2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU 2015)

Scottsdale, Arizona, USA
13-17 December 2015

IEEE

IEEE Catalog Number: CFP15SRW-POD
ISBN: 978-1-4799-7292-0

Copyright © 2015 by the Institute of Electrical and Electronic Engineers, Inc
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

******This publication is a representation of what appears in the IEEE Digital Libraries. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number: CFP15SRW-POD
ISBN (Print-On-Demand): 978-1-4799-7292-0
ISBN (Online): 978-1-4799-7291-3

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) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

M1: AUTOMATIC SPEECH RECOGNITION I

M1.1: DIFFERENT WORD REPRESENTATIONS AND THEIR COMBINATION FOR PROPER NAME RETRIEVAL FROM DIACHRONIC DOCUMENTS 1
Irina Illina, Dominique Fohr, LORIA-INRIA, France

M1.2: SPARSE NON-NEGATIVE MATRIX LANGUAGE MODELING FOR GEO-ANNOTATED QUERY SESSION DATA 8
Ciprian Chelba, Noam Shazeer, Google Inc, United States

M1.3: TRAINING DATA PSEUDO-SHUFFLING AND DIRECT DECODING FRAMEWORK FOR RECURRENT NEURAL NETWORK BASED ACOUSTIC MODELING 15
Naoyuki Kanda, Mitsuyoshi Tachimori, Xugang Lu, Hisashi Kawai, National Institute of Information and Communications Technology, Japan

M1.4: ON CONSTRUCTING AND ANALYSING AN INTERPRETABLE BRAIN MODEL FOR THE DNN BASED ON HIDDEN ACTIVITY PATTERNS 22
Khe Chai Sim, National University of Singapore, Singapore

M1.5: SPEAKER LOCATION AND MICROPHONE SPACING INVARIANT ACOUSTIC MODELING FROM RAW MULTICHANNEL WAVEFORMS 30
Tara Sainath, Ron Weiss, Kevin Wilson, Arun Narayanan, Michiel Bacchiani, Andrew Senior, Google Inc, United States

M1.6: HYBRID DNN-LATENT STRUCTURED SVM ACOUSTIC MODELS FOR CONTINUOUS SPEECH RECOGNITION 37
Suman Ravuri, International Computer Science Institute; University of California - Berkeley, United States

M1.7: DISCRIMINATIVE TRAINING OF CONTEXT-DEPENDENT LANGUAGE MODEL SCALING FACTORS AND INTERPOLATION WEIGHTS 45
Shuangyu Chang, Abhik Lahiri, Issac Alphonso, Barlas Oguz, Michael Levit, Microsoft Corporation, United States; Benoit Dumoulin, Facebook Inc., United States

M1.8: ACOUSTIC MODEL TRAINING BASED ON NODE-WISE WEIGHT BOUNDARY MODEL INCREASING SPEED OF DISCRETE NEURAL NETWORKS 52
Ryu Takeda, Kazunori Komatani, Osaka University, Japan; Kazuhiro Nakadai, Honda Research Institute Japan Co., Ltd., Japan

M1.9: TWO-STAGE ASGD FRAMEWORK FOR PARALLEL TRAINING OF DNN ACOUSTIC MODELS USING ETHERNET 59
Zhichao Wang, Xingyu Na, Xin Li, Jielin Pan, Yonghong Yan, Institute of Acoustics, Chinese Academy of Sciences, China

M1.10: RNNDROP: A NOVEL DROPOUT FOR RNNS IN ASR 65
Taesup Moon, Heeyoul Choi, Hoshik Lee, Inchul Song, Samsung Advanced Institute of Technology, Republic of Korea

M1.11: SPECTRAL LEARNING WITH NON NEGATIVE PROBABILITIES FOR FINITE STATE AUTOMATON 71
Hadrien Glaude, Thales Airborne Systems / University Lille 1, France; Cyrille Enderli, Thales Airborne Systems, France; Olivier Pietquin, University Lille 1, France

M1.12: DEEP BI-DIRECTIONAL RECURRENT NETWORKS OVER SPECTRAL WINDOWS 78
Abdel-Rahman Mohamed, Frank Seide, Dong Yu, Jasha Droppo, Andreas Stolcke, Geoffrey Zweig, Microsoft, United States; Gerald Penn, University of Toronto, Canada

M1.13: PERSONALIZING UNIVERSAL RECURRENT NEURAL NETWORK LANGUAGE MODEL WITH USER CHARACTERISTIC FEATURES BY SOCIAL NETWORK CROWDSOURCING 84
Bo-Hsiang Tseng, Hung-yi Lee, Lin-Shan Lee, National Taiwan University, Taiwan

M1.14: TIME DELAY DEEP NEURAL NETWORK-BASED UNIVERSAL BACKGROUND MODELS FOR SPEAKER RECOGNITION 92
David Snyder, Daniel Garcia-Romero, Daniel Povey, The Johns Hopkins University, United States

M2: TEXT-TO-SPEECH SYSTEMS

M2.1: AUTOMATIC PROSODY PREDICTION FOR CHINESE SPEECH SYNTHESIS USING BLSTM-RNN AND EMBEDDING FEATURES 98
Chuang Ding, Lei Xie, Jie Yan, Weini Zhang, Yang Liu, Northwestern Polytechnical University, China

M2.2: NATURALNESS AND RAPPORT IN A PITCH ADAPTIVE LEARNING COMPANION.....103
Nichola Lubold, Arizona State University, United States; Heather Pon-Barry, Mount Holyoke College, United States; Erin Walker, Arizona State University, United States

M2.3: LEARNING CONTINUOUS REPRESENTATION OF TEXT FOR PHONE DURATION MODELING IN STATISTICAL PARAMETRIC SPEECH SYNTHESIS 111
Sai Krishna Rallabandi, Sai Sirisha Rallabandi, Padmini Bandi, Suryakanth Gangashetty, International Institute of Information Technology- Hyderabad, India

M2.4: SPEAKER INTONATION ADAPTATION FOR TRANSFORMING TEXT-TO-SPEECH SYNTHESIS SPEAKER IDENTITY 116
Mahsa Sadat Elyasi Langarani, Jan van Santen, Oregon Health and Science University, United States

M3: AUTOMATIC SPEECH RECOGNITION II

M3.1: INVESTIGATING SPARSE DEEP NEURAL NETWORKS FOR SPEECH RECOGNITION 124
Gueorgui Pironkov, Stéphane Dupont, Thierry Dutoit, University of Mons, Belgium

M3.2: LATENT DIRICHLET ALLOCATION BASED ORGANISATION OF BROADCAST MEDIA ARCHIVES FOR DEEP NEURAL NETWORK ADAPTATION 130
Mortaza Doulaty, Oscar Saz, Raymond W. M. Ng, Thomas Hain, University of Sheffield, United Kingdom

M3.3: TOWARDS STRUCTURED DEEP NEURAL NETWORK FOR AUTOMATIC SPEECH RECOGNITION 137
Yi-Hsiu Liao, Graduate Institute of Electronic Engineering, National Taiwan University, Taiwan; Hung-Yi Lee, Graduate Institute of Electrical Engineering, National Taiwan University, Taiwan; Lin-shan Lee, Graduate Institute of Electronic Engineering, National Taiwan University, Taiwan

M3.4: LEARNING FACTORIZED FEATURE TRANSFORMS FOR SPEAKER NORMALIZATION 145
Lahiru Samarakoon, Khe Chai SIM, National University of Singapore, Singapore

M3.5: IMPROVING DATA SELECTION FOR LOW-RESOURCE STT AND KWS 153
Thiago Fraga da Silva, Antoine Laurent, Vocapia Research, France; Jean-Luc Gauvain, Lori Lamel, CNRS-LIMSI, France; Viet Bac Le, Abdel Messaoudi, Vocapia Research, France

M3.6: STRUCTURED DISCRIMINATIVE MODELS USING DEEP NEURAL-NETWORK FEATURES 160
Rogier van Dalen, Jingzhou Yang, Haipeng Wang, Anton Ragni, Chao Zhang, Mark J. F. Gales, University of Cambridge, United Kingdom

M3.7: EESEN: END-TO-END SPEECH RECOGNITION USING DEEP RNN MODELS AND WFST-BASED DECODING 167
Yajie Miao, Mohammad Gowayyed, Florian Metze, Carnegie Mellon University, United States

M3.8: STOCHASTIC GRADIENT VARIATIONAL BAYES FOR DEEP LEARNING-BASED ASR 175
Andros Tjandra, Universitas Indonesia, Indonesia; Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology, Japan; Mirna Adriani, Universitas Indonesia, Indonesia

M3.9: INVESTIGATION OF BACK-OFF BASED INTERPOLATION BETWEEN RECURRENT NEURAL NETWORK AND N-GRAM LANGUAGE MODELS 181
Xie Chen, Xunying Liu, Mark J. F. Gales, Philip C. Woodland, Cambridge University, United Kingdom

M3.10: LSTM TIME AND FREQUENCY RECURRENCE FOR AUTOMATIC SPEECH RECOGNITION 187
Jinyu Li, Abdel-Rahman Mohamed, Geoffrey Zweig, Yifan Gong, Microsoft, United States

M4: SPOKEN DOCUMENT RETRIEVAL, SPEECH SUMMARIZATION, AND APPLICATIONS

M4.1: INCORPORATING USER FEEDBACK TO RE-RANK KEYWORD SEARCH RESULTS 192
Scott Novotney, Kevin Jett, Owen Kimball, Raytheon BBN Technologies, United States

M4.2: COMBINATION OF SYLLABLE BASED N-GRAM SEARCH AND WORD SEARCH FOR SPOKEN TERM DETECTION THROUGH SPOKEN QUERIES AND IV/OOV CLASSIFICATION 200
Nagisa Sakamoto, Kazumasa Yamamoto, Seiichi Nakagawa, Toyohashi University of Technology, Japan

M4.3: INCORPORATING PARAGRAPH EMBEDDINGS AND DENSITY PEAKS CLUSTERING FOR SPOKEN DOCUMENT SUMMARIZATION 207
Kuan-Yu Chen, Academia Sinica, Taiwan; Kai-Wun Shih, National Taiwan Normal University, Taiwan; Shih-Hung Liu, Academia Sinica, Taiwan; Berlin Chen, National Taiwan Normal University, Taiwan; Hsin-Min Wang, Academia Sinica, Taiwan

M4.4: HIGH-PERFORMANCE SWAHILI KEYWORD SEARCH WITH VERY LIMITED LANGUAGE PACK: THE THUEE SYSTEM FOR THE OPENKWS15 EVALUATION 215
Meng Cai, Zhiqiang Lv, Cheng Lu, Jian Kang, Like Hui, Zhuo Zhang, Jia Liu, Tsinghua University, China

M4.5: PHONETIC UNIT SELECTION FOR CROSS-LINGUAL QUERY-BY-EXAMPLE SPOKEN TERM DETECTION 223
Paula Lopez-Otero, Laura Docio-Fernandez, Carmen Garcia-Mateo, Universidade de Vigo, Spain

M4.6: IMPROVED SYSTEM FUSION FOR KEYWORD SEARCH 230
Zhiqiang Lv, Meng Cai, Cheng Lu, Jian Kang, Like Hui, Wei-Qiang Zhang, Jia Liu, Tsinghua University, China

M4.7: DEEP MULTIMODAL SEMANTIC EMBEDDINGS FOR SPEECH AND IMAGES 237
David Harwath, James Glass, Massachusetts Institute of Technology, United States

M4.8: AN ITERATIVE DEEP LEARNING FRAMEWORK FOR UNSUPERVISED DISCOVERY OF SPEECH FEATURES AND LINGUISTIC UNITS WITH APPLICATIONS ON SPOKEN TERM DETECTION 245
Cheng-Tao Chung, Cheng-Yu Tsai, Hsiang-Hung Lu, Chia-Hsiang Liu, Hung-yi Lee, Lin-shan Lee, National Taiwan University, Taiwan

M4.9: INCREMENTAL SENTENCE COMPRESSION USING LSTM RECURRENT NETWORKS 252
Sakriani Sakti, Nara Institute of Science and Technology, Japan; Faiz Ilham, Bandung Institute of Technology, Indonesia; Graham Neubig, Tomoki Toda, Nara Institute of Science and Technology, Japan; Ayu Purwarianti, Bandung Institute of Technology, Indonesia; Satoshi Nakamura, Nara Institute of Science and Technology, Japan

M4.10: MULTILINGUAL REPRESENTATIONS FOR LOW RESOURCE SPEECH RECOGNITION AND KEYWORD SEARCH 259
Jia Cui, Brian Kingsbury, Bhuvana Ramabhadran, Abhinav Sethy, Kartik Audhkhasi, Xiaodong Cui, Ellen Kislal, Lidia Mangu, Markus Nussbaum-Thom, Michael Picheny, IBM T.J. Watson, United States; Zoltán Tüske, Pavel Golik, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany; Mark J. F. Gales, Kate M. Knill, Anton Ragni, Haipeng Wang, Philip C. Woodland, Cambridge University, United Kingdom

R1: ROBUSTNESS IN AUTOMATIC SPEECH RECOGNITION, SPEECH-TO-SPEECH TRANSLATION, AND SPONTANEOUS SPEECH PROCESSING

R1.1: SPOKEN LANGUAGE TRANSLATION GRAPHS RE-DECODING USING AUTOMATIC QUALITY ASSESSMENT 267
Laurent Besacier, Benjamin Lecouteux, LIG - Univ. Grenoble Alpes, France; Ngoc-Quang Luong, IDIAP - Switzerland, Switzerland; Le Ngoc-Tien, LIG - Univ. Grenoble Alpes, France

R1.2: THE DIRHA-ENGLISH CORPUS AND RELATED TASKS FOR DISTANT-SPEECH RECOGNITION IN DOMESTIC ENVIRONMENTS 275
Mirco Ravanelli, Luca Cristoforetti, Roberto Gretter, Marco Pellin, Alessandro Sosi, Maurizio Omologo, Fondazione Bruno Kessler, Italy

R1.3: UNCERTAINTY ESTIMATION OF DNN CLASSIFIERS 283
Sri Harish Mallidi, The Johns Hopkins University, United States; Tetsuji Ogawa, Waseda University, Japan; Hynek Hermansky, The Johns Hopkins University, United States

R1.4: TOWARDS UTTERANCE-BASED NEURAL NETWORK ADAPTATION IN ACOUSTIC MODELING 289
Ivan Himawan, Petr Motlicek, Marc Ferras Font, Srikanth Madikeri, Idiap Research Institute, Switzerland

R1.5: PHONETICALLY-ORIENTED WORD ERROR ALIGNMENT FOR SPEECH RECOGNITION ERROR ANALYSIS IN SPEECH TRANSLATION 296
Nicholas Ruiz, Marcello Federico, Fondazione Bruno Kessler, Italy

R1.6: UTTERANCE CLASSIFICATION IN SPEECH-TO-SPEECH TRANSLATION FOR ZERO-RESOURCE LANGUAGES IN THE HOSPITAL ADMINISTRATION DOMAIN 303
Lara J. Martin, Andrew Wilkinson, Sai Sumanth Miryala, Vivian Robison, Alan W. Black, Carnegie Mellon University, United States

R1.7: MULTI-TASK JOINT-LEARNING OF DEEP NEURAL NETWORKS FOR ROBUST SPEECH RECOGNITION 310
Yanmin Qian, Maofan Yin, Yongbin You, Kai Yu, Shanghai Jiao Tong University, China

R1.8: TIME-FREQUENCY CONVOLUTIONAL NETWORKS FOR ROBUST SPEECH RECOGNITION 317
Vikramjit Mitra, Horacio Franco, SRI International, United States

R1.9: NAME-AWARE LANGUAGE MODEL ADAPTATION AND SPARSE FEATURES FOR STATISTICAL MACHINE TRANSLATION 324
Wen Wang, SRI International, United States; Haibo Li, Nuance, United States; Heng Ji, Rensselaer Polytechnic Institute, United States

R1.10: AN I-VECTOR PLDA BASED GENDER IDENTIFICATION APPROACH FOR SEVERELY DISTORTED AND MULTILINGUAL DARPA RATS DATA 331
Shivesh Ranjan, Gang Liu, John H. L. Hansen, The University of Texas at Dallas, United States

R1.11: USING BIDIRECTIONAL LSTM RECURRENT NEURAL NETWORKS TO LEARN HIGH-LEVEL ABSTRACTIONS OF SEQUENTIAL FEATURES FOR AUTOMATED SCORING OF NON-NATIVE SPONTANEOUS SPEECH 338
Zhou Yu, Carnegie Mellon University, United States; Vikram Ramanarayanan, David Suendermann-Oeft, Xinhao Wang, Klaus Zechner, Lei Chen, Jidong Tao, Alexei V. Ivanou, Yao Qian, Educational Testing Service, United States

R2: SPOKEN LANGUAGE UNDERSTANDING

R2.1: TOPIC-SPACE BASED SETUP OF A NEURAL NETWORK FOR THEME IDENTIFICATION OF HIGHLY IMPERFECT TRANSCRIPTIONS 346
Mohamed Morchid, Richard Dufour, Georges Linarès, LIA - University of Avignon, France

R2.2: SEMI-SUPERVISED SLOT TAGGING IN SPOKEN LANGUAGE UNDERSTANDING USING RECURRENT TRANSDUCTIVE SUPPORT VECTOR MACHINES 353
Yangyang Shi, Microsoft, China; Kaisheng Yao, Microsoft Research, United States; Hu Chen, Yi-Cheng Pan, Mei-Yuh Hwang, Microsoft, China

R2.3: A UNIVERSAL MODEL FOR FLEXIBLE ITEM SELECTION IN CONVERSATIONAL DIALOGS 361
Asli Celikyilmaz, Zhaleh Feizollahi, Dilek Hakkani-Tur, Ruhi Sarikaya, Microsoft, United States

R2.4: A COMPARATIVE STUDY OF NEURAL NETWORK MODELS FOR LEXICAL INTENT CLASSIFICATION 368
Suman Ravuri, International Computer Science Institute; University of California - Berkeley, United States; Andreas Stolcke, Microsoft Research; International Computer Science Institute, United States

R2.5: DETECTING ACTIONABLE ITEMS IN MEETINGS BY CONVOLUTIONAL DEEP STRUCTURED SEMANTIC MODELS 375
Yun-Nung Chen, Carnegie Mellon University, United States; Dilek Hakkani-Tur, Xiaodong He, Microsoft Research, United States

R2.6: MULTIMODAL EMBEDDING FUSION FOR ROBUST SPEAKER ROLE RECOGNITION IN VIDEO BROADCAST 383
Mickael Rouvier, Sebastien Delecraz, Benoit Favre, Meriem Bendris, Frederic Béchet, Aix-Marseille Université, France

R2.7: RECENT IMPROVEMENTS TO NEUROCRFS FOR NAMED ENTITY RECOGNITION 390
Marc-Antoine Rondeau, McGill University, Canada; Yi Su, Nuance Communications, Inc., Canada

R2.8: NATURAL LANGUAGE UNDERSTANDING FOR PARTIAL QUERIES 397
Xiaohu Liu, Asli Celikyilmaz, Ruhi Sarikaya, Microsoft, United States

T1: AUTOMATIC SPEECH RECOGNITION IN REVERBERANT ENVIRONMENTS (ASPIRE)

T1.1: ANALYSIS OF FACTORS AFFECTING SYSTEM PERFORMANCE IN THE ASPIRE CHALLENGE 512
Jennifer Melot, Nicolas Malyska, Jessica Ray, Wade Shen, MIT Lincoln Laboratory, United States

T1.2: SINGLE AND MULTI-CHANNEL APPROACHES FOR DISTANT SPEECH RECOGNITION UNDER NOISY REVERBERANT CONDITIONS: I2R'S SYSTEM DESCRIPTION FOR THE ASPIRE CHALLENGE 518
Jonathan Dennis, Huy Dat Tran, Institute For Infocomm Research, Singapore

T1.3: IMPROVING ROBUSTNESS AGAINST REVERBERATION FOR AUTOMATIC SPEECH RECOGNITION 525
Vikramjit Mitra, Julien Van Hout, Wen Wang, Martin Graciarena, Mitchell McLaren, Horacio Franco, Dimitra Vergyri, SRI International, United States

T1.4: ROBUST SPEECH RECOGNITION IN UNKNOWN REVERBERANT AND NOISY CONDITIONS 533
Roger Hsiao, Jeff Ma, William Hartmann, Raytheon BBN Technologies, United States; Martin Karafiat, Frantisek Grezl, Lukas Burget, Igor Szoke, Jan Honza Cernocky, Brno University of Technology, Czech Republic; Shinji Watanabe, Zhuo Chen, Mitsubishi Electric Research Laboratories, United States; Sri Harish Mallidi, Hynek Hermansky, Johns Hopkins University, United States; Stavros Tsakalidis, Richard Schwartz, Raytheon BBN Technologies, United States

T1.5: JHU ASPIRE SYSTEM : ROBUST LVCSR WITH TDNNS, IVECTOR ADAPTATION AND RNN-LMS 539
Vijayaditya Peddinti, Guoguo Chen, Vimal Manohar, Johns Hopkins University, United States; Tom Ko, Huawei, China; Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, United States

T1.6: THE AUTOMATIC SPEECH RECOGITION IN REVERBERANT ENVIRONMENTS (ASPIRE) CHALLENGE 547
Mary Harper, IARPA, United States

T2: 3RD CHIME SPEECH SEPARATION AND RECOGNITION CHALLENGE

T2.1: ADAPTIVE BEAMFORMING AND ADAPTIVE TRAINING OF DNN ACOUSTIC MODELS FOR ENHANCED MULTICHANNEL NOISY SPEECH RECOGNITION 401
Alexey Prudnikov, Speech Technology Center Inc., Russian Federation; Maxim Korenevsky, Sergei Aleinik, ITMO University, Russian Federation

T2.2: BOOSTED ACOUSTIC MODEL LEARNING AND HYPOTHESES RESCORING ON THE CHIME-3 TASK 409
Shahab Jalalvand, University of Trento, Italy; Daniele Falavigna, Marco Matassoni, Piergiorgio Svaizer, Maurizio Omologo, Fondazione Bruno Kessler, Italy

T2.3: UNIFIED ASR SYSTEM USING LGM-BASED SOURCE SEPARATION, NOISE-ROBUST FEATURE EXTRACTION, AND WORD HYPOTHESIS SELECTION 416
Yusuke Fujita, Ryoichi Takashima, Takeshi Homma, Rintaro Ikeshita, Yohei Kawaguchi, Takashi Sumiyoshi, Takashi Endo, Masahito Togami, Hitachi, Ltd, Japan

T2.4: SPEECH ENHANCEMENT USING BEAMFORMING AND NON NEGATIVE MATRIX FACTORIZATION FOR ROBUST SPEECH RECOGNITION IN THE CHIME-3 CHALLENGE 423
Thanh T. Vu, Benjamin Bigot, Eng Siong Chng, Nanyang Technological University, Singapore

T2.5: AN INFORMATION FUSION APPROACH TO RECOGNIZING MICROPHONE ARRAY SPEECH IN THE CHIME-3 CHALLENGE BASED ON A DEEP LEARNING FRAMEWORK 430
Jun Du, Qing Wang, Yan-Hui Tu, Xiao Bao, Li-Rong Dai, University of Science and Technology of China, China; Chin-Hui Lee, Georgia Institute of Technology, United States

T2.6: THE NTT CHIME-3 SYSTEM: ADVANCES IN SPEECH ENHANCEMENT AND RECOGNITION FOR MOBILE MULTI-MICROPHONE DEVICES 436
Takuya Yoshioka, Nobutaka Ito, Marc Delcroix, Atsunori Ogawa, Keisuke Kinoshita, Masakiyo Fujimoto, NTT Corporation, Japan; Chengzhu Yu, The University of Texas at Dallas, United States; Wojciech Fabian, Miquel Espi, Takuya Higuchi, Shoko Araki, Tomohiro Nakatani, NTT Corporation, Japan

T2.7: BLSTM SUPPORTED GEV BEAMFORMER FRONT-END FOR THE 3RD CHIME CHALLENGE 444
Jahn Heymann, Lukas Drude, Aleksej Chinaev, Reinhold Haeb-Umbach, University of Paderborn, Germany

T2.8: MULTI-CHANNEL SPEECH PROCESSING ARCHITECTURES FOR NOISE ROBUST SPEECH RECOGNITION: 3RD CHIME CHALLENGE RESULTS 452
Lukas Pfeifenberger, Tobias Schrank, Matthias Zöhrer, Martin Hagmüller, Franz Pernkopf, Graz University of Technology, Austria

T2.9: ROBUST SPEECH RECOGNITION USING BEAMFORMING WITH ADAPTIVE MICROPHONE GAINS AND MULTICHANNEL NOISE REDUCTION 460
Shengkui Zhao, Advanced Digital Sciences Center, Singapore; Xiong Xiao, Nanyang Technological University, Singapore; Zhaofeng Zhang, Nagaoka University of Technology, Japan; Thi Ngoc Tho Nguyen, Advanced Digital Sciences Center, Singapore; Xionghu Zhong, Nanyang Technological University, Singapore; Bo Ren, Longbiao Wang, Nagaoka University of Technology, Japan; Douglas L. Jones, Advanced Digital Sciences Center, Singapore; Eng Siong Chng, Nanyang Technological University, Singapore; Haizhou Li, Institute For Infocomm Research, Singapore

T2.10: A CHIME-3 CHALLENGE SYSTEM: LONG-TERM ACOUSTIC FEATURES FOR NOISE ROBUST AUTOMATIC SPEECH RECOGNITION 468
Niko Moritz, Stephan Gerlach, Fraunhofer IDMT, Project Group for Hearing, Speech, and Audio Technology, Germany; Kamil Adiloglu, Hörtech gGmbH, Germany; Jörn Anemüller, Birger Kollmeier, University of Oldenburg, Germany; Stefan Goetze, Fraunhofer IDMT, Project Group for Hearing, Speech, and Audio Technology, Germany

T2.11: THE MERL/SRI SYSTEM FOR THE 3RD CHIME CHALLENGE USING BEAMFORMING, ROBUST FEATURE EXTRACTION, AND ADVANCED SPEECH RECOGNITION 475
Takaaki Hori, Mitsubishi Electric Research Laboratories, United States; Zhuo Chen, Columbia University, United States; Hakan Erdogan, Sabanci University, Turkey; John Hershey, Jonathan Le Roux, Mitsubishi Electric Research Laboratories, United States; Vikramjit Mitra, SRI International, United States; Shinji Watanabe, Mitsubishi Electric Research Laboratories, United States

T2.12: ROBUST ASR USING NEURAL NETWORK BASED SPEECH ENHANCEMENT AND FEATURE SIMULATION 482
Sunit Sivasankaran, Aditya Arie Nugraha, Emmanuel Vincent, Juan A. Morales-Cordovilla, Siddharth Dalmia, Irina Illina, Antoine Liutkus, INRIA, France

T2.13: EXPLOITING SYNCHRONY SPECTRA AND DEEP NEURAL NETWORKS FOR NOISE-ROBUST AUTOMATIC SPEECH RECOGNITION 490
Ning Ma, Ricard Marxer, Jon Barker, Guy J. Brown, University of Sheffield, United Kingdom

T2.14: COMBINING SPECTRAL FEATURE MAPPING AND MULTI-CHANNEL MODEL-BASED SOURCE SEPARATION FOR NOISE-ROBUST AUTOMATIC SPEECH RECOGNITION 496
Deblin Bagchi, Michael Mandel, Zhongqiu Wang, Yanzhang He, Andrew Plummer, Eric Fosler-Lussier, The Ohio State University, United States

T2.15: THE THIRD `CHIME' SPEECH SEPARATION AND RECOGNITION CHALLENGE: DATASET, TASK AND BASELINES 504
Jon Barker, Ricard Marxer, University of Sheffield, United Kingdom; Emmanuel Vincent, INRIA, France; Shinji Watanabe, Mitsubishi Electric Research Laboratories, United States

T3: AUTOMATIC SPEECH RECOGNITION III

T3.1: DEEP BOTTLENECK FEATURES FOR I-VECTOR BASED TEXT-INDEPENDENT SPEAKER VERIFICATION 555
Sina Hamidi Ghalehjegh, Richard C. Rose, McGill University, Canada

T3.2: DISCRIMINATIVE SEGMENTAL CASCADES FOR FEATURE-RICH PHONE RECOGNITION 561
Hao Tang, Weiran Wang, Kevin Gimpel, Karen Livescu, Toyota Technological Institute at Chicago, United States

T3.3: HILBERT SPECTRAL ANALYSIS OF VOWELS USING INTRINSIC MODE FUNCTIONS 569
Steven Sandoval, Arizona State University, United States; Phillip L. De Leon, New Mexico State University, United States; Julie Liss, Arizona State University, United States

T3.4: MULTI-REFERENCE WER FOR EVALUATING ASR FOR LANGUAGES WITH NO ORTHOGRAPHIC RULES 576
Ahmed Ali, Walid Magdy, Qatar Computing Research Institute, Qatar; Steve Renals, Peter Bell, University of Edinburgh, United Kingdom

T3.5: ACOUSTIC MODELING WITH NEURAL GRAPH EMBEDDINGS 581
Yuzong Liu, Katrin Kirchhoff, University of Washington, United States

T3.6: MULTITASK LEARNING AND SYSTEM COMBINATION FOR AUTOMATIC SPEECH RECOGNITION 589
Olivier Siohan, David Rybach, Google Inc, United States

T3.7: SPEAKER ADAPTIVE JOINT TRAINING OF GAUSSIAN MIXTURE MODELS AND BOTTLENECK FEATURES 596
Zoltán Tüske, Pavel Golik, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

T3.8: ACOUSTIC MODELLING WITH CD-CTC-SMBR LSTM RNNS 604
Andrew Senior, Hasim Sak, Felix de Chaumont Quitry, Tara Sainath, Kanishka Rao, Google Inc, United States

T3.9: AUTOMATION OF SYSTEM BUILDING FOR STATE-OF-THE-ART LARGE VOCABULARY SPEECH RECOGNITION USING EVOLUTION STRATEGY 610
Takafumi Moriya, Tomohiro Tanaka, Takahiro Shinozaki, Tokyo Institute of Technology, Japan; Shinji Watanabe, Mitsubishi Electric Research Laboratories, United States; Kevin Duh, Nara Institute of Science and Technology, Japan

T3.10: IMPROVING THE INTERPRETABILITY OF DEEP NEURAL NETWORKS WITH STIMULATED LEARNING 617
Shawn Tan, Khe Chai Sim, National University of Singapore, Singapore; Mark J. F. Gales, University of Cambridge, United Kingdom

T4: THE MGB CHALLENGE - RECOGNITION OF MULTI-GENRE BROADCAST DATA

T4.1: THE 2015 SHEFFIELD SYSTEM FOR TRANSCRIPTION OF MULTI-GENRE BROADCAST MEDIA 624
Oscar Saz, Mortaza Doulaty, Salil Deena, Rosanna Milner, Raymond W. M. Ng, Madina Hasan, Yulan Liu, Thomas Hain, University of Sheffield, United Kingdom

T4.2: THE 2015 SHEFFIELD SYSTEM FOR LONGITUDINAL DIARISATION OF BROADCAST MEDIA 632
Rosanna Milner, Oscar Saz, Salil Deena, Mortaza Doulaty, Raymond W. M. Ng, Thomas Hain, University of Sheffield, United Kingdom

T4.3: CAMBRIDGE UNIVERSITY TRANSCRIPTION SYSTEMS FOR THE MULTI-GENRE BROADCAST CHALLENGE 639
Philip C. Woodland, Xunying Liu, Yanmin Qian, Chao Zhang, Mark J. F. Gales, Penny Karanasou, Pierre Lanchantin, Linlin Wang, University of Cambridge, United Kingdom

T4.4: THE DEVELOPMENT OF THE CAMBRIDGE UNIVERSITY ALIGNMENT SYSTEMS FOR THE MULTI-GENRE BROADCAST CHALLENGE 647
Pierre Lanchantin, Mark J. F. Gales, Penny Karanasou, Xunying Liu, Yanmin Qian, Linlin Wang, Philip C. Woodland, Chao Zhang, University of Cambridge, United Kingdom

T4.5: THE NAIST ASR SYSTEM FOR THE 2015 MULTI-GENRE BROADCAST CHALLENGE: ON COMBINATION OF DEEP LEARNING SYSTEMS USING A RANK-SCORE FUNCTION 654
Quoc Truong Do, Michael Heck, Sakriani Sakti, Graham Neubig, Tomoki Toda, Satoshi Nakamura, Nara Institute of Science and Technology, Japan

T4.6: SPEAKER DIARISATION AND LONGITUDINAL LINKING IN MULTI-GENRE BROADCAST DATA 660
Penny Karanasou, Mark J. F. Gales, Pierre Lanchantin, Xunying Liu, Yanmin Qian, Linlin Wang, Philip C. Woodland, Chao Zhang, University of Cambridge, United Kingdom

T4.7: VARIATIONAL BAYESIAN PLDA FOR SPEAKER DIARIZATION IN THE MGB CHALLENGE 667
Jesus Villalba, Alfonso Ortega, Antonio Miguel, Eduardo Lleida, Universidad de Zaragoza, Spain

T4.8: A SYSTEM FOR AUTOMATIC ALIGNMENT OF BROADCAST MEDIA CAPTIONS USING WEIGHTED FINITE-STATE TRANSDUCERS 675
Peter Bell, Steve Renals, University of Edinburgh, United Kingdom

T4.9: CRIM AND LIUM APPROACHES FOR MULTI-GENRE BROADCAST MEDIA TRANSCRIPTION 681
Vishwa Gupta, Centre de Recherche Informatique de Montreal (CRIM), Canada; Paul Deleglise, LIUM - University of Le Mans, France; Gilles Boulianne, Centre de Recherche Informatique de Montreal (CRIM), Canada; Yannick Esteve, Sylvain Meignier, Anthony Rousseau, LIUM - University of Le Mans, France

T4.10: THE MGB CHALLENGE: EVALUATING MULTI-GENRE BROADCAST MEDIA RECOGNITION 687
Peter Bell, University of Edinburgh, United Kingdom; Mark J. F. Gales, University of Cambridge, United Kingdom; Thomas Hain, University of Sheffield, United Kingdom; Jonathan Kilgour, University of Edinburgh, United Kingdom; Pierre Lanchantin, Xunying Liu, University of Cambridge, United Kingdom; Andrew McParland, BBC, United Kingdom; Steve Renals, University of Edinburgh, United Kingdom; Oscar Saz, University of Sheffield, United Kingdom; Mirjam Wester, University of Edinburgh, United Kingdom; Philip C. Woodland, University of Cambridge, United Kingdom

W2: SPOKEN DIALOG SYSTEMS

W2.1: INCREMENTAL LSTM-BASED DIALOG STATE TRACKER 757
Lukas Zilka, Filip Jurcicek, Charles University in Prague, Czech Republic

W2.2: MULTI-DOMAIN DIALOGUE SUCCESS CLASSIFIERS FOR POLICY TRAINING.....763
David Vandyke, Pei-Hao Su, Milica Gasic, Nikola Mrksic, Tsung-Hsien Wen, Steve Young, University of Cambridge, United Kingdom

W2.3: OPEN-DOMAIN PERSONALIZED DIALOG SYSTEM USING 771
USER-INTERESTED TOPICS IN SYSTEM RESPONSES
Jeesoo Bang, Sangdo Han, Kyusong Lee, Gary Geunbae Lee, Pohang University of Science and Technology, Republic of Korea

W2.4: A STUDY OF SOCIAL-AFFECTIVE COMMUNICATION: AUTOMATIC 777
PREDICTION OF EMOTION TRIGGERS AND RESPONSES IN TELEVISION TALK SHOWS
Nurul Lubis, Sakriani Sakti, Graham Neubig, Koichiro Yoshino, Tomoki Toda, Satoshi Nakamura, Nara Institute of Science and Technology, Japan

W2.5: ADAPTIVE SELECTION FROM MULTIPLE RESPONSE CANDIDATES IN 784
EXAMPLE-BASED DIALOGUE
Masahiro Mizukami, Graduate School of Information Science, Nara Institute of Science and Technology, Japan; Hideaki Kizuki, Toshio Nomura, SHARP Corporation, Japan; Graham Neubig, Koichiro Yoshino, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura, Graduate School of Information Science, Nara Institute of Science and Technology, Japan

W2.6: OPTIMIZING HUMAN-INTERPRETABLE DIALOG MANAGEMENT POLICY 791
USING GENETIC ALGORITHM
Hang Ren, Weiqun Xu, Yonghong Yan, Institute of Acoustics, Chinese Academy of Sciences, China

W2.7: IMPLEMENTATION OF GENERIC POSITIVE-NEGATIVE TRACKER IN 798
EXTENSIBLE DIALOG SYSTEM
Sangjun Koo, Seonghan Ryu, Gary Geunbae Lee, Pohang University of Science and Technology, Republic of Korea

W2.8: POLICY COMMITTEE FOR ADAPTATION IN MULTI-DOMAIN SPOKEN 806
DIALOGUE SYSTEMS
Milica Gasic, Nikola Mrksic, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen, Steve Young, University of Cambridge, United Kingdom

W2.9: APPLYING DEEP LEARNING TO ANSWER SELECTION: A STUDY AND AN 813
OPEN TASK
Minwei Feng, Bing Xiang, Michael Glass, Lidan Wang, Bowen Zhou, IBM T.J. Watson, United States