

21st Annual Conference of the International Speech Communication Association (INTERSPEECH 2020)

Cognitive Intelligence for Speech
Processing

Held online due to COVID-19

Shanghai, China
25 – 29 October 2020

Volume 1 of 7

ISBN: 978-1-7138-2069-7

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© (2020) by International Speech Communication Association
All rights reserved.

Printed with permission by Curran Associates, Inc. (2020)

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

International Speech Communication Association
c/o Mme Emmanuelle FOXONET
4 Rue des Fauvettes - Lous Tourils
F-66390 Baixas, France

Phone: 49 228 735 643
Fax: 33 468 385 827

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: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

VOLUME 1

INTERSPEECH 2020 : ASR NEURAL NETWORK ARCHITECTURES I

ON THE COMPARISON OF POPULAR END-TO-END MODELS FOR LARGE SCALE SPEECH RECOGNITION.....	1
<i>Jinyu Li, Yu Wu, Yashesh Gaur, Chengyi Wang, Rui Zhao, Shujie Liu</i>	
SAN-M: MEMORY EQUIPPED SELF-ATTENTION FOR END-TO-END SPEECH RECOGNITION.....	6
<i>Zhifu Gao, Shiliang Zhang, Ming Lei, Ian McLoughlin</i>	
CONTEXTUAL RNN-T FOR OPEN DOMAIN ASR.....	11
<i>Mahaveer Jain, Gil Keren, Jay Mahadeokar, Geoffrey Zweig, Florian Metz, Yatharth Saraf</i>	
ASAPP-ASR: MULTISTREAM CNN AND SELF-ATTENTIVE SRU FOR SOTA SPEECH RECOGNITION.....	16
<i>Jing Pan, Joshua Shapiro, Jeremy Wohlwend, Kyu J. Han, Tao Lei, Tao Ma</i>	
COMPRESSING LSTM NETWORKS WITH HIERARCHICAL COARSE-GRAIN SPARSITY.....	21
<i>Deepak Kadetotad, Jian Meng, Visar Berisha, Chaitali Chakrabarti, Jae-Sun Seo</i>	
BLSTM-DRIVEN STREAM FUSION FOR AUTOMATIC SPEECH RECOGNITION: NOVEL METHODS AND A MULTI-SIZE WINDOW FUSION EXAMPLE.....	26
<i>Timo Lohrenz, Tim Fingscheidt</i>	
RELATIVE POSITIONAL ENCODING FOR SPEECH RECOGNITION AND DIRECT TRANSLATION.....	31
<i>Ngoc-Quan Pham, Thanh-Le Ha, Tuan-Nam Nguyen, Thai-Son Nguyen, Elizabeth Salesky, Sebastian Stuker, Jan Niehues, Alex Waibel</i>	
JOINT SPEAKER COUNTING, SPEECH RECOGNITION, AND SPEAKER IDENTIFICATION FOR OVERLAPPED SPEECH OF ANY NUMBER OF SPEAKERS.....	36
<i>Naoyuki Kanda, Yashesh Gaur, Xiaofei Wang, Zhong Meng, Zhuo Chen, Tianyan Zhou, Takuya Yoshioka</i>	
IMPLICIT TRANSFER OF PRIVILEGED ACOUSTIC INFORMATION IN A GENERALIZED KNOWLEDGE DISTILLATION FRAMEWORK.....	41
<i>Takashi Fukuda, Samuel Thomas</i>	
EFFECT OF ADDING POSITIONAL INFORMATION ON CONVOLUTIONAL NEURAL NETWORKS FOR END-TO-END SPEECH RECOGNITION.....	46
<i>Jinhwan Park, Wonyong Sung</i>	

INTERSPEECH 2020 : MULTI-CHANNEL SPEECH ENHANCEMENT

DEEP NEURAL NETWORK-BASED GENERALIZED SIDELobe CANCELLER FOR ROBUST MULTI-CHANNEL SPEECH RECOGNITION.....	51
<i>Guanjun Li, Shan Liang, Shuai Nie, Wenju Liu, Zhanlei Yang, Longshuai Xiao</i>	

NEURAL SPATIO-TEMPORAL BEAMFORMER FOR TARGET SPEECH SEPARATION.....	56
<i>Yong Xu, Meng Yu, Shi-Xiong Zhang, Lianwu Chen, Chao Weng, Jianming Liu, Dong Yu</i>	
ONLINE DIRECTIONAL SPEECH ENHANCEMENT USING GEOMETRICALLY CONSTRAINED INDEPENDENT VECTOR ANALYSIS.....	61
<i>Li Li, Kazuhito Koishida, Shoji Makino</i>	
END-TO-END MULTI-LOOK KEYWORD SPOTTING	66
<i>Meng Yu, Xuan Ji, Bo Wu, Dan Su, Dong Yu</i>	
DIFFERENTIAL BEAMFORMING FOR UNIFORM CIRCULAR ARRAY WITH DIRECTIONAL MICROPHONES.....	71
<i>Weilong Huang, Jinwei Feng</i>	
EXPLORING DEEP HYBRID TENSOR-TO-VECTOR NETWORK ARCHITECTURES FOR REGRESSION BASED SPEECH ENHANCEMENT	76
<i>Jun Qi, Hu Hu, Yannan Wang, Chao-Han Huck Yang, Sabato Marco Siniscalchi, Chin-Hui Lee</i>	
AN END-TO-END ARCHITECTURE OF ONLINE MULTI-CHANNEL SPEECH SEPARATION	81
<i>Jian Wu, Zhuo Chen, Jinyu Li, Takuya Yoshioka, Zhili Tan, Edward Lin, Yi Luo, Lei Xie</i>	
MENTORING-REVERSE MENTORING FOR UNSUPERVISED MULTI-CHANNEL SPEECH SOURCE SEPARATION	86
<i>Yu Nakagome, Masahito Togami, Tetsuji Ogawa, Tetsunori Kobayashi</i>	
COMPUTATIONALLY EFFICIENT AND VERSATILE FRAMEWORK FOR JOINT OPTIMIZATION OF BLIND SPEECH SEPARATION AND DEREVERBERATION.....	91
<i>Tomohiro Nakatani, Rintaro Ikeshita, Keisuke Kinoshita, Hiroshi Sawada, Shoko Araki</i>	
A SPACE-AND-SPEAKER-AWARE ITERATIVE MASK ESTIMATION APPROACH TO MULTI-CHANNEL SPEECH RECOGNITION IN THE CHIME-6 CHALLENGE	96
<i>Yan-Hui Tu, Jun Du, Lei Sun, Feng Ma, Jia Pan, Chin-Hui Lee</i>	
<u>INTERSPEECH 2020 : SPEECH PROCESSING IN THE BRAIN</u>	
IDENTIFYING CAUSAL RELATIONSHIPS BETWEEN BEHAVIOR AND LOCAL BRAIN ACTIVITY DURING NATURAL CONVERSATION.....	101
<i>Hmamouche Youssef, Prevot Laurent, Ochs Magalie, Chaminade Thierry</i>	
NEURAL ENTRAINMENT TO NATURAL SPEECH ENVELOPE BASED ON SUBJECT ALIGNED EEG SIGNALS.....	106
<i>Di Zhou, Gaoyan Zhang, Jianwu Dang, Shuang Wu, Zhuo Zhang</i>	
DOES LEXICAL RETRIEVAL DETERIORATE IN PATIENTS WITH MILD COGNITIVE IMPAIRMENT? ANALYSIS OF BRAIN FUNCTIONAL NETWORK WILL TELL	111
<i>Chongyuan Lian, Tianqi Wang, Mingxiao Gu, Manwa L. Ng, Feiqi Zhu, Lan Wang, Nan Yan</i>	
CONGRUENT AUDIOVISUAL SPEECH ENHANCES CORTICAL ENVELOPE TRACKING DURING AUDITORY SELECTIVE ATTENTION	116
<i>Zhen Fu, Jing Chen</i>	
CONTRIBUTION OF RMS-LEVEL-BASED SPEECH SEGMENTS TO TARGET SPEECH DECODING UNDER NOISY CONDITIONS	121
<i>Lei Wang, Ed X. Wu, Fei Chen</i>	

CORTICAL OSCILLATORY HIERARCHY FOR NATURAL SENTENCE PROCESSING 125
Bin Zhao, Jianwu Dang, Gaoyan Zhang, Masashi Unoki

COMPARING EEG ANALYSES WITH DIFFERENT EPOCH ALIGNMENTS IN AN
AUDITORY LEXICAL DECISION EXPERIMENT 130
Louis Ten Bosch, Kimberley Mulder, Lou Boves

DETECTION OF SUBCLINICAL MILD TRAUMATIC BRAIN INJURY (MTBI) THROUGH
SPEECH AND GAIT 135
*Tanya Talkar, Sophia Yuditskaya, James R. Williamson, Adam C. Lammert, Hrishikesh Rao,
Daniel Hannon, Anne O'Brien, Gloria Vergara-Diaz, Richard Delaura, Douglas Sturim,
Gregory Ciccarelli, Ross Zafonte, Jeffrey Palmer, Paolo Bonato, Thomas F. Quatieri*

INTERSPEECH 2020 : SPEECH SIGNAL REPRESENTATION

TOWARDS LEARNING A UNIVERSAL NON-SEMANTIC REPRESENTATION OF SPEECH 140
*Joel Shor, Aren Jansen, Ronnie Maor, Oran Lang, Omry Tuval, Felix De Chaumont Quitry,
Marco Tagliasacchi, Ira Shavitt, Dotan Emanuel, Yinnon Haviv*

POETIC METER CLASSIFICATION USING I-VECTOR-MTF FUSION 145
Rajeev Rajan, Aiswarya Vinod Kumar, Ben P. Babu

FORMANT TRACKING USING DILATED CONVOLUTIONAL NETWORKS THROUGH
DENSE CONNECTION WITH GATING MECHANISM 150
Wang Dai, Jinsong Zhang, Yingming Gao, Wei Wei, Dengfeng Ke, Binghui Lin, Yanlu Xie

AUTOMATIC ANALYSIS OF SPEECH PROSODY IN DUTCH 155
Na Hu, Berit Janssen, Judith Hanssen, Carlos Gussenhoven, Aoju Chen

LEARNING VOICE REPRESENTATION USING KNOWLEDGE DISTILLATION FOR
AUTOMATIC VOICE CASTING 160
Adrien Gresse, Mathias Quillot, Richard Dufour, Jean-Francois Bonastre

ENHANCING FORMANT INFORMATION IN SPECTROGRAPHIC DISPLAY OF SPEECH 165
B. Yegnanarayana, Anand Joseph, Vishala Pannala

UNSUPERVISED METHODS FOR EVALUATING SPEECH REPRESENTATIONS 170
Michael Gump, Wei-Ning Hsu, James Glass

ROBUST PITCH REGRESSION WITH VOICED/UNVOICED CLASSIFICATION IN
NONSTATIONARY NOISE ENVIRONMENTS 175
Dung N. Tran, Uros Batricevic, Kazuhito Koishida

NONLINEAR ISA WITH AUXILIARY VARIABLES FOR LEARNING SPEECH
REPRESENTATIONS 180
Amrith Setlur, Barnabas Poczós, Alan W. Black

HARMONIC LOWERING FOR ACCELERATING HARMONIC CONVOLUTION FOR
AUDIO SIGNALS 185
Hirotochi Takeuchi, Kunio Kashino, Yasunori Ohishi, Hiroshi Saruwatari

INTERSPEECH 2020 : SPEECH SYNTHESIS: NEURAL WAVEFORM GENERATION I

KNOWLEDGE-AND-DATA-DRIVEN AMPLITUDE SPECTRUM PREDICTION FOR HIERARCHICAL NEURAL VOCODERS.....	190
<i>Yang Ai, Zhen-Hua Ling</i>	
FEATHERWAVE: AN EFFICIENT HIGH-FIDELITY NEURAL VOCODER WITH MULTI-BAND LINEAR PREDICTION	195
<i>Qiao Tian, Zewang Zhang, Heng Lu, Ling-Hui Chen, Shan Liu</i>	
VOCGAN: A HIGH-FIDELITY REAL-TIME VOCODER WITH A HIERARCHICALLY-NESTED ADVERSARIAL NETWORK.....	200
<i>Jinhyeok Yang, Junmo Lee, Youngik Kim, Hoon-Young Cho, Injung Kim</i>	
LIGHTWEIGHT LPCNET-BASED NEURAL VOCODER WITH TENSOR DECOMPOSITION	205
<i>Hiroki Kanagawa, Yusuke Ijima</i>	
WG-WAVENET: REAL-TIME HIGH-FIDELITY SPEECH SYNTHESIS WITHOUT GPU	210
<i>Po-Chun Hsu, Hung-Yi Lee</i>	
WHAT THE FUTURE BRINGS: INVESTIGATING THE IMPACT OF LOOKAHEAD FOR INCREMENTAL NEURAL TTS	215
<i>Brooke Stephenson, Laurent Besacier, Laurent Girin, Thomas Hueber</i>	
FAST AND LIGHTWEIGHT ON-DEVICE TTS WITH TACOTRON2 AND LPCNET.....	220
<i>Vadim Popov, Stanislav Kamenev, Mikhail Kudinov, Sergey Repevsky, Tasnima Sadekova, Vitalii Bushaev, Vladimir Kryzhanovskiy, Denis Parkhomenko</i>	
EFFICIENT WAVEGLOW: AN IMPROVED WAVEGLOW VOCODER WITH ENHANCED SPEED.....	225
<i>Wei Song, Guanghui Xu, Zhengchen Zhang, Chao Zhang, Xiaodong He, Bowen Zhou</i>	
CAN AUDITORY NERVE MODELS TELL US WHAT'S DIFFERENT ABOUT WAVENET VOCODED SPEECH?.....	230
<i>Sebastien Le Maguer, Naomi Harte</i>	
SPEAKER CONDITIONAL WEVERNN: TOWARDS UNIVERSAL NEURAL VOCODER FOR UNSEEN SPEAKER AND RECORDING CONDITIONS.....	235
<i>Dipjyoti Paul, Yannis Pantazis, Yannis Stylianou</i>	
NEURAL HOMOMORPHIC VOCODER	240
<i>Zhijun Liu, Kuan Chen, Kai Yu</i>	

INTERSPEECH 2020 : AUTOMATIC SPEECH RECOGNITION FOR NON-NATIVE CHILDREN'S SPEECH

OVERVIEW OF THE INTERSPEECH TLT2020 SHARED TASK ON ASR FOR NON-NATIVE CHILDREN'S SPEECH	245
<i>Roberto Gretter, Marco Matassoni, Daniele Falavigna, Keelan Evanini, Chee Wee Leong</i>	
THE NTNU SYSTEM AT THE INTERSPEECH 2020 NON-NATIVE CHILDREN'S SPEECH ASR CHALLENGE	250
<i>Tien-Hong Lo, Fu-An Chao, Shi-Yan Weng, Berlin Chen</i>	

NON-NATIVE CHILDREN'S AUTOMATIC SPEECH RECOGNITION: THE INTERSPEECH 2020 SHARED TASK ALTA SYSTEMS	255
<i>Kate M. Knill, Linlin Wang, Yu Wang, Xixin Wu, Mark J. F. Gales</i>	

DATA AUGMENTATION USING PROSODY AND FALSE STARTS TO RECOGNIZE NON- NATIVE CHILDREN'S SPEECH	260
<i>Hemant Kathania, Mittul Singh, Tamas Grosz, Mikko Kurimo</i>	

UNSW SYSTEM DESCRIPTION FOR THE SHARED TASK ON AUTOMATIC SPEECH RECOGNITION FOR NON-NATIVE CHILDREN'S SPEECH.....	265
<i>Mostafa Shahin, Renee Lu, Julien Epps, Beena Ahmed</i>	

INTERSPEECH 2020 : SPEAKER DIARIZATION

END-TO-END SPEAKER DIARIZATION FOR AN UNKNOWN NUMBER OF SPEAKERS WITH ENCODER-DECODER BASED ATTRACTORS.....	269
<i>Shota Horiguchi, Yusuke Fujita, Shinji Watanabe, Yawen Xue, Kenji Nagamatsu</i>	

TARGET-SPEAKER VOICE ACTIVITY DETECTION: A NOVEL APPROACH FOR MULTI- SPEAKER DIARIZATION IN A DINNER PARTY SCENARIO.....	274
<i>Ivan Medennikov, Maxim Korenevsky, Tatiana Prisyach, Yuri Khokhlov, Mariya Korenevskaya, Ivan Sorokin, Tatiana Timofeeva, Anton Mitrofanov, Andrei Andrusenko, Ivan Podluzhny, Aleksandr Laptev, Aleksei Romanenko</i>	

NEW ADVANCES IN SPEAKER DIARIZATION.....	279
<i>Hagai Aronowitz, Weizhong Zhu, Masayuki Suzuki, Gakuto Kurata, Ron Hoory</i>	

SELF-ATTENTIVE SIMILARITY MEASUREMENT STRATEGIES IN SPEAKER DIARIZATION	284
<i>Qingjian Lin, Yu Hou, Ming Li</i>	

SPEAKER ATTRIBUTION WITH VOICE PROFILES BY GRAPH-BASED SEMI- SUPERVISED LEARNING.....	289
<i>Jixuan Wang, Xiong Xiao, Jian Wu, Ranjani Ramamurthy, Frank Rudzicz, Michael Brudno</i>	

DEEP SELF-SUPERVISED HIERARCHICAL CLUSTERING FOR SPEAKER DIARIZATION	294
<i>Prachi Singh, Sriram Ganapathy</i>	

SPOT THE CONVERSATION: SPEAKER DIARISATION IN THE WILD	299
<i>Joon Son Chung, Jaesung Huh, Arsha Nagrani, Triantafyllos Afouras, Andrew Zisserman</i>	

INTERSPEECH 2020 : NOISE ROBUST AND DISTANT SPEECH RECOGNITION

LEARNING CONTEXTUAL LANGUAGE EMBEDDINGS FOR MONAURAL MULTI- TALKER SPEECH RECOGNITION	304
<i>Wangyou Zhang, Yanmin Qian</i>	

DOUBLE ADVERSARIAL NETWORK BASED MONAURAL SPEECH ENHANCEMENT FOR ROBUST SPEECH RECOGNITION.....	309
<i>Zhihao Du, Jiqing Han, Xueliang Zhang</i>	

ANTI-ALIASING REGULARIZATION IN STACKING LAYERS.....	314
<i>Antoine Bruguier, Ananya Misra, Arun Narayanan, Rohit Prabhavalkar</i>	

TOWARDS A COMPETITIVE END-TO-END SPEECH RECOGNITION FOR CHIME-6 DINNER PARTY TRANSCRIPTION.....	319
<i>Andrei Andrusenko, Aleksandr Laptev, Ivan Medennikov</i>	
END-TO-END FAR-FIELD SPEECH RECOGNITION WITH UNIFIED DEREVERBERATION AND BEAMFORMING	324
<i>Wangyou Zhang, Aswin Shanmugam Subramanian, Xuankai Chang, Shinji Watanabe, Yanmin Qian</i>	
QUATERNION NEURAL NETWORKS FOR MULTI-CHANNEL DISTANT SPEECH RECOGNITION.....	329
<i>Xinchi Qiu, Titouan Parcollet, Mirco Ravanelli, Nicholas D. Lane, Mohamed Morchid</i>	
IMPROVED GUIDED SOURCE SEPARATION INTEGRATED WITH A STRONG BACK- END FOR THE CHIME-6 DINNER PARTY SCENARIO	334
<i>Hangting Chen, Pengyuan Zhang, Qian Shi, Zuozhen Liu</i>	
NEURAL SPEECH SEPARATION USING SPATIALLY DISTRIBUTED MICROPHONES	339
<i>Dongmei Wang, Zhuo Chen, Takuya Yoshioka</i>	
UTTERANCE-WISE MEETING TRANSCRIPTION SYSTEM USING ASYNCHRONOUS DISTRIBUTED MICROPHONES	344
<i>Shota Horiguchi, Yusuke Fujita, Kenji Nagamatsu</i>	
SIMULATING REALISTICALLY-SPATIALISED SIMULTANEOUS SPEECH USING VIDEO- DRIVEN SPEAKER DETECTION AND THE CHIME-5 DATASET	349
<i>Jack Deadman, Jon Barker</i>	

INTERSPEECH 2020 : SPEECH IN MULTIMODALITY

TOWARD SILENT PARALINGUISTICS: SPEECH-TO-EMG --- RETRIEVING ARTICULATORY MUSCLE ACTIVITY FROM SPEECH	354
<i>Catarina Botelho, Lorenz Diener, Dennis Kuster, Kevin Scheck, Shahin Amiriparian, Bjorn W. Schuller, Tanja Schultz, Alberto Abad, Isabel Trancoso</i>	
MULTIMODAL DECEPTION DETECTION USING AUTOMATICALLY EXTRACTED ACOUSTIC, VISUAL, AND LEXICAL FEATURES.....	359
<i>Jiaxuan Zhang, Sarah Ita Levitan, Julia Hirschberg</i>	
MULTI-MODAL ATTENTION FOR SPEECH EMOTION RECOGNITION.....	364
<i>Zexu Pan, Zhaojie Luo, Jichen Yang, Haizhou Li</i>	
WISE: WORD-LEVEL INTERACTION-BASED MULTIMODAL FUSION FOR SPEECH EMOTION RECOGNITION	369
<i>Guang Shen, Riwei Lai, Rui Chen, Yu Zhang, Kejia Zhang, Qilong Han, Hongtao Song</i>	
A MULTI-SCALE FUSION FRAMEWORK FOR BIMODAL SPEECH EMOTION RECOGNITION.....	374
<i>Ming Chen, Xudong Zhao</i>	
GROUP GATED FUSION ON ATTENTION-BASED BIDIRECTIONAL ALIGNMENT FOR MULTIMODAL EMOTION RECOGNITION	379
<i>Pengfei Liu, Kun Li, Helen Meng</i>	

MULTI-MODAL EMBEDDINGS USING MULTI-TASK LEARNING FOR EMOTION RECOGNITION.....	384
<i>Aparna Khare, Srinivas Parthasarathy, Shiva Sundaram</i>	

USING SPEAKER-ALIGNED GRAPH MEMORY BLOCK IN MULTIMODALLY ATTENTIVE EMOTION RECOGNITION NETWORK.....	389
<i>Jeng-Lin Li, Chi-Chun Lee</i>	

CONTEXT-DEPENDENT DOMAIN ADVERSARIAL NEURAL NETWORK FOR MULTIMODAL EMOTION RECOGNITION	394
<i>Zheng Lian, Jianhua Tao, Bin Liu, Jian Huang, Zhanlei Yang, Rongjun Li</i>	

INTERSPEECH 2020 : SPEECH, LANGUAGE, AND MULTIMODAL RESOURCES

ATCSPEECH: A MULTILINGUAL PILOT-CONTROLLER SPEECH CORPUS FROM REAL AIR TRAFFIC CONTROL ENVIRONMENT.....	399
<i>Bo Yang, Xianlong Tan, Zhengmao Chen, Bing Wang, Min Ruan, Dan Li, Zhongping Yang, Xiping Wu, Yi Lin</i>	

DEVELOPING AN OPEN-SOURCE CORPUS OF YORUBA SPEECH.....	404
<i>Alexander Gutkin, Isin Demirsahin, Oddur Kjartansson, Clara Rivera, Kola Tubosun</i>	

CLOVACALL: KOREAN GOAL-ORIENTED DIALOG SPEECH CORPUS FOR AUTOMATIC SPEECH RECOGNITION OF CONTACT CENTERS	409
<i>Jung-Woo Ha, Kihyun Nam, Jingu Kang, Sang-Woo Lee, Sohee Yang, Hyunhoon Jung, Hyeji Kim, Eunmi Kim, Soojin Kim, Hyun Ah Kim, Kyoungtae Doh, Chan Kyu Lee, Nako Sung, Sunghun Kim</i>	

LAIX CORPUS OF CHINESE LEARNER ENGLISH: TOWARDS A BENCHMARK FOR L2 ENGLISH ASR	414
<i>Yanhong Wang, Huan Luan, Jiahong Yuan, Bin Wang, Hui Lin</i>	

DESIGN AND DEVELOPMENT OF A HUMAN-MACHINE DIALOG CORPUS FOR THE AUTOMATED ASSESSMENT OF CONVERSATIONAL ENGLISH PROFICIENCY	419
<i>Vikram Ramanarayanan</i>	

CUCHILD: A LARGE-SCALE CANTONESE CORPUS OF CHILD SPEECH FOR PHONOLOGY AND ARTICULATION ASSESSMENT.....	424
<i>Si-Ioi Ng, Cymie Wing-Yee Ng, Jiarui Wang, Tan Lee, Kathy Yuet-Sheung Lee, Michael Chi-Fai Tong</i>	

FINCHAT: CORPUS AND EVALUATION SETUP FOR FINNISH CHAT CONVERSATIONS ON EVERYDAY TOPICS.....	429
<i>Katri Leino, Juho Leinonen, Mittul Singh, Sami Virpioja, Mikko Kurimo</i>	

DIPCO --- DINNER PARTY CORPUS.....	434
<i>Maarten Van Segbroeck, Ahmed Zaid, Ksenia Kutsenko, Cirenia Huerta, Tinh Nguyen, Xuewen Luo, Bjorn Hoffmeister, Jan Trmal, Maurizio Omologo, Roland Maas</i>	

LEARNING TO DETECT BIPOLAR DISORDER AND BORDERLINE PERSONALITY DISORDER WITH LANGUAGE AND SPEECH IN NON-CLINICAL INTERVIEWS	437
<i>Bo Wang, Yue Wu, Niall Taylor, Terry Lyons, Maria Liakata, Alejo J. Nevado-Holgado, Kate E. A. Saunders</i>	

FT SPEECH: DANISH PARLIAMENT SPEECH CORPUS.....	442
<i>Andreas Kirkedal, Marija Stepanovic, Barbara Plank</i>	

INTERSPEECH 2020 : LANGUAGE RECOGNITION

METRIC LEARNING LOSS FUNCTIONS TO REDUCE DOMAIN MISMATCH IN THE X-VECTOR SPACE FOR LANGUAGE RECOGNITION.....	447
<i>Raphael Duroselle, Denis Jouviet, Irina Illina</i>	
THE XMUSPEECH SYSTEM FOR THE AP19-OLR CHALLENGE.....	452
<i>Zheng Li, Miao Zhao, Jing Li, Yiming Zhi, Lin Li, Qingyang Hong</i>	
ON THE USAGE OF MULTI-FEATURE INTEGRATION FOR SPEAKER VERIFICATION AND LANGUAGE IDENTIFICATION	457
<i>Zheng Li, Miao Zhao, Jing Li, Lin Li, Qingyang Hong</i>	
WHAT DOES AN END-TO-END DIALECT IDENTIFICATION MODEL LEARN ABOUT NON-DIALECTAL INFORMATION?	462
<i>Shammur A. Chowdhury, Ahmed Ali, Suwon Shon, James Glass</i>	
RELEASING A TOOLKIT AND COMPARING THE PERFORMANCE OF LANGUAGE EMBEDDINGS ACROSS VARIOUS SPOKEN LANGUAGE IDENTIFICATION DATASETS	467
<i>Matias Lindgren, Tommi Jauhiainen, Mikko Kurimo</i>	
LEARNING INTONATION PATTERN EMBEDDINGS FOR ARABIC DIALECT IDENTIFICATION	472
<i>Aitor Arronte Alvarez, Elsayed Sabry Abdelaal Issa</i>	
CROSS-DOMAIN ADAPTATION OF SPOKEN LANGUAGE IDENTIFICATION FOR RELATED LANGUAGES: THE CURIOUS CASE OF SLAVIC LANGUAGES	477
<i>Badr M. Abdullah, Tania Avgustinova, Bernd Mobius, Dietrich Klakow</i>	

INTERSPEECH 2020 : SPEECH PROCESSING AND ANALYSIS

ICE-TALK: AN INTERFACE FOR A CONTROLLABLE EXPRESSIVE TALKING MACHINE	482
<i>Noe Tits, Kevin El Haddad, Thierry Dutoit</i>	
KALDI-WEB: AN INSTALLATION-FREE, ON-DEVICE SPEECH RECOGNITION SYSTEM.....	484
<i>Mathieu Hu, Laurent Pierron, Emmanuel Vincent, Denis Jouviet</i>	
SOAPBOX LABS VERIFICATION PLATFORM FOR CHILD SPEECH.....	486
<i>Amelia C. Kelly, Eleni Karamichali, Armin Saeb, Karel Vesely, Nicholas Parslow, Agape Deng, Arnaud Letondor, Robert O'Regan, Qiru Zhou</i>	
SOAPBOX LABS FLUENCY ASSESSMENT PLATFORM FOR CHILD SPEECH.....	488
<i>Amelia C. Kelly, Eleni Karamichali, Armin Saeb, Karel Vesely, Nicholas Parslow, Gloria Montoya Gomez, Agape Deng, Arnaud Letondor, Niall Mullally, Adrian Hempel, Robert O'Regan, Qiru Zhou</i>	
CATOTRON --- A NEURAL TEXT-TO-SPEECH SYSTEM IN CATALAN.....	490
<i>Baybars Kulebi, Alp Oktem, Alex Peiro-Lilja, Santiago Pascual, Mireia Farrus</i>	
TOWARD REMOTE PATIENT MONITORING OF SPEECH, VIDEO, COGNITIVE AND RESPIRATORY BIOMARKERS USING MULTIMODAL DIALOG TECHNOLOGY	492
<i>Vikram Ramanarayanan, Oliver Roesler, Michael Neumann, David Pautler, Doug Habberstad, Andrew Cornish, Hardik Kothare, Vignesh Murali, Jackson Liscombe, Dirk Schnelle-Walka, Patrick Lange, David Suendermann-Oeft</i>	

VOICEID ON THE FLY: A SPEAKER RECOGNITION SYSTEM THAT LEARNS FROM SCRATCH	494
<i>Baihan Lin, Xinxin Zhang</i>	

INTERSPEECH 2020 : SPEECH EMOTION RECOGNITION I

ENHANCING TRANSFERABILITY OF BLACK-BOX ADVERSARIAL ATTACKS VIA LIFELONG LEARNING FOR SPEECH EMOTION RECOGNITION MODELS	496
<i>Zhao Ren, Jing Han, Nicholas Cummins, Bjorn W. Schuller</i>	

END-TO-END SPEECH EMOTION RECOGNITION COMBINED WITH ACOUSTIC-TO-WORD ASR MODEL	501
<i>Han Feng, Sei Ueno, Tatsuya Kawahara</i>	

IMPROVING SPEECH EMOTION RECOGNITION USING GRAPH ATTENTIVE BI-DIRECTIONAL GATED RECURRENT UNIT NETWORK	506
<i>Bo-Hao Su, Chun-Min Chang, Yun-Shao Lin, Chi-Chun Lee</i>	

AN INVESTIGATION OF CROSS-CULTURAL SEMI-SUPERVISED LEARNING FOR CONTINUOUS AFFECT RECOGNITION	511
<i>Adria Mallo-Ragolta, Nicholas Cummins, Bjorn W. Schuller</i>	

ENSEMBLE OF STUDENTS TAUGHT BY PROBABILISTIC TEACHERS TO IMPROVE SPEECH EMOTION RECOGNITION	516
<i>Kusha Sridhar, Carlos Busso</i>	

AUGMENTING GENERATIVE ADVERSARIAL NETWORKS FOR SPEECH EMOTION RECOGNITION	521
<i>Siddique Latif, Muhammad Asim, Rajib Rana, Sara Khalifa, Raja Jurdak, Bjorn W. Schuller</i>	

SPEECH EMOTION RECOGNITION 'IN THE WILD' USING AN AUTOENCODER	526
<i>Vipula Dissanayake, Haimo Zhang, Mark Billingham, Suranga Nanayakkara</i>	

EMOTION PROFILE REFINERY FOR SPEECH EMOTION CLASSIFICATION	531
<i>Shuiyang Mao, P. C. Ching, Tan Lee</i>	

SPEECH REPRESENTATION LEARNING FOR EMOTION RECOGNITION USING END-TO-END ASR WITH FACTORIZED ADAPTATION	536
<i>Sung-Lin Yeh, Yun-Shao Lin, Chi-Chun Lee</i>	

INTERSPEECH 2020 : ASR NEURAL NETWORK ARCHITECTURES AND TRAINING I

FAST AND SLOW ACOUSTIC MODEL	541
<i>Kshitiz Kumar, Emilian Stoimenov, Hosam Khalil, Jian Wu</i>	

SELF-DISTILLATION FOR IMPROVING CTC-TRANSFORMER-BASED ASR SYSTEMS	546
<i>Takafumi Moriya, Tsubasa Ochiai, Shigeki Karita, Hiroshi Sato, Tomohiro Tanaka, Takanori Ashihara, Ryo Masumura, Yusuke Shinohara, Marc Delcroix</i>	

SINGLE HEADED ATTENTION BASED SEQUENCE-TO-SEQUENCE MODEL FOR STATE-OF-THE-ART RESULTS ON SWITCHBOARD	551
<i>Zoltan Tuske, George Saon, Kartik Audhkhasi, Brian Kingsbury</i>	

IMPROVING SPEECH RECOGNITION USING GAN-BASED SPEECH SYNTHESIS AND CONTRASTIVE UNSPOKEN TEXT SELECTION	556
<i>Zhehuai Chen, Andrew Rosenberg, Yu Zhang, Gary Wang, Bhuvana Ramabhadran, Pedro J. Moreno</i>	
PYCHAIN: A FULLY PARALLELIZED PYTORCH IMPLEMENTATION OF LF-MMI FOR END-TO-END ASR.....	561
<i>Yiwen Shao, Yiming Wang, Daniel Povey, Sanjeev Khudanpur</i>	
CAT: A CTC-CRF BASED ASR TOOLKIT BRIDGING THE HYBRID AND THE END-TO-END APPROACHES TOWARDS DATA EFFICIENCY AND LOW LATENCY	566
<i>Keyu An, Hongyu Xiang, Zhijian Ou</i>	
CTC-SYNCHRONOUS TRAINING FOR MONOTONIC ATTENTION MODEL	571
<i>Hirofumi Inaguma, Masato Mimura, Tatsuya Kawahara</i>	
CONTINUAL LEARNING FOR MULTI-DIALECT ACOUSTIC MODELS.....	576
<i>Brady Houston, Katrin Kirchhoff</i>	
SPECSWAP: A SIMPLE DATA AUGMENTATION METHOD FOR END-TO-END SPEECH RECOGNITION.....	581
<i>Xingchen Song, Zhiyong Wu, Yiheng Huang, Dan Su, Helen Meng</i>	
<u>INTERSPEECH 2020 : EVALUATION OF SPEECH TECHNOLOGY SYSTEMS AND METHODS FOR RESOURCE CONSTRUCTION AND ANNOTATION</u>	
RECOAPY: DATA RECORDING, PRE-PROCESSING AND PHONETIC TRANSCRIPTION FOR END-TO-END SPEECH-BASED APPLICATIONS.....	586
<i>Adriana Stan</i>	
ANALYZING THE QUALITY AND STABILITY OF A STREAMING END-TO-END ON-DEVICE SPEECH RECOGNIZER	591
<i>Yuan Shangquan, Kate Knister, Yanzhang He, Ian McGraw, Francoise Beaufays</i>	
STATISTICAL TESTING ON ASR PERFORMANCE VIA BLOCKWISE BOOTSTRAP	596
<i>Zhe Liu, Fuchun Peng</i>	
SENTENCE LEVEL ESTIMATION OF PSYCHOLINGUISTIC NORMS USING JOINT MULTIDIMENSIONAL ANNOTATIONS	601
<i>Anil Ramakrishna, Shrikanth Narayanan</i>	
NEURAL ZERO-INFLATED QUALITY ESTIMATION MODEL FOR AUTOMATIC SPEECH RECOGNITION SYSTEM	606
<i>Kai Fan, Bo Li, Jiayi Wang, Shiliang Zhang, Boxing Chen, Niyu Ge, Zhijie Yan</i>	
CONFIDENCE MEASURES IN ENCODER-DECODER MODELS FOR SPEECH RECOGNITION.....	611
<i>Alejandro Woodward, Clara Bonnin, Issey Masuda, David Varas, Elisenda Bou-Balust, Juan Carlos Riveiro</i>	
WORD ERROR RATE ESTIMATION WITHOUT ASR OUTPUT: E-WER2	616
<i>Ahmed Ali, Steve Renals</i>	
AN EVALUATION OF MANUAL AND SEMI-AUTOMATIC LAUGHTER ANNOTATION	621
<i>Bogdan Ludusan, Petra Wagner</i>	

UNDERSTANDING RACIAL DISPARITIES IN AUTOMATIC SPEECH RECOGNITION: THE CASE OF HABITUAL "BE"	626
<i>Joshua L. Martin, Kevin Tang</i>	

INTERSPEECH 2020 : PHONETICS AND PHONOLOGY

SECONDARY PHONETIC CUES IN THE PRODUCTION OF THE NASAL SHORT-A SYSTEM IN CALIFORNIA ENGLISH	631
<i>Georgia Zellou, Rebecca Scarborough, Renee Kemp</i>	
ACOUSTIC PROPERTIES OF STRIDENT FRICATIVES AT THE EDGES: IMPLICATIONS FOR CONSONANT DISCRIMINATION.....	636
<i>Louis-Marie Lorin, Lorenzo Maselli, Leo Varnet, Maria Giavazzi</i>	
PROCESSES AND CONSEQUENCES OF CO-ARTICULATION IN MANDARIN V_1N.(C_2)V_2 CONTEXT: PHONOLOGY AND PHONETICS	641
<i>Mingqiong Luo</i>	
VOICING DISTINCTION OF OBSTRUENTS IN THE HANGZHOU WU CHINESE DIALECT.....	646
<i>Yang Yue, Fang Hu</i>	
THE PHONOLOGY AND PHONETICS OF KAIFENG MANDARIN VOWELS	651
<i>Lei Wang</i>	
MICROPROSODIC VARIABILITY IN PLOSIVES IN GERMAN AND AUSTRIAN GERMAN.....	656
<i>Margaret Zellers, Barbara Schuppler</i>	
ER-SUFFIXATION IN SOUTHWESTERN MANDARIN: AN EMA AND ULTRASOUND STUDY.....	661
<i>Jing Huang, Feng-Fan Hsieh, Yueh-Chin Chang</i>	
ELECTROGLOTTOGRAPHIC-PHONETIC STUDY ON KOREAN PHONATION INDUCED BY TRIPARTITE PLOSIVES IN YANBIAN KOREAN	666
<i>Yinghao Li, Jinghua Zhang</i>	
MODELING GLOBAL BODY CONFIGURATIONS IN AMERICAN SIGN LANGUAGE.....	671
<i>Nicholas Wilkins, Max Cordes Galbraith, Ifeoma Nwogu</i>	

INTERSPEECH 2020 : TOPICS IN ASR I

AUGMENTING TURN-TAKING PREDICTION WITH WEARABLE EYE ACTIVITY DURING CONVERSATION.....	676
<i>Hang Li, Siyuan Chen, Julien Epps</i>	
CAM: UNINTERESTING SPEECH DETECTOR.....	681
<i>Weiyi Lu, Yi Xu, Peng Yang, Belinda Zeng</i>	
MIXED CASE CONTEXTUAL ASR USING CAPITALIZATION MASKS.....	686
<i>Diamantino Caseiro, Pat Rondon, Quoc-Nam Le The, Petar Aleksic</i>	
SPEECH RECOGNITION AND MULTI-SPEAKER DIARIZATION OF LONG CONVERSATIONS.....	691
<i>Huanru Henry Mao, Shuyang Li, Julian McAuley, Garrison W. Cottrell</i>	

INVESTIGATION OF DATA AUGMENTATION TECHNIQUES FOR DISORDERED SPEECH RECOGNITION.....	696
<i>Mengzhe Geng, Xurong Xie, Shansong Liu, Jianwei Yu, Shoukang Hu, Xunying Liu, Helen Meng</i>	
A REAL-TIME ROBOT-BASED AUXILIARY SYSTEM FOR RISK EVALUATION OF COVID-19 INFECTION.....	701
<i>Wenqi Wei, Jianzong Wang, Jiteng Ma, Ning Cheng, Jing Xiao</i>	
AN UTTERANCE VERIFICATION SYSTEM FOR WORD NAMING THERAPY IN APHASIA	706
<i>David S. Barbera, Mark Huckvale, Victoria Fleming, Emily Upton, Henry Coley-Fisher, Ian Shaw, William Latham, Alexander P. Leff, Jenny Crinion</i>	
EXPLOITING CROSS-DOMAIN VISUAL FEATURE GENERATION FOR DISORDERED SPEECH RECOGNITION.....	711
<i>Shansong Liu, Xurong Xie, Jianwei Yu, Shoukang Hu, Mengzhe Geng, Rongfeng Su, Shi-Xiong Zhang, Xunying Liu, Helen Meng</i>	
JOINT PREDICTION OF PUNCTUATION AND DISFLUENCY IN SPEECH TRANSCRIPTS	716
<i>Binghuai Lin, Liyuan Wang</i>	
FOCAL LOSS FOR PUNCTUATION PREDICTION.....	721
<i>Jiangyan Yi, Jianhua Tao, Zhengkun Tian, Ye Bai, Cunhang Fan</i>	

VOLUME 2

INTERSPEECH 2020 : LARGE-SCALE EVALUATION OF SHORT-DURATION SPEAKER VERIFICATION

IMPROVING X-VECTOR AND PLDA FOR TEXT-DEPENDENT SPEAKER VERIFICATION	726
<i>Zhuxin Chen, Yue Lin</i>	
SDSV CHALLENGE 2020: LARGE-SCALE EVALUATION OF SHORT-DURATION SPEAKER VERIFICATION	731
<i>Hossein Zeinali, Kong Aik Lee, Jahangir Alam, Lukas Burget</i>	
THE XMUSPEECH SYSTEM FOR SHORT-DURATION SPEAKER VERIFICATION CHALLENGE 2020	736
<i>Tao Jiang, Miao Zhao, Lin Li, Qingyang Hong</i>	
ROBUST TEXT-DEPENDENT SPEAKER VERIFICATION VIA CHARACTER-LEVEL INFORMATION PRESERVATION FOR THE SDSV CHALLENGE 2020.....	741
<i>Sung Hwan Mun, Woo Hyun Kang, Min Hyun Han, Nam Soo Kim</i>	
THE TALTECH SYSTEMS FOR THE SHORT-DURATION SPEAKER VERIFICATION CHALLENGE 2020	746
<i>Tanel Alumae, Jorgen Valk</i>	
INVESTIGATION OF NICT SUBMISSION FOR SHORT-DURATION SPEAKER VERIFICATION CHALLENGE 2020	751
<i>Peng Shen, Xugang Lu, Hisashi Kawai</i>	
CROSS-LINGUAL SPEAKER VERIFICATION WITH DOMAIN-BALANCED HARD PROTOTYPE MINING AND LANGUAGE-DEPENDENT SCORE NORMALIZATION.....	756
<i>Jenthe Thienpondt, Brecht Desplanques, Kris Demuynck</i>	

BUT TEXT-DEPENDENT SPEAKER VERIFICATION SYSTEM FOR SDSV CHALLENGE 2020.....	761
<i>Alicia Lozano-Diez, Anna Silnova, Bhargav Pulugundla, Johan Rohdin, Karel Vesely, Lukas Burget, Oldrich Plchot, Ondrej Glembek, Ondrej Novotny, Pavel Matejka</i>	

EXPLORING THE USE OF AN UNSUPERVISED AUTOREGRESSIVE MODEL AS A SHARED ENCODER FOR TEXT-DEPENDENT SPEAKER VERIFICATION	766
<i>Vijay Ravi, Ruchao Fan, Amber Afshan, Huanhua Lu, Abeer Alwan</i>	

INTERSPEECH 2020 : VOICE CONVERSION AND ADAPTATION I

RECOGNITION-SYNTHESIS BASED NON-PARALLEL VOICE CONVERSION WITH ADVERSARIAL LEARNING.....	771
<i>Jing-Xuan Zhang, Zhen-Hua Ling, Li-Rong Dai</i>	

IMPROVING THE SPEAKER IDENTITY OF NON-PARALLEL MANY-TO-MANY VOICE CONVERSION WITH ADVERSARIAL SPEAKER RECOGNITION	776
<i>Shaojin Ding, Guanlong Zhao, Ricardo Gutierrez-Osuna</i>	

NON-PARALLEL MANY-TO-MANY VOICE CONVERSION WITH PSR-STARGAN	781
<i>Yanping Li, Dongxiang Xu, Yan Zhang, Yang Wang, Binbin Chen</i>	

TTS SKINS: SPEAKER CONVERSION VIA ASR	786
<i>Adam Polyak, Lior Wolf, Yaniv Taigman</i>	

GAZEV: GAN-BASED ZERO-SHOT VOICE CONVERSION OVER NON-PARALLEL SPEECH CORPUS.....	791
<i>Zining Zhang, Bingsheng He, Zhenjie Zhang</i>	

SPOKEN CONTENT AND VOICE FACTORIZATION FOR FEW-SHOT SPEAKER ADAPTATION	796
<i>Tao Wang, Jianhua Tao, Ruibo Fu, Jiangyan Yi, Zhengqi Wen, Rongxiu Zhong</i>	

UNSUPERVISED CROSS-DOMAIN SINGING VOICE CONVERSION.....	801
<i>Adam Polyak, Lior Wolf, Yossi Adi, Yaniv Taigman</i>	

ATTENTION-BASED SPEAKER EMBEDDINGS FOR ONE-SHOT VOICE CONVERSION	806
<i>Tatsuma Ishihara, Daisuke Saito</i>	

DATA EFFICIENT VOICE CLONING FROM NOISY SAMPLES WITH DOMAIN ADVERSARIAL TRAINING.....	811
<i>Jian Cong, Shan Yang, Lei Xie, Guoqiao Yu, Guanglu Wan</i>	

INTERSPEECH 2020 : ACOUSTIC EVENT DETECTION

GATED MULTI-HEAD ATTENTION POOLING FOR WEAKLY LABELLED AUDIO TAGGING.....	816
<i>Sixin Hong, Yuexian Zou, Wenwu Wang</i>	

ENVIRONMENTAL SOUND CLASSIFICATION WITH PARALLEL TEMPORAL-SPECTRAL ATTENTION	821
<i>Helin Wang, Yuexian Zou, Dading Chong, Wenwu Wang</i>	

CONTRASTIVE PREDICTIVE CODING OF AUDIO WITH AN ADVERSARY	826
<i>Luyu Wang, Kazuya Kawakami, Aaron Van Den Oord</i>	

MEMORY CONTROLLED SEQUENTIAL SELF ATTENTION FOR SOUND RECOGNITION.....	831
<i>Arjun Pankajakshan, Helen L. Bear, Vinod Subramanian, Emmanouil Benetos</i>	
DUAL STAGE LEARNING BASED DYNAMIC TIME-FREQUENCY MASK GENERATION FOR AUDIO EVENT CLASSIFICATION	836
<i>Donghyeon Kim, Jaihyun Park, David K. Han, Hanseok Ko</i>	
AN EFFECTIVE PERTURBATION BASED SEMI-SUPERVISED LEARNING METHOD FOR SOUND EVENT DETECTION	841
<i>Xu Zheng, Yan Song, Jie Yan, Li-Rong Dai, Ian McLoughlin, Lin Liu</i>	
A JOINT FRAMEWORK FOR AUDIO TAGGING AND WEAKLY SUPERVISED ACOUSTIC EVENT DETECTION USING DENSENET WITH GLOBAL AVERAGE POOLING	846
<i>Chieh-Chi Kao, Bowen Shi, Ming Sun, Chao Wang</i>	
INTRA-UTTERANCE SIMILARITY PRESERVING KNOWLEDGE DISTILLATION FOR AUDIO TAGGING	851
<i>Chun-Chieh Chang, Chieh-Chi Kao, Ming Sun, Chao Wang</i>	
TWO-STAGE POLYPHONIC SOUND EVENT DETECTION BASED ON FASTER R-CNN- LSTM WITH MULTI-TOKEN CONNECTIONIST TEMPORAL CLASSIFICATION	856
<i>Inyoung Park, Hong Kook Kim</i>	
SPEECHMIX --- AUGMENTING DEEP SOUND RECOGNITION USING HIDDEN SPACE INTERPOLATIONS	861
<i>Amit Jindal, Narayanan Elavathur Ranganatha, Aniket Didolkar, Arijit Ghosh Chowdhury, Di Jin, Ramit Sawhney, Rajiv Ratn Shah</i>	
<u>INTERSPEECH 2020 : SPOKEN LANGUAGE UNDERSTANDING I</u>	
END-TO-END NEURAL TRANSFORMER BASED SPOKEN LANGUAGE UNDERSTANDING	866
<i>Martin Radfar, Athanasios Mouchtaris, Siegfried Kunzmann</i>	
JOINTLY ENCODING WORD CONFUSION NETWORK AND DIALOGUE CONTEXT WITH BERT FOR SPOKEN LANGUAGE UNDERSTANDING.....	871
<i>Chen Liu, Su Zhu, Zijian Zhao, Ruisheng Cao, Lu Chen, Kai Yu</i>	
SPEECH TO SEMANTICS: IMPROVE ASR AND NLU JOINTLY VIA ALL-NEURAL INTERFACES.....	876
<i>Milind Rao, Anirudh Raju, Pranav Dheram, Bach Bui, Ariya Rastrow</i>	
PRETRAINED SEMANTIC SPEECH EMBEDDINGS FOR END-TO-END SPOKEN LANGUAGE UNDERSTANDING VIA CROSS-MODAL TEACHER-STUDENT LEARNING	881
<i>Pavel Denisov, Ngoc Thang Vu</i>	
CONTEXT DEPENDENT RNNLM FOR AUTOMATIC TRANSCRIPTION OF CONVERSATIONS.....	886
<i>Srikanth Raj Chetupalli, Sriram Ganapathy</i>	
IMPROVING END-TO-END SPEECH-TO-INTENT CLASSIFICATION WITH REPTILE.....	891
<i>Yusheng Tian, Philip John Gorinski</i>	
SPEECH TO TEXT ADAPTATION: TOWARDS AN EFFICIENT CROSS-MODAL DISTILLATION	896
<i>Won Ik Cho, Donghyun Kwak, Ji Won Yoon, Nam Soo Kim</i>	

TOWARDS AN ASR ERROR ROBUST SPOKEN LANGUAGE UNDERSTANDING SYSTEM..... 901
Weitong Ruan, Yaroslav Nechaev, Luoxin Chen, Chengwei Su, Imre Kiss

END-TO-END SPOKEN LANGUAGE UNDERSTANDING WITHOUT FULL TRANSCRIPTS 906
Hong-Kwang J. Kuo, Zoltan Tuske, Samuel Thomas, Yinghui Huang, Kartik Audhkhasi, Brian Kingsbury, Gakuto Kurata, Zvi Kons, Ron Hoory, Luis Lastras

ARE NEURAL OPEN-DOMAIN DIALOG SYSTEMS ROBUST TO SPEECH RECOGNITION
ERRORS IN THE DIALOG HISTORY? AN EMPIRICAL STUDY 911
Karthik Gopalakrishnan, Behnam Hedayatnia, Longshaokan Wang, Yang Liu, Dilek Hakkani-Tur

INTERSPEECH 2020 : DNN ARCHITECTURES FOR SPEAKER RECOGNITION

AUTOSPEECH: NEURAL ARCHITECTURE SEARCH FOR SPEAKER RECOGNITION..... 916
Shaojin Ding, Tianlong Chen, Xinyu Gong, Weiwei Zha, Zhangyang Wang

DENSELY CONNECTED TIME DELAY NEURAL NETWORK FOR SPEAKER
VERIFICATION 921
Ya-Qi Yu, Wu-Jun Li

PHONETICALLY-AWARE COUPLED NETWORK FOR SHORT DURATION TEXT-
INDEPENDENT SPEAKER VERIFICATION 926
Siqi Zheng, Yun Lei, Hongbin Suo

MULTI-TASK NETWORK FOR NOISE-ROBUST KEYWORD SPOTTING AND SPEAKER
VERIFICATION USING CTC-BASED SOFT VAD AND GLOBAL QUERY ATTENTION..... 931
Myunghun Jung, Youngmoon Jung, Jahyun Goo, Hoirin Kim

VECTOR-BASED ATTENTIVE POOLING FOR TEXT-INDEPENDENT SPEAKER
VERIFICATION 936
Yanfeng Wu, Chenkai Guo, Hongcan Gao, Xiaolei Hou, Jing Xu

SELF-ATTENTION ENCODING AND POOLING FOR SPEAKER RECOGNITION..... 941
Pooyan Safari, Miquel India, Javier Hernando

ARET: AGGREGATED RESIDUAL EXTENDED TIME-DELAY NEURAL NETWORKS FOR
SPEAKER VERIFICATION 946
Ruiteng Zhang, Jianguo Wei, Wenhuan Lu, Longbiao Wang, Meng Liu, Lin Zhang, Jiayu Jin, Junhai Xu

ADVERSARIAL SEPARATION NETWORK FOR SPEAKER RECOGNITION..... 951
Hanyi Zhang, Longbiao Wang, Yunchun Zhang, Meng Liu, Kong Aik Lee, Jianguo Wei

TEXT-INDEPENDENT SPEAKER VERIFICATION WITH DUAL ATTENTION NETWORK..... 956
Jingyu Li, Tan Lee

EVOLUTIONARY ALGORITHM ENHANCED NEURAL ARCHITECTURE SEARCH FOR
TEXT-INDEPENDENT SPEAKER VERIFICATION 961
Xiaoyang Qu, Jianzong Wang, Jing Xiao

INTERSPEECH 2020 : ASR MODEL TRAINING AND STRATEGIES

MINIMUM BAYES RISK TRAINING OF RNN-TRANSDUCER FOR END-TO-END SPEECH RECOGNITION.....	966
<i>Chao Weng, Chengzhu Yu, Jia Cui, Chunlei Zhang, Dong Yu</i>	
SEMANTIC MASK FOR TRANSFORMER BASED END-TO-END SPEECH RECOGNITION.....	971
<i>Chengyi Wang, Yu Wu, Yujiao Du, Jinyu Li, Shujie Liu, Liang Lu, Shuo Ren, Guoli Ye, Sheng Zhao, Ming Zhou</i>	
FASTER, SIMPLER AND MORE ACCURATE HYBRID ASR SYSTEMS USING WORDPIECES	976
<i>Frank Zhang, Yongqiang Wang, Xiaohui Zhang, Chunxi Liu, Yatharth Saraf, Geoffrey Zweig</i>	
A FEDERATED APPROACH IN TRAINING ACOUSTIC MODELS	981
<i>Dimitrios Dimitriadis, Kenichi Kumatani, Robert Gmyr, Yashesh Gaur, Sefik Emre Eskimez</i>	
ON SEMI-SUPERVISED LF-MMI TRAINING OF ACOUSTIC MODELS WITH LIMITED DATA.....	986
<i>Imran Sheikh, Emmanuel Vincent, Irina Illina</i>	
ON FRONT-END GAIN INVARIANT MODELING FOR WAKE WORD SPOTTING.....	991
<i>Yixin Gao, Noah D. Stein, Chieh-Chi Kao, Yunliang Cai, Ming Sun, Tao Zhang, Shiv Naga Prasad Vitaladevuni</i>	
UNSUPERVISED REGULARIZATION-BASED ADAPTIVE TRAINING FOR SPEECH RECOGNITION.....	996
<i>Fenglin Ding, Wu Guo, Bin Gu, Zhen-Hua Ling, Jun Du</i>	
ON THE ROBUSTNESS AND TRAINING DYNAMICS OF RAW WAVEFORM MODELS	1001
<i>Erfan Loweimi, Peter Bell, Steve Renals</i>	
ITERATIVE PSEUDO-LABELING FOR SPEECH RECOGNITION.....	1006
<i>Qiantong Xu, Tatiana Likhomanenko, Jacob Kahn, Awni Hannun, Gabriel Synnaeve, Ronan Collobert</i>	

INTERSPEECH 2020 : SPEECH ANNOTATION AND SPEECH ASSESSMENT

SMART TUBE: A BIOFEEDBACK SYSTEM FOR VOCAL TRAINING AND THERAPY THROUGH TUBE PHONATION.....	1011
<i>Naoko Kawamura, Tatsuya Kitamura, Kenta Hamada</i>	
VCTUBE : A LIBRARY FOR AUTOMATIC SPEECH DATA ANNOTATION	1013
<i>Seong Choi, Seunghoon Jeong, Jeewoo Yoon, Migyeong Yang, Minsam Ko, Eunil Park, Jinyoung Han, Munyoung Lee, Seonghee Lee</i>	
A MANDARIN L2 LEARNING APP WITH MISPRONUNCIATION DETECTION AND FEEDBACK.....	1015
<i>Yanlu Xie, Xiaoli Feng, Boxue Li, Jinsong Zhang, Yujia Jin</i>	
RAPID ENHANCEMENT OF NLP SYSTEMS BY ACQUISITION OF DATA IN CORRELATED DOMAINS	1017
<i>Tejas Udayakumar, Kinnera Saranu, Mayuresh Sanjay Oak, Ajit Ashok Saunshikar, Sandip Shriram Bapat</i>	

COMPUTER-ASSISTED LANGUAGE LEARNING SYSTEM: AUTOMATIC SPEECH
EVALUATION FOR CHILDREN LEARNING MALAY AND TAMIL 1019
*Ke Shi, Kye Min Tan, Richeng Duan, Siti Umairah Md. Salleh, Nur Farah Ain Suhaimi, Rajan
Vellu, Ngoc Thuy Huong Helen Thai, Nancy F. Chen*

REAL-TIME, FULL-BAND, ONLINE DNN-BASED VOICE CONVERSION SYSTEM USING
A SINGLE CPU 1021
Takaaki Saeki, Yuki Saito, Shinnosuke Takamichi, Hiroshi Saruwatari

A DYNAMIC 3D PRONUNCIATION TEACHING MODEL BASED ON PRONUNCIATION
ATTRIBUTES AND ANATOMY 1023
Xiaoli Feng, Yanlu Xie, Yayue Deng, Boxue Li

END-TO-END DEEP LEARNING SPEECH RECOGNITION MODEL FOR SILENT SPEECH
CHALLENGE 1025
Naoki Kimura, Zixiong Su, Takaaki Saeki

INTERSPEECH 2020 : CROSS/MULTI-LINGUAL AND CODE-SWITCHED SPEECH RECOGNITION

AUTOSEGMENTAL NEURAL NETS: SHOULD PHONES AND TONES BE SYNCHRONOUS
OR ASYNCHRONOUS? 1027
Jialu Li, Mark Hasegawa-Johnson

DEVELOPMENT OF MULTILINGUAL ASR USING GLOBALPHONE FOR LESS-
RESOURCED LANGUAGES: THE CASE OF ETHIOPIAN LANGUAGES 1032
Martha Yifiru Tachbelie, Solomon Teferra Abate, Tanja Schultz

LARGE-SCALE END-TO-END MULTILINGUAL SPEECH RECOGNITION AND
LANGUAGE IDENTIFICATION WITH MULTI-TASK LEARNING 1037
Wenxin Hou, Yue Dong, Bairong Zhuang, Longfei Yang, Jiatong Shi, Takahiro Shinozaki

MULTI-ENCODER-DECODER TRANSFORMER FOR CODE-SWITCHING SPEECH
RECOGNITION 1042
Xinyuan Zhou, Emre Yilmaz, Yanhua Long, Yijie Li, Haizhou Li

MULTILINGUAL ACOUSTIC AND LANGUAGE MODELING FOR ETHIO-SEMITIC
LANGUAGES 1047
Solomon Teferra Abate, Martha Yifiru Tachbelie, Tanja Schultz

MULTILINGUAL JOINTLY TRAINED ACOUSTIC AND WRITTEN WORD EMBEDDINGS 1052
Yushi Hu, Shane Settle, Karen Livescu

IMPROVING CODE-SWITCHING LANGUAGE MODELING WITH ARTIFICIALLY
GENERATED TEXTS USING CYCLE-CONSISTENT ADVERSARIAL NETWORKS 1057
Chia-Yu Li, Ngoc Thang Vu

DATA AUGMENTATION FOR CODE-SWITCH LANGUAGE MODELING BY FUSING
MULTIPLE TEXT GENERATION METHODS 1062
Xinhui Hu, Qi Zhang, Lei Yang, Binbin Gu, Xinkang Xu

A 43 LANGUAGE MULTILINGUAL PUNCTUATION PREDICTION NEURAL NETWORK
MODEL 1067
Xinxing Li, Edward Lin

EXPLORING LEXICON-FREE MODELING UNITS FOR END-TO-END KOREAN AND KOREAN-ENGLISH CODE-SWITCHING SPEECH RECOGNITION.....	1072
<i>Jisung Wang, Jihwan Kim, Sangki Kim, Yeha Lee</i>	

INTERSPEECH 2020 : ANTI-SPOOFING AND LIVENESS DETECTION

MULTI-TASK SIAMESE NEURAL NETWORK FOR IMPROVING REPLAY ATTACK DETECTION.....	1076
<i>Patrick Von Platen, Fei Tao, Gokhan Tur</i>	

POCO: A VOICE SPOOFING AND LIVENESS DETECTION CORPUS BASED ON POP NOISE.....	1081
<i>Kosuke Akimoto, Seng Pei Liew, Sakiko Mishima, Ryo Mizushima, Kong Aik Lee</i>	

DUAL-ADVERSARIAL DOMAIN ADAPTATION FOR GENERALIZED REPLAY ATTACK DETECTION.....	1086
<i>Hongji Wang, Heinrich Dinkel, Shuai Wang, Yanmin Qian, Kai Yu</i>	

SELF-SUPERVISED PRE-TRAINING WITH ACOUSTIC CONFIGURATIONS FOR REPLAY SPOOFING DETECTION.....	1091
<i>Hye-Jin Shim, Hee-Soo Heo, Jee-Weon Jung, Ha-Jin Yu</i>	

COMPETENCY EVALUATION IN VOICE MIMICKING USING ACOUSTIC CUES.....	1096
<i>Abhijith G., Adharsh S., Akshay P. L., Rajeev Rajan</i>	

LIGHT CONVOLUTIONAL NEURAL NETWORK WITH FEATURE GENUINIZATION FOR DETECTION OF SYNTHETIC SPEECH ATTACKS.....	1101
<i>Zhenzong Wu, Rohan Kumar Das, Jichen Yang, Haizhou Li</i>	

SPOOFING ATTACK DETECTION USING THE NON-LINEAR FUSION OF SUB-BAND CLASSIFIERS.....	1106
<i>Hemlata Tak, Jose Patino, Andreas Nautsch, Nicholas Evans, Massimiliano Todisco</i>	

INVESTIGATING LIGHT-RESNET ARCHITECTURE FOR SPOOFING DETECTION UNDER MISMATCHED CONDITIONS.....	1111
<i>Prasanth Parasu, Julien Epps, Kaavya Sriskandaraja, Gajan Suthokumar</i>	

SIAMESE CONVOLUTIONAL NEURAL NETWORK USING GAUSSIAN PROBABILITY FEATURE FOR SPOOFING SPEECH DETECTION.....	1116
<i>Zhenchun Lei, Yingen Yang, Changhong Liu, Jihua Ye</i>	

INTERSPEECH 2020 : NOISE REDUCTION AND INTELLIGIBILITY

LIGHTWEIGHT ONLINE NOISE REDUCTION ON EMBEDDED DEVICES USING HIERARCHICAL RECURRENT NEURAL NETWORKS.....	1121
<i>H. Schroter, T. Rosenkranz, A. N. Escalante-B., P. Zobel, Andreas Maier</i>	

SEANET: A MULTI-MODAL SPEECH ENHANCEMENT NETWORK.....	1126
<i>Marco Tagliasacchi, Yunpeng Li, Karolis Misiunas, Dominik Roblek</i>	

LITE AUDIO-VISUAL SPEECH ENHANCEMENT.....	1131
<i>Shang-Yi Chuang, Yu Tsao, Chen-Chou Lo, Hsin-Min Wang</i>	

ORCA-CLEAN: A DEEP DENOISING TOOLKIT FOR KILLER WHALE COMMUNICATION.....	1136
<i>Christian Bergler, Manuel Schmitt, Andreas Maier, Simeon Smeele, Volker Barth, Elmar Noth</i>	
A DEEP LEARNING APPROACH TO ACTIVE NOISE CONTROL	1141
<i>Hao Zhang, Deliang Wang</i>	
IMPROVING SPEECH INTELLIGIBILITY THROUGH SPEAKER DEPENDENT AND INDEPENDENT SPECTRAL STYLE CONVERSION.....	1146
<i>Tuan Dinh, Alexander Kain, Kris Tjaden</i>	
END-TO-END SPEECH INTELLIGIBILITY PREDICTION USING TIME-DOMAIN FULLY CONVOLUTIONAL NEURAL NETWORKS.....	1151
<i>Mathias B. Pedersen, Morten Kolbaek, Asger H. Andersen, Soren H. Jensen, Jesper Jensen</i>	
PREDICTING INTELLIGIBILITY OF ENHANCED SPEECH USING POSTERIOBS DERIVED FROM DNN-BASED ASR SYSTEM	1156
<i>Kenichi Arai, Shoko Araki, Atsunori Ogawa, Keisuke Kinoshita, Tomohiro Nakatani, Toshio Irino</i>	
AUTOMATIC ESTIMATION OF INTELLIGIBILITY MEASURE FOR CONSONANTS IN SPEECH	1161
<i>Ali Abavisani, Mark Hasegawa-Johnson</i>	
LARGE SCALE EVALUATION OF IMPORTANCE MAPS IN AUTOMATIC SPEECH RECOGNITION.....	1166
<i>Viet Anh Trinh, Michael I. Mandel</i>	
 <u>INTERSPEECH 2020 : ACOUSTIC SCENE CLASSIFICATION</u>	
NEURAL ARCHITECTURE SEARCH ON ACOUSTIC SCENE CLASSIFICATION.....	1171
<i>Jixiang Li, Chuming Liang, Bo Zhang, Zhao Wang, Fei Xiang, Xiangxiang Chu</i>	
ACOUSTIC SCENE CLASSIFICATION USING AUDIO TAGGING	1176
<i>Jee-Weon Jung, Hye-Jin Shim, Ju-Ho Kim, Seung-Bin Kim, Ha-Jin Yu</i>	
ATRESN-NET: CAPTURING ATTENTIVE TEMPORAL RELATIONS IN SEMANTIC NEIGHBORHOOD FOR ACOUSTIC SCENE CLASSIFICATION.....	1181
<i>Liwen Zhang, Jiqing Han, Ziqiang Shi</i>	
ENVIRONMENT SOUND CLASSIFICATION USING MULTIPLE FEATURE CHANNELS AND ATTENTION BASED DEEP CONVOLUTIONAL NEURAL NETWORK.....	1186
<i>Jivitesh Sharma, Ole-Christoffer Granmo, Morten Goodwin</i>	
ACOUSTIC SCENE ANALYSIS WITH MULTI-HEAD ATTENTION NETWORKS	1191
<i>Weimin Wang, Weiran Wang, Ming Sun, Chao Wang</i>	
RELATIONAL TEACHER STUDENT LEARNING WITH NEURAL LABEL EMBEDDING FOR DEVICE ADAPTATION IN ACOUSTIC SCENE CLASSIFICATION	1196
<i>Hu Hu, Sabato Marco Siniscalchi, Yannan Wang, Chin-Hui Lee</i>	
AN ACOUSTIC SEGMENT MODEL BASED SEGMENT UNIT SELECTION APPROACH TO ACOUSTIC SCENE CLASSIFICATION WITH PARTIAL UTTERANCES	1201
<i>Hu Hu, Sabato Marco Siniscalchi, Yannan Wang, Xue Bai, Jun Du, Chin-Hui Lee</i>	

ATTENTION-DRIVEN PROJECTIONS FOR SOUNDSCAPE CLASSIFICATION 1206
Dhanunjaya Varma Devalraju, Muralikrishna H., Padmanabhan Rajan, Dileep Aroor Dinesh

COMPUTER AUDITION FOR CONTINUOUS RAINFOREST OCCUPANCY MONITORING:
THE CASE OF BORNEAN GIBBONS' CALL DETECTION 1211
Panagiotis Tzirakis, Alexander Shiarella, Robert Ewers, Bjorn W. Schuller

DEEP LEARNING BASED OPEN SET ACOUSTIC SCENE CLASSIFICATION..... 1216
Zuzanna Kwiatkowska, Benjamin Kalinowski, Michal Kosmider, Krzysztof Rykaczewski

INTERSPEECH 2020 : SINGING VOICE COMPUTING AND PROCESSING IN MUSIC

SINGING SYNTHESIS: WITH A LITTLE HELP FROM MY ATTENTION 1221
Orazio Angelini, Alexis Moinet, Kayoko Yanagisawa, Thomas Drugman

PEKING OPERA SYNTHESIS VIA DURATION INFORMED ATTENTION NETWORK..... 1226
Yusong Wu, Shengchen Li, Chengzhu Yu, Heng Lu, Chao Weng, Liqiang Zhang, Dong Yu

DURIAN-SC: DURATION INFORMED ATTENTION NETWORK BASED SINGING VOICE
CONVERSION SYSTEM..... 1231
*Liqiang Zhang, Chengzhu Yu, Heng Lu, Chao Weng, Chunlei Zhang, Yusong Wu, Xiang Xie,
Zijin Li, Dong Yu*

TRANSFER LEARNING FOR IMPROVING SINGING-VOICE DETECTION IN
POLYPHONIC INSTRUMENTAL MUSIC 1236
Yuanbo Hou, Frank K. Soong, Jian Luan, Shengchen Li

CHANNEL-WISE SUBBAND INPUT FOR BETTER VOICE AND ACCOMPANIMENT
SEPARATION ON HIGH RESOLUTION MUSIC 1241
Haohe Liu, Lei Xie, Jian Wu, Geng Yang

INTERSPEECH 2020 : ACOUSTIC MODEL ADAPTATION FOR ASR

CONTINUAL LEARNING IN AUTOMATIC SPEECH RECOGNITION..... 1246
Samik Sadhu, Hynek Hermansky

SPEAKER ADAPTIVE TRAINING FOR SPEECH RECOGNITION BASED ON ATTENTION-
OVER-ATTENTION MECHANISM..... 1251
Genshun Wan, Jia Pan, Qingran Wang, Jianqing Gao, Zhongfu Ye

RAPID RNN-T ADAPTATION USING PERSONALIZED SPEECH SYNTHESIS AND
NEURAL LANGUAGE GENERATOR..... 1256
Yan Huang, Jinyu Li, Lei He, Wenning Wei, William Gale, Yifan Gong

SPEECH TRANSFORMER WITH SPEAKER AWARE PERSISTENT MEMORY 1261
Yingzhu Zhao, Chongjia Ni, Cheung-Chi Leung, Shafiq Joty, Eng Siong Chng, Bin Ma

ADAPTIVE SPEAKER NORMALIZATION FOR CTC-BASED SPEECH RECOGNITION 1266
Fenglin Ding, Wu Guo, Bin Gu, Zhen-Hua Ling, Jun Du

UNSUPERVISED DOMAIN ADAPTATION UNDER LABEL SPACE MISMATCH FOR
SPEECH CLASSIFICATION..... 1271
Akhil Mathur, Nadia Berthouze, Nicholas D. Lane

LEARNING FAST ADAPTATION ON CROSS-ACCENTED SPEECH RECOGNITION..... 1276
Genta Indra Winata, Samuel Cahyawijaya, Zihan Liu, Zhaojiang Lin, Andrea Madotto, Peng Xu, Pascale Fung

BLACK-BOX ADAPTATION OF ASR FOR ACCENTED SPEECH..... 1281
Kartik Khandelwal, Preethi Jyothi, Abhijeet Awasthi, Sunita Sarawagi

ACHIEVING MULTI-ACCENT ASR VIA UNSUPERVISED ACOUSTIC MODEL
ADAPTATION 1286
M. A. Tugtekin Turan, Emmanuel Vincent, Denis Jouviet

FRAME-WISE ONLINE UNSUPERVISED ADAPTATION OF DNN-HMM ACOUSTIC
MODEL FROM PERSPECTIVE OF ROBUST ADAPTIVE FILTERING..... 1291
Ryu Takeda, Kazunori Komatani

INTERSPEECH 2020 : SINGING AND MULTIMODAL SYNTHESIS

ADVERSARIALY TRAINED MULTI-SINGER SEQUENCE-TO-SEQUENCE SINGING
SYNTHESIZER 1296
Jie Wu, Jian Luan

PREDICTION OF HEAD MOTION FROM SPEECH WAVEFORMS WITH A CANONICAL-
CORRELATION-CONSTRAINED AUTOENCODER..... 1301
Jinhong Lu, Hiroshi Shimodaira

XIAOICESING: A HIGH-QUALITY AND INTEGRATED SINGING VOICE SYNTHESIS
SYSTEM 1306
Peiling Lu, Jie Wu, Jian Luan, Xu Tan, Li Zhou

STOCHASTIC TALKING FACE GENERATION USING LATENT DISTRIBUTION
MATCHING 1311
Ravindra Yadav, Ashish Sardana, Vinay P. Namboodiri, Rajesh M. Hegde

SPEECH-TO-SINGING CONVERSION BASED ON BOUNDARY EQUILIBRIUM GAN 1316
Da-Yi Wu, Yi-Hsuan Yang

FACE2SPEECH: TOWARDS MULTI-SPEAKER TEXT-TO-SPEECH SYNTHESIS USING AN
EMBEDDING VECTOR PREDICTED FROM A FACE IMAGE 1321
Shunsuke Goto, Kotaro Onishi, Yuki Saito, Kentaro Tachibana, Koichiro Mori

SPEECH DRIVEN TALKING HEAD GENERATION VIA ATTENTIONAL LANDMARKS
BASED REPRESENTATION 1326
Wentao Wang, Yan Wang, Jianqing Sun, Qingsong Liu, Jiaen Liang, Teng Li

INTERSPEECH 2020 : INTELLIGIBILITY-ENHANCING SPEECH MODIFICATION

OPTIMIZATION AND EVALUATION OF AN INTELLIGIBILITY-IMPROVING SIGNAL
PROCESSING APPROACH (IISPA) FOR THE HURRICANE CHALLENGE 2.0 WITH FADE..... 1331
Marc Rene Schadler

IMETRICGAN: INTELLIGIBILITY ENHANCEMENT FOR SPEECH-IN-NOISE USING
GENERATIVE ADVERSARIAL NETWORK-BASED METRIC LEARNING 1336
Haoyu Li, Szu-Wei Fu, Yu Tsao, Junichi Yamagishi

INTELLIGIBILITY-ENHANCING SPEECH MODIFICATIONS --- THE HURRICANE CHALLENGE 2.0	1341
<i>Jan RENNIES, Henning Schepker, Cassia Valentini-Botinhao, Martin Cooke</i>	
EXPLORING LISTENERS' SPEECH RATE PREFERENCES	1346
<i>Olympia Simantiraki, Martin Cooke</i>	
ADAPTIVE COMPRESSIVE ONSET-ENHANCEMENT FOR IMPROVED SPEECH INTELLIGIBILITY IN NOISE AND REVERBERATION	1351
<i>Felicitas Bederna, Henning Schepker, Christian Rollwage, Simon Doclo, Arne Pusch, Jorg Bitzer, Jan RENNIES</i>	
A SOUND ENGINEERING APPROACH TO NEAR END LISTENING ENHANCEMENT	1356
<i>Carol Chermaz, Simon King</i>	
ENHANCING SPEECH INTELLIGIBILITY IN TEXT-TO-SPEECH SYNTHESIS USING SPEAKING STYLE CONVERSION	1361
<i>Dipjyoti Paul, Muhammed P. V. Shifas, Yannis Pantazis, Yannis Stylianou</i>	

INTERSPEECH 2020 : HUMAN SPEECH PRODUCTION I

TWO DIFFERENT MECHANISMS OF MOVABLE MANDIBLE FOR VOCAL-TRACT MODEL WITH FLEXIBLE TONGUE	1366
<i>Takayuki Arai</i>	
IMPROVING THE PERFORMANCE OF ACOUSTIC-TO-ARTICULATORY INVERSION BY REMOVING THE TRAINING LOSS OF NONCRITICAL PORTIONS OF ARTICULATORY CHANNELS DYNAMICALLY	1371
<i>Qiang Fang</i>	
SPEAKER CONDITIONED ACOUSTIC-TO-ARTICULATORY INVERSION USING X-VECTORS.....	1376
<i>Aravind Illa, Prasanta Kumar Ghosh</i>	
COARTICULATION AS SYNCHRONISED SEQUENTIAL TARGET APPROXIMATION: AN EMA STUDY	1381
<i>Zirui Liu, Yi Xu, Feng-Fan Hsieh</i>	
IMPROVED MODEL FOR VOCAL FOLDS WITH A POLYP WITH POTENTIAL APPLICATION.....	1386
<i>Jonatas Santos, Jugurta Montalvao, Israel Santos</i>	
REGIONAL RESONANCE OF THE LOWER VOCAL TRACT AND ITS CONTRIBUTION TO SPEAKER CHARACTERISTICS	1391
<i>Lin Zhang, Kiyoshi Honda, Jianguo Wei, Seiji Adachi</i>	
AIR-TISSUE BOUNDARY SEGMENTATION IN REAL TIME MAGNETIC RESONANCE IMAGING VIDEO USING 3-D CONVOLUTIONAL NEURAL NETWORK	1396
<i>Renuka Mannem, Navaneetha Gaddam, Prasanta Kumar Ghosh</i>	
AN INVESTIGATION OF THE VIRTUAL LIP TRAJECTORIES DURING THE PRODUCTION OF BILABIAL STOPS AND NASAL AT DIFFERENT SPEAKING RATES	1401
<i>Tilak Purohit, Prasanta Kumar Ghosh</i>	

INTERSPEECH 2020 : TARGETED SOURCE SEPARATION

SPEX+: A COMPLETE TIME DOMAIN SPEAKER EXTRACTION NETWORK	1406
<i>Meng Ge, Chenglin Xu, Longbiao Wang, Eng Siong Chng, Jianwu Dang, Haizhou Li</i>	
ATSS-NET: TARGET SPEAKER SEPARATION VIA ATTENTION-BASED NEURAL NETWORK	1411
<i>Tingle Li, Qingjian Lin, Yuanyuan Bao, Ming Li</i>	
MULTIMODAL TARGET SPEECH SEPARATION WITH VOICE AND FACE REFERENCES.....	1416
<i>Leyuan Qu, Cornelius Weber, Stefan Wermter</i>	
X-TASNET: ROBUST AND ACCURATE TIME-DOMAIN SPEAKER EXTRACTION NETWORK	1421
<i>Zining Zhang, Bingsheng He, Zhenjie Zhang</i>	
LISTEN, WATCH AND UNDERSTAND AT THE COCKTAIL PARTY: AUDIO-VISUAL-CONTEXTUAL SPEECH SEPARATION.....	1426
<i>Chenda Li, Yanmin Qian</i>	
A UNIFIED FRAMEWORK FOR LOW-LATENCY SPEAKER EXTRACTION IN COCKTAIL PARTY ENVIRONMENTS	1431
<i>Yunzhe Hao, Jiaming Xu, Jing Shi, Peng Zhang, Lei Qin, Bo Xu</i>	
TIME-DOMAIN TARGET-SPEAKER SPEECH SEPARATION WITH WAVEFORM-BASED SPEAKER EMBEDDING	1436
<i>Jianshu Zhao, Shengzhou Gao, Takahiro Shinozaki</i>	
LISTEN TO WHAT YOU WANT: NEURAL NETWORK-BASED UNIVERSAL SOUND SELECTOR.....	1441
<i>Tsubasa Ochiai, Marc Delcroix, Yuma Koizumi, Hiroaki Ito, Keisuke Kinoshita, Shoko Araki</i>	
CROSSMODAL SOUND RETRIEVAL BASED ON SPECIFIC TARGET CO-OCCURRENCE DENOTED WITH WEAK LABELS.....	1446
<i>Masahiro Yasuda, Yasunori Ohishi, Yuma Koizumi, Noboru Harada</i>	

VOLUME 3

SPEAKER-AWARE MONAURAL SPEECH SEPARATION	1451
<i>Jiahao Xu, Kun Hu, Chang Xu, Duc Chung Tran, Zhiyong Wang</i>	

INTERSPEECH 2020 : SPEECH TRANSLATION AND MULTILINGUAL/MULTIMODAL LEARNING

A DNN-HMM-DNN HYBRID MODEL FOR DISCOVERING WORD-LIKE UNITS FROM SPOKEN CAPTIONS AND IMAGE REGIONS	1456
<i>Liming Wang, Mark Hasegawa-Johnson</i>	
EFFICIENT WAIT-K MODELS FOR SIMULTANEOUS MACHINE TRANSLATION	1461
<i>Maha Elbayad, Laurent Besacier, Jakob Verbeek</i>	
INVESTIGATING SELF-SUPERVISED PRE-TRAINING FOR END-TO-END SPEECH TRANSLATION	1466
<i>Ha Nguyen, Fethi Bougares, N. Tomashenko, Yannick Esteve, Laurent Besacier</i>	

CONTEXTUALIZED TRANSLATION OF AUTOMATICALLY SEGMENTED SPEECH	1471
<i>Marco Gaido, Mattia A. Di Gangi, Matteo Negri, Mauro Cettolo, Marco Turchi</i>	
SELF-TRAINING FOR END-TO-END SPEECH TRANSLATION.....	1476
<i>Juan Pino, Qiantong Xu, Xutai Ma, Mohammad Javad Dousti, Yun Tang</i>	
EVALUATING AND OPTIMIZING PROSODIC ALIGNMENT FOR AUTOMATIC DUBBING	1481
<i>Marcello Federico, Yogesh Virkar, Robert Enyedi, Roberto Barra-Chicote</i>	
PAIR EXPANSION FOR LEARNING MULTILINGUAL SEMANTIC EMBEDDINGS USING DISJOINT VISUALLY-GROUNDED SPEECH AUDIO DATASETS	1486
<i>Yasunori Ohishi, Akisato Kimura, Takahito Kawanishi, Kunio Kashino, David Harwath, James Glass</i>	
SELF-SUPERVISED REPRESENTATIONS IMPROVE END-TO-END SPEECH TRANSLATION	1491
<i>Anne Wu, Changhan Wang, Juan Pino, Jiatao Gu</i>	
<u>INTERSPEECH 2020 : SPEAKER RECOGNITION I</u>	
IMPROVED RAWNET WITH FEATURE MAP SCALING FOR TEXT-INDEPENDENT SPEAKER VERIFICATION USING RAW WAVEFORMS.....	1496
<i>Jee-Weon Jung, Seung-Bin Kim, Hye-Jin Shim, Ju-Ho Kim, Ha-Jin Yu</i>	
IMPROVING MULTI-SCALE AGGREGATION USING FEATURE PYRAMID MODULE FOR ROBUST SPEAKER VERIFICATION OF VARIABLE-DURATION UTTERANCES	1501
<i>Youngmoon Jung, Seong Min Kye, Yeunju Choi, Myunghun Jung, Hoirin Kim</i>	
AN ADAPTIVE X-VECTOR MODEL FOR TEXT-INDEPENDENT SPEAKER VERIFICATION	1506
<i>Bin Gu, Wu Guo, Fenglin Ding, Zhen-Hua Ling, Jun Du</i>	
SHOUTED SPEECH COMPENSATION FOR SPEAKER VERIFICATION ROBUST TO VOCAL EFFORT CONDITIONS	1511
<i>Santi Prieto, Alfonso Ortega, Ivan Lopez-Espejo, Eduardo Lleida</i>	
SUM-PRODUCT NETWORKS FOR ROBUST AUTOMATIC SPEAKER IDENTIFICATION.....	1516
<i>Aaron Nicolson, Kuldip K. Paliwal</i>	
SEGMENT AGGREGATION FOR SHORT UTTERANCES SPEAKER VERIFICATION USING RAW WAVEFORMS	1521
<i>Seung-Bin Kim, Jee-Weon Jung, Hye-Jin Shim, Ju-Ho Kim, Ha-Jin Yu</i>	
SIAMESE X-VECTOR RECONSTRUCTION FOR DOMAIN ADAPTED SPEAKER RECOGNITION.....	1526
<i>Shai Rozenberg, Hagai Aronowitz, Ron Hoory</i>	
SPEAKER RE-IDENTIFICATION WITH SPEAKER DEPENDENT SPEECH ENHANCEMENT	1530
<i>Yanpei Shi, Qiang Huang, Thomas Hain</i>	
BLIND SPEECH SIGNAL QUALITY ESTIMATION FOR SPEAKER VERIFICATION SYSTEMS.....	1535
<i>Galina Lavrentyeva, Marina Volkova, Anastasia Avdeeva, Sergey Novoselov, Artem Gorlanov, Tseren Andzhukaev, Artem Ivanov, Alexander Kozlov</i>	

INVESTIGATING ROBUSTNESS OF ADVERSARIAL SAMPLES DETECTION FOR AUTOMATIC SPEAKER VERIFICATION.....	1540
<i>Xu Li, Na Li, Jinghua Zhong, Xixin Wu, Xunying Liu, Dan Su, Dong Yu, Helen Meng</i>	

INTERSPEECH 2020 : SPOKEN LANGUAGE UNDERSTANDING II

MODELING ASR AMBIGUITY FOR NEURAL DIALOGUE STATE TRACKING	1545
<i>Vaishali Pal, Fabien Guillot, Manish Shrivastava, Jean-Michel Renders, Laurent Besacier</i>	

ASR ERROR CORRECTION WITH AUGMENTED TRANSFORMER FOR ENTITY RETRIEVAL.....	1550
<i>Haoyu Wang, Shuyan Dong, Yue Liu, James Logan, Ashish Kumar Agrawal, Yang Liu</i>	

LARGE-SCALE TRANSFER LEARNING FOR LOW-RESOURCE SPOKEN LANGUAGE UNDERSTANDING	1555
<i>Xueli Jia, Jianzong Wang, Zhiyong Zhang, Ning Cheng, Jing Xiao</i>	

DATA BALANCING FOR BOOSTING PERFORMANCE OF LOW-FREQUENCY CLASSES IN SPOKEN LANGUAGE UNDERSTANDING	1560
<i>Judith Gaspers, Quynh Do, Fabian Triefenbach</i>	

AN INTERACTIVE ADVERSARIAL REWARD LEARNING-BASED SPOKEN LANGUAGE UNDERSTANDING SYSTEM	1565
<i>Yu Wang, Yilin Shen, Hongxia Jin</i>	

STYLE ATTUNED PRE-TRAINING AND PARAMETER EFFICIENT FINE-TUNING FOR SPOKEN LANGUAGE UNDERSTANDING	1570
<i>Jin Cao, Jun Wang, Wael Hamza, Kelly Vanee, Shang-Wen Li</i>	

UNSUPERVISED DOMAIN ADAPTATION FOR DIALOGUE SEQUENCE LABELING BASED ON HIERARCHICAL ADVERSARIAL TRAINING	1575
<i>Shota Orihashi, Mana Ihori, Tomohiro Tanaka, Ryo Masumura</i>	

DEEP F-MEASURE MAXIMIZATION FOR END-TO-END SPEECH UNDERSTANDING.....	1580
<i>Leda Sari, Mark Hasegawa-Johnson</i>	

AN EFFECTIVE DOMAIN ADAPTIVE POST-TRAINING METHOD FOR BERT IN RESPONSE SELECTION	1585
<i>Taesun Whang, Dongyub Lee, Chanhee Lee, Kisu Yang, Dongsuk Oh, Heuiseok Lim</i>	

CONFIDENCE MEASURE FOR SPEECH-TO-CONCEPT END-TO-END SPOKEN LANGUAGE UNDERSTANDING.....	1590
<i>Antoine Caubriere, Yannick Esteve, Antoine Laurent, Emmanuel Morin</i>	

INTERSPEECH 2020 : HUMAN SPEECH PROCESSING

ATTENTION TO INDEXICAL INFORMATION IMPROVES VOICE RECALL	1595
<i>Grant L. McGuire, Molly Babel</i>	

CATEGORIZATION OF WHISTLED CONSONANTS BY FRENCH SPEAKERS	1600
<i>Anais Tran Ngoc, Julien Meyer, Fanny Meunier</i>	

WHISTLED VOWEL IDENTIFICATION BY FRENCH LISTENERS	1605
<i>Anais Tran Ngoc, Julien Meyer, Fanny Meunier</i>	

F0 SLOPE AND MEAN: CUES TO SPEECH SEGMENTATION IN FRENCH	1610
<i>Maria Del Mar Cordero, Fanny Meunier, Nicolas Grimault, Stephane Pota, Elsa Spinelli</i>	
DOES FRENCH LISTENERS' ABILITY TO USE ACCENTUAL INFORMATION AT THE WORD LEVEL DEPEND ON THE EAR OF PRESENTATION?.....	1615
<i>Amandine Michelas, Sophie Dufour</i>	
A PERCEPTUAL STUDY OF THE FIVE LEVEL TONES IN HMU (XINZHAI VARIETY).....	1620
<i>Wen Liu</i>	
MANDARIN AND ENGLISH ADULTS' CUE-WEIGHTING OF LEXICAL STRESS	1624
<i>Zhen Zeng, Karen Mattock, Liquan Liu, Varghese Peter, Alba Tuninetti, Feng-Ming Tsao</i>	
AGE-RELATED DIFFERENCES OF TONE PERCEPTION IN MANDARIN-SPEAKING SENIORS	1629
<i>Yan Feng, Gang Peng, William Shi-Yuan Wang</i>	
SOCIAL AND FUNCTIONAL PRESSURES IN VOCAL ALIGNMENT: DIFFERENCES FOR HUMAN AND VOICE-AI INTERLOCUTORS	1634
<i>Georgia Zellou, Michelle Cohn</i>	
IDENTIFYING IMPORTANT TIME-FREQUENCY LOCATIONS IN CONTINUOUS SPEECH UTTERANCES	1639
<i>Hassan Salami Kavaki, Michael I. Mandel</i>	
 <u>INTERSPEECH 2020 : FEATURE EXTRACTION AND DISTANT ASR</u>	
RAW SIGN AND MAGNITUDE SPECTRA FOR MULTI-HEAD ACOUSTIC MODELLING	1644
<i>Erfan Loweimi, Peter Bell, Steve Renals</i>	
ROBUST RAW WAVEFORM SPEECH RECOGNITION USING RELEVANCE WEIGHTED REPRESENTATIONS	1649
<i>Purvi Agrawal, Sriram Ganapathy</i>	
A DEEP 2D CONVOLUTIONAL NETWORK FOR WAVEFORM-BASED SPEECH RECOGNITION	1654
<i>Dino Oglic, Zoran Cvetkovic, Peter Bell, Steve Renals</i>	
LIGHTWEIGHT END-TO-END SPEECH RECOGNITION FROM RAW AUDIO DATA USING SINC-CONVOLUTIONS	1659
<i>Ludwig Kurzinger, Nicolas Lindae, Palle Klewitz, Gerhard Rigoll</i>	
AN ALTERNATIVE TO MFCCS FOR ASR.....	1664
<i>Pegah Ghahramani, Hossein Hadian, Daniel Povey, Hynek Hermansky, Sanjeev Khudanpur</i>	
PHASE BASED SPECTRO-TEMPORAL FEATURES FOR BUILDING A ROBUST ASR SYSTEM	1668
<i>Anirban Dutta, G. Ashishkumar, Ch. V. Rama Rao</i>	
DEEP SCATTERING POWER SPECTRUM FEATURES FOR ROBUST SPEECH RECOGNITION	1673
<i>Neethu M. Joy, Dino Oglic, Zoran Cvetkovic, Peter Bell, Steve Renals</i>	
FUSIONRNN: SHARED NEURAL PARAMETERS FOR MULTI-CHANNEL DISTANT SPEECH RECOGNITION	1678
<i>Titouan Parcollet, Xinchu Qiu, Nicholas D. Lane</i>	

BANDPASS NOISE GENERATION AND AUGMENTATION FOR UNIFIED ASR..... 1683
Kshitiz Kumar, Bo Ren, Yifan Gong, Jian Wu

DEEP LEARNING BASED DEREVERBERATION OF TEMPORAL ENVELOPES FOR
ROBUST SPEECH RECOGNITION 1688
Anurenjan Purushothaman, Anirudh Sreeram, Rohit Kumar, Sriram Ganapathy

INTERSPEECH 2020 : VOICE PRIVACY CHALLENGE

INTRODUCING THE VOICEPRIVACY INITIATIVE..... 1693
*N. Tomashenko, Brij Mohan Lal Srivastava, Xin Wang, Emmanuel Vincent, Andreas Nautsch,
Junichi Yamagishi, Nicholas Evans, Jose Patino, Jean-Francois Bonastre, Paul-Gauthier
Noe, Massimiliano Todisco*

THE PRIVACY ZEBRA: ZERO EVIDENCE BIOMETRIC RECOGNITION ASSESSMENT 1698
*Andreas Nautsch, Jose Patino, N. Tomashenko, Junichi Yamagishi, Paul-Gauthier Noe, Jean-
Francois Bonastre, Massimiliano Todisco, Nicholas Evans*

X-VECTOR SINGULAR VALUE MODIFICATION AND STATISTICAL-BASED
DECOMPOSITION WITH ENSEMBLE REGRESSION MODELING FOR SPEAKER
ANONYMIZATION SYSTEM 1703
Candy Olivia Mawalim, Kasorn Galajit, Jessada Karnjana, Masashi Unoki

A COMPARATIVE STUDY OF SPEECH ANONYMIZATION METRICS 1708
*Mohamed Maouche, Brij Mohan Lal Srivastava, Nathalie Vauquier, Aurelien Bellet, Marc
Tommasi, Emmanuel Vincent*

DESIGN CHOICES FOR X-VECTOR BASED SPEAKER ANONYMIZATION 1713
*Brij Mohan Lal Srivastava, N. Tomashenko, Xin Wang, Emmanuel Vincent, Junichi
Yamagishi, Mohamed Maouche, Aurelien Bellet, Marc Tommasi*

SPEECH PSEUDONYMISATION ASSESSMENT USING VOICE SIMILARITY MATRICES 1718
*Paul-Gauthier Noe, Jean-Francois Bonastre, Driss Matrouf, N. Tomashenko, Andreas
Nautsch, Nicholas Evans*

INTERSPEECH 2020 : SPEECH SYNTHESIS: TEXT PROCESSING, DATA AND EVALUATION

G2PM: A NEURAL GRAPHEME-TO-PHONEME CONVERSION PACKAGE FOR
MANDARIN CHINESE BASED ON A NEW OPEN BENCHMARK DATASET 1723
Kyubyong Park, Seanie Lee

A MASK-BASED MODEL FOR MANDARIN CHINESE POLYPHONIC DISAMBIGUATION 1728
Haiteng Zhang, Huashan Pan, Xiulin Li

PERCEPTION OF CONCATENATIVE VS. NEURAL TEXT-TO-SPEECH (TTS):
DIFFERENCES IN INTELLIGIBILITY IN NOISE AND LANGUAGE ATTITUDES 1733
Michelle Cohn, Georgia Zellou

ENHANCING SEQUENCE-TO-SEQUENCE TEXT-TO-SPEECH WITH MORPHOLOGY 1738
Jason Taylor, Korin Richmond

DEEP MOS PREDICTOR FOR SYNTHETIC SPEECH USING CLUSTER-BASED MODELING 1743
Yeunju Choi, Youngmoon Jung, Hoirin Kim

DEEP LEARNING BASED ASSESSMENT OF SYNTHETIC SPEECH NATURALNESS.....	1748
<i>Gabriel Mittag, Sebastian Moller</i>	
DISTANT SUPERVISION FOR POLYPHONE DISAMBIGUATION IN MANDARIN CHINESE.....	1753
<i>Jiawen Zhang, Yuanyuan Zhao, Jiaqi Zhu, Jinba Xiao</i>	
AN UNSUPERVISED METHOD TO SELECT A SPEAKER SUBSET FROM LARGE MULTI-SPEAKER SPEECH SYNTHESIS DATASETS.....	1758
<i>Pilar Oplustil Gallegos, Jennifer Williams, Joanna Rownicka, Simon King</i>	
UNDERSTANDING THE EFFECT OF VOICE QUALITY AND ACCENT ON TALKER SIMILARITY.....	1763
<i>Anurag Das, Guanlong Zhao, John Levis, Evgeny Chukharev-Hudilainen, Ricardo Gutierrez-Osuna</i>	

INTERSPEECH 2020 : SEARCH FOR SPEECH RECOGNITION

ROBUST BEAM SEARCH FOR ENCODER-DECODER ATTENTION BASED SPEECH RECOGNITION WITHOUT LENGTH BIAS.....	1768
<i>Wei Zhou, Ralf Schluter, Hermann Ney</i>	
TRANSFORMER WITH BIDIRECTIONAL DECODER FOR SPEECH RECOGNITION.....	1773
<i>Xi Chen, Songyang Zhang, Dandan Song, Peng Ouyang, Shouyi Yin</i>	
AN INVESTIGATION OF PHONE-BASED SUBWORD UNITS FOR END-TO-END SPEECH RECOGNITION.....	1778
<i>Weiran Wang, Guangsen Wang, Aadyot Bhatnagar, Yingbo Zhou, Caiming Xiong, Richard Socher</i>	
COMBINATION OF END-TO-END AND HYBRID MODELS FOR SPEECH RECOGNITION.....	1783
<i>Jeremy H. M. Wong, Yashesh Gaur, Rui Zhao, Liang Lu, Eric Sun, Jinyu Li, Yifan Gong</i>	
EVOLVED SPEECH-TRANSFORMER: APPLYING NEURAL ARCHITECTURE SEARCH TO END-TO-END AUTOMATIC SPEECH RECOGNITION.....	1788
<i>Jihwan Kim, Jisung Wang, Sangki Kim, Yeha Lee</i>	
HIERARCHICAL MULTI-STAGE WORD-TO-GRAPHEME NAMED ENTITY CORRECTOR FOR AUTOMATIC SPEECH RECOGNITION.....	1793
<i>Abhinav Garg, Ashutosh Gupta, Dhananjaya Gowda, Shatrughan Singh, Chanwoo Kim</i>	
LVCSR WITH TRANSFORMER LANGUAGE MODELS.....	1798
<i>Eugen Beck, Ralf Schluter, Hermann Ney</i>	
DARTS-ASR: DIFFERENTIABLE ARCHITECTURE SEARCH FOR MULTILINGUAL SPEECH RECOGNITION AND ADAPTATION.....	1803
<i>Yi-Chen Chen, Jui-Yang Hsu, Cheng-Kuang Lee, Hung-Yi Lee</i>	

INTERSPEECH 2020 : COMPUTATIONAL PARALINGUISTICS I

UNCERTAINTY-AWARE MACHINE SUPPORT FOR PAPER REVIEWING ON THE INTERSPEECH 2019 SUBMISSION CORPUS.....	1808
<i>Lukas Stappen, Georgios Rizos, Madina Hasan, Thomas Hain, Bjorn W. Schuller</i>	

INDIVIDUAL VARIATION IN LANGUAGE ATTITUDES TOWARD VOICE-AI: THE ROLE OF LISTENERS' AUTISTIC-LIKE TRAITS	1813
<i>Michelle Cohn, Melina Sarian, Kristin Predeck, Georgia Zellou</i>	
DIFFERENCES IN GRADIENT EMOTION PERCEPTION: HUMAN VS. ALEXA VOICES	1818
<i>Michelle Cohn, Eran Raveh, Kristin Predeck, Iona Gessinger, Bernd Mobius, Georgia Zellou</i>	
THE MSP-CONVERSATION CORPUS.....	1823
<i>Luz Martinez-Lucas, Mohammed Abdelwahab, Carlos Busso</i>	
SPOTTING THE TRACES OF DEPRESSION IN READ SPEECH: AN APPROACH BASED ON COMPUTATIONAL PARALINGUISTICS AND SOCIAL SIGNAL PROCESSING	1828
<i>Fuxiang Tao, Anna Esposito, Alessandro Vinciarelli</i>	
SPEECH SENTIMENT AND CUSTOMER SATISFACTION ESTIMATION IN SOCIALBOT CONVERSATIONS.....	1833
<i>Yelin Kim, Joshua Levy, Yang Liu</i>	
PARDON THE INTERRUPTION: AN ANALYSIS OF GENDER AND TURN-TAKING IN U.S. SUPREME COURT ORAL ARGUMENTS.....	1838
<i>Haley Lepp, Gina-Anne Levow</i>	
ARE GERMANS BETTER HATERS THAN DANES? LANGUAGE-SPECIFIC IMPLICIT PROSODIES OF TYPES OF HATE SPEECH AND HOW THEY RELATE TO PERCEIVED SEVERITY AND SOCIETAL RULES	1843
<i>Jana Neitsch, Oliver Niebuhr</i>	
AN OBJECTIVE VOICE GENDER SCORING SYSTEM AND IDENTIFICATION OF THE SALIENT ACOUSTIC MEASURES	1848
<i>Fuling Chen, Roberto Togneri, Murray Maybery, Diana Tan</i>	
HOW ORDINAL ARE YOUR DATA?.....	1853
<i>Sadari Jayawardena, Julien Epps, Zhaocheng Huang</i>	
<u>INTERSPEECH 2020 : ACOUSTIC PHONETICS AND PROSODY</u>	
CORRELATING CEPSTRA WITH FORMANT FREQUENCIES: IMPLICATIONS FOR PHONETICALLY-INFORMED FORENSIC VOICE COMPARISON	1858
<i>Vincent Hughes, Frantz Clermont, Philip Harrison</i>	
PROSODY AND BREATHING: A COMPARISON BETWEEN RHETORICAL AND INFORMATION-SEEKING QUESTIONS IN GERMAN AND BRAZILIAN PORTUGUESE.....	1863
<i>Jana Neitsch, Plinio A. Barbosa, Oliver Niebuhr</i>	
SCALING PROCESSES OF CLAUSE CHAINS IN PITJANTJATJARA.....	1868
<i>Rebecca Defina, Catalina Torres, Hywel Stoakes</i>	
NEUTRALIZATION OF VOICING DISTINCTION OF STOPS IN TOHOKU DIALECTS OF JAPANESE: FIELD WORK AND ACOUSTIC MEASUREMENTS	1873
<i>Ai Mizoguchi, Ayako Hashimoto, Sanae Matsui, Setsuko Imatomi, Ryunosuke Kobayashi, Mafuyu Kitahara</i>	
CORRELATION BETWEEN PROSODY AND PRAGMATICS: CASE STUDY OF DISCOURSE MARKERS IN FRENCH AND ENGLISH	1878
<i>Lou Lee, Denis Jouviet, Katarina Bartkova, Yvon Keromnes, Mathilde Dargnat</i>	

AN ANALYSIS OF PROSODIC PROMINENCE CUES TO INFORMATION STRUCTURE IN EGYPTIAN ARABIC	1883
<i>Dina El Zarka, Anneliese Kelterer, Barbara Schuppler</i>	
LEXICAL STRESS IN URDU	1888
<i>Benazir Mumtaz, Tina Bogel, Miriam Butt</i>	
VOCAL MARKERS FROM SUSTAINED PHONATION IN HUNTINGTON'S DISEASE.....	1893
<i>Rachid Riad, Hadrien Titeux, Laurie Lemoine, Justine Montillot, Jennifer Hamet Bagnou, Xuan-Nga Cao, Emmanuel Dupoux, Anne-Catherine Bachoud-Levi</i>	
HOW RHYTHM AND TIMBRE ENCODE MOORE LANGUAGE IN BENDRE DRUMMED SPEECH	1898
<i>Laure Dentel, Julien Meyer</i>	

INTERSPEECH 2020 : TONAL ASPECTS OF ACOUSTIC PHONETICS AND PROSODY

INTERACTION OF TONE AND VOICING IN MIZO.....	1903
<i>Wendy Lalhminghlu, Priyankoo Sarmah</i>	
MANDARIN LEXICAL TONES: A CORPUS-BASED STUDY OF WORD LENGTH, SYLLABLE POSITION AND PROSODIC POSITION ON DURATION.....	1908
<i>Yaru Wu, Martine Adda-Decker, Lori Lamel</i>	
AN INVESTIGATION OF THE TARGET APPROXIMATION MODEL FOR TONE MODELING AND RECOGNITION IN CONTINUOUS MANDARIN SPEECH	1913
<i>Yingming Gao, Xinyu Zhang, Yi Xu, Jinsong Zhang, Peter Birkholz</i>	
INTEGRATING THE APPLICATION AND REALIZATION OF MANDARIN 3RD TONE SANDHI IN THE RESOLUTION OF SENTENCE AMBIGUITY	1918
<i>Wei Lai, Aini Li</i>	
NEUTRAL TONE IN CHANGDE MANDARIN	1923
<i>Zhenrui Zhang, Fang Hu</i>	
PITCH DECLINATION AND FINAL LOWERING IN NORTHEASTERN MANDARIN	1928
<i>Ping Cui, Jianjing Kuang</i>	
VARIATION IN SPECTRAL SLOPE AND INTERHARMONIC NOISE IN CANTONESE TONES	1933
<i>Phil Rose</i>	
THE ACOUSTIC REALIZATION OF MANDARIN TONES IN FAST SPEECH.....	1938
<i>Ping Tang, Shanpeng Li</i>	

INTERSPEECH 2020 : SPEECH CLASSIFICATION

DO FACE MASKS INTRODUCE BIAS IN SPEECH TECHNOLOGIES? THE CASE OF AUTOMATED SCORING OF SPEAKING PROFICIENCY.....	1942
<i>Anastassia Loukina, Keelan Evanini, Matthew Mulholland, Ian Blood, Klaus Zechner</i>	
A LOW LATENCY ASR-FREE END TO END SPOKEN LANGUAGE UNDERSTANDING SYSTEM	1947
<i>Mohamed Mhiri, Samuel Myer, Vikrant Singh Tomar</i>	

AN AUDIO-BASED WAKEWORD-INDEPENDENT VERIFICATION SYSTEM.....	1952
<i>Joe Wang, Rajath Kumar, Mike Rodehorst, Brian Kulis, Shiv Naga Prasad Vitaladevuni</i>	
LEARNABLE SPECTRO-TEMPORAL RECEPTIVE FIELDS FOR ROBUST VOICE TYPE DISCRIMINATION.....	1957
<i>Tyler Vuong, Yangyang Xia, Richard M. Stern</i>	
LOW LATENCY SPEECH RECOGNITION USING END-TO-END PREFETCHING	1962
<i>Shuo-Yiin Chang, Bo Li, David Rybach, Yanzhang He, Wei Li, Tara N. Sainath, Trevor Strohman</i>	
AUTOSPEECH 2020: THE SECOND AUTOMATED MACHINE LEARNING CHALLENGE FOR SPEECH CLASSIFICATION	1967
<i>Jingsong Wang, Tom Ko, Zhen Xu, Xiawei Guo, Souxiang Liu, Wei-Wei Tu, Lei Xie</i>	
BUILDING A ROBUST WORD-LEVEL WAKEWORD VERIFICATION NETWORK	1972
<i>Rajath Kumar, Mike Rodehorst, Joe Wang, Jiacheng Gu, Brian Kulis</i>	
A TRANSFORMER-BASED AUDIO CAPTIONING MODEL WITH KEYWORD ESTIMATION.....	1977
<i>Yuma Koizumi, Ryo Masumura, Kyosuke Nishida, Masahiro Yasuda, Shoichiro Saito</i>	
NEURAL ARCHITECTURE SEARCH FOR KEYWORD SPOTTING.....	1982
<i>Tong Mo, Yakun Yu, Mohammad Salameh, Di Niu, Shangling Jui</i>	
SMALL-FOOTPRINT KEYWORD SPOTTING WITH MULTI-SCALE TEMPORAL CONVOLUTION	1987
<i>Ximin Li, Xiaodong Wei, Xiaowei Qin</i>	

INTERSPEECH 2020 : SPEECH SYNTHESIS PARADIGMS AND METHODS I

USING CYCLIC NOISE AS THE SOURCE SIGNAL FOR NEURAL SOURCE-FILTER- BASED SPEECH WAVEFORM MODEL.....	1992
<i>Xin Wang, Junichi Yamagishi</i>	
UNCONDITIONAL AUDIO GENERATION WITH GENERATIVE ADVERSARIAL NETWORKS AND CYCLE REGULARIZATION	1997
<i>Jen-Yu Liu, Yu-Hua Chen, Yin-Cheng Yeh, Yi-Hsuan Yang</i>	
COMPLEX-VALUED VARIATIONAL AUTOENCODER: A NOVEL DEEP GENERATIVE MODEL FOR DIRECT REPRESENTATION OF COMPLEX SPECTRA.....	2002
<i>Toru Nakashika</i>	
ATTENTRON: FEW-SHOT TEXT-TO-SPEECH UTILIZING ATTENTION-BASED VARIABLE-LENGTH EMBEDDING.....	2007
<i>Seungwoo Choi, Seungju Han, Dongyoung Kim, Sungjoo Ha</i>	
REFORMER-TTS: NEURAL SPEECH SYNTHESIS WITH REFORMER NETWORK	2012
<i>Hyeong Rae Ihm, Joun Yeop Lee, Byoung Jin Choi, Sung Jun Cheon, Nam Soo Kim</i>	
CYCLEGAN-VC3: EXAMINING AND IMPROVING CYCLEGAN-VCS FOR MEL- SPECTROGRAM CONVERSION.....	2017
<i>Takuhiro Kaneko, Hirokazu Kameoka, Kou Tanaka, Nobukatsu Hojo</i>	

HIGH QUALITY STREAMING SPEECH SYNTHESIS WITH LOW, SENTENCE-LENGTH-INDEPENDENT LATENCY.....	2022
<i>Nikolaos Ellinas, Georgios Vamvoukakis, Konstantinos Markopoulos, Aimilios Chalamandaris, Georgia Maniati, Panos Kakoulidis, Spyros Raptis, June Sig Sung, Hyoungmin Park, Pirros Tsiakoulis</i>	
DURIAN: DURATION INFORMED ATTENTION NETWORK FOR SPEECH SYNTHESIS.....	2027
<i>Chengzhu Yu, Heng Lu, Na Hu, Meng Yu, Chao Weng, Kun Xu, Peng Liu, Deyi Tuo, Shiyin Kang, Guangzhi Lei, Dan Su, Dong Yu</i>	
MULTI-SPEAKER TEXT-TO-SPEECH SYNTHESIS USING DEEP GAUSSIAN PROCESSES	2032
<i>Kentaro Mitsui, Tomoki Koriyama, Hiroshi Saruwatari</i>	
A HYBRID HMM-WAVEGLOW BASED TEXT-TO-SPEECH SYNTHESIZER USING HISTOGRAM EQUALIZATION FOR LOW RESOURCE INDIAN LANGUAGES.....	2037
<i>Mano Ranjith Kumar M., Sudhanshu Srivastava, Anusha Prakash, Hema A. Murthy</i>	

INTERSPEECH 2020 : THE INTERSPEECH 2020 COMPUTATIONAL PARALINGUISTICS CHALLENGE (COMPARE)

THE INTERSPEECH 2020 COMPUTATIONAL PARALINGUISTICS CHALLENGE: ELDERLY EMOTION, BREATHING & MASKS.....	2042
<i>Bjorn W. Schuller, Anton Batliner, Christian Bergler, Eva-Maria Messner, Antonia Hamilton, Shahin Amiriparian, Alice Baird, Georgios Rizos, Maximilian Schmitt, Lukas Stappen, Harald Baumeister, Alexis Deighton Macintyre, Simone Hantke</i>	
LEARNING HIGHER REPRESENTATIONS FROM PRE-TRAINED DEEP MODELS WITH DATA AUGMENTATION FOR THE COMPARE 2020 CHALLENGE MASK TASK	2047
<i>Tomoya Koike, Kun Qian, Bjorn W. Schuller, Yoshiharu Yamamoto</i>	
SURGICAL MASK DETECTION WITH CONVOLUTIONAL NEURAL NETWORKS AND DATA AUGMENTATIONS ON SPECTROGRAMS	2052
<i>Steffen Illium, Robert Muller, Andreas Sedlmeier, Claudia Linnhoff-Popien</i>	
SURGICAL MASK DETECTION WITH DEEP RECURRENT PHONETIC MODELS	2057
<i>Philipp Klumpp, Tomas Arias-Vergara, Juan Camilo Vasquez-Correa, Paula Andrea Perez-Toro, Florian Honig, Elmar Noth, Juan Rafael Orozco-Arroyave</i>	
PHONETIC, FRAME CLUSTERING AND INTELLIGIBILITY ANALYSES FOR THE INTERSPEECH 2020 COMPARE CHALLENGE	2062
<i>Claude Montacie, Marie-Jose Caraty</i>	
EXPLORING TEXT AND AUDIO EMBEDDINGS FOR MULTI-DIMENSION ELDERLY EMOTION RECOGNITION	2067
<i>Mariana Juliao, Alberto Abad, Helena Moniz</i>	
ENSEMBLING END-TO-END DEEP MODELS FOR COMPUTATIONAL PARALINGUISTICS TASKS: COMPARE 2020 MASK AND BREATHING SUB-CHALLENGES.....	2072
<i>Maxim Markitantov, Denis Dresvyanskiy, Danila Mamontov, Heysem Kaya, Wolfgang Minker, Alexey Karpov</i>	
ANALYZING BREATH SIGNALS FOR THE INTERSPEECH 2020 COMPARE CHALLENGE.....	2077
<i>John Mendonca, Francisco Teixeira, Isabel Trancoso, Alberto Abad</i>	

DEEP ATTENTIVE END-TO-END CONTINUOUS BREATH SENSING FROM SPEECH	2082
<i>Alexis Deighton Macintyre, Georgios Rizos, Anton Batliner, Alice Baird, Shahin Amiriparian, Antonia Hamilton, Bjorn W. Schuller</i>	
PARALINGUISTIC CLASSIFICATION OF MASK WEARING BY IMAGE CLASSIFIERS AND FUSION	2087
<i>Jeno Szep, Salim Hariri</i>	
EXPLORATION OF ACOUSTIC AND LEXICAL CUES FOR THE INTERSPEECH 2020 COMPUTATIONAL PARALINGUISTIC CHALLENGE	2092
<i>Ziqing Yang, Zifan An, Zehao Fan, Chengye Jing, Houwei Cao</i>	
IS EVERYTHING FINE, GRANDMA? ACOUSTIC AND LINGUISTIC MODELING FOR ROBUST ELDERLY SPEECH EMOTION RECOGNITION	2097
<i>Gizem Sogancioglu, Oxana Verkholyak, Heysem Kaya, Dmitrii Fedotov, Tobias Cadee, Albert Ali Salah, Alexey Karpov</i>	
ARE YOU WEARING A MASK? IMPROVING MASK DETECTION FROM SPEECH USING AUGMENTATION BY CYCLE-CONSISTENT GANS	2102
<i>Nicolae-Catalin Ristea, Radu Tudor Ionescu</i>	
<u>INTERSPEECH 2020 : STREAMING ASR</u>	
1-D ROW-CONVOLUTION LSTM: FAST STREAMING ASR AT ACCURACY PARITY WITH LC-BLSTM	2107
<i>Kshitiz Kumar, Chaojun Liu, Yifan Gong, Jian Wu</i>	
LOW LATENCY END-TO-END STREAMING SPEECH RECOGNITION WITH A SCOUT NETWORK	2112
<i>Chengyi Wang, Yu Wu, Liang Lu, Shujie Liu, Jinyu Li, Guoli Ye, Ming Zhou</i>	
KNOWLEDGE DISTILLATION FROM OFFLINE TO STREAMING RNN TRANSDUCER FOR END-TO-END SPEECH RECOGNITION	2117
<i>Gakuto Kurata, George Saon</i>	
PARALLEL RESCORING WITH TRANSFORMER FOR STREAMING ON-DEVICE SPEECH RECOGNITION	2122
<i>Wei Li, James Qin, Chung-Cheng Chiu, Ruoming Pang, Yanzhang He</i>	
IMPROVED HYBRID STREAMING ASR WITH TRANSFORMER LANGUAGE MODELS	2127
<i>Pau Baquero-Arnal, Javier Jorge, Adria Gimenez, Joan Albert Silvestre-Cerda, Javier Iranzo-Sanchez, Albert Sanchis, Jorge Civera, Alfons Juan</i>	
STREAMING TRANSFORMER-BASED ACOUSTIC MODELS USING SELF-ATTENTION WITH AUGMENTED MEMORY	2132
<i>Chunyang Wu, Yongqiang Wang, Yangyang Shi, Ching-Feng Yeh, Frank Zhang</i>	
ENHANCING MONOTONIC MULTIHEAD ATTENTION FOR STREAMING ASR	2137
<i>Hirofumi Inaguma, Masato Mimura, Tatsuya Kawahara</i>	
STREAMING CHUNK-AWARE MULTIHEAD ATTENTION FOR ONLINE END-TO-END SPEECH RECOGNITION	2142
<i>Shiliang Zhang, Zhifu Gao, Haoneng Luo, Ming Lei, Jie Gao, Zhijie Yan, Lei Xie</i>	

HIGH PERFORMANCE SEQUENCE-TO-SEQUENCE MODEL FOR STREAMING SPEECH RECOGNITION.....	2147
<i>Thai-Son Nguyen, Ngoc-Quan Pham, Sebastian Stuker, Alex Waibel</i>	

TRANSFER LEARNING APPROACHES FOR STREAMING END-TO-END SPEECH RECOGNITION SYSTEM.....	2152
<i>Vikas Joshi, Rui Zhao, Rupesh R. Mehta, Kshitiz Kumar, Jinyu Li</i>	

**INTERSPEECH 2020 : ALZHEIMER'S DEMENTIA RECOGNITION THROUGH
SPONTANEOUS SPEECH**

TACKLING THE ADDRESS CHALLENGE: A MULTIMODAL APPROACH TO THE AUTOMATED RECOGNITION OF ALZHEIMER'S DEMENTIA.....	2157
<i>Matej Martinc, Senja Pollak</i>	

DISFLUENCIES AND FINE-TUNING PRE-TRAINED LANGUAGE MODELS FOR DETECTION OF ALZHEIMER'S DISEASE.....	2162
<i>Jiahong Yuan, Yuchen Bian, Xingyu Cai, Jiayi Huang, Zheng Ye, Kenneth Church</i>	

TO BERT OR NOT TO BERT: COMPARING SPEECH AND LANGUAGE-BASED APPROACHES FOR ALZHEIMER'S DISEASE DETECTION.....	2167
<i>Aparna Balagopalan, Benjamin Eyre, Frank Rudzicz, Jekaterina Novikova</i>	

ALZHEIMER'S DEMENTIA RECOGNITION THROUGH SPONTANEOUS SPEECH: THE ADDRESS CHALLENGE.....	2172
<i>Saturnino Luz, Fasih Haider, Sofia De La Fuente, Davida Fromm, Brian Macwhinney</i>	

USING STATE OF THE ART SPEAKER RECOGNITION AND NATURAL LANGUAGE PROCESSING TECHNOLOGIES TO DETECT ALZHEIMER'S DISEASE AND ASSESS ITS SEVERITY.....	2177
<i>Raghavendra Pappagari, Jaejin Cho, Laureano Moro-Velazquez, Najim Dehak</i>	

VOLUME 4

A COMPARISON OF ACOUSTIC AND LINGUISTICS METHODOLOGIES FOR ALZHEIMER'S DEMENTIA RECOGNITION.....	2182
<i>Nicholas Cummins, Yilin Pan, Zhao Ren, Julian Fritsch, Venkata Srikanth Nallanthighal, Heidi Christensen, Daniel Blackburn, Bjorn W. Schuller, Mathew Magimai-Doss, Helmer Strik, Aki Harma</i>	

MULTI-MODAL FUSION WITH GATING USING AUDIO, LEXICAL AND DISFLUENCY FEATURES FOR ALZHEIMER'S DEMENTIA RECOGNITION FROM SPONTANEOUS SPEECH.....	2187
<i>Morteza Rohanian, Julian Hough, Matthew Purver</i>	

COMPARING NATURAL LANGUAGE PROCESSING TECHNIQUES FOR ALZHEIMER'S DEMENTIA PREDICTION IN SPONTANEOUS SPEECH.....	2192
<i>Thomas Searle, Zina Ibrahim, Richard Dobson</i>	

MULTISCALE SYSTEM FOR ALZHEIMER'S DEMENTIA RECOGNITION THROUGH SPONTANEOUS SPEECH.....	2197
<i>Erik Edwards, Charles Dognin, Bajibabu Bollepalli, Maneesh Singh</i>	

THE INESC-ID MULTI-MODAL SYSTEM FOR THE ADDRESS 2020 CHALLENGE.....	2202
<i>Anna Pompili, Thomas Rolland, Alberto Abad</i>	

EXPLORING MMSE SCORE PREDICTION USING VERBAL AND NON-VERBAL CUES	2207
<i>Shahla Farzana, Natalie Parde</i>	
MULTIMODAL INDUCTIVE TRANSFER LEARNING FOR DETECTION OF ALZHEIMER'S DEMENTIA AND ITS SEVERITY	2212
<i>Utkarsh Sarawgi, Wazeer Zulfikar, Nouran Soliman, Pattie Maes</i>	
EXPLOITING MULTI-MODAL FEATURES FROM PRE-TRAINED NETWORKS FOR ALZHEIMER'S DEMENTIA RECOGNITION	2217
<i>Junghyun Koo, Jie Hwan Lee, Jaewoo Pyo, Yujin Jo, Kyogu Lee</i>	
AUTOMATED SCREENING FOR ALZHEIMER'S DEMENTIA THROUGH SPONTANEOUS SPEECH	2222
<i>Muhammad Shehram Shah Syed, Zafi Sherhan Syed, Margaret Lech, Elena Pirogova</i>	
<u>INTERSPEECH 2020 : SPEAKER RECOGNITION CHALLENGES AND APPLICATIONS</u>	
NEC-TT SPEAKER VERIFICATION SYSTEM FOR SRE'19 CTS CHALLENGE	2227
<i>Kong Aik Lee, Koji Okabe, Hitoshi Yamamoto, Qiongqiong Wang, Ling Guo, Takafumi Koshinaka, Jiacen Zhang, Keisuke Ishikawa, Koichi Shinoda</i>	
THUEE SYSTEM FOR NIST SRE19 CTS CHALLENGE	2232
<i>Ruyun Li, Tianyu Liang, Dandan Song, Yi Liu, Yangcheng Wu, Can Xu, Peng Ouyang, Xianwei Zhang, Xianhong Chen, Wei-Qiang Zhang, Shouyi Yin, Liang He</i>	
AUTOMATIC QUALITY ASSESSMENT FOR AUDIO-VISUAL VERIFICATION SYSTEMS. THE LOVE SUBMISSION TO NIST SRE CHALLENGE 2019	2237
<i>Grigory Antipov, Nicolas Gengembre, Olivier Le Blouch, Gael Le Lan</i>	
AUDIO-VISUAL SPEAKER RECOGNITION WITH A CROSS-MODAL DISCRIMINATIVE NETWORK	2242
<i>Ruijie Tao, Rohan Kumar Das, Haizhou Li</i>	
MULTIMODAL ASSOCIATION FOR SPEAKER VERIFICATION	2247
<i>Suwon Shon, James Glass</i>	
MULTI-MODALITY MATTERS: A PERFORMANCE LEAP ON VOXCELEB	2252
<i>Zhengyang Chen, Shuai Wang, Yanmin Qian</i>	
CROSS-DOMAIN ADAPTATION WITH DISCREPANCY MINIMIZATION FOR TEXT- INDEPENDENT FORENSIC SPEAKER VERIFICATION	2257
<i>Zhenyu Wang, Wei Xia, John H. L. Hansen</i>	
OPEN-SET SHORT UTTERANCE FORENSIC SPEAKER VERIFICATION USING TEACHER-STUDENT NETWORK WITH EXPLICIT INDUCTIVE BIAS	2262
<i>Mufan Sang, Wei Xia, John H. L. Hansen</i>	
JUKEBOX: A MULTILINGUAL SINGER RECOGNITION DATASET	2267
<i>Anurag Chowdhury, Austin Cozzo, Arun Ross</i>	
SPEAKER IDENTIFICATION FOR HOUSEHOLD SCENARIOS WITH SELF-ATTENTION AND ADVERSARIAL TRAINING	2272
<i>Ruirui Li, Jyun-Yu Jiang, Xian Wu, Chu-Cheng Hsieh, Andreas Stolcke</i>	

INTERSPEECH 2020 : APPLICATIONS OF ASR

STREAMING KEYWORD SPOTTING ON MOBILE DEVICES.....	2277
<i>Oleg Rybakov, Natasha Kononenko, Niranjana Subrahmanya, Mirko Visontai, Stella Laurenzo</i>	
METADATA-AWARE END-TO-END KEYWORD SPOTTING.....	2282
<i>Hongyi Liu, Apurva Abhyankar, Yuriy Mishchenko, Thibaud Senechal, Gengshen Fu, Brian Kulis, Noah D. Stein, Anish Shah, Shiv Naga Prasad Vitaladevuni</i>	
ADVERSARIAL AUDIO: A NEW INFORMATION HIDING METHOD	2287
<i>Yehao Kong, Jiliang Zhang</i>	
S2IGAN: SPEECH-TO-IMAGE GENERATION VIA ADVERSARIAL LEARNING	2292
<i>Xinsheng Wang, Tingting Qiao, Jihua Zhu, Alan Hanjalic, Odette Scharenborg</i>	
AUTOMATIC SPEECH RECOGNITION BENCHMARK FOR AIR-TRAFFIC COMMUNICATIONS	2297
<i>Juan Zuluaga-Gomez, Petr Motlicek, Qingran Zhan, Karel Vesely, Rudolf Braun</i>	
WHISPER AUGMENTED END-TO-END/HYBRID SPEECH RECOGNITION SYSTEM --- CYCLEGAN APPROACH	2302
<i>Prithvi R. R. Gudepu, Gowtham P. Vadiseti, Abhishek Niranjana, Kinnera Saranu, Raghava Sarma, M. Ali Basha Shaik, Periyasamy Paramasivam</i>	
RISK FORECASTING FROM EARNINGS CALLS ACOUSTICS AND NETWORK CORRELATIONS.....	2307
<i>Ramit Sawhney, Arshiya Aggarwal, Piyush Khanna, Puneet Mathur, Taru Jain, Rajiv Ratn Shah</i>	
SPECMARK: A SPECTRAL WATERMARKING FRAMEWORK FOR IP PROTECTION OF SPEECH RECOGNITION SYSTEMS	2312
<i>Huili Chen, Bitu Darvish, Farinaz Koushanfar</i>	
EVALUATING AUTOMATICALLY GENERATED PHONEME CAPTIONS FOR IMAGES	2317
<i>Justin Van Der Hout, Zoltan D'Haese, Mark Hasegawa-Johnson, Odette Scharenborg</i>	

INTERSPEECH 2020 : SPEECH EMOTION RECOGNITION II

AN EFFICIENT TEMPORAL MODELING APPROACH FOR SPEECH EMOTION RECOGNITION BY MAPPING VARIED DURATION SENTENCES INTO FIXED NUMBER OF CHUNKS	2322
<i>Wei-Cheng Lin, Carlos Busso</i>	
DEEP ARCHITECTURE ENHANCING ROBUSTNESS TO NOISE, ADVERSARIAL ATTACKS, AND CROSS-CORPUS SETTING FOR SPEECH EMOTION RECOGNITION	2327
<i>Siddique Latif, Rajib Rana, Sara Khalifa, Raja Jurdak, Bjorn W. Schuller</i>	
META-LEARNING FOR SPEECH EMOTION RECOGNITION CONSIDERING AMBIGUITY OF EMOTION LABELS	2332
<i>Takuya Fujioka, Takeshi Homma, Kenji Nagamatsu</i>	
TEMPORAL ATTENTION CONVOLUTIONAL NETWORK FOR SPEECH EMOTION RECOGNITION WITH LATENT REPRESENTATION	2337
<i>Jiaxing Liu, Zhilei Liu, Longbiao Wang, Yuan Gao, Lili Guo, Jianwu Dang</i>	

RECONCILIATION OF MULTIPLE CORPORA FOR SPEECH EMOTION RECOGNITION BY MULTIPLE CLASSIFIERS WITH AN ADVERSARIAL CORPUS DISCRIMINATOR.....	2342
<i>Zhi Zhu, Yoshinao Sato</i>	
CONVERSATIONAL EMOTION RECOGNITION USING SELF-ATTENTION MECHANISMS AND GRAPH NEURAL NETWORKS.....	2347
<i>Zheng Lian, Jianhua Tao, Bin Liu, Jian Huang, Zhanlei Yang, Rongjun Li</i>	
EIGENEMO: SPECTRAL UTTERANCE REPRESENTATION USING DYNAMIC MODE DECOMPOSITION FOR SPEECH EMOTION CLASSIFICATION	2352
<i>Shuiyang Mao, P. C. Ching, Tan Lee</i>	
ADVANCING MULTIPLE INSTANCE LEARNING WITH ATTENTION MODELING FOR CATEGORICAL SPEECH EMOTION RECOGNITION.....	2357
<i>Shuiyang Mao, P. C. Ching, C.-C. Jay Kuo, Tan Lee</i>	
 <u>INTERSPEECH 2020 : BI- AND MULTILINGUALITY</u>	
THE EFFECT OF LANGUAGE PROFICIENCY ON THE PERCEPTION OF SEGMENTAL FOREIGN ACCENT	2362
<i>Ruben Perez-Ramon, Maria Luisa Garcia Lecumberri, Martin Cooke</i>	
THE EFFECT OF LANGUAGE DOMINANCE ON THE SELECTIVE ATTENTION OF SEGMENTS AND TONES IN URDU-CANTONESE SPEAKERS	2367
<i>Yi Liu, Jinghong Ning</i>	
THE EFFECT OF INPUT ON THE PRODUCTION OF ENGLISH TENSE AND LAX VOWELS BY CHINESE LEARNERS: EVIDENCE FROM AN ELEMENTARY SCHOOL IN CHINA.....	2372
<i>Mengrou Li, Ying Chen, Jie Cui</i>	
EXPLORING THE USE OF AN ARTIFICIAL ACCENT OF ENGLISH TO ASSESS PHONETIC LEARNING IN MONOLINGUAL AND BILINGUAL SPEAKERS.....	2377
<i>Laura Spinu, Jiwon Hwang, Nadya Pincus, Mariana Vasilita</i>	
EFFECTS OF DIALECTAL CODE-SWITCHING ON SPEECH MODULES: A STUDY USING EGYPTIAN ARABIC BROADCAST SPEECH	2382
<i>Shammur A. Chowdhury, Younes Samih, Mohamed Eldesouki, Ahmed Ali</i>	
BILINGUAL ACOUSTIC VOICE VARIATION IS SIMILARLY STRUCTURED ACROSS LANGUAGES.....	2387
<i>Khia A. Johnson, Molly Babel, Robert A. Fuhrman</i>	
MONOLINGUAL DATA SELECTION ANALYSIS FOR ENGLISH-MANDARIN HYBRID CODE-SWITCHING SPEECH RECOGNITION.....	2392
<i>Haobo Zhang, Haihua Xu, Van Tung Pham, Hao Huang, Eng Siong Chng</i>	
PERCEPTION AND PRODUCTION OF MANDARIN INITIAL STOPS BY NATIVE URDU SPEAKERS	2397
<i>Dan Du, Xianjin Zhu, Zhu Li, Jinsong Zhang</i>	
NOW YOU'RE SPEAKING MY LANGUAGE: VISUAL LANGUAGE IDENTIFICATION.....	2402
<i>Triantafyllos Afouras, Joon Son Chung, Andrew Zisserman</i>	
THE DIFFERENT ENHANCEMENT ROLES OF COVARYING CUES IN THAI AND MANDARIN TONES	2407
<i>Nari Rhee, Jianjing Kuang</i>	

INTERSPEECH 2020 : SINGLE-CHANNEL SPEECH ENHANCEMENT I

SINGING VOICE EXTRACTION WITH ATTENTION-BASED SPECTROGRAMS FUSION	2412
<i>Hao Shi, Longbiao Wang, Sheng Li, Chenchen Ding, Meng Ge, Nan Li, Jianwu Dang, Hiroshi Seki</i>	
INCORPORATING BROAD PHONETIC INFORMATION FOR SPEECH ENHANCEMENT	2417
<i>Yen-Ju Lu, Chien-Feng Liao, Xugang Lu, Jieih-Weih Hung, Yu Tsao</i>	
A RECURSIVE NETWORK WITH DYNAMIC ATTENTION FOR MONAURAL SPEECH ENHANCEMENT.....	2422
<i>Andong Li, Chengshi Zheng, Cunhang Fan, Renhua Peng, Xiaodong Li</i>	
CONSTRAINED RATIO MASK FOR SPEECH ENHANCEMENT USING DNN	2427
<i>Hongjiang Yu, Wei-Ping Zhu, Yuhong Yang</i>	
SERIL: NOISE ADAPTIVE SPEECH ENHANCEMENT USING REGULARIZATION-BASED INCREMENTAL LEARNING	2432
<i>Chi-Chang Lee, Yu-Chen Lin, Hsuan-Tien Lin, Hsin-Min Wang, Yu Tsao</i>	
ADAPTIVE NEURAL SPEECH ENHANCEMENT WITH A DENOISING VARIATIONAL AUTOENCODER	2437
<i>Yoshiaki Bando, Kouhei Sekiguchi, Kazuyoshi Yoshii</i>	
LOW-LATENCY SINGLE CHANNEL SPEECH DEREVERBERATION USING U-NET CONVOLUTIONAL NEURAL NETWORKS.....	2442
<i>Ahmet E. Bulut, Kazuhito Koishida</i>	
SINGLE-CHANNEL SPEECH ENHANCEMENT BY SUBSPACE AFFINITY MINIMIZATION	2447
<i>Dung N. Tran, Kazuhito Koishida</i>	
NOISE TOKENS: LEARNING NEURAL NOISE TEMPLATES FOR ENVIRONMENT-AWARE SPEECH ENHANCEMENT	2452
<i>Haoyu Li, Junichi Yamagishi</i>	
NAAGN: NOISE-AWARE ATTENTION-GATED NETWORK FOR SPEECH ENHANCEMENT.....	2457
<i>Feng Deng, Tao Jiang, Xiao-Rui Wang, Chen Zhang, Yan Li</i>	

INTERSPEECH 2020 : DEEP NOISE SUPPRESSION CHALLENGE

ONLINE MONAURAL SPEECH ENHANCEMENT USING DELAYED SUBBAND LSTM.....	2462
<i>Xiaofei Li, Radu Horaud</i>	
INTERSPEECH 2020 DEEP NOISE SUPPRESSION CHALLENGE: A FULLY CONVOLUTIONAL RECURRENT NETWORK (FCRN) FOR JOINT DEREVERBERATION AND DENOISING.....	2467
<i>Maximilian Strake, Bruno Defraene, Kristoff Fluyt, Wouter Tirry, Tim Fingscheidt</i>	
DCCRN: DEEP COMPLEX CONVOLUTION RECURRENT NETWORK FOR PHASE-AWARE SPEECH ENHANCEMENT	2472
<i>Yanxin Hu, Yun Liu, Shubo Lv, Mengtao Xing, Shimin Zhang, Yihui Fu, Jian Wu, Bihong Zhang, Lei Xie</i>	

DUAL-SIGNAL TRANSFORMATION LSTM NETWORK FOR REAL-TIME NOISE SUPPRESSION.....	2477
<i>Nils L. Westhausen, Bernd T. Meyer</i>	
A PERCEPTUALLY-MOTIVATED APPROACH FOR LOW-COMPLEXITY, REAL-TIME ENHANCEMENT OF FULLBAND SPEECH.....	2482
<i>Jean-Marc Valin, Umut Isik, Neerad Phansalkar, Ritwik Giri, Karim Helwani, Arvindh Krishnaswamy</i>	
POCONET: BETTER SPEECH ENHANCEMENT WITH FREQUENCY-POSITIONAL EMBEDDINGS, SEMI-SUPERVISED CONVERSATIONAL DATA, AND BIASED LOSS	2487
<i>Umut Isik, Ritwik Giri, Neerad Phansalkar, Jean-Marc Valin, Karim Helwani, Arvindh Krishnaswamy</i>	
THE INTERSPEECH 2020 DEEP NOISE SUPPRESSION CHALLENGE: DATASETS, SUBJECTIVE TESTING FRAMEWORK, AND CHALLENGE RESULTS.....	2492
<i>Chandan K. A. Reddy, Vishak Gopal, Ross Cutler, Ebrahim Beyrami, Roger Cheng, Harishchandra Dubey, Sergiy Matushevych, Robert Aichner, Ashkan Aazami, Sebastian Braun, Puneet Rana, Sriram Srinivasan, Johannes Gehrke</i>	
<u>INTERSPEECH 2020 : VOICE AND HEARING DISORDERS</u>	
THE IMPLICATION OF SOUND LEVEL ON SPATIAL SELECTIVE AUDITORY ATTENTION FOR COCHLEAR IMPLANT USERS: BEHAVIORAL AND ELECTROPHYSIOLOGICAL MEASUREMENT	2497
<i>Sara Akbarzadeh, Sungmin Lee, Chin-Tuan Tan</i>	
ENHANCING THE INTERAURAL TIME DIFFERENCE OF BILATERAL COCHLEAR IMPLANTS WITH THE TEMPORAL LIMITS ENCODER.....	2502
<i>Yangyang Wan, Huali Zhou, Qinglin Meng, Nengheng Zheng</i>	
SPEECH CLARITY IMPROVEMENT BY VOCAL SELF-TRAINING USING A HEARING IMPAIRMENT SIMULATOR AND ITS CORRELATION WITH AN AUDITORY MODULATION INDEX.....	2507
<i>Toshio Irino, Soichi Higashiyama, Hanako Yoshigi</i>	
INVESTIGATION OF PHASE DISTORTION ON PERCEIVED SPEECH QUALITY FOR HEARING-IMPAIRED LISTENERS.....	2512
<i>Zhuohuang Zhang, Donald S. Williamson, Yi Shen</i>	
EEG-BASED SHORT-TIME AUDITORY ATTENTION DETECTION USING MULTI-TASK DEEP LEARNING.....	2517
<i>Zhuo Zhang, Gaoyan Zhang, Jianwu Dang, Shuang Wu, Di Zhou, Longbiao Wang</i>	
TOWARDS INTERPRETING DEEP LEARNING MODELS TO UNDERSTAND LOSS OF SPEECH INTELLIGIBILITY IN SPEECH DISORDERS --- STEP 1: CNN MODEL-BASED PHONE CLASSIFICATION	2522
<i>Sondes Abderrazek, Corinne Fredouille, Alain Ghio, Muriel Lalain, Christine Meunier, Virginie Woisard</i>	
IMPROVING COGNITIVE IMPAIRMENT CLASSIFICATION BY GENERATIVE NEURAL NETWORK-BASED FEATURE AUGMENTATION.....	2527
<i>Bahman Mirheidari, Daniel Blackburn, Ronan O'Malley, Annalena Venneri, Traci Walker, Markus Reuber, Heidi Christensen</i>	

UNCOMMONVOICE: A CROWDSOURCED DATASET OF DYSPHONIC SPEECH.....	2532
<i>Meredith Moore, Piyush Papreja, Michael Saxon, Visar Berisha, Sethuraman Panchanathan</i>	
TOWARDS AUTOMATIC ASSESSMENT OF VOICE DISORDERS: A CLINICAL APPROACH.....	2537
<i>Purva Barche, Krishna Gurugubelli, Anil Kumar Vuppala</i>	
BLABLA: LINGUISTIC FEATURE EXTRACTION FOR CLINICAL ANALYSIS IN MULTIPLE LANGUAGES.....	2542
<i>Abhishek Shivkumar, Jack Weston, Raphael Lenain, Emil Fristed</i>	

INTERSPEECH 2020 : SPOKEN TERM DETECTION

DEPTHWISE SEPARABLE CONVOLUTIONAL RESNET WITH SQUEEZE-AND-EXCITATION BLOCKS FOR SMALL-FOOTPRINT KEYWORD SPOTTING	2547
<i>Menglong Xu, Xiao-Lei Zhang</i>	
PREDICTING DETECTION FILTERS FOR SMALL FOOTPRINT OPEN-VOCABULARY KEYWORD SPOTTING	2552
<i>Theodore Bluche, Thibault Gisselbrecht</i>	
DEEP CONVOLUTIONAL SPIKING NEURAL NETWORKS FOR KEYWORD SPOTTING.....	2557
<i>Emre Yilmaz, Ozgur Bora Gevrek, Jibin Wu, Yuxiang Chen, Xuanbo Meng, Haizhou Li</i>	
DOMAIN AWARE TRAINING FOR FAR-FIELD SMALL-FOOTPRINT KEYWORD SPOTTING.....	2562
<i>Haiwei Wu, Yan Jia, Yuanfei Nie, Ming Li</i>	
RE-WEIGHTED INTERVAL LOSS FOR HANDLING DATA IMBALANCE PROBLEM OF END-TO-END KEYWORD SPOTTING	2567
<i>Kun Zhang, Zhiyong Wu, Daode Yuan, Jian Luan, Jia Jia, Helen Meng, Binheng Song</i>	
DEEP TEMPLATE MATCHING FOR SMALL-FOOTPRINT AND CONFIGURABLE KEYWORD SPOTTING	2572
<i>Peng Zhang, Xueliang Zhang</i>	
MULTI-SCALE CONVOLUTION FOR ROBUST KEYWORD SPOTTING.....	2577
<i>Chen Yang, Xue Wen, Liming Song</i>	
AN INVESTIGATION OF FEW-SHOT LEARNING IN SPOKEN TERM CLASSIFICATION	2582
<i>Yangbin Chen, Tom Ko, Lifeng Shang, Xiao Chen, Xin Jiang, Qing Li</i>	
END-TO-END KEYWORD SEARCH BASED ON ATTENTION AND ENERGY SCORER FOR LOW RESOURCE LANGUAGES.....	2587
<i>Zeyu Zhao, Wei-Qiang Zhang</i>	
STACKED 1D CONVOLUTIONAL NETWORKS FOR END-TO-END SMALL FOOTPRINT VOICE TRIGGER DETECTION	2592
<i>Takuya Higuchi, Mohammad Ghasemzadeh, Kisun You, Chandra Dhir</i>	

INTERSPEECH 2020 : THE FEARLESS STEPS CHALLENGE PHASE-02

STATISTICAL AND NEURAL NETWORK BASED SPEECH ACTIVITY DETECTION IN NON-STATIONARY ACOUSTIC ENVIRONMENTS	2597
<i>Jens Heitkaemper, Joerg Schmalenstroer, Reinhold Haeb-Umbach</i>	

SPEAKER DIARIZATION SYSTEM BASED ON DPCA ALGORITHM FOR FEARLESS STEPS CHALLENGE PHASE-2.....	2602
<i>Xueshuai Zhang, Wenchao Wang, Pengyuan Zhang</i>	
THE DKU SPEECH ACTIVITY DETECTION AND SPEAKER IDENTIFICATION SYSTEMS FOR FEARLESS STEPS CHALLENGE PHASE-02.....	2607
<i>Qingjian Lin, Tingle Li, Ming Li</i>	
“THIS IS HOUSTON. SAY AGAIN, PLEASE”. THE BEHAVOX SYSTEM FOR THE APOLLO-11 FEARLESS STEPS CHALLENGE (PHASE II)	2612
<i>Arseniy Gorin, Daniil Kulko, Steven Grima, Alex Glasman</i>	
FEARLESS STEPS CHALLENGE (FS-2): SUPERVISED LEARNING WITH MASSIVE NATURALISTIC APOLLO DATA	2617
<i>Aditya Joglekar, John H. L. Hansen, Meena Chandra Shekar, Abhijeet Sangwan</i>	

INTERSPEECH 2020 : MONAURAL SOURCE SEPARATION

SEPARATING VARYING NUMBERS OF SOURCES WITH AUXILIARY AUTOENCODING LOSS	2622
<i>Yi Luo, Nima Mesgarani</i>	
ON SYNTHESIS FOR SUPERVISED MONAURAL SPEECH SEPARATION IN TIME DOMAIN	2627
<i>Jingjing Chen, Qirong Mao, Dong Liu</i>	
LEARNING BETTER SPEECH REPRESENTATIONS BY WORSENING INTERFERENCE	2632
<i>Jun Wang</i>	
ASTEROID: THE PYTORCH-BASED AUDIO SOURCE SEPARATION TOOLKIT FOR RESEARCHERS.....	2637
<i>Manuel Pariente, Samuele Cornell, Joris Cosentino, Sunit Sivasankaran, Efthymios Tzinis, Jens Heitkaemper, Michel Olvera, Fabian-Robert Stoter, Mathieu Hu, Juan M. Martin-Donas, David Ditter, Ariel Frank, Antoine Deleforge, Emmanuel Vincent</i>	
DUAL-PATH TRANSFORMER NETWORK: DIRECT CONTEXT-AWARE MODELING FOR END-TO-END MONAURAL SPEECH SEPARATION	2642
<i>Jingjing Chen, Qirong Mao, Dong Liu</i>	
CONV-TASSAN: SEPARATIVE ADVERSARIAL NETWORK BASED ON CONV-TASNET	2647
<i>Chengyun Deng, Yi Zhang, Shiqian Ma, Yongtao Sha, Hui Song, Xiangang Li</i>	
MULTI-PATH RNN FOR HIERARCHICAL MODELING OF LONG SEQUENTIAL DATA AND ITS APPLICATION TO SPEAKER STREAM SEPARATION.....	2652
<i>Keisuke Kinoshita, Thilo Von Neumann, Marc Delcroix, Tomohiro Nakatani, Reinhold Haeb-Umbach</i>	
UNSUPERVISED AUDIO SOURCE SEPARATION USING GENERATIVE PRIORS.....	2657
<i>Vivek Narayanaswamy, Jayaraman J. Thiagarajan, Rushil Anirudh, Andreas Spanias</i>	

INTERSPEECH 2020 : SINGLE-CHANNEL SPEECH ENHANCEMENT II

ADVERSARIAL LATENT REPRESENTATION LEARNING FOR SPEECH ENHANCEMENT	2662
<i>Yuanhang Qiu, Ruili Wang</i>	

AN NMF-HMM SPEECH ENHANCEMENT METHOD BASED ON KULLBACK-LEIBLER DIVERGENCE	2667
<i>Yang Xiang, Liming Shi, Jesper Lisby Hojvang, Morten Hojfeldt Rasmussen, Mads Graesboll Christensen</i>	
MULTI-SCALE TCN: EXPLORING BETTER TEMPORAL DNN MODEL FOR CAUSAL SPEECH ENHANCEMENT	2672
<i>Lu Zhang, Mingjiang Wang</i>	
VOICEFILTER-LITE: STREAMING TARGETED VOICE SEPARATION FOR ON-DEVICE SPEECH RECOGNITION	2677
<i>Quan Wang, Ignacio Lopez Moreno, Mert Saglam, Kevin Wilson, Alan Chiao, Renjie Liu, Yanzhang He, Wei Li, Jason Pelecanos, Marily Nika, Alexander Gruenstein</i>	
SPEECH SEPARATION BASED ON MULTI-STAGE ELABORATED DUAL-PATH DEEP BILSTM WITH AUXILIARY IDENTITY LOSS	2682
<i>Ziqiang Shi, Rujie Liu, Jiqing Han</i>	
SUB-BAND KNOWLEDGE DISTILLATION FRAMEWORK FOR SPEECH ENHANCEMENT	2687
<i>Xiang Hao, Shixue Wen, Xiangdong Su, Yun Liu, Guanglai Gao, Xiaofei Li</i>	
A DEEP LEARNING-BASED KALMAN FILTER FOR SPEECH ENHANCEMENT	2692
<i>Sujan Kumar Roy, Aaron Nicolson, Kuldip K. Paliwal</i>	
SUBBAND KALMAN FILTERING WITH DNN ESTIMATED PARAMETERS FOR SPEECH ENHANCEMENT	2697
<i>Hongjiang Yu, Wei-Ping Zhu, Benoit Champagne</i>	
BIDIRECTIONAL LSTM NETWORK WITH ORDERED NEURONS FOR SPEECH ENHANCEMENT	2702
<i>Xiaoqi Li, Yaxing Li, Yuanjie Dong, Shan Xu, Zhihui Zhang, Dan Wang, Shengwu Xiong</i>	
SPEAKER-CONDITIONAL CHAIN MODEL FOR SPEECH SEPARATION AND EXTRACTION	2707
<i>Jing Shi, Jiaming Xu, Yusuke Fujita, Shinji Watanabe, Bo Xu</i>	
 <u>INTERSPEECH 2020 : TOPICS IN ASR II</u>	
UNSUPERVISED VS. TRANSFER LEARNING FOR MULTIMODAL ONE-SHOT MATCHING OF SPEECH AND IMAGES	2712
<i>Leanne Nortje, Herman Kamper</i>	
MULTIMODAL SPEECH EMOTION RECOGNITION USING CROSS ATTENTION WITH ALIGNED AUDIO AND TEXT	2717
<i>Yoonhyung Lee, Seunghyun Yoon, Kyomin Jung</i>	
SPEAKER DEPENDENT ARTICULATORY-TO-ACOUSTIC MAPPING USING REAL-TIME MRI OF THE VOCAL TRACT	2722
<i>Tamas Gabor Csapo</i>	
ULTRASOUND-BASED ARTICULATORY-TO-ACOUSTIC MAPPING WITH WAVEGLOW SPEECH SYNTHESIS	2727
<i>Tamas Gabor Csapo, Csaba Zainko, Laszlo Toth, Gabor Gosztolya, Alexandra Marko</i>	

UNSUPERVISED SUBWORD MODELING USING AUTOREGRESSIVE PRETRAINING
AND CROSS-LINGUAL PHONE-AWARE MODELING 2732
Siyuan Feng, Odette Scharenborg

GENERATIVE ADVERSARIAL TRAINING DATA ADAPTATION FOR VERY LOW-
RESOURCE AUTOMATIC SPEECH RECOGNITION 2737
Kohei Matsuura, Masato Mimura, Shinsuke Sakai, Tatsuya Kawahara

NEURAL SPEECH COMPLETION 2742
Kazuki Tsunematsu, Johannes Effendi, Sakriani Sakti, Satoshi Nakamura

IMPROVING UNSUPERVISED SPARSE SPEECH ACOUSTIC MODELS WITH
CATEGORICAL REPARAMETERIZATION 2747
Benjamin Milde, Chris Biemann

MULTIMODAL SIGN LANGUAGE RECOGNITION VIA TEMPORAL DEFORMABLE
CONVOLUTIONAL SEQUENCE LEARNING 2752
Katerina Papadimitriou, Gerasimos Potamianos

MLS: A LARGE-SCALE MULTILINGUAL DATASET FOR SPEECH RESEARCH 2757
Vineel Pratap, Qiantong Xu, Anuroop Sriram, Gabriel Synnaeve, Ronan Collobert

INTERSPEECH 2020 : NEURAL SIGNALS FOR SPOKEN COMMUNICATION

COMBINING AUDIO AND BRAIN ACTIVITY FOR PREDICTING SPEECH QUALITY 2762
Ivan Halim Parmonangan, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura

THE "SOUND OF SILENCE" IN EEG --- COGNITIVE VOICE ACTIVITY DETECTION 2767
Rini A. Sharon, Hema A. Murthy

LOW LATENCY AUDITORY ATTENTION DETECTION WITH COMMON SPATIAL
PATTERN ANALYSIS OF EEG SIGNALS 2772
Siqi Cai, Enze Su, Yonghao Song, Longhan Xie, Haizhou Li

SPEECH SPECTROGRAM ESTIMATION FROM INTRACRANIAL BRAIN ACTIVITY
USING A QUANTIZATION APPROACH 2777
Miguel Angrick, Christian Herff, Garrett Johnson, Jerry Shih, Dean Krusienski, Tanja Schultz

NEURAL SPEECH DECODING FOR AMYOTROPHIC LATERAL SCLEROSIS 2782
Debadatta Dash, Paul Ferrari, Angel Hernandez, Daragh Heitzman, Sara G. Austin, Jun Wang

INTERSPEECH 2020 : TRAINING STRATEGIES FOR ASR

SEMI-SUPERVISED ASR BY END-TO-END SELF-TRAINING 2787
Yang Chen, Weiran Wang, Chao Wang

IMPROVED TRAINING STRATEGIES FOR END-TO-END SPEECH RECOGNITION IN
DIGITAL VOICE ASSISTANTS 2792
Hitesh Tulsiani, Ashutosh Sapru, Harish Arshikere, Surabhi Punjabi, Sri Garimella

SERIALIZED OUTPUT TRAINING FOR END-TO-END OVERLAPPED SPEECH
RECOGNITION 2797
Naoyuki Kanda, Yashesh Gaur, Xiaofei Wang, Zhong Meng, Takuya Yoshioka

SEMI-SUPERVISED LEARNING WITH DATA AUGMENTATION FOR END-TO-END ASR.....	2802
<i>Felix Weninger, Franco Mana, Roberto Gemello, Jesus Andres-Ferrer, Puming Zhan</i>	
EFFICIENT MINIMUM WORD ERROR RATE TRAINING OF RNN-TRANSDUCER FOR END-TO-END SPEECH RECOGNITION.....	2807
<i>Jinxi Guo, Gautam Tiwari, Jasha Droppo, Maarten Van Segbroeck, Che-Wei Huang, Andreas Stolcke, Roland Maas</i>	
A NEW TRAINING PIPELINE FOR AN IMPROVED NEURAL TRANSDUCER.....	2812
<i>Albert Zeyer, Andre Merboldt, Ralf Schluter, Hermann Ney</i>	
IMPROVED NOISY STUDENT TRAINING FOR AUTOMATIC SPEECH RECOGNITION	2817
<i>Daniel S. Park, Yu Zhang, Ye Jia, Wei Han, Chung-Cheng Chiu, Bo Li, Yonghui Wu, Quoc V. Le</i>	
PHONEME-TO-GRAPHEME CONVERSION BASED LARGE-SCALE PRE-TRAINING FOR END-TO-END AUTOMATIC SPEECH RECOGNITION.....	2822
<i>Ryo Masumura, Naoki Makishima, Mana Ihori, Akihiko Takashima, Tomohiro Tanaka, Shota Orihashi</i>	
UTTERANCE INVARIANT TRAINING FOR HYBRID TWO-PASS END-TO-END SPEECH RECOGNITION.....	2827
<i>Dhananjaya Gowda, Ankur Kumar, Kwangyoun Kim, Hejung Yang, Abhinav Garg, Sachin Singh, Jiyeon Kim, Mehul Kumar, Sichen Jin, Shatrughan Singh, Chanwoo Kim</i>	
SCADA: STOCHASTIC, CONSISTENT AND ADVERSARIAL DATA AUGMENTATION TO IMPROVE ASR	2832
<i>Gary Wang, Andrew Rosenberg, Zhehuai Chen, Yu Zhang, Bhuvana Ramabhadran, Pedro J. Moreno</i>	
<u>INTERSPEECH 2020 : SPEECH TRANSMISSION & CODING</u>	
FUNDAMENTAL FREQUENCY MODEL FOR POSTFILTERING AT LOW BITRATES IN A TRANSFORM-DOMAIN SPEECH AND AUDIO CODEC	2837
<i>Sneha Das, Tom Backstrom, Guillaume Fuchs</i>	
HEARING-IMPAIRED BIO-INSPIRED COCHLEAR MODELS FOR REAL-TIME AUDITORY APPLICATIONS.....	2842
<i>Arthur Van Den Broucke, Deepak Baby, Sarah Verhulst</i>	
IMPROVING OPUS LOW BIT RATE QUALITY WITH NEURAL SPEECH SYNTHESIS	2847
<i>Jan Skoglund, Jean-Marc Valin</i>	
A DIFFERENTIABLE PERCEPTUAL AUDIO METRIC LEARNED FROM JUST NOTICEABLE DIFFERENCES.....	2852
<i>Pranay Manocha, Adam Finkelstein, Richard Zhang, Nicholas J. Bryan, Gautham J. Mysore, Zeyu Jin</i>	
STORIR: STOCHASTIC ROOM IMPULSE RESPONSE GENERATION FOR AUDIO DATA AUGMENTATION.....	2857
<i>Piotr Masztalski, Mateusz Matuszewski, Karol Piaskowski, Michal Romaniuk</i>	
AN OPEN SOURCE IMPLEMENTATION OF ITU-T RECOMMENDATION P.808 WITH VALIDATION	2862
<i>Babak Naderi, Ross Cutler</i>	

DNN NO-REFERENCE PSTN SPEECH QUALITY PREDICTION..... 2867
*Gabriel Mittag, Ross Cutler, Yasaman Hosseinkashi, Michael Revow, Sriram Srinivasan,
Naglakshmi Chande, Robert Aichner*

NON-INTRUSIVE DIAGNOSTIC MONITORING OF FULLBAND SPEECH QUALITY..... 2872
Sebastian Moller, Tobias Hubschen, Thilo Michael, Gabriel Mittag, Gerhard Schmidt

INTERSPEECH 2020 : BIOACOUSTICS AND ARTICULATION

TRANSFER LEARNING OF ARTICULATORY INFORMATION THROUGH PHONE
INFORMATION..... 2877
*Abdolreza Sabzi Shahrehabaki, Negar Olfati, Sabato Marco Siniscalchi, Giampiero Salvi,
Torbjorn Svendsen*

SEQUENCE-TO-SEQUENCE ARTICULATORY INVERSION THROUGH TIME
CONVOLUTION OF SUB-BAND FREQUENCY SIGNALS 2882
*Abdolreza Sabzi Shahrehabaki, Sabato Marco Siniscalchi, Giampiero Salvi, Torbjorn
Svendsen*

DISCRIMINATIVE SINGULAR SPECTRUM ANALYSIS FOR BIOACOUSTIC
CLASSIFICATION..... 2887
*Bernardo B. Gatto, Eulanda M. Dos Santos, Juan G. Colonna, Naoya Sogi, Lincon S. Souza,
Kazuhiro Fukui*

SPEECH RATE TASK-SPECIFIC REPRESENTATION LEARNING FROM ACOUSTIC-
ARTICULATORY DATA..... 2892
Renuka Mannem, Hima Jyothi R., Aravind Illa, Prasanta Kumar Ghosh

DYSARTHRIA DETECTION AND SEVERITY ASSESSMENT USING RHYTHM-BASED
METRICS..... 2897
Abner Hernandez, Eun Jung Yeo, Sunhee Kim, Minhwa Chung

LUNGRN+NL: AN IMPROVED ADVENTITIOUS LUNG SOUND CLASSIFICATION USING
NON-LOCAL BLOCK RESNET NEURAL NETWORK WITH MIXUP DATA
AUGMENTATION..... 2902
Yi Ma, Xinzi Xu, Yongfu Li

VOLUME 5

ATTENTION AND ENCODER-DECODER BASED MODELS FOR TRANSFORMING
ARTICULATORY MOVEMENTS AT DIFFERENT SPEAKING RATES 2907
Abhayjeet Singh, Aravind Illa, Prasanta Kumar Ghosh

ADVENTITIOUS RESPIRATORY CLASSIFICATION USING ATTENTIVE RESIDUAL
NEURAL NETWORKS..... 2912
Zijiang Yang, Shuo Liu, Meishu Song, Emilia Parada-Cabaleiro, Bjorn W. Schuller

SURFBOARD: AUDIO FEATURE EXTRACTION FOR MODERN MACHINE LEARNING 2917
Raphael Lenain, Jack Weston, Abhishek Shivkumar, Emil Fristed

WHISPER ACTIVITY DETECTION USING CNN-LSTM BASED ATTENTION POOLING
NETWORK TRAINED FOR A SPEAKER IDENTIFICATION TASK 2922
Abinay Reddy Naini, Malla Satyapriya, Prasanta Kumar Ghosh

INTERSPEECH 2020 : SPEECH SYNTHESIS: MULTILINGUAL AND CROSS-LINGUAL APPROACHES

TOWARDS NATURAL BILINGUAL AND CODE-SWITCHED SPEECH SYNTHESIS BASED ON MIX OF MONOLINGUAL RECORDINGS AND CROSS-LINGUAL VOICE CONVERSION	2927
<i>Shengkui Zhao, Trung Hieu Nguyen, Hao Wang, Bin Ma</i>	
MULTI-LINGUAL MULTI-SPEAKER TEXT-TO-SPEECH SYNTHESIS FOR VOICE CLONING WITH ONLINE SPEAKER ENROLLMENT	2932
<i>Zhaoyu Liu, Brian Mak</i>	
DYNAMIC SOFT WINDOWING AND LANGUAGE DEPENDENT STYLE TOKEN FOR CODE-SWITCHING END-TO-END SPEECH SYNTHESIS	2937
<i>Ruibo Fu, Jianhua Tao, Zhengqi Wen, Jiangyan Yi, Chunyu Qiang, Tao Wang</i>	
PHONOLOGICAL FEATURES FOR 0-SHOT MULTILINGUAL SPEECH SYNTHESIS	2942
<i>Marlene Staib, Tian Huey Teh, Alexandra Torresquintero, Devang S. Ram Mohan, Lorenzo Foglianti, Raphael Lenain, Jiameng Gao</i>	
CROSS-LINGUAL TEXT-TO-SPEECH SYNTHESIS VIA DOMAIN ADAPTATION AND PERCEPTUAL SIMILARITY REGRESSION IN SPEAKER SPACE	2947
<i>Detai Xin, Yuki Saito, Shinnosuke Takamichi, Tomoki Koriyama, Hiroshi Saruwatari</i>	
TONE LEARNING IN LOW-RESOURCE BILINGUAL TTS	2952
<i>Ruolan Liu, Xue Wen, Chunhui Lu, Xiao Chen</i>	
ON IMPROVING CODE MIXED SPEECH SYNTHESIS WITH MIXLINGUAL GRAPHEME-TO-PHONEME MODEL	2957
<i>Shubham Bansal, Arijit Mukherjee, Sandeepkumar Satpal, Rupeshkumar Mehta</i>	
GENERIC INDIC TEXT-TO-SPEECH SYNTHESISERS WITH RAPID ADAPTATION IN AN END-TO-END FRAMEWORK	2962
<i>Anusha Prakash, Hema A. Murthy</i>	
EFFICIENT NEURAL SPEECH SYNTHESIS FOR LOW-RESOURCE LANGUAGES THROUGH MULTILINGUAL MODELING	2967
<i>Marcel De Korte, Jaebok Kim, Esther Klabbers</i>	
ONE MODEL, MANY LANGUAGES: META-LEARNING FOR MULTILINGUAL TEXT-TO-SPEECH	2972
<i>Tomas Nekvinda, Ondrej Dusek</i>	

INTERSPEECH 2020 : LEARNING TECHNIQUES FOR SPEAKER RECOGNITION I

IN DEFENCE OF METRIC LEARNING FOR SPEAKER RECOGNITION	2977
<i>Joon Son Chung, Jaesung Huh, Seongkyu Mun, Minjae Lee, Hee-Soo Heo, Soyeon Choe, Chiheon Ham, Sunghwan Jung, Bong-Jin Lee, Icksang Han</i>	
META-LEARNING FOR SHORT UTTERANCE SPEAKER RECOGNITION WITH IMBALANCE LENGTH PAIRS	2982
<i>Seong Min Kye, Youngmoon Jung, Hae Beom Lee, Sung Ju Hwang, Hoirin Kim</i>	

SEGMENT-LEVEL EFFECTS OF GENDER, NATIONALITY AND EMOTION INFORMATION ON TEXT-INDEPENDENT SPEAKER VERIFICATION	2987
<i>Kai Li, Masato Akagi, Yibo Wu, Jianwu Dang</i>	
WEAKLY SUPERVISED TRAINING OF HIERARCHICAL ATTENTION NETWORKS FOR SPEAKER IDENTIFICATION	2992
<i>Yanpei Shi, Qiang Huang, Thomas Hain</i>	
MULTI-TASK LEARNING FOR VOICE RELATED RECOGNITION TASKS.....	2997
<i>Ana Montalvo, Jose R. Calvo, Jean-Francois Bonastre</i>	
UNSUPERVISED TRAINING OF SIAMESE NETWORKS FOR SPEAKER VERIFICATION.....	3002
<i>Umair Khan, Javier Hernando</i>	
AN EFFECTIVE SPEAKER RECOGNITION METHOD BASED ON JOINT IDENTIFICATION AND VERIFICATION SUPERVISIONS	3007
<i>Ying Liu, Yan Song, Yiheng Jiang, Ian McLoughlin, Lin Liu, Li-Rong Dai</i>	
SPEAKER-AWARE LINEAR DISCRIMINANT ANALYSIS IN SPEAKER VERIFICATION.....	3012
<i>Naijun Zheng, Xixin Wu, Jinghua Zhong, Xunying Liu, Helen Meng</i>	
ADVERSARIAL DOMAIN ADAPTATION FOR SPEAKER VERIFICATION USING PARTIALLY SHARED NETWORK.....	3017
<i>Zhengyang Chen, Shuai Wang, Yanmin Qian</i>	

INTERSPEECH 2020 : PRONUNCIATION

AUTOMATIC SCORING AT MULTI-GRANULARITY FOR L2 PRONUNCIATION	3022
<i>Binghuai Lin, Liyuan Wang, Xiaoli Feng, Jinsong Zhang</i>	
AN EFFECTIVE END-TO-END MODELING APPROACH FOR MISPRONUNCIATION DETECTION.....	3027
<i>Tien-Hong Lo, Shi-Yan Weng, Hsiu-Jui Chang, Berlin Chen</i>	
AN END-TO-END MISPRONUNCIATION DETECTION SYSTEM FOR L2 ENGLISH SPEECH LEVERAGING NOVEL ANTI-PHONE MODELING	3032
<i>Bi-Cheng Yan, Meng-Che Wu, Hsiao-Tsung Hung, Berlin Chen</i>	
UNSUPERVISED FEATURE ADAPTATION USING ADVERSARIAL MULTI-TASK TRAINING FOR AUTOMATIC EVALUATION OF CHILDREN'S SPEECH.....	3037
<i>Richeng Duan, Nancy F. Chen</i>	
PRONUNCIATION ERRONEOUS TENDENCY DETECTION WITH LANGUAGE ADVERSARIAL REPRESENT LEARNING	3042
<i>Longfei Yang, Kaiqi Fu, Jinsong Zhang, Takahiro Shinozaki</i>	
ASR-FREE PRONUNCIATION ASSESSMENT	3047
<i>Sitong Cheng, Zhixin Liu, Lantian Li, Zhiyuan Tang, Dong Wang, Thomas Fang Zheng</i>	
AUTOMATIC DETECTION OF ACCENT AND LEXICAL PRONUNCIATION ERRORS IN SPONTANEOUS NON-NATIVE ENGLISH SPEECH.....	3052
<i>Konstantinos Kyriakopoulos, Kate M. Knill, Mark J. F. Gales</i>	
CONTEXT-AWARE GOODNESS OF PRONUNCIATION FOR COMPUTER-ASSISTED PRONUNCIATION TRAINING	3057
<i>Jiatong Shi, Nan Huo, Qin Jin</i>	

RECOGNIZE MISPRONUNCIATIONS TO IMPROVE NON-NATIVE ACOUSTIC MODELING THROUGH A PHONE DECODER BUILT FROM ONE EDIT DISTANCE FINITE STATE AUTOMATON	3062
<i>Wei Chu, Yang Liu, Jianwei Zhou</i>	

INTERSPEECH 2020 : DIARIZATION

PARTIAL AUC OPTIMISATION USING RECURRENT NEURAL NETWORKS FOR MUSIC DETECTION WITH LIMITED TRAINING DATA	3067
<i>Pablo Gimeno, Victoria Mingote, Alfonso Ortega, Antonio Miguel, Eduardo Lleida</i>	
AN OPEN-SOURCE VOICE TYPE CLASSIFIER FOR CHILD-CENTERED DAYLONG RECORDINGS	3072
<i>Marvin Lavechin, Ruben Bousbib, Herve Bredin, Emmanuel Dupoux, Alejandrina Cristia</i>	
COMPETING SPEAKER COUNT ESTIMATION ON THE FUSION OF THE SPECTRAL AND SPATIAL EMBEDDING SPACE	3077
<i>Chao Peng, Xihong Wu, Tianshu Qu</i>	
AUDIO-VISUAL MULTI-SPEAKER TRACKING BASED ON THE GLMB FRAMEWORK	3082
<i>Shoufeng Lin, Xinyuan Qian</i>	
TOWARDS SPEECH ROBUSTNESS FOR ACOUSTIC SCENE CLASSIFICATION	3087
<i>Shuo Liu, Andreas Triantafyllopoulos, Zhao Ren, Bjorn W. Schuller</i>	
IDENTIFY SPEAKERS IN COCKTAIL PARTIES WITH END-TO-END ATTENTION	3092
<i>Junzhe Zhu, Mark Hasegawa-Johnson, Leda Sari</i>	
MULTI-TALKER ASR FOR AN UNKNOWN NUMBER OF SOURCES: JOINT TRAINING OF SOURCE COUNTING, SEPARATION AND ASR	3097
<i>Thilo Von Neumann, Christoph Boeddeker, Lukas Drude, Keisuke Kinoshita, Marc Delcroix, Tomohiro Nakatani, Reinhold Haeb-Umbach</i>	
ATTENTIVE CONVOLUTIONAL RECURRENT NEURAL NETWORK USING PHONEME- LEVEL ACOUSTIC REPRESENTATION FOR RARE SOUND EVENT DETECTION	3102
<i>Shreya G. Upadhyay, Bo-Hao Su, Chi-Chun Lee</i>	
DETECTING AND COUNTING OVERLAPPING SPEAKERS IN DISTANT SPEECH SCENARIOS	3107
<i>Samuele Cornell, Maurizio Omologo, Stefano Squartini, Emmanuel Vincent</i>	
ALL-IN-ONE TRANSFORMER: UNIFYING SPEECH RECOGNITION, AUDIO TAGGING, AND EVENT DETECTION	3112
<i>Niko Moritz, Gordon Wichern, Takaaki Hori, Jonathan Le Roux</i>	

INTERSPEECH 2020 : COMPUTATIONAL PARALINGUISTICS II

TOWARDS SILENT PARALINGUISTICS: DERIVING SPEAKING MODE AND SPEAKER ID FROM ELECTROMYOGRAPHIC SIGNALS	3117
<i>Lorenz Diener, Shahin Amiriparian, Catarina Botelho, Kevin Scheck, Dennis Kuster, Isabel Trancoso, Bjorn W. Schuller, Tanja Schultz</i>	
PREDICTING COLLABORATIVE TASK PERFORMANCE USING GRAPH INTERLOCUTOR ACOUSTIC NETWORK IN SMALL GROUP INTERACTION	3122
<i>Shun-Chang Zhong, Bo-Hao Su, Wei Huang, Yi-Ching Liu, Chi-Chun Lee</i>	

VERY SHORT-TERM CONFLICT INTENSITY ESTIMATION USING FISHER VECTORS	3127
<i>Gabor Gosztolya</i>	
GAMING CORPUS FOR STUDYING SOCIAL SCREAMS	3132
<i>Hiroki Mori, Yuki Kikuchi</i>	
SPEAKER DISCRIMINATION IN HUMANS AND MACHINES: EFFECTS OF SPEAKING STYLE VARIABILITY	3136
<i>Amber Afshan, Jody Kreiman, Abeer Alwan</i>	
AUTOMATIC PREDICTION OF CONFIDENCE LEVEL FROM CHILDREN'S ORAL READING RECORDINGS.....	3141
<i>Kamini Sabu, Preeti Rao</i>	
TOWARDS A COMPREHENSIVE ASSESSMENT OF SPEECH INTELLIGIBILITY FOR PATHOLOGICAL SPEECH	3146
<i>W. Xue, V. Mendoza Ramos, W. Harmsen, Catia Cucchiarini, R. W. N. M. Van Hout, Helmer Strik</i>	
EFFECTS OF COMMUNICATION CHANNELS AND ACTOR'S GENDER ON EMOTION IDENTIFICATION BY NATIVE MANDARIN SPEAKERS	3151
<i>Yi Lin, Hongwei Ding</i>	
DETECTION OF VOICING AND PLACE OF ARTICULATION OF FRICATIVES WITH DEEP LEARNING IN A VIRTUAL SPEECH AND LANGUAGE THERAPY TUTOR	3156
<i>Ivo Anjos, Maxine Eskenazi, Nuno Marques, Margarida Grilo, Isabel Guimaraes, Joao Magalhaes, Sofia Cavaco</i>	
<u>INTERSPEECH 2020 : SPEECH SYNTHESIS PARADIGMS AND METHODS II</u>	
UNSUPERVISED LEARNING FOR SEQUENCE-TO-SEQUENCE TEXT-TO-SPEECH FOR LOW-RESOURCE LANGUAGES	3161
<i>Haitong Zhang, Yue Lin</i>	
CONDITIONAL SPOKEN DIGIT GENERATION WITH STYLEGAN	3166
<i>Kasper Palkama, Lauri Juvela, Alexander Ilin</i>	
TOWARDS UNIVERSAL TEXT-TO-SPEECH.....	3171
<i>Jingzhou Yang, Lei He</i>	
SPEAKER-INDEPENDENT MEL-CEPSTRUM ESTIMATION FROM ARTICULATOR MOVEMENTS USING D-VECTOR INPUT	3176
<i>Kouichi Katsurada, Korin Richmond</i>	
ENHANCING MONOTONICITY FOR ROBUST AUTOREGRESSIVE TRANSFORMER TTS.....	3181
<i>Xiangyu Liang, Zhiyong Wu, Runnan Li, Yanqing Liu, Sheng Zhao, Helen Meng</i>	
INCREMENTAL TEXT TO SPEECH FOR NEURAL SEQUENCE-TO-SEQUENCE MODELS USING REINFORCEMENT LEARNING	3186
<i>Devang S. Ram Mohan, Raphael Lenain, Lorenzo Foglianti, Tian Huey Teh, Marlene Staib, Alexandra Torresquintero, Jiameng Gao</i>	
SEMI-SUPERVISED LEARNING FOR MULTI-SPEAKER TEXT-TO-SPEECH SYNTHESIS USING DISCRETE SPEECH REPRESENTATION	3191
<i>Tao Tu, Yuan-Jui Chen, Alexander H. Liu, Hung-Yi Lee</i>	

LEARNING JOINT ARTICULATORY-ACOUSTIC REPRESENTATIONS WITH
NORMALIZING FLOWS 3196
Pramit Saha, Sidney Fels

INVESTIGATING EFFECTIVE ADDITIONAL CONTEXTUAL FACTORS IN DNN-BASED
SPONTANEOUS SPEECH SYNTHESIS 3201
*Yuki Yamashita, Tomoki Koriyama, Yuki Saito, Shinnosuke Takamichi, Yusuke Ijima, Ryo
Masumura, Hiroshi Saruwatari*

HIDER-FINDER-COMBINER: AN ADVERSARIAL ARCHITECTURE FOR GENERAL
SPEECH SIGNAL MODIFICATION 3206
Jacob J. Webber, Olivier Perrotin, Simon King

INTERSPEECH 2020 : SPEAKER EMBEDDING

WAV2SPK: A SIMPLE DNN ARCHITECTURE FOR LEARNING SPEAKER EMBEDDINGS
FROM WAVEFORMS 3211
Weiwei Lin, Man-Wai Mak

HOW DOES LABEL NOISE AFFECT THE QUALITY OF SPEAKER EMBEDDINGS? 3216
Minh Pham, Zeqian Li, Jacob Whitehill

A COMPARATIVE RE-ASSESSMENT OF FEATURE EXTRACTORS FOR DEEP SPEAKER
EMBEDDINGS 3221
Xuechen Liu, Md. Sahidullah, Tomi Kinnunen

SPEAKER REPRESENTATION LEARNING USING GLOBAL CONTEXT GUIDED
CHANNEL AND TIME-FREQUENCY TRANSFORMATIONS 3226
Wei Xia, John H. L. Hansen

INTRA-CLASS VARIATION REDUCTION OF SPEAKER REPRESENTATION IN
DISENTANGLEMENT FRAMEWORK 3231
Yoohwan Kwon, Soo-Whan Chung, Hong-Goo Kang

COMPACT SPEAKER EMBEDDING: LRX-VECTOR 3236
Munir Georges, Jonathan Huang, Tobias Bocklet

COSINE-DISTANCE VIRTUAL ADVERSARIAL TRAINING FOR SEMI-SUPERVISED
SPEAKER-DISCRIMINATIVE ACOUSTIC EMBEDDINGS 3241
Florian L. Kreyssig, Philip C. Woodland

DEEP SPEAKER EMBEDDING WITH LONG SHORT TERM CENTROID LEARNING FOR
TEXT-INDEPENDENT SPEAKER VERIFICATION 3246
Junyi Peng, Rongzhi Gu, Yuexian Zou

NEURAL DISCRIMINANT ANALYSIS FOR DEEP SPEAKER EMBEDDING 3251
Lantian Li, Dong Wang, Thomas Fang Zheng

LEARNING SPEAKER EMBEDDING FROM TEXT-TO-SPEECH 3256
Jaejin Cho, Piotr Zelasko, Jesus Villalba, Shinji Watanabe, Najim Dehak

INTERSPEECH 2020 : SINGLE-CHANNEL SPEECH ENHANCEMENT III

NOISY-REVERBERANT SPEECH ENHANCEMENT USING DENSEUNET WITH TIME-FREQUENCY ATTENTION.....	3261
<i>Yan Zhao, Deliang Wang</i>	
ON LOSS FUNCTIONS AND RECURRENCE TRAINING FOR GAN-BASED SPEECH ENHANCEMENT SYSTEMS.....	3266
<i>Zhuohuang Zhang, Chengyun Deng, Yi Shen, Donald S. Williamson, Yongtao Sha, Yi Zhang, Hui Song, Xiangang Li</i>	
SELF-SUPERVISED ADVERSARIAL MULTI-TASK LEARNING FOR VOCODER-BASED MONAURAL SPEECH ENHANCEMENT.....	3271
<i>Zhihao Du, Ming Lei, Jiqing Han, Shiliang Zhang</i>	
DEEP SPEECH INPAINTING OF TIME-FREQUENCY MASKS.....	3276
<i>Mikolaj Kegler, Pierre Beckmann, Milos Cernak</i>	
REAL-TIME SINGLE-CHANNEL DEEP NEURAL NETWORK-BASED SPEECH ENHANCEMENT ON EDGE DEVICES.....	3281
<i>Nikhil Shankar, Gautam Shreedhar Bhat, Issa M. S. Panahi</i>	
IMPROVED SPEECH ENHANCEMENT USING A TIME-DOMAIN GAN WITH MASK LEARNING.....	3286
<i>Ju Lin, Sufeng Niu, Adriaan J. Van Wijngaarden, Jerome L. McClendon, Melissa C. Smith, Kuang-Ching Wang</i>	
REAL TIME SPEECH ENHANCEMENT IN THE WAVEFORM DOMAIN	3291
<i>Alexandre Defossez, Gabriel Synnaeve, Yossi Adi</i>	
EFFICIENT LOW-LATENCY SPEECH ENHANCEMENT WITH MOBILE AUDIO STREAMING NETWORKS.....	3296
<i>Michal Romaniuk, Piotr Masztalski, Karol Piaskowski, Mateusz Matuszewski</i>	

INTERSPEECH 2020 : MULTI-CHANNEL AUDIO AND EMOTION RECOGNITION

MULTI-STREAM ATTENTION-BASED BLSTM WITH FEATURE SEGMENTATION FOR SPEECH EMOTION RECOGNITION.....	3301
<i>Yuya Chiba, Takashi Nose, Akinori Ito</i>	
MICROPHONE ARRAY POST-FILTER FOR TARGET SPEECH ENHANCEMENT WITHOUT A PRIOR INFORMATION OF POINT INTERFERERS.....	3306
<i>Guanjun Li, Shan Liang, Shuai Nie, Wenju Liu, Zhanlei Yang, Longshuai Xiao</i>	
SIMILARITY-AND-INDEPENDENCE-AWARE BEAMFORMER: METHOD FOR TARGET SOURCE EXTRACTION USING MAGNITUDE SPECTROGRAM AS REFERENCE.....	3311
<i>Atsuo Hiroe</i>	
THE METHOD OF RANDOM DIRECTIONS OPTIMIZATION FOR STEREO AUDIO SOURCE SEPARATION	3316
<i>Oleg Golokolenko, Gerald Schuller</i>	
GATED RECURRENT FUSION OF SPATIAL AND SPECTRAL FEATURES FOR MULTI-CHANNEL SPEECH SEPARATION WITH DEEP EMBEDDING REPRESENTATIONS.....	3321
<i>Cunhang Fan, Jianhua Tao, Bin Liu, Jiangyan Yi, Zhengqi Wen</i>	

GENERALIZED MINIMAL DISTORTION PRINCIPLE FOR BLIND SOURCE SEPARATION 3326
Robin Scheibler

A LIGHTWEIGHT MODEL BASED ON SEPARABLE CONVOLUTION FOR SPEECH
EMOTION RECOGNITION 3331
Ying Zhong, Ying Hu, Hao Huang, Wushour Silamu

META MULTI-TASK LEARNING FOR SPEECH EMOTION RECOGNITION 3336
Ruichu Cai, Kaibin Guo, Boyan Xu, Xiaoyan Yang, Zhenjie Zhang

GEV BEAMFORMING SUPPORTED BY DOA-BASED MASKS GENERATED ON PAIRS OF
MICROPHONES 3341
Francois Grondin, Jean-Samuel Lauzon, Jonathan Vincent, Francois Michaud

INTERSPEECH 2020 : COMPUTATIONAL RESOURCE CONSTRAINED SPEECH RECOGNITION

ACCURATE DETECTION OF WAKE WORD START AND END USING A CNN 3346
*Christin Jose, Yuriy Mishchenko, Thibaud Senechal, Anish Shah, Alex Escott, Shiv Naga
Prasad Vitaladevuni*

HYBRID TRANSFORMER/CTC NETWORKS FOR HARDWARE EFFICIENT VOICE
TRIGGERING..... 3351
Saurabh Adya, Vineet Garg, Siddharth Sigtia, Pramod Simha, Chandra Dhir

MATCHBOXNET: 1D TIME-CHANNEL SEPARABLE CONVOLUTIONAL NEURAL
NETWORK ARCHITECTURE FOR SPEECH COMMANDS RECOGNITION..... 3356
Somshubra Majumdar, Boris Ginsburg

ITERATIVE COMPRESSION OF END-TO-END ASR MODEL USING AUTOML 3361
*Abhinav Mehrotra, Lukasz Dudziak, Jinsu Yeo, Young-Yoon Lee, Ravichander Vippera,
Mohamed S. Abdelfattah, Sourav Bhattacharya, Samin Ishtiaq, Alberto Gil C. P. Ramos,
Sangjeong Lee, Daehyun Kim, Nicholas D. Lane*

QUANTIZATION AWARE TRAINING WITH ABSOLUTE-COSINE REGULARIZATION
FOR AUTOMATIC SPEECH RECOGNITION 3366
Hieu Duy Nguyen, Anastasios Alexandridis, Athanasios Mouchtaris

STREAMING ON-DEVICE END-TO-END ASR SYSTEM FOR PRIVACY-SENSITIVE
VOICE-TYPING..... 3371
*Abhinav Garg, Gowtham P. Vadiseti, Dhananjaya Gowda, Sichen Jin, Aditya Jayasimha,
Youngho Han, Jiyeon Kim, Junmo Park, Kwangyoun Kim, Sooyeon Kim, Young-Yoon Lee,
Kyungbo Min, Chanwoo Kim*

SCALING UP ONLINE SPEECH RECOGNITION USING CONVNETS..... 3376
*Vineel Pratap, Qiantong Xu, Jacob Kahn, Gilad Avidov, Tatiana Likhomanenko, Awni
Hannun, Vitaliy Liptchinsky, Gabriel Synnaeve, Ronan Collobert*

LISTEN ATTENTIVELY, AND SPELL ONCE: WHOLE SENTENCE GENERATION VIA A
NON-AUTOREGRESSIVE ARCHITECTURE FOR LOW-LATENCY SPEECH
RECOGNITION..... 3381
Ye Bai, Jiangyan Yi, Jianhua Tao, Zhengkun Tian, Zhengqi Wen, Shuai Zhang

RESCORE IN A FLASH: COMPACT, CACHE EFFICIENT HASHING DATA STRUCTURES
FOR N-GRAM LANGUAGE MODELS..... 3386
Grant P. Strimel, Ariya Rastrow, Gautam Tiwari, Adrien Pierard, Jon Webb

INTERSPEECH 2020 : SPEECH SYNTHESIS: PROSODY AND EMOTION

MULTI-SPEAKER EMOTION CONVERSION VIA LATENT VARIABLE REGULARIZATION AND A CHAINED ENCODER-DECODER-PREDICTOR NETWORK.....	3391
<i>Ravi Shankar, Hsi-Wei Hsieh, Nicolas Charon, Archana Venkataraman</i>	
NON-PARALLEL EMOTION CONVERSION USING A DEEP-GENERATIVE HYBRID NETWORK AND AN ADVERSARIAL PAIR DISCRIMINATOR	3396
<i>Ravi Shankar, Jacob Sager, Archana Venkataraman</i>	
LAUGHTER SYNTHESIS: COMBINING SEQ2SEQ MODELING WITH TRANSFER LEARNING.....	3401
<i>Noe Tits, Kevin El Haddad, Thierry Dutoit</i>	
NONPARALLEL EMOTIONAL SPEECH CONVERSION USING VAE-GAN.....	3406
<i>Yuexin Cao, Zhengchen Liu, Minchuan Chen, Jun Ma, Shaojun Wang, Jing Xiao</i>	
PRINCIPAL STYLE COMPONENTS: EXPRESSIVE STYLE CONTROL AND CROSS- SPEAKER TRANSFER IN NEURAL TTS.....	3411
<i>Alexander Sorin, Slava Shechtman, Ron Hoory</i>	
CONVERTING ANYONE'S EMOTION: TOWARDS SPEAKER-INDEPENDENT EMOTIONAL VOICE CONVERSION.....	3416
<i>Kun Zhou, Berrak Sisman, Mingyang Zhang, Haizhou Li</i>	
CONTROLLING THE STRENGTH OF EMOTIONS IN SPEECH-LIKE EMOTIONAL SOUND GENERATED BY WAVENET	3421
<i>Kento Matsumoto, Sunao Hara, Masanobu Abe</i>	
LEARNING SYLLABLE-LEVEL DISCRETE PROSODIC REPRESENTATION FOR EXPRESSIVE SPEECH GENERATION.....	3426
<i>Guangyan Zhang, Ying Qin, Tan Lee</i>	
SIMULTANEOUS CONVERSION OF SPEAKER IDENTITY AND EMOTION BASED ON MULTIPLE-DOMAIN ADAPTIVE RBM.....	3431
<i>Takuya Kishida, Shin Tsukamoto, Toru Nakashika</i>	
EXPLOITING DEEP SENTENTIAL CONTEXT FOR EXPRESSIVE END-TO-END SPEECH SYNTHESIS	3436
<i>Fengyu Yang, Shan Yang, Qinghua Wu, Yujun Wang, Lei Xie</i>	
HIERARCHICAL MULTI-GRAINED GENERATIVE MODEL FOR EXPRESSIVE SPEECH SYNTHESIS	3441
<i>Yukiya Hono, Kazuna Tsuboi, Kei Sawada, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, Keiichi Tokuda</i>	
GAN-BASED DATA GENERATION FOR SPEECH EMOTION RECOGNITION	3446
<i>Sefik Emre Eskimez, Dimitrios Dimitriadis, Robert Gmyr, Kenichi Kumanati</i>	
THE PHONETIC BASES OF VOCAL EXPRESSED EMOTION: NATURAL VERSUS ACTED	3451
<i>Hira Dhamyal, Shahan Ali Memon, Bhiksha Raj, Rita Singh</i>	

INTERSPEECH 2020 : THE INTERSPEECH 2020 FAR FIELD SPEAKER VERIFICATION CHALLENGE

THE INTERSPEECH 2020 FAR-FIELD SPEAKER VERIFICATION CHALLENGE	3456
<i>Xiaoyi Qin, Ming Li, Hui Bu, Wei Rao, Rohan Kumar Das, Shrikanth Narayanan, Haizhou Li</i>	
DEEP EMBEDDING LEARNING FOR TEXT-DEPENDENT SPEAKER VERIFICATION	3461
<i>Peng Zhang, Peng Hu, Xueliang Zhang</i>	
STC-INNOVATION SPEAKER RECOGNITION SYSTEMS FOR FAR-FIELD SPEAKER VERIFICATION CHALLENGE 2020	3466
<i>Aleksei Gusev, Vladimir Volokhov, Alisa Vinogradova, Tseren Andzhukaev, Andrey Shulipa, Sergey Novoselov, Timur Pekhovsky, Alexander Kozlov</i>	
NPU SPEAKER VERIFICATION SYSTEM FOR INTERSPEECH 2020 FAR-FIELD SPEAKER VERIFICATION CHALLENGE	3471
<i>Li Zhang, Jian Wu, Lei Xie</i>	
THE JD AI SPEAKER VERIFICATION SYSTEM FOR THE FFSVC 2020 CHALLENGE	3476
<i>Ying Tong, Wei Xue, Shanluo Huang, Lu Fan, Chao Zhang, Guohong Ding, Xiaodong He</i>	

INTERSPEECH 2020 : MULTIMODAL SPEECH PROCESSING

FACEFILTER: AUDIO-VISUAL SPEECH SEPARATION USING STILL IMAGES	3481
<i>Soo-Whan Chung, Soyeon Choe, Joon Son Chung, Hong-Goo Kang</i>	
SEEING VOICES AND HEARING VOICES: LEARNING DISCRIMINATIVE EMBEDDINGS USING CROSS-MODAL SELF-SUPERVISION	3486
<i>Soo-Whan Chung, Hong-Goo Kang, Joon Son Chung</i>	
FUSION ARCHITECTURES FOR WORD-BASED AUDIOVISUAL SPEECH RECOGNITION.....	3491
<i>Michael Wand, Jurgen Schmidhuber</i>	
AUDIO-VISUAL MULTI-CHANNEL RECOGNITION OF OVERLAPPED SPEECH.....	3496
<i>Jianwei Yu, Bo Wu, Rongzhi Gu, Shi-Xiong Zhang, Lianwu Chen, Yong Xu, Meng Yu, Dan Su, Dong Yu, Xunying Liu, Helen Meng</i>	
TMT: A TRANSFORMER-BASED MODAL TRANSLATOR FOR IMPROVING MULTIMODAL SEQUENCE REPRESENTATIONS IN AUDIO VISUAL SCENE-AWARE DIALOG.....	3501
<i>Wubo Li, Dongwei Jiang, Wei Zou, Xiangang Li</i>	
SHOULD WE HARD-CODE THE RECURRENCE CONCEPT OR LEARN IT INSTEAD ? EXPLORING THE TRANSFORMER ARCHITECTURE FOR AUDIO-VISUAL SPEECH RECOGNITION	3506
<i>George Sterpu, Christian Saam, Naomi Harte</i>	
RESOURCE-ADAPTIVE DEEP LEARNING FOR VISUAL SPEECH RECOGNITION.....	3510
<i>Alexandros Koumparoulis, Gerasimos Potamianos, Samuel Thomas, Edmilson Da Silva Morais</i>	
SPEECH-IMAGE SEMANTIC ALIGNMENT DOES NOT DEPEND ON ANY PRIOR CLASSIFICATION TASKS	3515
<i>Masood S. Mortazavi</i>	

LIP GRAPH ASSISTED AUDIO-VISUAL SPEECH RECOGNITION USING BIDIRECTIONAL SYNCHRONOUS FUSION..... 3520
Hong Liu, Zhan Chen, Bing Yang

CAPTION ALIGNMENT FOR LOW RESOURCE AUDIO-VISUAL DATA..... 3525
Vighnesh Reddy Konda, Mayur Warialani, Rakesh Prasanth Achari, Varad Bhatnagar, Jayaprakash Akula, Preethi Jyothi, Ganesh Ramakrishnan, Gholamreza Haffari, Pankaj Singh

INTERSPEECH 2020 : SPEECH SYNTHESIS: NEURAL WAVEFORM GENERATION II

VOCODER-BASED SPEECH SYNTHESIS FROM SILENT VIDEOS..... 3530
Daniel Michelsanti, Olga Slizovskaia, Gloria Haro, Emilia Gomez, Zheng-Hua Tan, Jesper Jensen

QUASI-PERIODIC PARALLEL WAVEGAN VOCODER: A NON-AUTOREGRESSIVE PITCH-DEPENDENT DILATED CONVOLUTION MODEL FOR PARAMETRIC SPEECH GENERATION 3535
Yi-Chiao Wu, Tomoki Hayashi, Takuma Okamoto, Hisashi Kawai, Tomoki Toda

A CYCLICAL POST-FILTERING APPROACH TO MISMATCH REFINEMENT OF NEURAL VOCODER FOR TEXT-TO-SPEECH SYSTEMS 3540
Yi-Chiao Wu, Patrick Lumban Tobing, Kazuki Yasuhara, Noriyuki Matsunaga, Yamato Ohtani, Tomoki Toda

AUDIO DEQUANTIZATION FOR HIGH FIDELITY AUDIO GENERATION IN FLOW-BASED NEURAL VOCODER..... 3545
Hyun-Wook Yoon, Sang-Hoon Lee, Hyeong-Rae Noh, Seong-Whan Lee

STRAWNET: SELF-TRAINING WAVENET FOR TTS IN LOW-DATA REGIMES 3550
Manish Sharma, Tom Kenter, Rob Clark

AN EFFICIENT SUBBAND LINEAR PREDICTION FOR LPCNET-BASED NEURAL SYNTHESIS 3555
Yang Cui, Xi Wang, Lei He, Frank K. Soong

REVERBERATION MODELING FOR SOURCE-FILTER-BASED NEURAL VOCODER..... 3560
Yang Ai, Xin Wang, Junichi Yamagishi, Zhen-Hua Ling

BUNCHED LPCNET: VOCODER FOR LOW-COST NEURAL TEXT-TO-SPEECH SYSTEMS 3565
Ravichander Vippera, Sangjun Park, Kihyun Choo, Samin Ishtiaq, Kyoungbo Min, Sourav Bhattacharya, Abhinav Mehrotra, Alberto Gil C. P. Ramos, Nicholas D. Lane

NEURAL TEXT-TO-SPEECH WITH A MODELING-BY-GENERATION EXCITATION VOCODER..... 3570
Eunwoo Song, Min-Jae Hwang, Ryuichi Yamamoto, Jin-Seob Kim, Ohsung Kwon, Jae-Min Kim

SPEEDYSPEECH: EFFICIENT NEURAL SPEECH SYNTHESIS..... 3575
Jan Vainer, Ondrej Dusek

INTERSPEECH 2020 : ASR NEURAL NETWORK ARCHITECTURES AND TRAINING II

SEMI-SUPERVISED END-TO-END ASR VIA TEACHER-STUDENT LEARNING WITH CONDITIONAL POSTERIOR DISTRIBUTION	3580
<i>Zi-Qiang Zhang, Yan Song, Jian-Shu Zhang, Ian McLoughlin, Li-Rong Dai</i>	
LEVERAGING UNLABELED SPEECH FOR SEQUENCE DISCRIMINATIVE TRAINING OF ACOUSTIC MODELS.....	3585
<i>Ashtosh Sapru, Sri Garimella</i>	
DEVELOPING RNN-T MODELS SURPASSING HIGH-PERFORMANCE HYBRID MODELS WITH CUSTOMIZATION CAPABILITY	3590
<i>Jinyu Li, Rui Zhao, Zhong Meng, Yanqing Liu, Wenning Wei, Sarangarajan Parthasarathy, Vadim Mazalov, Zhenghao Wang, Lei He, Sheng Zhao, Yifan Gong</i>	
END-TO-END ASR WITH ADAPTIVE SPAN SELF-ATTENTION	3595
<i>Xuankai Chang, Aswin Shanmugam Subramanian, Pengcheng Guo, Shinji Watanabe, Yuya Fujita, Motoi Omachi</i>	
SUBWORD REGULARIZATION: AN ANALYSIS OF SCALABILITY AND GENERALIZATION FOR END-TO-END AUTOMATIC SPEECH RECOGNITION.....	3600
<i>Egor Lakomkin, Jahn Heymann, Ilya Sklyar, Simon Wiesler</i>	
EARLY STAGE LM INTEGRATION USING LOCAL AND GLOBAL LOG-LINEAR COMBINATION.....	3605
<i>Wilfried Michel, Ralf Schluter, Hermann Ney</i>	
CONTEXTNET: IMPROVING CONVOLUTIONAL NEURAL NETWORKS FOR AUTOMATIC SPEECH RECOGNITION WITH GLOBAL CONTEXT	3610
<i>Wei Han, Zhengdong Zhang, Yu Zhang, Jiahui Yu, Chung-Cheng Chiu, James Qin, Anmol Gulati, Ruoming Pang, Yonghui Wu</i>	
EMITTING WORD TIMINGS WITH END-TO-END MODELS	3615
<i>Tara N. Sainath, Ruoming Pang, David Rybach, Basi Garcia, Trevor Strohman</i>	
LOW-LATENCY SEQUENCE-TO-SEQUENCE SPEECH RECOGNITION AND TRANSLATION BY PARTIAL HYPOTHESIS SELECTION.....	3620
<i>Danni Liu, Gerasimos Spanakis, Jan Niehues</i>	

INTERSPEECH 2020 : NEURAL NETWORKS FOR LANGUAGE MODELING

NEURAL LANGUAGE MODELING WITH IMPLICIT CACHE POINTERS.....	3625
<i>Ke Li, Daniel Povey, Sanjeev Khudanpur</i>	
FINNISH ASR WITH DEEP TRANSFORMER MODELS.....	3630
<i>Abhilash Jain, Aku Rouhe, Stig-Arne Gronroos, Mikko Kurimo</i>	

VOLUME 6

DISTILLING THE KNOWLEDGE OF BERT FOR SEQUENCE-TO-SEQUENCE ASR.....	3635
<i>Hayato Futami, Hirofumi Inaguma, Sei Ueno, Masato Mimura, Shinsuke Sakai, Tatsuya Kawahara</i>	

STOCHASTIC CONVOLUTIONAL RECURRENT NETWORKS FOR LANGUAGE MODELING.....	3640
<i>Jen-Tzung Chien, Yu-Min Huang</i>	
INVESTIGATION OF LARGE-MARGIN SOFTMAX IN NEURAL LANGUAGE MODELING	3645
<i>Jingjing Huo, Yingbo Gao, Weiyue Wang, Ralf Schluter, Hermann Ney</i>	
CONTEXTUALIZING ASR LATTICE RESCORING WITH HYBRID POINTER NETWORK LANGUAGE MODEL.....	3650
<i>Da-Rong Liu, Chunxi Liu, Frank Zhang, Gabriel Synnaeve, Yatharth Saraf, Geoffrey Zweig</i>	
MASK CTC: NON-AUTOREGRESSIVE END-TO-END ASR WITH CTC AND MASK PREDICT	3655
<i>Yosuke Higuchi, Shinji Watanabe, Nanxin Chen, Tetsuji Ogawa, Tetsunori Kobayashi</i>	
INSERTION-BASED MODELING FOR END-TO-END AUTOMATIC SPEECH RECOGNITION.....	3660
<i>Yuya Fujita, Shinji Watanabe, Motoi Omachi, Xuankai Chang</i>	
 <u>INTERSPEECH 2020 : PHONETIC EVENT DETECTION AND SEGMENTATION</u>	
VOICE ACTIVITY DETECTION IN THE WILD VIA WEAKLY SUPERVISED SOUND EVENT DETECTION.....	3665
<i>Yefei Chen, Heinrich Dinkel, Mengyue Wu, Kai Yu</i>	
DUAL ATTENTION IN TIME AND FREQUENCY DOMAIN FOR VOICE ACTIVITY DETECTION.....	3670
<i>Joohyung Lee, Youngmoon Jung, Hoirin Kim</i>	
POLISHING THE CLASSICAL LIKELIHOOD RATIO TEST BY SUPERVISED LEARNING FOR VOICE ACTIVITY DETECTION	3675
<i>Tianjiao Xu, Hui Zhang, Xueliang Zhang</i>	
A NOISE ROBUST TECHNIQUE FOR DETECTING VOWELS IN SPEECH SIGNALS.....	3680
<i>Avinash Kumar, S. Shahnawazuddin, Waquar Ahmad</i>	
END-TO-END DOMAIN-ADVERSARIAL VOICE ACTIVITY DETECTION	3685
<i>Marvin Lavechin, Marie-Philippe Gill, Ruben Bousbib, Herve Bredin, Leibny Paola Garcia-Perera</i>	
VOP DETECTION IN VARIABLE SPEECH RATE CONDITION	3690
<i>Ayush Agarwal, Jagabandhu Mishra, S. R. Mahadeva Prasanna</i>	
MLNET: AN ADAPTIVE MULTIPLE RECEPTIVE-FIELD ATTENTION NEURAL NETWORK FOR VOICE ACTIVITY DETECTION	3695
<i>Zhenpeng Zheng, Jianzong Wang, Ning Cheng, Jian Luo, Jing Xiao</i>	
SELF-SUPERVISED CONTRASTIVE LEARNING FOR UNSUPERVISED PHONEME SEGMENTATION.....	3700
<i>Felix Kreuk, Joseph Keshet, Yossi Adi</i>	
THAT SOUNDS FAMILIAR: AN ANALYSIS OF PHONETIC REPRESENTATIONS TRANSFER ACROSS LANGUAGES.....	3705
<i>Piotr Zelasko, Laureano Moro-Velazquez, Mark Hasegawa-Johnson, Odette Scharenborg, Najim Dehak</i>	

ANALYZING READ ALOUD SPEECH BY PRIMARY SCHOOL PUPILS: INSIGHTS FOR RESEARCH AND DEVELOPMENT	3710
<i>S. Limonard, Catia Cucchiarini, R. W. N. M. Van Hout, Helmer Strik</i>	

INTERSPEECH 2020 : HUMAN SPEECH PRODUCTION II

DISCOVERING ARTICULATORY SPEECH TARGETS FROM SYNTHESIZED RANDOM BABBLE	3715
<i>Heikki Rasilo, Yannick Jadoul</i>	
SPEAKER DEPENDENT ACOUSTIC-TO-ARTICULATORY INVERSION USING REAL-TIME MRI OF THE VOCAL TRACT	3720
<i>Tamas Gabor Csapo</i>	
ACOUSTIC-TO-ARTICULATORY INVERSION WITH DEEP AUTOREGRESSIVE ARTICULATORY-WAVENET	3725
<i>Narjes Bozorg, Michael T. Johnson</i>	
USING SILENCE MR IMAGE TO SYNTHESISE DYNAMIC MRI VOCAL TRACT DATA OF CV	3730
<i>Ioannis K. Douros, Ajinkya Kulkarni, Chrysanthi Dourou, Yu Xie, Jacques Felblinger, Karyna Isaieva, Pierre-Andre Vuissoz, Yves Laprie</i>	
QUANTIFICATION OF TRANSDUCER MISALIGNMENT IN ULTRASOUND TONGUE IMAGING	3735
<i>Tamas Gabor Csapo, Kele Xu</i>	
INDEPENDENT AND AUTOMATIC EVALUATION OF SPEAKER-INDEPENDENT ACOUSTIC-TO-ARTICULATORY RECONSTRUCTION	3740
<i>Maud Parrot, Juliette Millet, Ewan Dunbar</i>	
CSL-EMG_ARRAY: AN OPEN ACCESS CORPUS FOR EMG-TO-SPEECH CONVERSION	3745
<i>Lorenz Diener, Mehrdad Roustay Vishkasougheh, Tanja Schultz</i>	
LINKS BETWEEN PRODUCTION AND PERCEPTION OF GLOTTALISATION IN INDIVIDUAL AUSTRALIAN ENGLISH SPEAKER/LISTENERS	3750
<i>Joshua Penney, Felicity Cox, Anita Szakay</i>	

INTERSPEECH 2020 : NEW TRENDS IN SELF-SUPERVISED SPEECH PROCESSING

JOINTLY FINE-TUNING ``BERT-LIKE" SELF SUPERVISED MODELS TO IMPROVE MULTIMODAL SPEECH EMOTION RECOGNITION.....	3755
<i>Shamane Siriwardhana, Andrew Reis, Rivindu Weerasekera, Suranga Nanayakkara</i>	
VECTOR-QUANTIZED AUTOREGRESSIVE PREDICTIVE CODING	3760
<i>Yu-An Chung, Hao Tang, James Glass</i>	
SPEECH-XLNET: UNSUPERVISED ACOUSTIC MODEL PRETRAINING FOR SELF-ATTENTION NETWORKS	3765
<i>Xingchen Song, Guangsen Wang, Yiheng Huang, Zhiyong Wu, Dan Su, Helen Meng</i>	
LARGE SCALE WEAKLY AND SEMI-SUPERVISED LEARNING FOR LOW-RESOURCE VIDEO ASR.....	3770
<i>Kritika Singh, Vimal Manohar, Alex Xiao, Sergey Edunov, Ross Girshick, Vitaliy Liptchinsky, Christian Fuegen, Yatharth Saraf, Geoffrey Zweig, Abdelrahman Mohamed</i>	

SEQUENCE-LEVEL SELF-LEARNING WITH MULTIPLE HYPOTHESES	3775
<i>Kenichi Kumatani, Dimitrios Dimitriadis, Yashesh Gaur, Robert Gmyr, Sefik Emre Eskimez, Jinyu Li, Michael Zeng</i>	
DEFENSE FOR BLACK-BOX ATTACKS ON ANTI-SPOOFING MODELS BY SELF-SUPERVISED LEARNING.....	3780
<i>Haibin Wu, Andy T. Liu, Hung-Yi Lee</i>	
UNDERSTANDING SELF-ATTENTION OF SELF-SUPERVISED AUDIO TRANSFORMERS.....	3785
<i>Shu-Wen Yang, Andy T. Liu, Hung-Yi Lee</i>	
A CONVOLUTIONAL DEEP MARKOV MODEL FOR UNSUPERVISED SPEECH REPRESENTATION LEARNING.....	3790
<i>Sameer Khurana, Antoine Laurent, Wei-Ning Hsu, Jan Chorowski, Adrian Lancucki, Ricard Marxer, James Glass</i>	
AUTOMATIC SPEECH RECOGNITION FOR ILSE-INTERVIEWS: LONGITUDINAL CONVERSATIONAL SPEECH RECORDINGS COVERING AGING AND COGNITIVE DECLINE.....	3795
<i>Ayimunishagu Abulimiti, Jochen Weiner, Tanja Schultz</i>	

INTERSPEECH 2020 : LEARNING TECHNIQUES FOR SPEAKER RECOGNITION II

DYNAMIC MARGIN SOFTMAX LOSS FOR SPEAKER VERIFICATION.....	3800
<i>Dao Zhou, Longbiao Wang, Kong Aik Lee, Yibo Wu, Meng Liu, Jianwu Dang, Jianguo Wei</i>	
ON PARAMETER ADAPTATION IN SOFTMAX-BASED CROSS-ENTROPY LOSS FOR IMPROVED CONVERGENCE SPEED AND ACCURACY IN DNN-BASED SPEAKER RECOGNITION.....	3805
<i>Magdalena Rybicka, Konrad Kowalczyk</i>	
TRAINING SPEAKER ENROLLMENT MODELS BY NETWORK OPTIMIZATION	3810
<i>Victoria Mingote, Antonio Miguel, Alfonso Ortega, Eduardo Lleida</i>	
SUPERVISED DOMAIN ADAPTATION FOR TEXT-INDEPENDENT SPEAKER VERIFICATION USING LIMITED DATA.....	3815
<i>Seyyed Saeed Sarfjoo, Srikanth Madikeri, Petr Motlicek, Sebastien Marcel</i>	
ANGULAR MARGIN CENTROID LOSS FOR TEXT-INDEPENDENT SPEAKER RECOGNITION.....	3820
<i>Yuheng Wei, Junzhao Du, Hui Liu</i>	
DOMAIN-INVARIANT SPEAKER VECTOR PROJECTION BY MODEL-AGNOSTIC META-LEARNING.....	3825
<i>Jiawen Kang, Ruiqi Liu, Lantian Li, Yunqi Cai, Dong Wang, Thomas Fang Zheng</i>	
ECAPA-TDNN: EMPHASIZED CHANNEL ATTENTION, PROPAGATION AND AGGREGATION IN TDNN BASED SPEAKER VERIFICATION	3830
<i>Brecht Desplanques, Jenthe Thienpondt, Kris Demuynck</i>	
LENGTH- AND NOISE-AWARE TRAINING TECHNIQUES FOR SHORT-UTTERANCE SPEAKER RECOGNITION	3835
<i>Wenda Chen, Jonathan Huang, Tobias Bocklet</i>	

INTERSPEECH 2020 : SPOKEN LANGUAGE EVALUATION

SPOKEN LANGUAGE 'GRAMMATICAL ERROR CORRECTION'.....	3840
<i>Yiting Lu, Mark J. F. Gales, Yu Wang</i>	
MIXTURES OF DEEP NEURAL EXPERTS FOR AUTOMATED SPEECH SCORING	3845
<i>Sara Papi, Edmondo Trentin, Roberto Gretter, Marco Matassoni, Daniele Falavigna</i>	
TARGETED CONTENT FEEDBACK IN SPOKEN LANGUAGE LEARNING AND ASSESSMENT	3850
<i>Xinhao Wang, Klaus Zechner, Christopher Hamill</i>	
UNIVERSAL ADVERSARIAL ATTACKS ON SPOKEN LANGUAGE ASSESSMENT SYSTEMS	3855
<i>Vyas Raina, Mark J. F. Gales, Kate M. Knill</i>	
ENSEMBLE APPROACHES FOR UNCERTAINTY IN SPOKEN LANGUAGE ASSESSMENT	3860
<i>Xixin Wu, Kate M. Knill, Mark J. F. Gales, Andrey Malinin</i>	
SHADOWABILITY ANNOTATION WITH FINE GRANULARITY ON L2 UTTERANCES AND ITS IMPROVEMENT WITH NATIVE LISTENERS' SCRIPT-SHADOWING	3865
<i>Zhenchao Lin, Ryo Takashima, Daisuke Saito, Nobuaki Minematsu, Noriko Nakanishi</i>	
ASR-BASED EVALUATION AND FEEDBACK FOR INDIVIDUALIZED READING PRACTICE.....	3870
<i>Yu Bai, Ferdy Hubers, Catia Cucchiarini, Helmer Strik</i>	
DOMAIN ADVERSARIAL NEURAL NETWORKS FOR DYSARTHIC SPEECH RECOGNITION.....	3875
<i>Dominika Woszczyk, Stavros Petridis, David Millard</i>	
AUTOMATIC ESTIMATION OF PATHOLOGICAL VOICE QUALITY BASED ON RECURRENT NEURAL NETWORK USING AMPLITUDE AND PHASE SPECTROGRAM.....	3880
<i>Shunsuke Hidaka, Yogaku Lee, Kohei Wakamiya, Takashi Nakagawa, Tokihiko Kaburagi</i>	

INTERSPEECH 2020 : SPOKEN DIALOGUE SYSTEM

STOCHASTIC CURIOSITY EXPLORATION FOR DIALOGUE SYSTEMS	3885
<i>Jen-Tzung Chien, Po-Chien Hsu</i>	
CONDITIONAL RESPONSE AUGMENTATION FOR DIALOGUE USING KNOWLEDGE DISTILLATION	3890
<i>Myeongho Jeong, Seungtaek Choi, Hojae Han, Kyungho Kim, Seung-Won Hwang</i>	
PROTOTYPICAL Q NETWORKS FOR AUTOMATIC CONVERSATIONAL DIAGNOSIS AND FEW-SHOT NEW DISEASE ADAPTION.....	3895
<i>Hongyin Luo, Shang-Wen Li, James Glass</i>	
END-TO-END TASK-ORIENTED DIALOG SYSTEM THROUGH TEMPLATE SLOT VALUE GENERATION	3900
<i>Teakgyu Hong, Oh-Woog Kwon, Young-Kil Kim</i>	
TASK-ORIENTED DIALOG GENERATION WITH ENHANCED ENTITY REPRESENTATION	3905
<i>Zhenhao He, Jiachun Wang, Jian Chen</i>	

END-TO-END SPEECH-TO-DIALOG-ACT RECOGNITION	3910
<i>Viet-Trung Dang, Tianyu Zhao, Sei Ueno, Hirofumi Inaguma, Tatsuya Kawahara</i>	
DISCRIMINATIVE TRANSFER LEARNING FOR OPTIMIZING ASR AND SEMANTIC LABELING IN TASK-ORIENTED SPOKEN DIALOG.....	3915
<i>Yao Qian, Yu Shi, Michael Zeng</i>	
DATASETS AND BENCHMARKS FOR TASK-ORIENTED LOG DIALOGUE RANKING TASK	3920
<i>Xinnuo Xu, Yizhe Zhang, Lars Liden, Sungjin Lee</i>	

INTERSPEECH 2020 : DEREVERBERATION AND ECHO CANCELLATION

A SEMI-BLIND SOURCE SEPARATION APPROACH FOR SPEECH DEREVERBERATION.....	3925
<i>Ziteng Wang, Yueyue Na, Zhang Liu, Yun Li, Biao Tian, Qiang Fu</i>	
VIRTUAL ACOUSTIC CHANNEL EXPANSION BASED ON NEURAL NETWORKS FOR WEIGHTED PREDICTION ERROR-BASED SPEECH DEREVERBERATION.....	3930
<i>Joon-Young Yang, Joon-Hyuk Chang</i>	
SKIPCONVNET: SKIP CONVOLUTIONAL NEURAL NETWORK FOR SPEECH DEREVERBERATION USING OPTIMALLY SMOOTHED SPECTRAL MAPPING.....	3935
<i>Vinay Kothapally, Wei Xia, Shahram Ghorbani, John H. L. Hansen, Wei Xue, Jing Huang</i>	
A ROBUST AND CASCADED ACOUSTIC ECHO CANCELLATION BASED ON DEEP LEARNING.....	3940
<i>Chenggang Zhang, Xueliang Zhang</i>	
GENERATIVE ADVERSARIAL NETWORK BASED ACOUSTIC ECHO CANCELLATION.....	3945
<i>Yi Zhang, Chengyun Deng, Shiqian Ma, Yongtao Sha, Hui Song, Xiangang Li</i>	
NONLINEAR RESIDUAL ECHO SUPPRESSION USING A RECURRENT NEURAL NETWORK	3950
<i>Lukas Pfeifenberger, Franz Pernkopf</i>	
INDEPENDENT ECHO PATH MODELING FOR STEREOPHONIC ACOUSTIC ECHO CANCELLATION	3955
<i>Yi Gao, Ian Liu, J. Zheng, Cheng Luo, Bin Li</i>	
NONLINEAR RESIDUAL ECHO SUPPRESSION BASED ON MULTI-STREAM CONV-TASNET.....	3959
<i>Hongsheng Chen, Teng Xiang, Kai Chen, Jing Lu</i>	
IMPROVING PARTITION-BLOCK-BASED ACOUSTIC ECHO CANCELER IN UNDER-MODELING SCENARIOS.....	3964
<i>Wenzhi Fan, Jing Lu</i>	
ATTENTION WAVE-U-NET FOR ACOUSTIC ECHO CANCELLATION	3969
<i>Jung-Hee Kim, Joon-Hyuk Chang</i>	

INTERSPEECH 2020 : SPEECH SYNTHESIS: TOWARD END-TO-END SYNTHESIS

FROM SPEAKER VERIFICATION TO MULTISPEAKER SPEECH SYNTHESIS, DEEP TRANSFER WITH FEEDBACK CONSTRAINT	3974
<i>Zexin Cai, Chuxiong Zhang, Ming Li</i>	

CAN SPEAKER AUGMENTATION IMPROVE MULTI-SPEAKER END-TO-END TTS?	3979
<i>Erica Cooper, Cheng-I Lai, Yusuke Yasuda, Junichi Yamagishi</i>	
NON-AUTOREGRESSIVE END-TO-END TTS WITH COARSE-TO-FINE DECODING	3984
<i>Tao Wang, Xuefei Liu, Jianhua Tao, Jiangyan Yi, Ruibo Fu, Zhengqi Wen</i>	
BI-LEVEL SPEAKER SUPERVISION FOR ONE-SHOT SPEECH SYNTHESIS	3989
<i>Tao Wang, Jianhua Tao, Ruibo Fu, Jiangyan Yi, Zhengqi Wen, Chunyu Qiang</i>	
NATURALNESS ENHANCEMENT WITH LINGUISTIC INFORMATION IN END-TO-END TTS USING UNSUPERVISED PARALLEL ENCODING	3994
<i>Alex Peiro-Lilja, Mireia Farrus</i>	
MOBOALIGNER: A NEURAL ALIGNMENT MODEL FOR NON-AUTOREGRESSIVE TTS WITH MONOTONIC BOUNDARY SEARCH	3999
<i>Naihan Li, Shujie Liu, Yanqing Liu, Sheng Zhao, Ming Liu, Ming Zhou</i>	
JDI-T: JOINTLY TRAINED DURATION INFORMED TRANSFORMER FOR TEXT-TO- SPEECH WITHOUT EXPLICIT ALIGNMENT	4004
<i>Dan Lim, Won Jang, Gyeonghwan O, Heayoung Park, Bongwan Kim, Jaesam Yoon</i>	
END-TO-END TEXT-TO-SPEECH SYNTHESIS WITH UNALIGNED MULTIPLE LANGUAGE UNITS BASED ON ATTENTION	4009
<i>Masashi Aso, Shinnosuke Takamichi, Hiroshi Saruwatari</i>	
ATTENTION FORCING FOR SPEECH SYNTHESIS	4014
<i>Qingyun Dou, Joshua Efiog, Mark J. F. Gales</i>	
TESTING THE LIMITS OF REPRESENTATION MIXING FOR PRONUNCIATION CORRECTION IN END-TO-END SPEECH SYNTHESIS	4019
<i>Jason Fong, Jason Taylor, Simon King</i>	
MULTISPEECH: MULTI-SPEAKER TEXT TO SPEECH WITH TRANSFORMER	4024
<i>Mingjian Chen, Xu Tan, Yi Ren, Jin Xu, Hao Sun, Sheng Zhao, Tao Qin</i>	
<u>INTERSPEECH 2020 : SPEECH ENHANCEMENT, BANDWIDTH EXTENSION AND HEARING AIDS</u>	
EXPLOITING CONIC AFFINITY MEASURES TO DESIGN SPEECH ENHANCEMENT SYSTEMS OPERATING IN UNSEEN NOISE CONDITIONS	4029
<i>Pavlos Papadopoulos, Shrikanth Narayanan</i>	
ADVERSARIAL DICTIONARY LEARNING FOR MONAURAL SPEECH ENHANCEMENT	4034
<i>Yunyun Ji, Longting Xu, Wei-Ping Zhu</i>	
SEMI-SUPERVISED SELF-PRODUCED SPEECH ENHANCEMENT AND SUPPRESSION BASED ON JOINT SOURCE MODELING OF AIR- AND BODY-CONDUCTED SIGNALS USING VARIATIONAL AUTOENCODER	4039
<i>Shogo Seki, Moe Takada, Tomoki Toda</i>	
SPATIAL COVARIANCE MATRIX ESTIMATION FOR REVERBERANT SPEECH WITH APPLICATION TO SPEECH ENHANCEMENT	4044
<i>Ran Weisman, Vladimir Tourbabin, Paul Calamia, Boaz Rafaely</i>	

A CROSS-CHANNEL ATTENTION-BASED WAVE-U-NET FOR MULTI-CHANNEL SPEECH ENHANCEMENT	4049
<i>Minh Tri Ho, Jinyoung Lee, Bong-Ki Lee, Dong Hoon Yi, Hong-Goo Kang</i>	
TINYLSTMS: EFFICIENT NEURAL SPEECH ENHANCEMENT FOR HEARING AIDS	4054
<i>Igor Fedorov, Marko Stamenovic, Carl Jensen, Li-Chia Yang, Ari Mandell, Yiming Gan, Matthew Mattina, Paul N. Whatmough</i>	
INTELLIGIBILITY ENHANCEMENT BASED ON SPEECH WAVEFORM MODIFICATION USING HEARING IMPAIRMENT	4059
<i>Shu Hikosaka, Shogo Seki, Tomoki Hayashi, Kazuhiro Kobayashi, Kazuya Takeda, Hideki Banno, Tomoki Toda</i>	
SPEAKER AND PHONEME-AWARE SPEECH BANDWIDTH EXTENSION WITH RESIDUAL DUAL-PATH NETWORK.....	4064
<i>Nana Hou, Chenglin Xu, Van Tung Pham, Joey Tianyi Zhou, Eng Siong Chng, Haizhou Li</i>	
MULTI-TASK LEARNING FOR END-TO-END NOISE-ROBUST BANDWIDTH EXTENSION	4069
<i>Nana Hou, Chenglin Xu, Joey Tianyi Zhou, Eng Siong Chng, Haizhou Li</i>	
PHASE-AWARE MUSIC SUPER-RESOLUTION USING GENERATIVE ADVERSARIAL NETWORKS	4074
<i>Shichao Hu, Bin Zhang, Beici Liang, Ethan Zhao, Simon Lui</i>	

INTERSPEECH 2020 : SPEECH EMOTION RECOGNITION III

LEARNING UTTERANCE-LEVEL REPRESENTATIONS WITH LABEL SMOOTHING FOR SPEECH EMOTION RECOGNITION.....	4079
<i>Jian Huang, Jianhua Tao, Bin Liu, Zheng Lian</i>	
REMOVING BIAS WITH RESIDUAL MIXTURE OF MULTI-VIEW ATTENTION FOR SPEECH EMOTION RECOGNITION.....	4084
<i>Md. Asif Jalal, Rosanna Milner, Thomas Hain, Roger K. Moore</i>	
ADAPTIVE DOMAIN-AWARE REPRESENTATION LEARNING FOR SPEECH EMOTION RECOGNITION.....	4089
<i>Weiquan Fan, Xiangmin Xu, Xiaofen Xing, Dongyan Huang</i>	
SPEECH EMOTION RECOGNITION WITH DISCRIMINATIVE FEATURE LEARNING.....	4094
<i>Huan Zhou, Kai Liu</i>	
USING SPEECH ENHANCEMENT PREPROCESSING FOR SPEECH EMOTION RECOGNITION IN REALISTIC NOISY CONDITIONS	4098
<i>Hengshun Zhou, Jun Du, Yan-Hui Tu, Chin-Hui Lee</i>	
COMPARISON OF GLOTTAL SOURCE PARAMETER VALUES IN EMOTIONAL VOWELS.....	4103
<i>Yongwei Li, Jianhua Tao, Bin Liu, Donna Erickson, Masato Akagi</i>	
LEARNING TO RECOGNIZE PER-RATER'S EMOTION PERCEPTION USING CO-RATER TRAINING STRATEGY WITH SOFT AND HARD LABELS	4108
<i>Huang-Cheng Chou, Chi-Chun Lee</i>	
EMPIRICAL INTERPRETATION OF SPEECH EMOTION PERCEPTION WITH ATTENTION BASED MODEL FOR SPEECH EMOTION RECOGNITION.....	4113
<i>Md. Asif Jalal, Rosanna Milner, Thomas Hain</i>	

INTERSPEECH 2020 : ACCOUSTIC PHONETICS OF L1-L2 AND OTHER INTERACTIONS

PHONETIC ACCOMMODATION OF L2 GERMAN SPEAKERS TO THE VIRTUAL LANGUAGE LEARNING TUTOR MIRABELLA	4118
<i>Iona Gessinger, Bernd Mobius, Bistra Andreeva, Eran Raveh, Ingmar Steiner</i>	
CHARACTERIZATION OF SINGAPOREAN CHILDREN'S ENGLISH: COMPARISONS TO AMERICAN AND BRITISH COUNTERPARTS USING ARCHETYPAL ANALYSIS.....	4123
<i>Yuling Gu, Nancy F. Chen</i>	
RHYTHMIC CONVERGENCE IN CANADIAN FRENCH VARIETIES?.....	4128
<i>Svetlana Kaminskaia</i>	
MALAYALAM-ENGLISH CODE-SWITCHED: GRAPHEME TO PHONEME SYSTEM.....	4133
<i>Sreeja Manghat, Sreeram Manghat, Tanja Schultz</i>	
ONGOING PHONOLOGIZATION OF WORD-FINAL VOICING ALTERNATIONS IN TWO ROMANCE LANGUAGES: ROMANIAN AND FRENCH.....	4138
<i>Mathilde Hutin, Adele Jatteau, Ioana Vasilescu, Lori Lamel, Martine Adda-Decker</i>	
CUES FOR PERCEPTION OF GENDER IN SYNTHETIC VOICES AND THE ROLE OF IDENTITY	4143
<i>Maxwell Hope, Jason Lilley</i>	
PHONETIC ENTRAINMENT IN COOPERATIVE DIALOGUES: A CASE OF RUSSIAN	4148
<i>Alla Menshikova, Daniil Kocharov, Tatiana Kachkovskaia</i>	
PROSODIC CHARACTERISTICS OF GENUINE AND MOCK (IM)POLITE MANDARIN UTTERANCES	4153
<i>Chengwei Xu, Wentao Gu</i>	
TONE VARIATIONS IN REGIONALLY ACCENTED MANDARIN	4158
<i>Yanping Li, Catherine T. Best, Michael D. Tyler, Denis Burnham</i>	
F0 PATTERNS IN MANDARIN STATEMENTS OF MANDARIN AND CANTONESE SPEAKERS	4163
<i>Yike Yang, Si Chen, Xi Chen</i>	

INTERSPEECH 2020 : CONVERSATIONAL SYSTEMS

SPEECHBERT: AN AUDIO-AND-TEXT JOINTLY LEARNED LANGUAGE MODEL FOR END-TO-END SPOKEN QUESTION ANSWERING.....	4168
<i>Yung-Sung Chuang, Chi-Liang Liu, Hung-Yi Lee, Lin-Shan Lee</i>	
AN AUDIO-ENRICHED BERT-BASED FRAMEWORK FOR SPOKEN MULTIPLE-CHOICE QUESTION ANSWERING	4173
<i>Chia-Chih Kuo, Shang-Bao Luo, Kuan-Yu Chen</i>	
ENTITY LINKING FOR SHORT TEXT USING STRUCTURED KNOWLEDGE GRAPH VIA MULTI-GRAINED TEXT MATCHING.....	4178
<i>Binxuan Huang, Han Wang, Tong Wang, Yue Liu, Yang Liu</i>	
SOUND-IMAGE GROUNDING BASED FOCUSING MECHANISM FOR EFFICIENT AUTOMATIC SPOKEN LANGUAGE ACQUISITION.....	4183
<i>Mingxin Zhang, Tomohiro Tanaka, Wenxin Hou, Shengzhou Gao, Takahiro Shinozaki</i>	

SEMI-SUPERVISED LEARNING FOR CHARACTER EXPRESSION OF SPOKEN DIALOGUE SYSTEMS	4188
<i>Kenta Yamamoto, Koji Inoue, Tatsuya Kawahara</i>	
DIMENSIONAL EMOTION PREDICTION BASED ON INTERACTIVE CONTEXT IN CONVERSATION.....	4193
<i>Xiaohan Shi, Sixia Li, Jianwu Dang</i>	
HRI-RNN: A USER-ROBOT DYNAMICS-ORIENTED RNN FOR ENGAGEMENT DECREASE DETECTION	4198
<i>Asma Atamna, Chloe Clavel</i>	
NEURAL REPRESENTATIONS OF DIALOGICAL HISTORY FOR IMPROVING UPCOMING TURN ACOUSTIC PARAMETERS PREDICTION	4203
<i>Simone Fuscone, Benoit Favre, Laurent Prevot</i>	
DETECTING DOMAIN-SPECIFIC CREDIBILITY AND EXPERTISE IN TEXT AND SPEECH.....	4208
<i>Shengli Hu</i>	

INTERSPEECH 2020 : THE ATTACKER'S PERSPECTIVE ON AUTOMATIC SPEAKER VERIFICATION

THE ATTACKER'S PERSPECTIVE ON AUTOMATIC SPEAKER VERIFICATION: AN OVERVIEW.....	4213
<i>Rohan Kumar Das, Xiaohai Tian, Tomi Kinnunen, Haizhou Li</i>	
EXTRAPOLATING FALSE ALARM RATES IN AUTOMATIC SPEAKER VERIFICATION	4218
<i>Alexey Sholokhov, Tomi Kinnunen, Ville Vestman, Kong Aik Lee</i>	
SELF-SUPERVISED SPOOFING AUDIO DETECTION SCHEME.....	4223
<i>Ziyue Jiang, Hongcheng Zhu, Li Peng, Wenbing Ding, Yanzhen Ren</i>	
INAUDIBLE ADVERSARIAL PERTURBATIONS FOR TARGETED ATTACK IN SPEAKER RECOGNITION.....	4228
<i>Qing Wang, Pengcheng Guo, Lei Xie</i>	
X-VECTORS MEET ADVERSARIAL ATTACKS: BENCHMARKING ADVERSARIAL ROBUSTNESS IN SPEAKER VERIFICATION.....	4233
<i>Jesus Villalba, Yuekai Zhang, Najim Dehak</i>	
BLACK-BOX ATTACKS ON SPOOFING COUNTERMEASURES USING TRANSFERABILITY OF ADVERSARIAL EXAMPLES.....	4238
<i>Yuekai Zhang, Ziyang Jiang, Jesus Villalba, Najim Dehak</i>	

INTERSPEECH 2020 : SUMMARIZATION, SEMANTIC ANALYSIS AND CLASSIFICATION

MULTIMODAL EMOTION RECOGNITION USING CROSS-MODAL ATTENTION AND 1D CONVOLUTIONAL NEURAL NETWORKS.....	4243
<i>Krishna D. N., Ankita Patil</i>	
ABSTRACTIVE SPOKEN DOCUMENT SUMMARIZATION USING HIERARCHICAL MODEL WITH MULTI-STAGE ATTENTION DIVERSITY OPTIMIZATION.....	4248
<i>Potsawee Manakul, Mark J. F. Gales, Linlin Wang</i>	

IMPROVED LEARNING OF WORD EMBEDDINGS WITH WORD DEFINITIONS AND SEMANTIC INJECTION	4253
<i>Yichi Zhang, Yinpei Dai, Zhijian Ou, Huixin Wang, Junlan Feng</i>	
WAKE WORD DETECTION WITH ALIGNMENT-FREE LATTICE-FREE MMI.....	4258
<i>Yiming Wang, Hang Lv, Daniel Povey, Lei Xie, Sanjeev Khudanpur</i>	
IMPROVING VIETNAMESE NAMED ENTITY RECOGNITION FROM SPEECH USING WORD CAPITALIZATION AND PUNCTUATION RECOVERY MODELS	4263
<i>Thai Binh Nguyen, Quang Minh Nguyen, Thi Thu Hien Nguyen, Quoc Truong Do, Chi Mai Luong</i>	
END-TO-END NAMED ENTITY RECOGNITION FROM ENGLISH SPEECH.....	4268
<i>Hemant Yadav, Sreyan Ghosh, Yi Yu, Rajiv Ratn Shah</i>	
SEMANTIC COMPLEXITY IN END-TO-END SPOKEN LANGUAGE UNDERSTANDING	4273
<i>Joseph P. McKenna, Samridhi Choudhary, Michael Saxon, Grant P. Strimel, Athanasios Mouchtaris</i>	
ANALYSIS OF DISFLUENCY IN CHILDREN'S SPEECH.....	4278
<i>Trang Tran, Morgan Tinkler, Gary Yeung, Abeer Alwan, Mari Ostendorf</i>	
REPRESENTATION BASED META-LEARNING FOR FEW-SHOT SPOKEN INTENT RECOGNITION.....	4283
<i>Ashish Mittal, Samarth Bharadwaj, Shreya Khare, Saneem Chemmengath, Karthik Sankaranarayanan, Brian Kingsbury</i>	
COMPLEMENTARY LANGUAGE MODEL AND PARALLEL BI-LRNN FOR FALSE TRIGGER MITIGATION.....	4288
<i>Rishika Agarwal, Xiaochuan Niu, Pranay Dighe, Srikanth Vishnubhotla, Sameer Badaskar, Devang Naik</i>	
<u>INTERSPEECH 2020 : SPEAKER RECOGNITION II</u>	
SPEAKER-UTTERANCE DUAL ATTENTION FOR SPEAKER AND UTTERANCE VERIFICATION	4293
<i>Tianchi Liu, Rohan Kumar Das, Maulik Madhavi, Shengmei Shen, Haizhou Li</i>	
ADVERSARIAL SEPARATION AND ADAPTATION NETWORK FOR FAR-FIELD SPEAKER VERIFICATION	4298
<i>Lu Yi, Man-Wai Mak</i>	
MIRNET: LEARNING MULTIPLE IDENTITIES REPRESENTATIONS IN OVERLAPPED SPEECH	4303
<i>Hyewon Han, Soo-Whan Chung, Hong-Goo Kang</i>	
STRATEGIES FOR END-TO-END TEXT-INDEPENDENT SPEAKER VERIFICATION.....	4308
<i>Weiwei Lin, Man-Wai Mak, Jen-Tzung Chien</i>	
WHY DID THE X-VECTOR SYSTEM MISS A TARGET SPEAKER? IMPACT OF ACOUSTIC MISMATCH UPON TARGET SCORE ON VOXCELEB DATA	4313
<i>Rosa Gonzalez Hautamaki, Tomi Kinnunen</i>	
VARIABLE FRAME RATE-BASED DATA AUGMENTATION TO HANDLE SPEAKING-STYLE VARIABILITY FOR AUTOMATIC SPEAKER VERIFICATION	4318
<i>Amber Afshan, Jinxi Guo, Soo Jin Park, Vijay Ravi, Alan McCree, Abeer Alwan</i>	

A MACHINE OF FEW WORDS: INTERACTIVE SPEAKER RECOGNITION WITH REINFORCEMENT LEARNING	4323
<i>Mathieu Seurin, Florian Strub, Philippe Preux, Olivier Pietquin</i>	
IMPROVING ON-DEVICE SPEAKER VERIFICATION USING FEDERATED LEARNING WITH PRIVACY	4328
<i>Filip Granqvist, Matt Seigel, Rogier Van Dalen, Aine Cahill, Stephen Shum, Matthias Paulik</i>	
NEURAL PLDA MODELING FOR END-TO-END SPEAKER VERIFICATION	4333
<i>Shreyas Ramoji, Prashant Krishnan, Sriram Ganapathy</i>	

INTERSPEECH 2020 : GENERAL TOPICS IN SPEECH RECOGNITION

STATE SEQUENCE POOLING TRAINING OF ACOUSTIC MODELS FOR KEYWORD SPOTTING.....	4338
<i>Kuba Lopatka, Tobias Bocklet</i>	
TRAINING KEYWORD SPOTTING MODELS ON NON-IID DATA WITH FEDERATED LEARNING.....	4343
<i>Andrew Hard, Kurt Partridge, Cameron Nguyen, Niranjan Subrahmanya, Aishanee Shah, Pai Zhu, Ignacio Lopez Moreno, Rajiv Mathews</i>	
CLASS LM AND WORD MAPPING FOR CONTEXTUAL BIASING IN END-TO-END ASR	4348
<i>Rongqing Huang, Ossama Abdel-Hamid, Xinwei Li, Gunnar Evermann</i>	
DO END-TO-END SPEECH RECOGNITION MODELS CARE ABOUT CONTEXT?	4352
<i>Lasse Borgholt, Jakob D. Havtorn, Zeljko Agic, Anders Sogaard, Lars Maaloe, Christian Igel</i>	
UTTERANCE CONFIDENCE MEASURE FOR END-TO-END SPEECH RECOGNITION WITH APPLICATIONS TO DISTRIBUTED SPEECH RECOGNITION SCENARIOS	4357
<i>Ankur Kumar, Sachin Singh, Dhananjaya Gowda, Abhinav Garg, Shatrughan Singh, Chanwoo Kim</i>	

VOLUME 7

SPEAKER CODE BASED SPEAKER ADAPTIVE TRAINING USING MODEL AGNOSTIC META-LEARNING	4362
<i>Huaxin Wu, Genshun Wan, Jia Pan</i>	
DOMAIN ADAPTATION USING CLASS SIMILARITY FOR ROBUST SPEECH RECOGNITION.....	4367
<i>Han Zhu, Jiangjiang Zhao, Yuling Ren, Li Wang, Pengyuan Zhang</i>	
INCREMENTAL MACHINE SPEECH CHAIN TOWARDS ENABLING LISTENING WHILE SPEAKING IN REAL-TIME.....	4372
<i>Sashi Novitasari, Andros Tjandra, Tomoya Yanagita, Sakriani Sakti, Satoshi Nakamura</i>	
CONTEXT-DEPENDENT ACOUSTIC MODELING WITHOUT EXPLICIT PHONE CLUSTERING	4377
<i>Tina Raissi, Eugen Beck, Ralf Schluter, Hermann Ney</i>	
VOICE CONVERSION BASED DATA AUGMENTATION TO IMPROVE CHILDREN'S SPEECH RECOGNITION IN LIMITED DATA SCENARIO.....	4382
<i>S. Shahnawazuddin, Nagaraj Adiga, Kunal Kumar, Aayushi Poddar, Waquar Ahmad</i>	

INTERSPEECH 2020 : SPEECH SYNTHESIS: PROSODY MODELING

COPYCAT: MANY-TO-MANY FINE-GRAINED PROSODY TRANSFER FOR NEURAL TEXT-TO-SPEECH.....	4387
<i>Sri Karlapati, Alexis Moinet, Arnaud Joly, Viacheslav Klimkov, Daniel Saez-Trigueros, Thomas Drugman</i>	
JOINT DETECTION OF SENTENCE STRESS AND PHRASE BOUNDARY FOR PROSODY.....	4392
<i>Binghuai Lin, Liyuan Wang, Xiaoli Feng, Jinsong Zhang</i>	
TRANSFER LEARNING OF THE EXPRESSIVITY USING FLOW METRIC LEARNING IN MULTISPEAKER TEXT-TO-SPEECH SYNTHESIS	4397
<i>Ajinkya Kulkarni, Vincent Colotte, Denis Jouviet</i>	
SPEAKING SPEED CONTROL OF END-TO-END SPEECH SYNTHESIS USING SENTENCE-LEVEL CONDITIONING	4402
<i>Jae-Sung Bae, Hanbin Bae, Young-Sun Joo, Junmo Lee, Gyeong-Hoon Lee, Hoon-Young Cho</i>	
DYNAMIC PROSODY GENERATION FOR SPEECH SYNTHESIS USING LINGUISTICS-DRIVEN ACOUSTIC EMBEDDING SELECTION.....	4407
<i>Shubhi Tyagi, Marco Nicolis, Jonas Rohnke, Thomas Drugman, Jaime Lorenzo-Trueba</i>	
IMPROVING THE PROSODY OF RNN-BASED ENGLISH TEXT-TO-SPEECH SYNTHESIS BY INCORPORATING A BERT MODEL.....	4412
<i>Tom Kenter, Manish Sharma, Rob Clark</i>	
IMPROVED PROSODY FROM LEARNED F0 CODEBOOK REPRESENTATIONS FOR VQ-VAE SPEECH WAVEFORM RECONSTRUCTION.....	4417
<i>Yi Zhao, Haoyu Li, Cheng-I Lai, Jennifer Williams, Erica Cooper, Junichi Yamagishi</i>	
PROSODY LEARNING MECHANISM FOR SPEECH SYNTHESIS SYSTEM WITHOUT TEXT LENGTH LIMIT.....	4422
<i>Zhen Zeng, Jianzong Wang, Ning Cheng, Jing Xiao</i>	
DISCRIMINATIVE METHOD TO EXTRACT COARSE PROSODIC STRUCTURE AND ITS APPLICATION FOR STATISTICAL PHRASE/ACCENT COMMAND ESTIMATION	4427
<i>Yuma Shirahata, Daisuke Saito, Nobuaki Minematsu</i>	
CONTROLLABLE NEURAL TEXT-TO-SPEECH SYNTHESIS USING INTUITIVE PROSODIC FEATURES	4432
<i>Tuomo Raitio, Ramya Rasipuram, Dan Castellani</i>	
CONTROLLABLE NEURAL PROSODY SYNTHESIS	4437
<i>Max Morrison, Zeyu Jin, Justin Salamon, Nicholas J. Bryan, Gautham J. Mysore</i>	
MULTI-REFERENCE NEURAL TTS STYLIZATION WITH ADVERSARIAL CYCLE CONSISTENCY	4442
<i>Matt Whitehill, Shuang Ma, Daniel McDuff, Yale Song</i>	
INTERACTIVE TEXT-TO-SPEECH SYSTEM VIA JOINT STYLE ANALYSIS.....	4447
<i>Yang Gao, Weiyi Zheng, Zhaojun Yang, Thilo Kohler, Christian Fuegen, Qing He</i>	

INTERSPEECH 2020 : LANGUAGE LEARNING

MOBILE-ASSISTED PROSODY TRAINING FOR LIMITED ENGLISH PROFICIENCY: LEARNER BACKGROUND AND SPEECH LEARNING PATTERN	4452
<i>Kevin Hirschi, Okim Kang, Catia Cucchiarini, John H. L. Hansen, Keelan Evanini, Helmer Strik</i>	
FINDING INTELLIGIBLE CONSONANT-VOWEL SOUNDS USING HIGH-QUALITY ARTICULATORY SYNTHESIS.....	4457
<i>Daniel R. Van Niekerk, Anqi Xu, Branislav Gerazov, Paul K. Krug, Peter Birkholz, Yi Xu</i>	
AUDIOVISUAL CORRESPONDENCE LEARNING IN HUMANS AND MACHINES	4462
<i>Venkat Krishnamohan, Akshara Soman, Anshul Gupta, Sriram Ganapathy</i>	
PERCEPTION OF ENGLISH FRICATIVES AND AFFRICATES BY ADVANCED CHINESE LEARNERS OF ENGLISH	4467
<i>Yizhou Lan</i>	
PERCEPTION OF JAPANESE CONSONANT LENGTH BY NATIVE SPEAKERS OF KOREAN DIFFERING IN JAPANESE LEARNING EXPERIENCE	4471
<i>Kimiko Tsukada, Joo-Yeon Kim, Jeong-Im Han</i>	
AUTOMATIC DETECTION OF PHONOLOGICAL ERRORS IN CHILD SPEECH USING SIAMESE RECURRENT AUTOENCODER	4476
<i>Si-Ioi Ng, Tan Lee</i>	
A COMPARISON OF ENGLISH RHYTHM PRODUCED BY NATIVE AMERICAN SPEAKERS AND MANDARIN ESL PRIMARY SCHOOL LEARNERS	4481
<i>Hongwei Ding, Binghuai Lin, Liyuan Wang, Hui Wang, Ruomei Fang</i>	
CROSS-LINGUISTIC INTERACTION BETWEEN PHONOLOGICAL CATEGORIZATION AND ORTHOGRAPHY PREDICTS PROSODIC EFFECTS IN THE ACQUISITION OF PORTUGUESE LIQUIDS BY L1-MANDARIN LEARNERS.....	4486
<i>Chao Zhou, Silke Hamann</i>	
CROSS-LINGUISTIC PERCEPTION OF UTTERANCES WITH WILLINGNESS AND RELUCTANCE IN MANDARIN BY KOREAN L2 LEARNERS.....	4491
<i>Wenqian Li, Jung-Yueh Tu</i>	

INTERSPEECH 2020 : SPEECH ENHANCEMENT

SPEECH ENHANCEMENT BASED ON BEAMFORMING AND POST-FILTERING BY COMBINING PHASE INFORMATION.....	4496
<i>Rui Cheng, Changchun Bao</i>	
A NOISE-AWARE MEMORY-ATTENTION NETWORK ARCHITECTURE FOR REGRESSION-BASED SPEECH ENHANCEMENT	4501
<i>Yu-Xuan Wang, Jun Du, Li Chai, Chin-Hui Lee, Jia Pan</i>	
HIFI-GAN: HIGH-FIDELITY DENOISING AND DEREVERBERATION BASED ON SPEECH DEEP FEATURES IN ADVERSARIAL NETWORKS.....	4506
<i>Jiaqi Su, Zeyu Jin, Adam Finkelstein</i>	

LEARNING COMPLEX SPECTRAL MAPPING FOR SPEECH ENHANCEMENT WITH IMPROVED CROSS-CORPUS GENERALIZATION	4511
<i>Ashutosh Pandey, Deliang Wang</i>	
SPEECH ENHANCEMENT WITH STOCHASTIC TEMPORAL CONVOLUTIONAL NETWORKS.....	4516
<i>Julius Richter, Guillaume Carbajal, Timo Gerkmann</i>	
VISUAL SPEECH IN REAL NOISY ENVIRONMENTS (VISION): A NOVEL BENCHMARK DATASET AND DEEP LEARNING-BASED BASELINE SYSTEM.....	4521
<i>Mandar Gogate, Kia Dashtipour, Amir Hussain</i>	
SPARSE MIXTURE OF LOCAL EXPERTS FOR EFFICIENT SPEECH ENHANCEMENT	4526
<i>Aswin Sivaraman, Minje Kim</i>	
IMPROVED SPEECH ENHANCEMENT USING TCN WITH MULTIPLE ENCODER-DECODER LAYERS.....	4531
<i>Vinith Kishore, Nitya Tiwari, Periyasamy Paramasivam</i>	
JOINT TRAINING FOR SIMULTANEOUS SPEECH DENOISING AND DEREVERBERATION WITH DEEP EMBEDDING REPRESENTATIONS.....	4536
<i>Cunhang Fan, Jianhua Tao, Bin Liu, Jiangyan Yi, Zhengqi Wen</i>	
UNSUPERVISED ROBUST SPEECH ENHANCEMENT BASED ON ALPHA-STABLE FAST MULTICHANNEL NONNEGATIVE MATRIX FACTORIZATION	4541
<i>Mathieu Fontaine, Kouhei Sekiguchi, Aditya Arie Nugraha, Kazuyoshi Yoshii</i>	
 <u>INTERSPEECH 2020 : SPEECH IN HEALTH II</u>	
SQUEEZE FOR SNEEZE: COMPACT NEURAL NETWORKS FOR COLD AND FLU RECOGNITION.....	4546
<i>Merlin Albes, Zhao Ren, Bjorn W. Schuller, Nicholas Cummins</i>	
EXTENDED STUDY ON THE USE OF VOCAL TRACT VARIABLES TO QUANTIFY NEUROMOTOR COORDINATION IN DEPRESSION	4551
<i>Nadee Seneviratne, James R. Williamson, Adam C. Lammert, Thomas F. Quatieri, Carol Espy-Wilson</i>	
AFFECTIVE CONDITIONING ON HIERARCHICAL ATTENTION NETWORKS APPLIED TO DEPRESSION DETECTION FROM TRANSCRIBED CLINICAL INTERVIEWS.....	4556
<i>Danai Xezonaki, Georgios Paraskevopoulos, Alexandros Potamianos, Shrikanth Narayanan</i>	
DOMAIN ADAPTATION FOR ENHANCING SPEECH-BASED DEPRESSION DETECTION IN NATURAL ENVIRONMENTAL CONDITIONS USING DILATED CNNs	4561
<i>Zhaocheng Huang, Julien Epps, Dale Joachim, Brian Stasak, James R. Williamson, Thomas F. Quatieri</i>	
MAKING A DISTINCTION BETWEEN SCHIZOPHRENIA AND BIPOLAR DISORDER BASED ON TEMPORAL PARAMETERS IN SPONTANEOUS SPEECH.....	4566
<i>Gabor Gosztolya, Anita Bagi, Szilvia Szaloki, Istvan Szendi, Ildiko Hoffmann</i>	
PREDICTION OF SLEEPINESS RATINGS FROM VOICE BY MAN AND MACHINE.....	4571
<i>Mark Huckvale, Andras Beke, Mirei Ikushima</i>	

TONGUE AND LIP MOTION PATTERNS IN ALARYNGEAL SPEECH	4576
<i>Kristin J. Teplansky, Alan Wisler, Beiming Cao, Wendy Liang, Chad W. Whited, Ted Mau, Jun Wang</i>	
AUTOENCODER BOTTLENECK FEATURES WITH MULTI-TASK OPTIMISATION FOR IMPROVED CONTINUOUS DYSARTHIC SPEECH RECOGNITION	4581
<i>Zhengjun Yue, Heidi Christensen, Jon Barker</i>	
RAW SPEECH WAVEFORM BASED CLASSIFICATION OF PATIENTS WITH ALS, PARKINSON'S DISEASE AND HEALTHY CONTROLS USING CNN-BLSTM.....	4586
<i>Jhansi Mallela, Aravind Illa, Yamini Belur, Nalini Atchayaram, Ravi Yadav, Pradeep Reddy, Dipanjan Gope, Prasanta Kumar Ghosh</i>	
ASSESSMENT OF PARKINSON'S DISEASE MEDICATION STATE THROUGH AUTOMATIC SPEECH ANALYSIS.....	4591
<i>Anna Pompili, Ruben Solera-Urena, Alberto Abad, Rita Cardoso, Isabel Guimaraes, Margherita Fabbri, Isabel P. Martins, Joaquim Ferreira</i>	
<u>INTERSPEECH 2020 : SPEECH AND AUDIO QUALITY ASSESSMENT</u>	
IMPROVING REPLAY DETECTION SYSTEM WITH CHANNEL CONSISTENCY DENSENEXT FOR THE ASVSPOOF 2019 CHALLENGE	4596
<i>Chao Zhang, Junjie Cheng, Yanmei Gu, Huacan Wang, Jun Ma, Shaojun Wang, Jing Xiao</i>	
SUBJECTIVE QUALITY EVALUATION OF SPEECH SIGNALS TRANSMITTED VIA BPL-PLC WIRED SYSTEM.....	4601
<i>Przemyslaw Falkowski-Gilski, Grzegorz Debita, Marcin Habrych, Bogdan Miedzinski, Przemyslaw Jedlikowski, Bartosz Polnik, Jan Wandzio, Xin Wang</i>	
INVESTIGATING THE VISUAL LOMBARD EFFECT WITH GABOR BASED FEATURES	4606
<i>Waito Chiu, Yan Xu, Andrew Abel, Chun Lin, Zhengzheng Tu</i>	
EXPLORATION OF AUDIO QUALITY ASSESSMENT AND ANOMALY LOCALISATION USING ATTENTION MODELS	4611
<i>Qiang Huang, Thomas Hain</i>	
DEVELOPMENT OF A SPEECH QUALITY DATABASE UNDER UNCONTROLLED CONDITIONS.....	4616
<i>Alessandro Ragano, Emmanouil Benetos, Andrew Hines</i>	
EVALUATING THE RELIABILITY OF ACOUSTIC SPEECH EMBEDDINGS	4621
<i>Robin Algayres, Mohamed Salah Zaiem, Benoit Sagot, Emmanuel Dupoux</i>	
FRAME-LEVEL SIGNAL-TO-NOISE RATIO ESTIMATION USING DEEP LEARNING	4626
<i>Hao Li, Deliang Wang, Xueliang Zhang, Guanglai Gao</i>	
A PYRAMID RECURRENT NETWORK FOR PREDICTING CROWDSOURCED SPEECH-QUALITY RATINGS OF REAL-WORLD SIGNALS.....	4631
<i>Xuan Dong, Donald S. Williamson</i>	
EFFECT OF SPECTRAL COMPLEXITY REDUCTION AND NUMBER OF INSTRUMENTS ON MUSICAL ENJOYMENT WITH COCHLEAR IMPLANTS.....	4636
<i>Avamarie Brueggeman, John H. L. Hansen</i>	

SPECTRUM CORRECTION: ACOUSTIC SCENE CLASSIFICATION WITH MISMATCHED RECORDING DEVICES	4641
<i>Michal Kosmider</i>	

INTERSPEECH 2020 : PRIVACY AND SECURITY IN SPEECH COMMUNICATION

DISTRIBUTED SUMMATION PRIVACY FOR SPEECH ENHANCEMENT	4646
<i>Matt O'Connor, W. Bastiaan Kleijn</i>	
PERCEPTION OF PRIVACY MEASURED IN THE CROWD --- PAIRED COMPARISON ON THE EFFECT OF BACKGROUND NOISES	4651
<i>Anna Leschanowsky, Sneha Das, Tom Backstrom, Pablo Perez Zarazaga</i>	
HIDE AND SPEAK: TOWARDS DEEP NEURAL NETWORKS FOR SPEECH STEGANOGRAPHY	4656
<i>Felix Kreuk, Yossi Adi, Bhiksha Raj, Rita Singh, Joseph Keshet</i>	
DETECTING ADVERSARIAL EXAMPLES FOR SPEECH RECOGNITION VIA UNCERTAINTY QUANTIFICATION	4661
<i>Sina Daubener, Lea Schonherr, Asja Fischer, Dorothea Kolossa</i>	
PRIVACY GUARANTEES FOR DE-IDENTIFYING TEXT TRANSFORMATIONS	4666
<i>David Ifeoluwa Adelani, Ali Davody, Thomas Kleinbauer, Dietrich Klakow</i>	
DETECTING AUDIO ATTACKS ON ASR SYSTEMS WITH DROPOUT UNCERTAINTY	4671
<i>Tejas Jayashankar, Jonathan Le Roux, Pierre Moulin</i>	

INTERSPEECH 2020 : VOICE CONVERSION AND ADAPTATION II

VOICE TRANSFORMER NETWORK: SEQUENCE-TO-SEQUENCE VOICE CONVERSION USING TRANSFORMER WITH TEXT-TO-SPEECH PRETRAINING	4676
<i>Wen-Chin Huang, Tomoki Hayashi, Yi-Chiao Wu, Hirokazu Kameoka, Tomoki Toda</i>	
NONPARALLEL TRAINING OF EXEMPLAR-BASED VOICE CONVERSION SYSTEM USING INCA-BASED ALIGNMENT TECHNIQUE	4681
<i>Hitoshi Suda, Gaku Kotani, Daisuke Saito</i>	
ENHANCING INTELLIGIBILITY OF DYSARTHIC SPEECH USING GATED CONVOLUTIONAL-BASED VOICE CONVERSION SYSTEM	4686
<i>Chen-Yu Chen, Wei-Zhong Zheng, Syu-Siang Wang, Yu Tsao, Pei-Chun Li, Ying-Hui Lai</i>	
VQVC+: ONE-SHOT VOICE CONVERSION BY VECTOR QUANTIZATION AND U-NET ARCHITECTURE	4691
<i>Da-Yi Wu, Yen-Hao Chen, Hung-Yi Lee</i>	
COTATRON: TRANSCRIPTION-GUIDED SPEECH ENCODER FOR ANY-TO-MANY VOICE CONVERSION WITHOUT PARALLEL DATA	4696
<i>Seung-Won Park, Doo-Young Kim, Myun-Chul Joe</i>	
DYNAMIC SPEAKER REPRESENTATIONS ADJUSTMENT AND DECODER FACTORIZATION FOR SPEAKER ADAPTATION IN END-TO-END SPEECH SYNTHESIS	4701
<i>Ruibo Fu, Jianhua Tao, Zhengqi Wen, Jiangyan Yi, Tao Wang, Chunyu Qiang</i>	

ARVC: AN AUTO-REGRESSIVE VOICE CONVERSION SYSTEM WITHOUT PARALLEL TRAINING DATA.....	4706
<i>Zheng Lian, Zhengqi Wen, Xinyong Zhou, Songbai Pu, Shengkai Zhang, Jianhua Tao</i>	
IMPROVED ZERO-SHOT VOICE CONVERSION USING EXPLICIT CONDITIONING SIGNALS	4711
<i>Shahan Nercessian</i>	
NON-PARALLEL VOICE CONVERSION WITH FEWER LABELED DATA BY CONDITIONAL GENERATIVE ADVERSARIAL NETWORKS	4716
<i>Minchuan Chen, Weijian Hou, Jun Ma, Shaojun Wang, Jing Xiao</i>	
TRANSFERRING SOURCE STYLE IN NON-PARALLEL VOICE CONVERSION.....	4721
<i>Songxiang Liu, Yuewen Cao, Shiyin Kang, Na Hu, Xunying Liu, Dan Su, Dong Yu, Helen Meng</i>	
VOICE CONVERSION USING SPEECH-TO-SPEECH NEURO-STYLE TRANSFER.....	4726
<i>Ehab A. Albadawy, Siwei Lyu</i>	

INTERSPEECH 2020 : MULTILINGUAL AND CODE-SWITCHED ASR

IMPROVING CROSS-LINGUAL TRANSFER LEARNING FOR END-TO-END SPEECH RECOGNITION WITH SPEECH TRANSLATION.....	4731
<i>Changhan Wang, Juan Pino, Jiatao Gu</i>	
TRANSLITERATION BASED DATA AUGMENTATION FOR TRAINING MULTILINGUAL ASR ACOUSTIC MODELS IN LOW RESOURCE SETTINGS	4736
<i>Samuel Thomas, Kartik Audhkhasi, Brian Kingsbury</i>	
MULTILINGUAL SPEECH RECOGNITION WITH SELF-ATTENTION STRUCTURED PARAMETERIZATION.....	4741
<i>Yun Zhu, Parisa Haghani, Anshuman Tripathi, Bhuvana Ramabhadran, Brian Farris, Hainan Xu, Han Lu, Hasim Sak, Isabel Leal, Neeraj Gaur, Pedro J. Moreno, Qian Zhang</i>	
LATTICE-FREE MAXIMUM MUTUAL INFORMATION TRAINING OF MULTILINGUAL SPEECH RECOGNITION SYSTEMS	4746
<i>Srikanth Madikeri, Banriskhem K. Khonglah, Sibio Tong, Petr Motlicek, Herve Bourlard, Daniel Povey</i>	
MASSIVELY MULTILINGUAL ASR: 50 LANGUAGES, 1 MODEL, 1 BILLION PARAMETERS.....	4751
<i>Vineel Pratap, Anuroop Sriram, Paden Tomasello, Awni Hannun, Vitaliy Liptchinsky, Gabriel Synnaeve, Ronan Collobert</i>	
MULTILINGUAL SPEECH RECOGNITION USING LANGUAGE-SPECIFIC PHONEME RECOGNITION AS AUXILIARY TASK FOR INDIAN LANGUAGES	4756
<i>Hardik B. Sailor, Thomas Hain</i>	
STYLE VARIATION AS A VANTAGE POINT FOR CODE-SWITCHING.....	4761
<i>Khyathi Raghavi Chandu, Alan W. Black</i>	
BI-ENCODER TRANSFORMER NETWORK FOR MANDARIN-ENGLISH CODE-SWITCHING SPEECH RECOGNITION USING MIXTURE OF EXPERTS	4766
<i>Yizhou Lu, Mingkun Huang, Hao Li, Jiaqi Guo, Yanmin Qian</i>	

IMPROVING LOW RESOURCE CODE-SWITCHED ASR USING AUGMENTED CODE-SWITCHED TTS 4771
Yash Sharma, Basil Abraham, Karan Taneja, Preethi Jyothi

TOWARDS CONTEXT-AWARE END-TO-END CODE-SWITCHING SPEECH RECOGNITION 4776
Zimeng Qiu, Yiyuan Li, Xinjian Li, Florian Metze, William M. Campbell

INTERSPEECH 2020 : SPEECH AND VOICE DISORDERS

INCREASING THE INTELLIGIBILITY AND NATURALNESS OF ALARYNGEAL SPEECH USING VOICE CONVERSION AND SYNTHETIC FUNDAMENTAL FREQUENCY 4781
Tuan Dinh, Alexander Kain, Robin Samlan, Beiming Cao, Jun Wang

AUTOMATIC ASSESSMENT OF DYSARTHIC SEVERITY LEVEL USING AUDIO-VIDEO CROSS-MODAL APPROACH IN DEEP LEARNING 4786
Han Tong, Hamid Sharifzadeh, Ian McLoughlin

STAGED KNOWLEDGE DISTILLATION FOR END-TO-END DYSARTHIC SPEECH RECOGNITION AND SPEECH ATTRIBUTE TRANSCRIPTION 4791
Yuqin Lin, Longbiao Wang, Sheng Li, Jianwu Dang, Chenchen Ding

DYSARTHIC SPEECH RECOGNITION BASED ON DEEP METRIC LEARNING 4796
Yuki Takashima, Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Arika

AUTOMATIC GLOTTIS DETECTION AND SEGMENTATION IN STROBOSCOPIC VIDEOS USING CONVOLUTIONAL NETWORKS 4801
Divya Degala, Achuth Rao M. V., Rahul Krishnamurthy, Pebbili Gopikishore, Veeramani Priyadharshini, Prakash T. K., Prasanta Kumar Ghosh

ACOUSTIC FEATURE EXTRACTION WITH INTERPRETABLE DEEP NEURAL NETWORK FOR NEURODEGENERATIVE RELATED DISORDER CLASSIFICATION 4806
Yilin Pan, Bahman Mirheidari, Zehai Tu, Ronan O'Malley, Traci Walker, Annalena Venneri, Markus Reuber, Daniel Blackburn, Heidi Christensen

COSWARA --- A DATABASE OF BREATHING, COUGH, AND VOICE SOUNDS FOR COVID-19 DIAGNOSIS 4811
Neeraj Sharma, Prashant Krishnan, Rohit Kumar, Shreyas Ramoji, Srikanth Raj Chetupalli, Nirmala R., Prasanta Kumar Ghosh, Sriram Ganapathy

ACOUSTIC-BASED ARTICULATORY PHENOTYPES OF AMYOTROPHIC LATERAL SCLEROSIS AND PARKINSON'S DISEASE: TOWARDS AN INTERPRETABLE, HYPOTHESIS-DRIVEN FRAMEWORK OF MOTOR CONTROL 4816
Hannah P. Rowe, Sarah E. Gutz, Marc F. Maffei, Jordan R. Green

RECOGNISING EMOTIONS IN DYSARTHIC SPEECH USING TYPICAL SPEECH DATA 4821
Lubna Alhinti, Stuart Cunningham, Heidi Christensen

DETECTING AND ANALYSING SPONTANEOUS ORAL CANCER SPEECH IN THE WILD 4826
Bence Mark Halpern, Rob Van Son, Michiel Van Den Brekel, Odette Scharenborg

INTERSPEECH 2020 : THE ZERO RESOURCE SPEECH CHALLENGE 2020

THE ZERO RESOURCE SPEECH CHALLENGE 2020: DISCOVERING DISCRETE SUBWORD AND WORD UNITS.....	4831
<i>Ewan Dunbar, Julien Karadayi, Mathieu Bernard, Xuan-Nga Cao, Robin Algayres, Lucas Ondel, Laurent Besacier, Sakriani Sakti, Emmanuel Dupoux</i>	
VECTOR-QUANTIZED NEURAL NETWORKS FOR ACOUSTIC UNIT DISCOVERY IN THE ZEROSPEECH 2020 CHALLENGE.....	4836
<i>Benjamin Van Niekerk, Leanne Nortje, Herman Kamper</i>	
EXPLORATION OF END-TO-END SYNTHESISERS FOR ZERO RESOURCE SPEECH CHALLENGE 2020.....	4841
<i>Karthik Pandia D. S., Anusha Prakash, Mano Ranjith Kumar M., Hema A. Murthy</i>	
VECTOR QUANTIZED TEMPORALLY-AWARE CORRESPONDENCE SPARSE AUTOENCODERS FOR ZERO-RESOURCE ACOUSTIC UNIT DISCOVERY.....	4846
<i>Batuhan Gundogdu, Bolaji Yusuf, Mansur Yesilbursa, Murat Saraclar</i>	
TRANSFORMER VQ-VAE FOR UNSUPERVISED UNIT DISCOVERY AND SPEECH SYNTHESIS: ZEROSPEECH 2020 CHALLENGE.....	4851
<i>Andros Tjandra, Sakriani Sakti, Satoshi Nakamura</i>	
EXPLORING TTS WITHOUT T USING BIOLOGICALLY/PSYCHOLOGICALLY MOTIVATED NEURAL NETWORK MODULES (ZEROSPEECH 2020).....	4856
<i>Takashi Morita, Hiroki Koda</i>	
CYCLIC SPECTRAL MODELING FOR UNSUPERVISED UNIT DISCOVERY INTO VOICE CONVERSION WITH EXCITATION AND WAVEFORM MODELING.....	4861
<i>Patrick Lumban Tobing, Tomoki Hayashi, Yi-Chiao Wu, Kazuhiro Kobayashi, Tomoki Toda</i>	
UNSUPERVISED ACOUSTIC UNIT REPRESENTATION LEARNING FOR VOICE CONVERSION USING WAVENET AUTO-ENCODERS.....	4866
<i>Mingjie Chen, Thomas Hain</i>	
UNSUPERVISED DISCOVERY OF RECURRING SPEECH PATTERNS USING PROBABILISTIC ADAPTIVE METRICS.....	4871
<i>Okko Rasanen, Maria Andrea Cruz Blandon</i>	
SELF-EXPRESSING AUTOENCODERS FOR UNSUPERVISED SPOKEN TERM DISCOVERY.....	4876
<i>Saurabhchand Bhati, Jesus Villalba, Piotr Zelasko, Najim Dehak</i>	
PERCEPTIMATIC: A HUMAN SPEECH PERCEPTION BENCHMARK FOR UNSUPERVISED SUBWORD MODELLING.....	4881
<i>Juliette Millet, Ewan Dunbar</i>	
DECODING IMAGINED, HEARD, AND SPOKEN SPEECH: CLASSIFICATION AND REGRESSION OF EEG USING A 14-CHANNEL DRY-CONTACT MOBILE HEADSET.....	4886
<i>Jonathan Clayton, Scott Wellington, Cassia Valentini-Botinhao, Oliver Watts</i>	
GLOTTAL CLOSURE INSTANTS DETECTION FROM EGG SIGNAL BY CLASSIFICATION APPROACH.....	4891
<i>Gurunath Reddy M., K. Sreenivasa Rao, Partha Pratim Das</i>	

CLASSIFY IMAGINARY MANDARIN TONES WITH CORTICAL EEG SIGNALS..... 4896
Hua Li, Fei Chen

INTERSPEECH 2020 : LM ADAPTATION, LEXICAL UNITS AND PUNCTUATION

AUGMENTING IMAGES FOR ASR AND TTS THROUGH SINGLE-LOOP AND DUAL-
LOOP MULTIMODAL CHAIN FRAMEWORK..... 4901
Johanes Effendi, Andros Tjandra, Sakriani Sakti, Satoshi Nakamura

PUNCTUATION PREDICTION IN SPONTANEOUS CONVERSATIONS: CAN WE
MITIGATE ASR ERRORS WITH RETROFITTED WORD EMBEDDINGS?..... 4906
Lukasz Augustyniak, Piotr Szymanski, Mikolaj Morzy, Piotr Zelasko, Adrian Szymczak, Jan Mizgajski, Yishay Carmiel, Najim Dehak

MULTIMODAL SEMI-SUPERVISED LEARNING FRAMEWORK FOR PUNCTUATION
PREDICTION IN CONVERSATIONAL SPEECH 4911
Monica Sunkara, Srikanth Ronanki, Dhanush Bekal, Sravan Bodapati, Katrin Kirchhoff

EFFICIENT MDI ADAPTATION FOR N-GRAM LANGUAGE MODELS..... 4916
Ruizhe Huang, Ke Li, Ashish Arora, Daniel Povey, Sanjeev Khudanpur

IMPROVING TAIL PERFORMANCE OF A DELIBERATION E2E ASR MODEL USING A
LARGE TEXT CORPUS 4921
Cal Peyser, Sepand Mavandadi, Tara N. Sainath, James Apfel, Ruoming Pang, Shankar Kumar

LANGUAGE MODEL DATA AUGMENTATION BASED ON TEXT DOMAIN TRANSFER..... 4926
Atsunori Ogawa, Naohiro Tawara, Marc Delcroix

CONTEMPORARY POLISH LANGUAGE MODEL (VERSION 2) USING BIG DATA AND
SUB-WORD APPROACH 4931
Krzysztof Wolk

IMPROVING SPEECH RECOGNITION OF COMPOUND-RICH LANGUAGES..... 4936
Prabhat Pandey, Volker Leutnant, Simon Wiesler, Jahn Heymann, Daniel Willett

LANGUAGE MODELING FOR SPEECH ANALYTICS IN UNDER-RESOURCED
LANGUAGES..... 4941
Simone Wills, Pieter Uys, Charl Van Heerden, Etienne Barnard

INTERSPEECH 2020 : SPEECH IN HEALTH I

AN EARLY STUDY ON INTELLIGENT ANALYSIS OF SPEECH UNDER COVID-19:
SEVERITY, SLEEP QUALITY, FATIGUE, AND ANXIETY 4946
Jing Han, Kun Qian, Meishu Song, Zijiang Yang, Zhao Ren, Shuo Liu, Juan Liu, Huaiyuan Zheng, Wei Ji, Tomoya Koike, Xiao Li, Zixing Zhang, Yoshiharu Yamamoto, Bjorn W. Schuller

AN EVALUATION OF THE EFFECT OF ANXIETY ON SPEECH --- COMPUTATIONAL
PREDICTION OF ANXIETY FROM SUSTAINED VOWELS..... 4951
Alice Baird, Nicholas Cummins, Sebastian Schnieder, Jarek Krajewski, Bjorn W. Schuller

HYBRID NETWORK FEATURE EXTRACTION FOR DEPRESSION ASSESSMENT FROM SPEECH	4956
<i>Ziping Zhao, Qifei Li, Nicholas Cummins, Bin Liu, Haishuai Wang, Jianhua Tao, Bjorn W. Schuller</i>	
IMPROVING DETECTION OF ALZHEIMER'S DISEASE USING AUTOMATIC SPEECH RECOGNITION TO IDENTIFY HIGH-QUALITY SEGMENTS FOR MORE ROBUST FEATURE EXTRACTION.....	4961
<i>Yilin Pan, Bahman Mirheidari, Markus Reuber, Annalena Venneri, Daniel Blackburn, Heidi Christensen</i>	
CLASSIFICATION OF MANIFEST HUNTINGTON DISEASE USING VOWEL DISTORTION MEASURES.....	4966
<i>Amrit Romana, John Bandon, Noelle Carlozzi, Angela Roberts, Emily Mower Provost</i>	
PARKINSON'S DISEASE DETECTION FROM SPEECH USING SINGLE FREQUENCY FILTERING CEPSTRAL COEFFICIENTS	4971
<i>Sudarsana Reddy Kadiri, Rashmi Kethireddy, Paavo Alku</i>	
AUTOMATIC PREDICTION OF SPEECH INTELLIGIBILITY BASED ON X-VECTORS IN THE CONTEXT OF HEAD AND NECK CANCER	4976
<i>Sebastiao Quintas, Julie Maclair, Virginie Woisard, Julien Pinquier</i>	
SPECTRAL MOMENT AND DURATION OF BURST OF PLOSIVES IN SPEECH OF CHILDREN WITH HEARING IMPAIRMENT AND TYPICALLY DEVELOPING CHILDREN - -- A COMPARATIVE STUDY	4981
<i>Ajish K. Abraham, M. Pushpavathi, N. Sreedevi, A. Navya, C. M. Vikram, S. R. Mahadeva Prasanna</i>	
APHASIC SPEECH RECOGNITION USING A MIXTURE OF SPEECH INTELLIGIBILITY EXPERTS.....	4986
<i>Matthew Perez, Zakaria Aldeneh, Emily Mower Provost</i>	
AUTOMATIC DISCRIMINATION OF APRAXIA OF SPEECH AND DYSARTHRIA USING A MINIMALISTIC SET OF HANDCRAFTED FEATURES	4991
<i>Ina Kodrasi, Michaela Pernon, Marina Laganaro, Herve Bourlard</i>	
<u>INTERSPEECH 2020 : ASR NEURAL NETWORK ARCHITECTURES II --- TRANSFORMERS</u>	
WEAK-ATTENTION SUPPRESSION FOR TRANSFORMER BASED SPEECH RECOGNITION.....	4996
<i>Yangyang Shi, Yongqiang Wang, Chunyang Wu, Christian Fuegen, Frank Zhang, Duc Le, Ching-Feng Yeh, Michael L. Seltzer</i>	
CONV-TRANSFORMER TRANSDUCER: LOW LATENCY, LOW FRAME RATE, STREAMABLE END-TO-END SPEECH RECOGNITION	5001
<i>Wenyong Huang, Wenchao Hu, Yu Ting Yeung, Xiao Chen</i>	
IMPROVING TRANSFORMER-BASED SPEECH RECOGNITION WITH UNSUPERVISED PRE-TRAINING AND MULTI-TASK SEMANTIC KNOWLEDGE LEARNING	5006
<i>Song Li, Lin Li, Qingyang Hong, Lingling Liu</i>	
TRANSFORMER-BASED LONG-CONTEXT END-TO-END SPEECH RECOGNITION	5011
<i>Takaaki Hori, Niko Moritz, Chiori Hori, Jonathan Le Roux</i>	

SELF-AND-MIXED ATTENTION DECODER WITH DEEP ACOUSTIC STRUCTURE FOR TRANSFORMER-BASED LVCSR	5016
<i>Xinyuan Zhou, Grandee Lee, Emre Yilmaz, Yanhua Long, Jiaen Liang, Haizhou Li</i>	
UNIVERSAL SPEECH TRANSFORMER	5021
<i>Yingzhu Zhao, Chongjia Ni, Cheung-Chi Leung, Shafiq Joty, Eng Siong Chng, Bin Ma</i>	
SPIKE-TRIGGERED NON-AUTOREGRESSIVE TRANSFORMER FOR END-TO-END SPEECH RECOGNITION	5026
<i>Zhengkun Tian, Jiangyan Yi, Jianhua Tao, Ye Bai, Shuai Zhang, Zhengqi Wen</i>	
CROSS ATTENTION WITH MONOTONIC ALIGNMENT FOR SPEECH TRANSFORMER.....	5031
<i>Yingzhu Zhao, Chongjia Ni, Cheung-Chi Leung, Shafiq Joty, Eng Siong Chng, Bin Ma</i>	
CONFORMER: CONVOLUTION-AUGMENTED TRANSFORMER FOR SPEECH RECOGNITION.....	5036
<i>Anmol Gulati, James Qin, Chung-Cheng Chiu, Niki Parmar, Yu Zhang, Jiahui Yu, Wei Han, Shibo Wang, Zhengdong Zhang, Yonghui Wu, Ruoming Pang</i>	
EXPLORING TRANSFORMERS FOR LARGE-SCALE SPEECH RECOGNITION	5041
<i>Liang Lu, Changliang Liu, Jinyu Li, Yifan Gong</i>	
<u>INTERSPEECH 2020 : SPATIAL AUDIO</u>	
SPARSENESS-AWARE DOA ESTIMATION WITH MAJORIZATION MINIMIZATION	5046
<i>Masahito Togami, Robin Scheibler</i>	
SPATIAL RESOLUTION OF EARLY REFLECTION FOR SPEECH AND WHITE NOISE.....	5051
<i>Xiaoli Zhong, Hao Song, Xuejie Liu</i>	
EFFECT OF MICROPHONE POSITION MEASUREMENT ERROR ON RIR AND ITS IMPACT ON SPEECH INTELLIGIBILITY AND QUALITY	5056
<i>Aditya Raikar, Karan Nathwani, Ashish Panda, Sunil Kumar Koppurapu</i>	
ONLINE BLIND REVERBERATION TIME ESTIMATION USING CRNNS.....	5061
<i>Shuwen Deng, Wolfgang Mack, Emanuel A. P. Habets</i>	
SINGLE-CHANNEL BLIND DIRECT-TO-REVERBERATION RATIO ESTIMATION USING MASKING	5066
<i>Wolfgang Mack, Shuwen Deng, Emanuel A. P. Habets</i>	
THE IMPORTANCE OF TIME-FREQUENCY AVERAGING FOR BINAURAL SPEAKER LOCALIZATION IN REVERBERANT ENVIRONMENTS	5071
<i>Hanan Beit-On, Vladimir Tourbabin, Boaz Rafaely</i>	
ACOUSTIC SIGNAL ENHANCEMENT USING RELATIVE HARMONIC COEFFICIENTS: SPHERICAL HARMONICS DOMAIN APPROACH.....	5076
<i>Yonggang Hu, Prasanga N. Samarasinghe, Thushara D. Abhayapala</i>	
INSTANTANEOUS TIME DELAY ESTIMATION OF BROADBAND SIGNALS	5081
<i>B. H. V. S. Narayana Murthy, J. V. Satyanarayana, Nivedita Chennupati, B. Yegnanarayana</i>	
U-NET BASED DIRECT-PATH DOMINANCE TEST FOR ROBUST DIRECTION-OF-ARRIVAL ESTIMATION.....	5086
<i>Hao Wang, Kai Chen, Jing Lu</i>	

SOUND EVENT LOCALIZATION AND DETECTION BASED ON MULTIPLE DOA
BEAMFORMING AND MULTI-TASK LEARNING 5091
Wei Xue, Ying Tong, Chao Zhang, Guohong Ding, Xiaodong He, Bowen Zhou

Author Index