nußbaum-Thom, Markus / Schlüter, Ralf / Ney, Hermann: "RWTH LVCSR systems for quaero and EU-bridge: German, Polish, Spanish and Portuguese", 973-977.](#)

Source Separation and Computational Auditory Scene Analysis

[Zöhner, Matthias / Pernkopf, Franz: "Single channel source separation with general stochastic networks", 978-982.](#)

[Yeung, Yu Ting / Lee, Tan / Leung, Cheung-Chi: "Large-margin conditional random fields for single-microphone speech separation", 983-987.](#)

[Jafari, Ingrid / Togneri, Roberto / Nordholm, Sven: "On the use of the Watson mixture model for clustering-based under-determined blind source separation", 988-992.](#)

[Hsu, Chung-Chien / Chien, Jen-Tzung / Chi, Tai-Shih: "Binary mask estimation based on frequency modulations", 993-997.](#)

[Yang, Po-Kai / Hsu, Chung-Chien / Chien, Jen-Tzung: "Bayesian factorization and selection for speech and music separation", 998-1002.](#)

[Wohlmayr, Michael / Mohr, Ludwig / Pernkopf, Franz: "Self-adaption in single-channel source separation", 1003-1007.](#)

Speech Technologies for Ambient Assisted Living (Special Session)

[Vacher, Michel / Lecouteux, Benjamin / Portet, François: "Multichannel automatic recognition of voice command in a multi-room smart home: an experiment involving seniors and users with visual impairment", 1008-1012.](#)

[Walter, Oliver / Despotovic, Vladimir / Haeb-Umbach, Reinhold / Gemmeke, Jort F. / Ons, Bart / Van hamme, Hugo: "An evaluation of unsupervised acoustic model training for a dysarthric speech interface", 1013-1017.](#)

[Gonzalez, Jose A. / Cheah, Lam A. / Bai, Jie / Ell, Stephen R. / Gilbert, James M. / Moore, Roger K. / Green, Phil D.: "Analysis of phonetic similarity in a silent speech interface based on permanent magnetic articulography", 1018-1022.](#)

[Karpov, Alexey / Akarun, Lale / Yalçın, Hülya / Ronzhin, Alexander / Demiröz, Barış Evrim / Çoban, Aysun / Železný, Miloš: "Audio-visual signal processing in a multimodal assisted living environment", 1023-1027.](#)

[Ravanelli, Mirco / Omologo, Maurizio: "On the selection of the impulse responses for distant-speech recognition based on contaminated speech training", 1028-1032.](#)

[Casanueva, I. / Christensen, H. / Hain, Thomas / Green, Phil D.: "Adaptive speech recognition and dialogue management for users with speech disorders", 1033-1037.](#)

[Yu, Bea / Quatieri, Thomas F. / Williamson, James R. / Mundt, James C.: "Prediction of cognitive performance in an animal fluency task based on rate and articulatory markers", 1038-1042.](#)

[Ishi, Carlos / Hatano, Hiroaki / Hagita, Norihiro: "Analysis of laughter events in real science classes by using multiple environment sensor data", 1043-1047.](#)

DNN for ASR

[Sainath, Tara N. / Chung, I-hsin / Ramabhadran, Bhuvana / Picheny, Michael / Gunnels, John / Kingsbury, Brian / Saon, George / Austel, Vernon / Chaudhari, Upendra: "Parallel deep neural network training for LVCSR tasks using blue gene/Q", 1048-1052.](#)

[Bengio, Samy / Heigold, Georg: "Word embeddings for speech recognition", 1053-1057.](#)

[Seide, Frank / Fu, Hao / Droppo, Jasha / Li, Gang / Yu, Dong: "1-bit stochastic gradient descent and its application to data-parallel distributed training of speech DNNs", 1058-1062.](#)

[Takeda, Ryu / Kanda, Naoyuki / Nukaga, Nobuo: "Boundary contraction training for acoustic models based on discrete deep neural networks", 1063-1067.](#)

[Kubo, Yotaro / Suzuki, Jun / Hori, Takaaki / Nakamura, Atsushi: "Restructuring output layers of deep neural networks using minimum risk parameter clustering", 1068-1072.](#)

[Chan, William / Lane, Ian: "Distributed asynchronous optimization of convolutional neural networks", 1073-1077.](#)

[Tóth, László: "Convolutional deep maxout networks for phone recognition", 1078-1082.](#)

[Chen, Dongpeng / Mak, Brian / Sivadas, Sunil: "Joint sequence training of phone and grapheme acoustic model based on multi-task learning deep neural networks", 1083-1087.](#)

[Hsiao, Roger / Ng, Tim / Zhang, Le / Ranjan, Shivesh / Tsakalidis, Stavros / Nguyen, Long / Schwartz, Richard: "Improving semi-supervised deep neural network for keyword search in low resource languages", 1088-1091.](#)

[Liu, Chao / Zhang, Zhiyong / Wang, Dong: "Pruning deep neural networks by optimal brain damage", 1092-1095.](#)

Speaker Recognition — General Topics

[Avila, Anderson R. / Sarria-Paja, Milton / Fraga, Francisco J. / O'Shaughnessy, Douglas / Falk, Tiago H.: "Improving the performance of far-field speaker verification using multi-condition training: the case of GMM-UBM and i-vector systems", 1096-1100.](#)

[Lee, Hung-Shin / Tsao, Yu / Wang, Hsin-Min / Jeng, Shyh-Kang: "Clustering-based i-vector formulation for speaker recognition", 1101-1105.](#)

[Arsikere, Harish / Gupta, Hitesh Anand / Alwan, Abeer: "Speaker recognition via fusion of subglottal features and MFCCs", 1106-1110.](#)

[Sun, Hanwu / Ma, Bin: "The NIST SRE summed channel speaker recognition system", 1111-1114.](#)

[Gallardo, Laura Fernández / Wagner, Michael / Möller, Sebastian: "Advantages of wideband over narrowband channels for speaker verification employing MFCCs and LFCCs", 1115-1119.](#)

[Li, Ming / Liu, Wenbo: "Speaker verification and spoken language identification using a generalized i-vector framework with phonetic tokenizations and tandem features", 1120-1124.](#)

[Asha, T. / Saranya, M. S. / Pandia, D. S. Karthik / Madikeri, Srikanth / Murthy, Hema A.: "Feature Switching in the i-vector framework for speaker verification", 1125-1129.](#)

[Zhong, Jinghua / Jiang, Weiwu / Rao, Wei / Mak, Man-Wai / Meng, Helen: "PLDA modeling in the fisher voice subspace for speaker verification", 1130-1134.](#)

[Martin, Alvin F. / Greenberg, Craig S. / Stanford, Vincent M. / Howard, John M. / Doddington, George R. / Godfrey, John J.: "Performance factor analysis for the 2012 NIST speaker recognition evaluation", 1135-1138.](#)

[Fujimura, Hiroshi: "Simultaneous gender classification and voice activity detection using deep neural networks", 1139-1143.](#)

Speech Processing with Multi-Modalities

[Abdelaziz, Ahmed Hussen / Kolossa, Dorothea: "Dynamic stream weight estimation in coupled-HMM-based audio-visual speech recognition using multilayer perceptrons", 1144-1148.](#)

[Noda, Kuniaki / Yamaguchi, Yuki / Nakadai, Kazuhiro / Okuno, Hiroshi G. / Ogata, Tetsuya: "Lipreading using convolutional neural network", 1149-1153.](#)

[Tao, Fei / Busso, Carlos: "Lipreading approach for isolated digits recognition under whisper and neutral speech", 1154-1158.](#)

[Masaka, Kenta / Aihara, Ryo / Takiguchi, Tetsuya / Ariki, Yasuo: "Multimodal exemplar-based voice conversion using lip features in noisy environments", 1159-1163.](#)

[Deng, Yunbin / Heaton, James T. / Meltzner, Geoffrey S.: "Towards a practical silent speech recognition system", 1164-1168.](#)

[Freitas, João / Ferreira, Artur / Figueiredo, Mário / Teixeira, António / Dias, Miguel Sales: "Enhancing multimodal silent speech interfaces with feature selection", 1169-1173.](#)

[Katz, William / Campbell, Thomas F. / Wang, Jun / Farrar, Eric / Eubanks, J. Coleman / Balasubramanian, Arvind / Prabhakaran, Balakrishnan / Rennaker, Rob: "Opti-speech: a real-time, 3d visual feedback system for speech training", 1174-1178.](#)

[Wang, Jun / Samal, Ashok / Green, Jordan R.: "Across-speaker articulatory normalization for speaker-independent silent speech recognition", 1179-1183.](#)

[Zahner, Marlene / Janke, Matthias / Wand, Michael / Schultz, Tanja: "Conversion from facial myoelectric signals to speech: a unit selection approach", 1184-1188.](#)

[Wand, Michael / Schultz, Tanja: "Towards real-life application of EMG-based speech recognition by using unsupervised adaptation", 1189-1193.](#)

[Liang, Yuan / Iwano, Koji / Shinoda, Koichi: "Simple gesture-based error correction interface for smartphone speech recognition", 1194-1198.](#)

Normalization and Discriminative Training Methods

[Kumar, Kshitiz / Liu, Chaojun / Gong, Yifan: "Normalization of ASR confidence classifier scores via confidence mapping", 1199-1203.](#)

[Alumäe, Tanel: "Neural network phone duration model for speech recognition", 1204-1208.](#)

[Sak, Haşim / Vinyals, Oriol / Heigold, Georg / Senior, Andrew / McDermott, Erik / Monga, Rajat / Mao, Mark: "Sequence discriminative distributed training of long short-term memory recurrent neural networks", 1209-1213.](#)

[Huang, Zhen / Li, Jinyu / Weng, Chao / Lee, Chin-Hui: "Beyond cross-entropy: towards better frame-level objective functions for deep neural network training in automatic speech recognition", 1214-1218.](#)

[Tang, Hao / Gimpel, Kevin / Livescu, Karen: "A comparison of training approaches for discriminative segmental models", 1219-1223.](#)

[McDermott, Erik / Heigold, Georg / Moreno, Pedro J. / Senior, Andrew / Bacchiani, Michiel: "Asynchronous stochastic optimization for sequence training of deep neural networks: towards big data", 1224-1228.](#)

Paralinguistic and Extralinguistic Information

[Rao, Hrishikesh / Kim, Jonathan C. / Clements, Mark A. / Rozga, Agata / Messinger, Daniel S.: "Detection of children's paralinguistic events in interaction with caregivers", 1229-1233.](#)

[Pettorino, Massimo / Pellegrino, Elisa: "Age and rhythmic variations: a study on Italian", 1234-1237.](#)

[Cummins, Nicholas / Sethu, Vidhyasaharan / Epps, Julien / Krajewski, Jarek: "Probabilistic acoustic volume analysis for speech affected by depression", 1238-1242.](#)

[Bozkurt, Elif / Toledo-Ronen, Orith / Sorin, Alexander / Hoory, Ron: "Exploring modulation spectrum features for speech-based depression level classification", 1243-1247.](#)

[Hönig, Florian / Batliner, Anton / Nöth, Elmar / Schnieder, Sebastian / Krajewski, Jarek: "Automatic modelling of depressed speech: relevant features and relevance of gender", 1248-1252.](#)

[Gangamohan, P. / Kadiri, Sudarsana Reddy / Gangashetty, Suryakanth V. / Yegnanarayana, B.: "Excitation source features for discrimination of anger and happy emotions", 1253-1257.](#)

Text Processing for Speech Synthesis

[Wu, Ke / Allauzen, Cyril / Hall, Keith / Riley, Michael / Roark, Brian: "Encoding linear models as weighted finite-state transducers", 1258-1262.](#)

[Kubo, Keigo / Sakti, Sakriani / Neubig, Graham / Toda, Tomoki / Nakamura, Satoshi: "Structured soft margin confidence weighted learning for grapheme-to-phoneme conversion", 1263-1267.](#)

[Zhang, Wei / Clark, Robert A. J. / Wang, Yongyuan: "Unsupervised language filtering using the latent dirichlet allocation", 1268-1272.](#)

[Kolluru, BalaKrishna / Wan, Vincent / Latorre, Javier / Yanagisawa, Kayoko / Gales, Mark J. F.: "Generating multiple-accent pronunciations for TTS using joint sequence model interpolation", 1273-1277.](#)

[Mendonça, Gustavo / Aluisio, Sandra: "Using a hybrid approach to build a pronunciation dictionary for Brazilian Portuguese", 1278-1282.](#)

[Aylett, Matthew P. / Dall, Rasmus / Ghoshal, Arnab / Henter, Gustav Eje / Merritt, Thomas: "A flexible front-end for HTS", 1283-1287.](#)

Cross-language Perception and Production

[Tsukada, Kimiko / Cox, Felicity / Hajek, John: "Cross-language perception of Japanese singleton and geminate consonants: preliminary data from non-native learners of Japanese and native speakers of Italian and Australian English", 1288-1292.](#)

[Alispahic, Samra / Escudero, Paola / Mulak, Karen E.: "Difficulty in discriminating non-native vowels: are Dutch vowels easier for Australian English than Spanish listeners?", 1293-1296.](#)

[Yang, Jing / Fox, Robert Allen: "Acoustic properties of shared vowels in bilingual Mandarin-English children", 1297-1301.](#)

[Lecumberri, María Luisa García / Barra-Chicote, Roberto / Ramón, Rubén Pérez / Yamagishi, Junichi / Cooke, Martin: "Generating segmental foreign accent", 1302-1306.](#)

[Andreeva, Bistra / Demenko, Grażyna / Möbius, Bernd / Zimmerer, Frank / Jügler, Jeanin / Oleskowicz-Popiel, Magdalena: "Differences of pitch profiles in Germanic and slavic languages", 1307-1311.](#)

[Avanzi, Mathieu / Bordal, Guri / Nimbona, Gélase: "The obligatory contour principle in african and European varieties of French", 1312-1316.](#)

Text-Dependent Speaker Verification With Short Utterances (Special)

[Scheffer, Nicolas / Lei, Yun: "Content matching for short duration speaker recognition", 1317-1321.](#)

[Larcher, Anthony / Lee, Kong Aik / Martínez, Pablo L. Sordo / Nguyen, Trung Hieu / Ma, Bin / Li, Haizhou: "Extended RSR2015 for text-dependent speaker verification over VHF channel", 1322-1326.](#)

[Fu, Tianfan / Qian, Yanmin / Liu, Yuan / Yu, Kai: "Tandem deep features for text-dependent speaker verification", 1327-1331.](#)

[Kenny, Patrick / Stafylakis, Themis / Alam, M. J. / Ouellet, Pierre / Kockmann, Marcel: "In-domain versus out-of-domain training for text-dependent JFA", 1332-1336.](#)

[Aronowitz, Hagai / Rendel, Asaf: "Domain adaptation for text dependent speaker verification", 1337-1341.](#)

[Miguel, Antonio / Villalba, Jesús / Ortega, Alfonso / Lleida, Eduardo / Vaquero, Carlos: "Factor analysis with sampling methods for text dependent speaker recognition", 1342-1346.](#)

Speech and Audio Analysis

[Berg, Ewout van den / Ramabhadran, Bhuvana: "Dictionary-based pitch tracking with dynamic programming", 1347-1351.](#)

[Hu, Hongbing / Zahorian, Stephen A. / Guzewich, Peter / Wu, Jiang: "Acoustic features for robust classification of Mandarin tones", 1352-1356.](#)

[Karlsson, Anastasia / Lundström, Håkan / Svantesson, Jan-Olof: "Preservation of lexical tones in singing in a tone language", 1357-1360.](#)

[Yakoumaki, Theodora / Kafentzis, George P. / Stylianou, Yannis: "Emotional speech classification using adaptive sinusoidal modelling", 1361-1365.](#)

[Wang, Shengbei / Unoki, Masashi / Kim, Nam Soo: "Formant enhancement based speech watermarking for tampering detection", 1366-1370.](#)

[Barker, Tom / Van hamme, Hugo / Virtanen, Tuomas: "Modelling primitive streaming of simple tone sequences through factorisation of modulation pattern tensors", 1371-1375.](#)

[Sarma, Biswajit Dev / Prasanna, S. R. M.: "Detection of vowel onset points in voiced aspirated sounds of indian languages", 1376-1380.](#)

[Sasou, Akira: "Accuracy evaluation of esophageal voice analysis based on automatic topology generated-voicing source HMM", 1381-1385.](#)

[Zhang, Xuejun / Xie, Xiang: "Audio watermarking based on multiple echoes hiding for FM radio", 1386-1390.](#)

Cross-Lingual and Adaptive Language Modeling

[Motlicek, Petr / Imseng, David / Cernak, Milos / Kim, Namhoon: "Development of bilingual ASR system for MediaParl corpus", 1391-1394.](#)

[Li, Jie / Zheng, Rong / Xu, Bo: "Investigation of cross-lingual bottleneck features in hybrid ASR systems", 1395-1399.](#)

[Giwa, Oluwapelumi / Davel, Marelle H.: "Language identification of individual words with joint sequence models", 1400-1404.](#)

[Anquera, Xavier / Luque, Jordi / Gracia, Ciro: "Audio-to-text alignment for speech recognition with very limited resources", 1405-1409.](#)

[Ngo, Hoang Gia / Chen, Nancy F. / Sivadas, Sunil / Ma, Bin / Li, Haizhou: "A minimal-resource transliteration framework for vietnamese", 1410-1414.](#)

[Adel, Heike / Telaar, Dominic / Vu, Ngoc Thang / Kirchhoff, Katrin / Schultz, Tanja: "Combining recurrent neural networks and factored language models during decoding of code-switching speech", 1415-1419.](#)

[Tüske, Zoltán / Golik, Pavel / Nolden, David / Schlüter, Ralf / Ney, Hermann: "Data augmentation, feature combination, and multilingual neural networks to improve ASR and KWS performance for low-resource languages", 1420-1424.](#)

[Masumura, Ryo / Asami, Taichi / Oba, Takanobu / Masataki, Hirokazu / Sakauchi, Sumitaka: "Mixture of latent words language models for domain adaptation", 1425-1429.](#)

[Herms, Robert / Ritter, Marc / Wilhelm-Stein, Thomas / Eibl, Maximilian: "Improving spoken document retrieval by unsupervised language model adaptation using utterance-based web search", 1430-1433.](#)

[Chien, Jen-Tzung / Chang, Ying-Lan: "The nested indian buffet process for flexible topic modeling", 1434-1437.](#)

[Levin, K. / Ponomareva, I. / Bulusheva, A. / Chernykh, G. / Medennikov, I. / Merkin, N. / Prudnikov, A. / Tomashenko, Natalia: "Automated closed captioning for Russian live broadcasting", 1438-1442.](#)

Pronunciation Modeling and Learning

[Wang, Lei / Tong, Rong: "Pronunciation modeling of foreign words for Mandarin ASR by considering the effect of language transfer", 1443-1447.](#)

[Rutherford, Attapol T. / Peng, Fuchun / Beaufays, Françoise: "Pronunciation learning for named-entities through crowd-sourcing", 1448-1452.](#)

[Schuppler, Barbara / Adda-Decker, Martine / Morales-Cordovilla, Juan A.: "Pronunciation variation in read and conversational austrian German", 1453-1457.](#)

[Lehr, Maider / Gorman, Kyle / Shafran, Izhak: "Discriminative pronunciation modeling for dialectal speech recognition", 1458-1462.](#)

[Pellegrini, Thomas / Fontan, Lionel / Mauclair, Julie / Farinas, Jérôme / Robert, Marina: "The goodness of pronunciation algorithm applied to disordered speech", 1463-1467.](#)

[Metallinou, Angeliki / Cheng, Jian: "Using deep neural networks to improve proficiency assessment for children English language learners", 1468-1472.](#)

[Lu, Han / Shen, Sheng-syun / Shiang, Sz-Rung / Lee, Hung-yi / Lee, Lin-shan: "Alignment of spoken utterances with slide content for easier learning with recorded lectures using structured support vector machine \(SVM\)", 1473-1477.](#)

[Duan, Richeng / Zhang, Jinsong / Cao, Wen / Xie, Yanlu: "A preliminary study on ASR-based detection of Chinese mispronunciation by Japanese learners", 1478-1481.](#)

Show and Tell Session 1, 1

[Xu, Kele / Yang, Yin / Jaumard-Hakoun, A. / Adda-Decker, Martine / Amelot, A. / Kork, S. K. Al / Crevier-Buchman, L. / Chawah, P. / Dreyfus, G. / Fux, T. / Pillot-Loiseau, C. / Roussel, P. / Stone, M. / Denby, B.: "3d tongue motion visualization based on ultrasound image sequences", 1482-1483.](#)

[Derrick, Donald / Rybel, Tom De / O'Beirne, Greg A. / Hay, Jennifer: "Listen with your skin: aerotak speech perception enhancement system", 1484-1485.](#)

[Czap, László: "Speech assistant system", 1486-1487.](#)

[Banchs, Rafael E. / Kim, Seokhwan: "Spoken dialogue system for restaurant recommendation and reservation", 1488-1489.](#)

[Akira, Hayakawa / Campbell, Nick / Luz, Saturnino: "Interlingual map task corpus collection", 1490-1491.](#)

[Centelles, Jordi / Costa-jussà, Marta R. / Banchs, Rafael E.: "A client mobile application for Chinese-Spanish statistical machine translation", 1492-1493.](#)

[Benin, Alberto / Cosi, Piero / Leone, Giuseppe Riccardo / Paci, Giulio: "LuciawebGL: a new WebGL-Based talking head", 1494-1495.](#)

[Naderi, Babak / Polzehl, Tim / Beyer, André / Pilz, Tibor / Möller, Sebastian: "Crowdee: mobile crowdsourcing micro-task platform for celebrating the diversity of languages", 1496-1497.](#)

[Moore, Roger K.: "On the use of the `pure data' programming language for teaching and public outreach in speech processing", 1498-1499.](#)

[Dubinsky, Aleksandr: "Syncwords: a platform for semi-automated closed captioning and subtitles", 1500-1501.](#)

[Clark, Robert A. J.: "Simple⁴all", 1502-1503.](#)

[Chawah, P. / Kork, S. K. Al / Fux, T. / Adda-Decker, Martine / Amelot, A. / Audibert, N. / Denby, B. / Dreyfus, G. / Jaumard-Hakoun, A. / Pillot-Loiseau, C. / Roussel, P. / Stone, M. / Xu, Kele / Crevier-Buchman, L.: "An educational platform to capture, visualize and analyze rare singing", 2128-2129.](#)

[Jeon, Kwang Myung / Chun, Chan Jun / Seong, Woo Kyeong / Kim, Hong Kook / Choi, Myung Kyu: "Single-channel speech enhancement based on non-negative matrix factorization and online noise adaptation", 2130-2131.](#)

[Maurer, Dieter / Mok, Peggy / Friedrichs, Daniel / Dellwo, Volker: "Intelligibility of high-pitched vowel sounds in the singing and speaking of a female Cantonese opera singer", 2132-2133.](#)

[Mowlae, Pejman / Watanabe, Mario Kaoru / Saeidi, Rahim: "Iterative refinement of amplitude and phase in single-channel speech enhancement", 2134-2135.](#)

[Roekhaut, Sophie / Brognaux, Sandrine / Beaufort, Richard / Dutoit, Thierry: "elite-HTS: a NLP tool for French HMM-based speech synthesis", 2136-2137.](#)

[Niculescu, Andreea I. / Banchs, Rafael E. / Jiang, Ridong / Kim, Seokhwan / Yeo, Kheng Hui / Niswar, Arthur: "SARA — singapore's automated responsive assistant for the touristic domain", 2138-2139.](#)

[Plummer, Andrew / Riebling, Eric / Kumar, Anuj / Metze, Florian / Fosler-Lussier, Eric / Bates, Rebecca: "The speech recognition virtual kitchen: launch party", 2140-2141.](#)

[Marek-Spartz, Kyle / Knoll, Benjamin / Bill, Robert / Christie, Thomas / Pakhomov, Serguei: "System for automated speech and language analysis \(SALSA\)", 2142-2143.](#)

[Masuda-Katsuse, Ikuyo: "Pronunciation practice support system for children who have difficulty correctly pronouncing words", 2144-2145.](#)

[Driesen, Joris / Birch, Alexandra / Grimsey, Simon / Safarfashandi, Saeid / Gauthier, Juliet / Simpson, Matt / Renals, Steve: "Automated production of true-cased punctuated subtitles for weather and news broadcasts", 2146-2147.](#)

[Dong, Minghui / Lee, S. W. / Li, Haizhou / Chan, Paul / Peng, Xuejian / Ehnes, Jochen Walter / Huang, Dongyan: "I²r speech2singing perfects everyone's singing", 2148-2149.](#)

Statistical Parametric Speech Synthesis

[Henter, Gustav Eje / Merritt, Thomas / Shannon, Matt / Mayo, Catherine / King, Simon: "Measuring the perceptual effects of modelling assumptions in speech synthesis using stimuli constructed from repeated natural speech", 1504-1508.](#)

[Merritt, Thomas / Raitio, Tuomo / King, Simon: "Investigating source and filter contributions, and their interaction, to statistical parametric speech synthesis", 1509-1513.](#)

[Latorre, Javier / Wan, Vincent / Yanagisawa, Kayoko: "Voice expression conversion with factorised HMM-TTS models", 1514-1518.](#)

[Yanagisawa, Kayoko / Chen, Langzhou / Gales, Mark J. F.: "Noise-robust TTS speaker adaptation with statistics smoothing", 1519-1523.](#)

[Brognaux, Sandrine / Picart, Benjamin / Drugman, Thomas: "Speech synthesis in various communicative situations: impact of pronunciation variations", 1524-1528.](#)

[Cai, Ming-Qi / Ling, Zhen-Hua / Dai, Li-Rong: "Formant-controlled speech synthesis using hidden trajectory model", 1529-1533.](#)

Voice Activity Detection

[Zhang, Xiao-Lei / Wang, DeLiang: "Boosted deep neural networks and multi-resolution cochleagram features for voice activity detection", 1534-1538.](#)

[Prasad, Abhay / Ghosh, Prasanta Kumar / Narayanan, Shrikanth S.: "Selection of optimal vocal tract regions using real-time magnetic resonance imaging for robust voice activity detection", 1539-1543.](#)

[Ziaei, Ali / Kaushik, Lakshmesh / Sangwan, Abhijeet / Hansen, John H. L. / Oard, Douglas W.: "Speech activity detection for NASA apollo space missions: challenges and solutions", 1544-1548.](#)

[Tu, Ming / Xie, Xiang / Jiao, Yishan: "Towards improving statistical model based voice activity detection", 1549-1552.](#)

[McLoughlin, Ian Vince: "The use of low-frequency ultrasound for voice activity detection", 1553-1557.](#)

[Ma, Jeff: "Improving the speech activity detection for the DARPA RATS phase-3 evaluation", 1558-1562.](#)

Disordered Speech

[Le, Duc / Provost, Emily Mower: "Modeling pronunciation, rhythm, and intonation for automatic assessment of speech quality in aphasia rehabilitation", 1563-1567.](#)

[Strömbergsson, Sofia / Tännander, Christina / Edlund, Jens: "Ranking severity of speech errors by their phonological impact in context", 1568-1572.](#)

[Orozco-Arroyave, J. R. / Hönig, Florian / Arias-Londoño, J. D. / Vargas-Bonilla, J. F. / Skodda, S. / Ruzs, J. / Nöth, Elmar: "Automatic detection of parkinson's disease from words uttered in three different languages", 1573-1577.](#)

[Lilley, Jason / Nittrouer, Susan / Bunnell, H. Timothy: "Automating an objective measure of pediatric speech intelligibility", 1578-1582.](#)

[Shahin, Mostafa / Ahmed, Beena / McKechnie, Jacqueline / Ballard, Kirrie / Gutierrez-Osuna, Ricardo: "A comparison of GMM-HMM and DNN-HMM based](#)

[pronunciation verification techniques for use in the assessment of childhood apraxia of speech", 1583-1587.](#)

[Berry, Jeff / Kolb, Andrew / North, Cassandra / Johnson, Michael T.: "Acoustic and kinematic characteristics of vowel production through a virtual vocal tract in dysarthria", 1588-1592.](#)

Speech and Multimodal Resources

[Wand, Michael / Janke, Matthias / Schultz, Tanja: "The EMG-UKA corpus for electromyographic speech processing", 1593-1597.](#)

[Lee, Pei Xuan / Wee, Darren / Toh, Hilary Si Yin / Lim, Boon Pang / Chen, Nancy F. / Ma, Bin: "A whispered Mandarin corpus for speech technology applications", 1598-1602.](#)

[Gretter, Roberto: "Euronews: a multilingual benchmark for ASR and LID", 1603-1607.](#)

[Tsiami, Antigoni / Rodomagoulakis, Isidoros / Giannoulis, Panagiotis / Katsamanis, Athanasios / Potamianos, Gerasimos / Maragos, Petros: "ATHENA: a Greek multi-sensory database for home automation control uthor: isidoros rodomagoulakis \(NTUA, Greece\)", 1608-1612.](#)

[Matassoni, Marco / Astudillo, Ramón Fernandez / Katsamanis, Athanasios / Ravanelli, Mirco: "The DIRHA-GRID corpus: baseline and tools for multi-room distant speech recognition using distributed microphones", 1613-1617.](#)

[Henriques, Diogo / Trancoso, Isabel / Mendes, Daniel / Ferreira, Alfredo: "Verbal description of LEGO blocks", 1618-1622.](#)

Phase Importance in Speech Processing Applications (Special Session)

[Mowlae, Pejman / Saeidi, Rahim / Stylianou, Yannis: "Phase importance in speech processing applications", 1623-1627.](#)

[Cano, Estefanía / Plumbley, Mark / Dittmar, Christian: "Phase-based harmonic/percussive separation", 1628-1632.](#)

[Degottex, Gilles / Obin, Nicolas: "Phase distortion statistics as a representation of the glottal source: application to the classification of voice qualities", 1633-1637.](#)

[Degottex, Gilles / Erro, Daniel: "A measure of phase randomness for the harmonic model in speech synthesis", 1638-1642.](#)

[Jokinen, Emma / Takanen, Marko / Pulakka, Hannu / Alku, Paavo: "Enhancement of speech intelligibility in near-end noise conditions with phase modification", 1643-1647.](#)

[Shanmugam, S. Aswin / Murthy, Hema: "A hybrid approach to segmentation of speech using group delay processing and HMM based embedded reestimation", 1648-1652.](#)

[Koutsogiannaki, Maria / Simantiraki, Olympia / Degottex, Gilles / Stylianou, Yannis: "The importance of phase on voice quality assessment", 1653-1657.](#)

[Vijayan, Karthika / Kumar, Vinay / Murty, K. Sri Rama: "Feature extraction from analytic phase of speech signals for speaker verification", 1658-1662.](#)

[Sanchez, Jon / Saratxaga, Ibon / Hernaez, Inma / Navas, Eva / Erro, Daniel: "A cross-vocoder study of speaker independent synthetic speech detection using phase information", 1663-1667.](#)

Spoken Term Detection and Document Retrieval

[Yang, Peng / Leung, Cheung-Chi / Xie, Lei / Ma, Bin / Li, Haizhou: "Intrinsic spectral analysis based on temporal context features for query-by-example spoken term detection", 1722-1726.](#)

[Hout, Julien van / Mitra, Vikramjit / Lei, Yun / Vergyri, Dimitra / Graciarena, Martin / Mandal, Arindam / Franco, Horacio: "Recent improvements in SRI's keyword detection system for noisy audio", 1727-1731.](#)

[Makino, Mitsuaki / Yamamoto, Naoki / Kai, Atsuhiko: "Utilizing state-level distance vector representation for improved spoken term detection by text and spoken queries", 1732-1736.](#)

[Pappagari, Raghavendra Reddy / Nayak, Shekhar / Murty, K. Sri Rama: "Unsupervised spoken word retrieval using Gaussian-bernoulli restricted boltzmann machines", 1737-1741.](#)

[George, Basil / Saxena, Abhijeet / Mantena, Gautam / Prahallad, Kishore / Yegnanarayana, B.: "Unsupervised query-by-example spoken term detection using bag of acoustic words and non-segmental dynamic time warping", 1742-1746.](#)

[Li, Jie / Wang, Xiaorui / Xu, Bo: "An empirical study of multilingual and low-resource spoken term detection using deep neural networks", 1747-1751.](#)

[Schulam, Peter / Akbacak, Murat: "Diagnostic techniques for spoken keyword discovery", 1752-1756.](#)

[Kawasaki, Sho / Akiba, Tomoyosi: "Robust retrieval models for false positive errors in spoken documents", 1757-1761.](#)

[Liou, Yuan-ming / Fu, Yi-sheng / Lee, Hung-yi / Lee, Lin-shan: "Semantic retrieval of personal photos using matrix factorization and two-layer random walk fusing sparse speech annotations with visual features", 1762-1766.](#)

[Gravier, Guillaume / Souviraà-Labastie, Nathan / Campion, Sébastien / Bimbot, Frédéric: "Audio thumbnails for spoken content without transcription based on a maximum motif coverage criterion", 1767-1771.](#)

[García, Fernando / Sanchis, Emilio / Pla, Ferran: "Semantically based search in a social speech task", 1772-1776.](#)

Prosody and Paralinguistic Information

[Mittal, Vinay Kumar / Yegnanarayana, B.: "Study of changes in glottal vibration characteristics during laughter", 1777-1781.](#)

[Ntalampiras, Stavros / Potamitis, Ilyas: "On predicting the unpleasantness level of a sound event", 1782-1785.](#)

[Piot, Bilal / Pietquin, Olivier / Geist, Matthieu: "Predicting when to laugh with structured classification", 1786-1790.](#)

[Weiss, Benjamin / Schoenenberg, Katrin: "Conversational structures affecting auditory likeability", 1791-1795.](#)

[Avanzi, Mathieu / Christodoulides, George / Lolive, Damien / Delais-Roussarie, Elisabeth / Barbot, Nelly: "Towards the adaptation of prosodic models for expressive text-to-speech synthesis", 1796-1800.](#)

[Matsumiya, Sho / Sakti, Sakriani / Neubig, Graham / Toda, Tomoki / Nakamura, Satoshi: "Data-driven generation of text balloons based on linguistic and acoustic features of a comics-anime corpus", 1801-1805.](#)

[Tseng, Chiu-yu / Su, Chao-yu: "Learning L2 prosody is more difficult than you realize — F0 characteristics and chunking size of L1 English, TW L2 English and TW L1 Mandarin", 1806-1810.](#)

[Truong, Khiet P. / Trouvain, Jürgen: "Investigating prosodic relations between initiating and responding laughs", 1811-1815.](#)

[Prylipko, Dmytro / Egorow, Olga / Siegert, Ingo / Wendemuth, Andreas: "Application of image processing methods to filled pauses detection from spontaneous speech", 1816-1820.](#)

[Kakouros, Sofoklis / Räsänen, Okko: "Perception of sentence stress in English infant directed speech", 1821-1825.](#)

[Madzlan, Noor Alhusna / Han, JingGuang / Bonin, Francesca / Campbell, Nick: "Automatic recognition of attitudes in video blogs — prosodic and visual feature analysis", 1826-1830.](#)

[Katerenchuk, Denys / Brizan, David Guy / Rosenberg, Andrew: "'was that your mother on the phone?': classifying interpersonal relationships between dialog participants with lexical and acoustic properties", 1831-1835.](#)

Features and Robustness in Speaker and Language Recognition

[Das, Rohan Kumar / Abhiram, S. / Prasanna, S. R. M. / Ramakrishnan, A. G.: "Combining source and system information for limited data speaker verification", 1836-1840.](#)

[Diez, Mireia / Varona, Amparo / Penagarikano, Mikel / Rodriguez-Fuentes, Luis Javier / Bordel, German: "New insight into the use of phone log-likelihood ratios as features for language recognition", 1841-1845.](#)

[Ganapathy, Sriram / Han, Kyu / Thomas, Samuel / Omar, Mohamed / Segbroeck, Maarten Van / Narayanan, Shrikanth S.: "Robust language identification using convolutional neural network features", 1846-1850.](#)

[Yu, Chengzhu / Liu, Gang / Hansen, John H. L.: "Acoustic feature transformation using UBM-based LDA for speaker recognition", 1851-1854.](#)

[Mak, Man-Wai: "SNR-dependent mixture of PLDA for noise robust speaker verification", 1855-1859.](#)

[Sadjadi, Seyed Omid / Pelecanos, Jason / Zhu, Weizhong: "Nearest neighbor discriminant analysis for robust speaker recognition", 1860-1864.](#)

Topic Spotting and Summarization of Spoken Documents

[Liu, Shih-Hung / Chen, Kuan-Yu / Hsieh, Yu-Lun / Chen, Berlin / Wang, Hsin-Min / Yen, Hsu-Chun / Hsu, Wen-Lian: "Enhanced language modeling for extractive speech summarization with sentence relatedness information", 1865-1869.](#)

[Morchid, Mohamed / Bouallegue, Mohamed / Dufour, Richard / Linarès, Georges / Matrouf, Driss / Mori, Renato De: "I-vector based representation of highly imperfect automatic transcriptions", 1870-1874.](#)

[Lai, Catherine / Renals, Steve: "Incorporating lexical and prosodic information at different levels for meeting summarization", 1875-1879.](#)

[Bouallegue, Mohamed / Morchid, Mohamed / Dufour, Richard / Matrouf, Driss / Linarès, Georges / Mori, Renato De: "Subspace Gaussian mixture models for dialogues classification", 1880-1884.](#)

[Bouallegue, Mohamed / Morchid, Mohamed / Dufour, Richard / Matrouf, Driss / Linarès, Georges / Mori, Renato De: "Factor analysis based semantic variability compensation for automatic conversation representation", 1885-1889.](#)

[Bouчекif, Abdessalam / Damnati, Géraldine / Charlet, Delphine: "Speech cohesion for topic segmentation of spoken contents", 1890-1894.](#)

DNN Learning

[Huang, Yan / Yu, Dong / Liu, Chaojun / Gong, Yifan: "A comparative analytic study on the Gaussian mixture and context dependent deep neural network hidden Markov models", 1895-1899.](#)

[Bacchiani, Michiel / Senior, Andrew / Heigold, Georg: "Asynchronous, online, GMM-free training of a context dependent acoustic model for speech recognition", 1900-1904.](#)

[Jaitly, Navdeep / Vanhoucke, Vincent / Hinton, Geoffrey: "Autoregressive product of multi-frame predictions can improve the accuracy of hybrid models", 1905-1909.](#)

[Li, Jinyu / Zhao, Rui / Huang, Jui-Ting / Gong, Yifan: "Learning small-size DNN with output-distribution-based criteria", 1910-1914.](#)

[Deng, Li / Platt, John C.: "Ensemble deep learning for speech recognition", 1915-1919.](#)

[Zhou, Yucan / Hu, Qinghua / Liu, Jie / Jia, Yuan: "Learning conditional random field with hierarchical representations for dialogue act recognition", 1920-1923.](#)

Perception of Emotion and Prosody

[Hsu, Cristiane / Xu, Yi: "Can adolescents with autism perceive emotional prosody?", 1924-1928.](#)

[Schmidt, Juliane / Janse, Esther / Scharenborg, Odette: "Age, hearing loss and the perception of affective utterances in conversational speech", 1929-1933.](#)

[Yang, Zhaojun / Narayanan, Shrikanth S.: "Analysis of emotional effect on speech-body gesture interplay", 1934-1938.](#)

[Chappuis, Cyrielle / Grandjean, Didier: "When voices get emotional: a study of emotion-enhanced memory and impairment during emotional prosody exposure", 1939-1943.](#)

[Zellers, Margaret: "Perception of pitch tails at potential turn boundaries in Swedish", 1944-1948.](#)

[Fuchs, Robert: "Towards a perceptual model of speech rhythm: integrating the influence of f0 on perceived duration", 1949-1953.](#)

Deep Neural Networks for Speech Generation and Synthesis (Special

[Chen, Ling-Hui / Raitio, Tuomo / Valentini-Botinhao, Cassia / Yamagishi, Junichi / Ling, Zhen-Hua: "DNN-based stochastic postfilter for HMM-based speech synthesis", 1954-1958.](#)

[Kang, Shiyin / Meng, Helen: "Statistical parametric speech synthesis using weighted multi-distribution deep belief network", 1959-1963.](#)

[Fan, Yuchen / Qian, Yao / Xie, Feng-Long / Soong, Frank K.: "TTS synthesis with bidirectional LSTM based recurrent neural networks", 1964-1968.](#)

[Raitio, Tuomo / Suni, Antti / Juvela, Lauri / Vainio, Martti / Alku, Paavo: "Deep neural network based trainable voice source model for synthesis of speech with varying vocal effort", 1969-1973.](#)

[Yu, Dong / Eversole, Adam / Seltzer, Michael L. / Yao, Kaisheng / Guenter, Brian / Kuchaiev, Oleksii / Seide, Frank / Wang, Huaming / Droppo, Jasha / Huang, Zhiheng / Zweig, Geoff / Rossbach, Chris / Currey, Jon: "An introduction to computational networks and the computational network toolkit \(invited talk\)" \(abstract\). --- N/A](#)

[Fernandez, Raul / Rendel, Asaf / Ramabhadran, Bhuvana / Hoory, Ron: "Prosody contour prediction with long short-term memory, bi-directional, deep recurrent neural networks", 2268-2272.](#)

[Yin, Xiang / Lei, Ming / Qian, Yao / Soong, Frank K. / He, Lei / Ling, Zhen-Hua / Dai, Li-Rong: "Modeling DCT parameterized F0 trajectory at intonation phrase level with DNN or decision tree", 2273-2277.](#)

[Nakashika, Toru / Takiguchi, Tetsuya / Arika, Yasuo: "High-order sequence modeling using speaker-dependent recurrent temporal restricted boltzmann machines for voice conversion", 2278-2282.](#)

Xie, Feng-Long / Qian, Yao / Fan, Yuchen / Soong, Frank K. / Li, Haifeng: "Sequence error (SE) minimization training of neural network for voice conversion", 2283-2287.

Bocquelet, Florent / Hueber, Thomas / Girin, Laurent / Badin, Pierre / Yvert, Blaise: "Robust articulatory speech synthesis using deep neural networks for BCI applications", 2288-2292.

Speech Analysis and Perception

Xu, Shufang: "Acoustic investigation of /t^h/ lenition in brunei Mandarin", 1974-1977.

Wang, Ting / Ding, Hongwei / Kuang, Jianjing / Ma, Qiuwu: "Mapping emotions into acoustic space: the role of voice quality", 1978-1982.

Mahajan, Nagaraj / Mesgarani, Nima / Hermansky, Hynek: "Principal components of auditory spectro-temporal receptive fields", 1983-1987.

Thlithi, Marwa / Pellegrini, Thomas / Pinquier, Julien / André-Obrecht, Régine: "Segmentation in singer turns with the Bayesian information criterion", 1988-1992.

Watson, Catherine I.: "Mappings between vocal tract area functions, vocal tract resonances and speech formants for multiple speakers", 1993-1997.

Arndt, Sebastian / Wenzel, Markus / Antons, Jan-Niklas / Köster, Friedemann / Möller, Sebastian / Curio, Gabriel: "A next step towards measuring perceived quality of speech through physiology", 1998-2001.

Chen, Fei / Wong, Sharon W. K. / Wong, Lena L. N.: "Effect of spectral degradation to the intelligibility of vowel sentences", 2002-2005.

Berry, Jeff / IV, John Jaeger / Wiedenhoeft, Melissa / Bernal, Brittany / Johnson, Michael T.: "Consonant context effects on vowel sensorimotor adaptation", 2006-2010.

Bailly, Gérard / Martin, Amélie: "Assessing objective characterizations of phonetic convergence", 2011-2015.

Mandel, Michael I. / Yoho, Sarah E. / Healy, Eric W.: "Generalizing time-frequency importance functions across noises, talkers, and phonemes", 2016-2020.

Mahajan, Yatin / Kim, Jeesun / Davis, Chris: "Does elderly speech recognition in noise benefit from spectral and visual cues?", 2021-2025.

[Laskowski, Kornel: "On the conversant-specificity of stochastic turn-taking models", 2026-2030.](#)

Intelligibility Enhancement and Predictive Measures

[Sakano, Toshihiro / Kobayashi, Yosuke / Kondo, Kazuhiro: "Single-ended estimation of speech intelligibility using the ITU p.563 feature set", 2031-2035.](#)

[Jokinen, Emma / Remes, Ulpu / Takanen, Marko / Palomäki, Kalle / Kurimo, Mikko / Alku, Paavo: "Spectral tilt modelling with GMMs for intelligibility enhancement of narrowband telephone speech", 2036-2040.](#)

[Köster, Friedemann / Möller, Sebastian: "Analyzing perceptual dimensions of conversational speech quality", 2041-2045.](#)

[Aubanel, Vincent / Davis, Chris / Kim, Jeesun: "Interplay of informational content and energetic masking in speech perception in noise", 2046-2049.](#)

[Zorilă, Tudor-Cătălin / Stylianou, Yannis: "On spectral and time domain energy reallocation for speech-in-noise intelligibility enhancement", 2050-2054.](#)

[Chen, Fei / Hu, Yi: "Objective quality evaluation of noise-suppressed speech: effects of temporal envelope and fine-structure cues", 2055-2058.](#)

[Wang, Dongmei / Loizou, Philipos C. / Hansen, John H. L.: "Noisy speech enhancement based on long term harmonic model to improve speech intelligibility for hearing impaired listeners", 2059-2062.](#)

[Valentini-Botinhao, Cassia / Wester, Mirjam: "Using linguistic predictability and the lombard effect to increase the intelligibility of synthetic speech in noise", 2063-2067.](#)

[Dabel, Maryam Al / Barker, Jon: "Speech pre-enhancement using a discriminative microscopic intelligibility model", 2068-2072.](#)

[Harvilla, Mark J. / Stern, Richard M.: "Least squares signal declipping for robust speech recognition", 2073-2077.](#)

Speech and Language Processing — General Topics

[Xu, Haihua / Su, Hang / Chng, Eng Siong / Li, Haizhou: "Semi-supervised training for bottle-neck feature based DNN-HMM hybrid systems", 2078-2082.](#)

[Kapralova, Olga / Alex, John / Weinstein, Eugene / Moreno, Pedro J. / Siohan, Olivier: "A big data approach to acoustic model training corpus selection", 2083-2087.](#)

[Cardinal, Patrick / Ali, Ahmed / Dehak, Najim / Zhang, Yu / Hanai, Tuka Al / Zhang, Yifan / Glass, James R. / Vogel, Stephan: "Recent advances in ASR applied to an Arabic transcription system for Al-Jazeera", 2088-2092.](#)

[Sundermeyer, Martin / Schlüter, Ralf / Ney, Hermann: "rwthlm — the RWTH aachen university neural network language modeling toolkit", 2093-2097.](#)

[Cheng, Wei-Chen / Kok, Stanley / Pham, Hoai Vu / Chieu, Hai Leong / Chai, Kian Ming A.: "Language modeling with sum-product networks", 2098-2102.](#)

[Cui, Xiaodong / Kingsbury, Brian / Cui, Jia / Ramabhadran, Bhuvana / Rosenberg, Andrew / Rasooli, Mohammad Sadegh / Rambow, Owen / Habash, Nizar / Goel, Vaibhava: "Improving deep neural network acoustic modeling for audio corpus indexing under the IARPA babel program", 2103-2107.](#)

[Chowdhury, Shammur Absar / Ghosh, Arindam / Stepanov, Evgeny A. / Bayer, Ali Orkan / Riccardi, Giuseppe / Klasinas, Ioannis: "Cross-language transfer of semantic annotation via targeted crowdsourcing", 2108-2112.](#)

[Hakkani-Tür, Dilek / Celikyilmaz, Asli / Heck, Larry / Tur, Gokhan / Zweig, Geoff: "Probabilistic enrichment of knowledge graph entities for relation detection in conversational understanding", 2113-2117.](#)

[Garner, Philip N. / Imseng, David / Meyer, Thomas: "Automatic speech recognition and translation of a Swiss German dialect: Walliserdeutsch", 2118-2122.](#)

[Harrat, S. / Meftouh, K. / Abbas, M. / Smaili, K.: "Building resources for Algerian Arabic dialects", 2123-2127.](#)

Language, Dialect and Accent Recognition

[Ferrer, Luciana / Lei, Yun / McLaren, Mitchell / Scheffer, Nicolas: "Spoken language recognition based on senone posteriors", 2150-2154.](#)

[Gonzalez-Dominguez, Javier / Lopez-Moreno, Ignacio / Sak, Haşim / Gonzalez-Rodriguez, Joaquin / Moreno, Pedro J.: "Automatic language identification using long short-term memory recurrent neural networks", 2155-2159.](#)

[Desplanques, Brecht / Demuynck, Kris / Martens, Jean-Pierre: "Robust language recognition via adaptive language factor extraction", 2160-2164.](#)

[Behravan, Hamid / Hautamäki, Ville / Siniscalchi, Sabato Marco / Khoury, Elie / Kurki, Tommi / Kinnunen, Tomi / Lee, Chin-Hui: "Dialect levelling in Finnish: a universal speech attribute approach", 2165-2169.](#)

[Chen, Mingming / Yang, Zhanlei / Zheng, Hao / Liu, Wenju: "Improving native accent identification using deep neural networks", 2170-2174.](#)

[Kolly, Marie-José / Leemann, Adrian / Dellwo, Volker: "Foreign accent recognition based on temporal information contained in lowpass-filtered speech", 2175-2179.](#)

Adaptation 1, 2

[Karanasou, Penny / Wang, Yongqiang / Gales, Mark J. F. / Woodland, Philip C.: "Adaptation of deep neural network acoustic models using factorised i-vectors", 2180-2184.](#)

[Fukuda, Takashi / Ichikawa, Osamu / Nishimura, Masafumi / Rennie, Steven J. / Goel, Vaibhava: "Regularized feature-space discriminative adaptation for robust ASR", 2185-2188.](#)

[Miao, Yajie / Zhang, Hao / Metze, Florian: "Towards speaker adaptive training of deep neural network acoustic models", 2189-2193.](#)

[Gorin, Arseniy / Jouvét, Denis: "Component structuring and trajectory modeling for speech recognition", 2194-2198.](#)

[Doddipatla, Rama / Hasan, Madina / Hain, Thomas: "Speaker dependent bottleneck layer training for speaker adaptation in automatic speech recognition", 2199-2203.](#)

[You, Zhao / Xu, Bo: "Improving wideband acoustic models using mixed-bandwidth training data via DNN adaptation", 2204-2208.](#)

[Pellegrini, Thomas / Hedayati, Vahid / Trancoso, Isabel / Hämäläinen, Annika / Dias, Miguel Sales: "Speaker age estimation for elderly speech recognition in European Portuguese", 2962-2966.](#)

[Najafian, Maryam / DeMarco, Andrea / Cox, Stephen / Russell, Martin: "Unsupervised model selection for recognition of regional accented speech", 2967-2971.](#)

[Zhang, Wen-Lin / Qu, Dan / Zhang, Wei-Qiang / Li, Bi-Cheng: "Speaker adaptation based on sparse and low-rank eigenphone matrix estimation", 2972-2976.](#)

[Huang, Yan / Yu, Dong / Liu, Chaojun / Gong, Yifan: "Multi-accent deep neural network acoustic model with accent-specific top layer using the KLD-regularized model adaptation", 2977-2981.](#)

[Shahnawazuddin, S. / Sinha, Rohit: "A low complexity model adaptation approach involving sparse coding over multiple dictionaries", 2982-2986.](#)

[Kubota, Yuichi / Omachi, Motoi / Ogawa, Tetsuji / Kobayashi, Tetsunori / Nitta, Tsuneo: "Effect of frequency weighting on MLP-based speaker canonicalization", 2987-2991.](#)

[Huang, Zhen / Li, Jinyu / Siniscalchi, Sabato Marco / Chen, I-Fan / Weng, Chao / Lee, Chin-Hui: "Feature space maximum a posteriori linear regression for adaptation of deep neural networks", 2992-2996.](#)

[Tomashenko, Natalia / Khokhlov, Yuri: "Speaker adaptation of context dependent deep neural networks based on MAP-adaptation and GMM-derived feature processing", 2997-3001.](#)

[Karafiát, Martin / Grézl, František / Veselý, Karel / Hannemann, Mirko / Szőke, Igor / Černocký, Jan: "BUT 2014 Babel system: analysis of adaptation in NN based systems", 3002-3006.](#)

[Rouvier, Mickael / Favre, Benoit: "Speaker adaptation of DNN-based ASR with i-vectors: does it actually adapt models to speakers?", 3007-3011.](#)

Speaker Localization

[Singhal, Kushagra / Hegde, Rajesh M.: "A sparse reconstruction method for speech source localization using partial dictionaries over a spherical microphone array", 2209-2213.](#)

[Cui, Weiwei / Cho, Jaeyeon / Lee, Seungyeol: "A robust TDOA estimation method for in-car-noise environments", 2214-2217.](#)

[Netsch, Lorin / Stachurski, Jacek: "Robust low-resource sound localization in correlated noise", 2218-2222.](#)

[Ying, Dongwen / Zhou, Ruohua / Li, Junfeng / Pan, Jielin / Yan, Yonghong: "Direction-of-arrival estimation of multiple speakers using a planar array", 2223-2227.](#)

[Xue, Wei / Liang, Shan / Liu, Wenju: "Weighted spatial bispectrum correlation matrix for DOA estimation in the presence of interferences", 2228-2232.](#)

[Bouafif, Mariem / Lachiri, Zied: "Multi-sources separation for sound source localization", 2233-2237.](#)

Speech Representation, Detection and Classification

[Zhang, Chiyuan / Voinea, Stephen / Evangelopoulos, Georgios / Rosasco, Lorenzo / Poggio, Tomaso: "Phone classification by a hierarchy of invariant representation layers", 2346-2350.](#)

[Sinclair, Mark / Bell, Peter / Birch, Alexandra / McInnes, Fergus: "A semi-Markov model for speech segmentation with an utterance-break prior", 2351-2355.](#)

[Aneeja, G. / Yegnanarayana, B.: "Speech detection in transient noises", 2356-2360.](#)

[He, Yongjun / Sun, Guanglu / Zheng, Guibin / Han, Jiqing: "Evaluation of dictionary for sparse coding in speech processing", 2361-2364.](#)

[Vaz, Colin / Ramanarayanan, Vikram / Narayanan, Shrikanth S.: "Joint filtering and factorization for recovering latent structure from noisy speech data", 2365-2369.](#)

[Gallardo-Antolín, A. / Montero, J. M. / King, Simon: "A comparison of open-source segmentation architectures for dealing with imperfect data from the media in speech synthesis", 2370-2374.](#)

[Asami, Taichi / Masumura, Ryo / Masataki, Hirokazu / Sakauchi, Sumitaka: "Read and spontaneous speech classification based on variance of GMM supervectors", 2375-2379.](#)

[Shokouhi, Navid / Sadjadi, Seyed Omid / Hansen, John H. L.: "Co-channel speech detection via spectral analysis of frequency modulated sub-bands", 2380-2384.](#)

[Voinea, Stephen / Zhang, Chiyuan / Evangelopoulos, Georgios / Rosasco, Lorenzo / Poggio, Tomaso: "Word-level invariant representations from acoustic waveforms", 2385-2389.](#)

[Dalsgaard, Paul / Andersen, Ove: "On closed form calculation of line spectral frequencies \(LSF\)", 2390-2394.](#)

[Ouali, Chahid / Dumouchel, Pierre / Gupta, Vishwa: "Robust features for content-based audio copy detection", 2395-2399.](#)

[Jiang, Yi / Wang, DeLiang / Liu, RunSheng: "Binaural deep neural network classification for reverberant speech segregation", 2400-2404.](#)

Spoken Term Detection for Low-Resource Languages I, II

[Anguera, Xavier / Rodriguez-Fuentes, Luis Javier / Szóke, Igor / Buzo, Andi / Metze, Florian / Penagarikano, Mikel: "Query-by-example spoken term detection on multilingual unconstrained speech", 2459-2463.](#)

[Soto, Victor / Mangu, Lidia / Rosenberg, Andrew / Hirschberg, Julia: "A comparison of multiple methods for rescoring keyword search lists for low resource languages", 2464-2468.](#)

[Karakos, Damianos / Schwartz, Richard: "Subword and phonetic search for detecting out-of-vocabulary keywords", 2469-2473.](#)

[Wang, Yun / Metze, Florian: "An in-depth comparison of keyword specific thresholding and sum-to-one score normalization", 2474-2478.](#)

[Lee, Hung-yi / Zhang, Yu / Chuangsuwanich, Ekapol / Glass, James R.: "Graph-based re-ranking using acoustic feature similarity between search results for spoken term detection on low-resource languages", 2479-2483.](#)

[Le, Viet-Bac / Lamel, Lori / Messaoudi, Abdel / Hartmann, William / Gauvain, Jean-Luc / Woehrling, Cécile / Despres, Julien / Roy, Anindya: "Developing STT and KWS systems using limited language resources", 2484-2488.](#)

[Hartmann, William / Le, Viet-Bac / Messaoudi, Abdel / Lamel, Lori / Gauvain, Jean-Luc: "Comparing decoding strategies for subword-based keyword spotting in low-resourced languages", 2764-2768.](#)

[Ma, Min / Richards, Justin / Soto, Victor / Hirschberg, Julia / Rosenberg, Andrew: "Strategies for rescoring keyword search results using word-burst and acoustic features", 2769-2773.](#)

[Xu, Di / Metze, Florian: "Word-based probabilistic phonetic retrieval for low-resource spoken term detection", 2774-2778.](#)

[Chen, I-Fan / Chen, Nancy F. / Lee, Chin-Hui: "A keyword-boosted sMBR criterion to enhance keyword search performance in deep neural network based acoustic modeling", 2779-2783.](#)

[Chiu, Justin / Wang, Yun / Trmal, Jan / Povey, Daniel / Chen, Guoguo / Rudnicky, Alexander I.: "Combination of FST and CN search in spoken term detection", 2784-2788.](#)

[Liu, Chunxi / Jansen, Aren / Chen, Guoguo / Kintzley, Keith / Trmal, Jan / Khudanpur, Sanjeev: "Low-resource open vocabulary keyword search using point process models", 2789-2793.](#)

Voice Conversion

[Ohtani, Yamato / Tamura, Masatsune / Morita, Masahiro / Akamine, Masami: "GMM-based bandwidth extension using sub-band basis spectrum model", 2489-2493.](#)

[Nakamura, Kazuhiro / Hashimoto, Kei / Oura, Keiichiro / Nankaku, Yoshihiko / Tokuda, Keiichi: "A mel-cepstral analysis technique restoring high frequency components from low-sampling-rate speech", 2494-2498.](#)

[Lee, S. W. / Wu, Zhizheng / Dong, Minghui / Tian, Xiaohai / Li, Haizhou: "A comparative study of spectral transformation techniques for singing voice synthesis", 2499-2503.](#)

[Saito, Daisuke / Doi, Hidenobu / Minematsu, Nobuaki / Hirose, Keikichi: "Application of matrix variate Gaussian mixture model to statistical voice conversion", 2504-2508.](#)

[Wu, Zhizheng / Chng, Eng Siong / Li, Haizhou: "Joint nonnegative matrix factorization for exemplar-based voice conversion", 2509-2513.](#)

[Kobayashi, Kazuhiro / Toda, Tomoki / Neubig, Graham / Sakti, Sakriani / Nakamura, Satoshi: "Statistical singing voice conversion with direct waveform modification based on the spectrum differential", 2514-2518.](#)

Speech and Audio Segmentation and Classification

[Ellis, Daniel P. W. / Satoh, Hiroyuki / Chen, Zhuo: "Detecting proximity from personal audio recordings", 2519-2523.](#)

[Phan, Huy / Maaß, Marco / Mazur, Radoslaw / Mertins, Alfred: "Acoustic event detection and localization with regression forests", 2524-2528.](#)

[Ferràs, Marc / Boulard, Hervé: "Multi-source posteriors for speech activity detection on public talks", 2529-2532.](#)

[Dennis, Jonathan / Tran, Huy Dat / Chng, Eng Siong: "Analysis of spectrogram image methods for sound event classification", 2533-2537.](#)

[Satt, Aharon / Hoory, Ron / König, Alexandra / Aalten, Pauline / Robert, Philippe H.: "Speech-based automatic and robust detection of very early dementia", 2538-2542.](#)

[Raboshchuk, Ganna / Nadeu, Climent / Ghahabi, Omid / Solvez, Sergi / Mahamud, Blanca Muñoz / Veciana, Ana Riverola de / Hervas, Santiago Navarro: "On the acoustic environment of a neonatal intensive care unit: initial description, and detection of equipment alarms", 2543-2547.](#)

Language Acquisition

[Fox, Robert Allen / Jacewicz, Ewa / Hardjono, Florence: "Non-native perception of regionally accented speech in a multitalker context", 2548-2552.](#)

[Turco, Giuseppina / Delais-Roussarie, Elisabeth: "A crosslinguistic and acquisitional perspective on intonational rises in French", 2553-2557.](#)

[Tu, Jung-Yueh / Hsiung, Yuwen / Wu, Min-Da / Sung, Yao-Ting: "Error patterns of Mandarin disyllabic tones by Japanese learners", 2558-2562.](#)

[Leong, Victoria / Kalashnikova, Marina / Burnham, Denis / Goswami, Usha: "Infant-directed speech enhances temporal rhythmic structure in the envelope", 2563-2567.](#)

[Wewalaarachchi, Dilu / Singh, Leher: "Influences of tone sandhi on word recognition in preschool children", 2568-2571.](#)

[Goh, Hwee Hwee / Hu, Charlene / Yeo, Kheng Hui / Singh, Leher: "Lexical representation of consonant, vowels and tones in early childhood", 2572-2574.](#)

Speech Perception

[Francisco, Ana A. / Jesse, Alexandra / Groen, Margriet A. / McQueen, James M.: "Audiovisual temporal sensitivity in typical and dyslexic adult readers", 2575-2579.](#)

[Derrick, Donald / O'Beirne, Greg A. / Rybel, Tom De / Hay, Jennifer: "Aero-tactile integration in fricatives: converting audio to air flow information for speech perception enhancement", 2580-2584.](#)

[Mai, Guangting: "Relative importance of AM and FM cues for speech comprehension: effects of speaking rate and their implications for neurophysiological processing of speech", 2585-2589.](#)

[Stringer, Louise / Iverson, Paul: "The effect of regional and non-native accents on word recognition processes: a comparison of EEG responses in quiet to speech recognition in noise", 2590-2594.](#)

[Fong, Manson C. -M. / Minett, James W. / Blu, Thierry / Wang, William S. -Y.: "Towards a neural measure of perceptual distance — classification of electroencephalographic responses to synthetic vowels", 2595-2599.](#)

[Scharenborg, Odette / Sanders, Eric / Cranen, Bert: "Collecting a corpus of Dutch noise-induced `slips of the ear'", 2600-2604.](#)

Language and Lexical Modeling

[Hanai, Tuka Al / Glass, James R.: "Lexical modeling for Arabic ASR: a systematic approach", 2605-2609.](#)

[Orosanu, Luiza / Jouviet, Denis: "Hybrid language models for speech transcription", 2610-2614.](#)

[Gandhe, Ankur / Metze, Florian / Lane, Ian: "Neural network language models for low resource languages", 2615-2619.](#)

[Gangireddy, Siva Reddy / McInnes, Fergus / Renals, Steve: "Feed forward pre-training for recurrent neural network language models", 2620-2624.](#)

[Roy, Brandon C. / Vosoughi, Soroush / Roy, Deb: "Grounding language models in spatiotemporal context", 2625-2629.](#)

[Jalalvand, Shahab / Falavigna, Daniele: "Direct word graph rescoring using a* search and RNNLM", 2630-2634.](#)

[Chelba, Ciprian / Mikolov, Tomas / Schuster, Mike / Ge, Qi / Brants, Thorsten / Koehn, Phillip / Robinson, Tony: "One billion word benchmark for measuring progress in statistical language modeling", 2635-2639.](#)

[Schnall, Andrea / Heckmann, Martin: "Integrating sequence information in the audio-visual detection of word prominence in a human-machine interaction scenario", 2640-2644.](#)

[Biadsy, Fadi / Hall, Keith / Moreno, Pedro J. / Roark, Brian: "Backoff inspired features for maximum entropy language models", 2645-2649.](#)

[Telaar, Dominic / Wand, Michael / Gehrig, Dirk / Putze, Felix / Amma, Christoph / Heger, Dominic / Vu, Ngoc Thang / Erhardt, Mark / Schlippe, Tim / Janke, Matthias / Herff, Christian / Schultz, Tanja: "BioKIT — real-time decoder for biosignal processing", 2650-2654.](#)

[Harwath, David / Glass, James R.: "Speech recognition without a lexicon — bridging the gap between graphemic and phonetic systems", 2655-2659.](#)

Speech Enhancement (Single- and Multi-Channel) 1, 2

[Zhao, Shengkui / Jones, Douglas L.: "A new auxiliary-vector algorithm with conjugate orthogonality for speech enhancement", 2660-2664.](#)

[Jathar, Neehar / Rao, Preeti: "Acoustic characteristics of critical message utterances in noise applied to speech intelligibility enhancement", 2665-2669.](#)

[Xu, Yong / Du, Jun / Dai, Li-Rong / Lee, Chin-Hui: "Dynamic noise aware training for speech enhancement based on deep neural networks", 2670-2674.](#)

[Pertilä, Pasi / Nikunen, Joonas: "Microphone array post-filtering using supervised machine learning for speech enhancement", 2675-2679.](#)

[Mani, Senthil Kumar / Dhiman, Jitendra Kumar / Murty, K. Sri Rama: "Novel speech duration modifier for packet based communication system", 2680-2684.](#)

[Liu, Ding / Smaragdakis, Paris / Kim, Minje: "Experiments on deep learning for speech denoising", 2685-2689.](#)

[Mohammadiha, Nasser / Doclo, Simon: "Single-channel dynamic exemplar-based speech enhancement", 2690-2694.](#)

[Kato, Akihiro / Milner, Ben: "Using hidden Markov models for speech enhancement", 2695-2699.](#)

[Pfeifenberger, Lukas / Pernkopf, Franz: "Blind source extraction based on a direction-dependent a-priori SNR", 2700-2704.](#)

[Chacón, Carlos Eduardo Cancino / Mowlaei, Pejman: "Least squares phase estimation of mixed signals", 2705-2709.](#)

[Ming, Ji / Crookes, Danny: "Speech enhancement from additive noise and channel distortion — a corpus-based approach", 2710-2714.](#)

[Zhou, Zhiyuan / Ding, Zhaogui / Li, Weifeng / Wu, Zhiyong / Wang, Longbiao / Liao, Qingmin: "Multi-channel speech enhancement using sparse coding on local time-frequency structures", 2824-2827.](#)

[Mirsamadi, Seyedmahdad / Hansen, John H. L.: "Multichannel speech dereverberation based on convolutive nonnegative tensor factorization for ASR applications", 2828-2832.](#)

[Chen, Zhuo / McFee, Brian / Ellis, Daniel P. W.: "Speech enhancement by low-rank and convolutive dictionary spectrogram decomposition", 2833-2837.](#)

[Jaureguiberry, Xabier / Vincent, Emmanuel / Richard, Gaël: "Multiple-order non-negative matrix factorization for speech enhancement", 2838-2842.](#)

[Kang, Tae Gyoon / Kwon, Kisoo / Shin, Jong Won / Kim, Nam Soo: "NMF-based speech enhancement incorporating deep neural network", 2843-2846.](#)

[Sonowal, Sukanya / Kwon, Kisoo / Kim, Nam Soo / Shin, Jong Won: "A data-driven approach to speech enhancement using Gaussian process", 2847-2851.](#)

Speech Coding and Transmission

[Bäckström, Tom / Helmrich, Christian R.: "Decorrelated innovative codebooks for ACELP using factorization of autocorrelation matrix", 2794-2798.](#)

[Cernak, Milos / Lazaridis, Alexandros / Garner, Philip N. / Motlicek, Petr: "Stress and accent transmission in HMM-based syllable-context very low bit rate speech coding", 2799-2803.](#)

[Pulakka, Hannu / Rämö, Anssi / Myllylä, Ville / Toukoma, Henri / Alku, Paavo: "Subjective voice quality evaluation of artificial bandwidth extension: comparing different audio bandwidths and speech codecs", 2804-2808.](#)

[Fu, Zhong-Hua / Xie, Lei: "Stereo acoustic echo suppression using widely linear filtering in the frequency domain", 2809-2813.](#)

[Lee, Bong-Ki / Hwang, Inyoung / Park, Jihwan / Chang, Joon-Hyuk: "Enhanced muting method in packet loss concealment of ITU-t g.722 using sigmoid function with on-line optimized parameters", 2814-2818.](#)

[Wu, Chao / Jiang, Kaiyu / Guo, Yanmeng / Fu, Qiang / Yan, Yonghong: "A robust step-size control algorithm for frequency domain acoustic echo cancellation", 2819-2823.](#)

Unsupervised or Corrective Lexical Modeling

[Byambakhishig, E. / Tanaka, K. / Aihara, Ryo / Nakashika, Toru / Takiguchi, Tetsuya / Ariki, Yasuo: "Error correction of automatic speech recognition based on normalized web distance", 2852-2856.](#)

[Dikici, Erinc / Saraçlar, Murat: "Unsupervised training methods for discriminative language modeling", 2857-2861.](#)

[Qin, Long / Rudnicky, Alexander I.: "Building a vocabulary self-learning speech recognition system", 2862-2866.](#)

[Schlippe, Tim / Merz, Matthias / Schultz, Tanja: "Methods for efficient semi-automatic pronunciation dictionary bootstrapping", 2867-2871.](#)

[Akbacak, Murat / Hakkani-Tür, Dilek / Tur, Gokhan: "Rapidly building domain-specific entity-centric language models using semantic web knowledge sources", 2872-2876.](#)

[Lee, Ann / Glass, James R.: "Context-dependent pronunciation error pattern discovery with limited annotations", 2877-2881.](#)

Meta Data

[Sapru, Ashtosh / Boulard, Hervé: "Detecting speaker roles and topic changes in multiparty conversations using latent topic models", 2882-2886.](#)

[Xu, Chenglin / Xie, Lei / Huang, Guangpu / Xiao, Xiong / Chng, Eng Siong / Li, Haizhou: "A deep neural network approach for sentence boundary detection in broadcast news", 2887-2891.](#)

[Gupta, Rahul / Ananthkrishnan, Sankaranarayanan / Yang, Zhaojun / Narayanan, Shrikanth S.: "Variable Span disfluency detection in ASR transcripts", 2892-2896.](#)

[Dutrey, Camille / Clavel, Chloé / Rosset, Sophie / Vasilescu, Ioana / Adda-Decker, Martine: "A CRF-based approach to automatic disfluency detection in a French call-centre corpus", 2897-2901.](#)

[Hasan, Madina / Doddipatla, Rama / Hain, Thomas: "Multi-pass sentence-end detection of lecture speech", 2902-2906.](#)

[Zayats, Victoria / Ostendorf, Mari / Hajishirzi, Hannaneh: "Multi-domain disfluency and repair detection", 2907-2911.](#)

Language Recognition

[Jiang, Bing / Song, Yan / Wei, Si / McLoughlin, Ian Vince / Dai, Li-Rong: "Task-aware deep bottleneck features for spoken language identification", 3012-3016.](#)

[Tong, Rong / Ma, Bin / Li, Haizhou: "Virtual example for phonotactic language recognition", 3017-3021.](#)

[Liu, Wei-Wei / Zhang, Wei-Qiang / Liu, Jia: "Phonotactic language recognition based on time-gap-weighted lattice kernels", 3022-3026.](#)

[Segbroeck, Maarten van / Travadi, Ruchir / Narayanan, Shrikanth S.: "UBM fused total variability modeling for language identification", 3027-3031.](#)

[Diez, Mireia / Penagarikano, Mikel / Bordel, German / Varona, Amparo / Rodriguez-Fuentes, Luis Javier: "On the complementarity of short-time fourier analysis windows of different lengths for improved language recognition", 3032-3036.](#)

[Travadi, Ruchir / Segbroeck, Maarten Van / Narayanan, Shrikanth S.: "Modified-prior i-vector estimation for language identification of short duration utterances", 3037-3041.](#)

[D'Haro, Luis Fernando / Cordoba, Ricardo / Salamea, Christian / Ferreiros, Javier: "Language recognition using phonotactic-based shifted delta coefficients and multiple phone recognizers", 3042-3046.](#)

[Plchot, Oldřich / Diez, Mireia / Soufifar, Mehdi / Burget, Lukáš: "PLLR features in language recognition system for RATS", 3047-3051.](#)

[Yeong, Yin-Lai / Tan, Tien-Ping: "Language identification of code Switching sentences and multilingual sentences of under-resourced languages by using multi structural word information", 3052-3055.](#)