

2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP 2022)

**Singapore
11-14 December 2022**



**IEEE Catalog Number: CFP22583-POD
ISBN: 979-8-3503-9797-0**

**Copyright © 2022 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

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

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

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

IEEE Catalog Number:	CFP22583-POD
ISBN (Print-On-Demand):	979-8-3503-9797-0
ISBN (Online):	979-8-3503-9796-3

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

Table of Content

About the Conference	XIII
Message from the Conference Chairs	XIV
Message from the Program Chairs	XV
Conference Committees	XVII
Keynote Speeches	XXII
Tutorials	XXVI
Author Index	XXXIII
Papers	
Oral 1: Speech Recognition I	
OS1.1 An Ensemble Teacher-Student Learning Approach with Poisson Sub-sampling to Differential Privacy Preserving Speech Recognition <i>Chao-Han Huck Yang, Jun Qi, Sabato Marco Siniscalchi and Chin-Hui Lee</i>	1
OS1.2 Adaptive Attention Network with Domain Adversarial Training for Multi-Accent Speech Recognition <i>Yanbing Yang, Hao Shi, Yuqin Lin, Meng Ge, Longbiao Wang, Qingzhi Hou and Jianwu Dang</i>	6
OS1.3 Multilingual Zero Resource Speech Recognition Base on Self-Supervise Pre-Trained Acoustic Models <i>Haoyu Wang, Wei-Qiang Zhang, Hongbin Suo and Yulong Wan</i>	11
OS1.4 Towards Language-universal Mandarin-English Speech Recognition with Unsupervised Label Synchronous Adaptation <i>Song Li, Haoneng Luo, Wenxuan Hu, Yuan Liu, Shiliang Zhang, Lin Li and Qingyang Hong</i>	16
OS1.5 Sequence Distribution Matching for Unsupervised Domain Adaptation in ASR <i>Qingxuan Li, Han Zhu, Liuping Luo, Gaofeng Cheng, Pengyuan Zhang, Jiasong Sun and Yonghong Yan</i>	21
OS1.6 Improving Rare Words Recognition through Homophone Extension and Unified Writing for Low-resource Cantonese Speech Recognition <i>HoLam Chung, Junan Li, Pengfei Liu, Wai Kim Leung, Xixin Wu and Helen Meng</i>	26
Oral 2: Speech Production and Perception I	
OS2.1 Perception and production of Mandarin vowels by teenagers--blind and sighted <i>Moyu Chen, Jing Qi and Xiyu Wu</i>	31
OS2.2 The Production of Contrastive Focus by Children Learning Mandarin Chinese <i>Jing Lu and Ping Tang</i>	36

OS2.3 Production Characteristics of Vowels in Standard Chinese by Preschool Bilingual Teachers <i>Linjiao Pan and Yuan Jia</i>	41
OS2.4 Effects of Aspiration on Tone Production and Perception in Standard Chinese <i>Chong Cao and Aijun Li</i>	46
OS2.5 The Disyllabic Tone Production and Tone Context Effect in Mandarin-speaking Children with Cochlear Implants <i>Jingwen Cheng, Yingming Gao, Yuchen Yan, Xiaoli Feng, Binghuai Lin and Jinsong Zhang</i>	51
OS2.6 A preliminary ultrasonic investigation of tenseness in Northern Yi <i>Shuwen Chen</i>	56
Oral 3: Speech Synthesis	
OS3.1 Style-Label-Free: Cross-Speaker Style Transfer by Quantized VAE and Speaker-wise Normalization in Speech Synthesis <i>Chunyu Qiang, Peng Yang, Hao Che, Xiaorui Wang and Zhongyuan Wang</i>	61
OS3.2 Multi-speaker Multi-style Text-to-speech Synthesis with Single-speaker Single-style Training Data Scenarios <i>Qicong Xie, Tao Li, Xinsheng Wang, Zhichao Wang, Lei Xie, Guoqiao Yu and Guanglu Wan</i>	66
OS3.3 Robust MelGAN: A robust universal neural vocoder for high-fidelity TTS <i>Kun Song, Jian Cong, Xinsheng Wang, Yongmao Zhang, Lei Xie, Ning Jiang and Haiying Wu</i>	71
OS3.4 AccentSpeech: Learning Accent from Crowd-sourced Data for Target Speaker TTS with Accents <i>Yongmao Zhang, Zhichao Wang, Peiji Yang, Hongshen Sun, Zhisheng Wang and Lei Xie</i>	76
OS3.5 CorrectSpeech: A Fully Automated System for Speech Correction and Accent Reduction <i>Daxin Tan, Liqun Deng, Nianzu Zheng, Yu Ting Yeung, Xin Jiang, Xiao Chen and Tan Lee</i>	81
OS3.6 HILvoice: Human-in-the-Loop Style Selection for Elder-Facing Speech Synthesis <i>Xueyuan Chen, Qiaochu Huang, Xixin Wu, Zhiyong Wu and Helen Meng</i>	86
Oral 4: Voice Conversion & Spoofing Speech Detection	
OS4.1 End-to-End Voice Conversion with Information Perturbation <i>Qicong Xie, Shan Yang, Yi Lei, Lei Xie and Dan Su</i>	91
OS4.2 Mix-Guided VC: Any-to-many Voice Conversion by Combining ASR and TTS Bottleneck Features <i>Zeqing Zhao, Sifan Ma, Yan Jia, Jingyu Hou, Lin Yang and Junjie Wang</i>	96

OS4.3 A New Spoken Language Teaching Tech: Combining Multi-attention and AdaIN for One-shot Cross Language Voice Conversion <i>Dengfeng Ke, Wenhan Yao, Ruixin Hu, Qi Luo, Liangjie Huang, Qi Luo and Wentao Shu</i>	101
OS4.4 The Impact of Room Acoustics on Replay Speech Signal <i>Madhu R. Kamble and Hemant A. Patil</i>	105
OS4.5 Effect of Speaker-Microphone Proximity on Pop Noise: Continuous Wavelet Transform-Based Approach <i>Priyanka Gupta and Hemant A. Patil</i>	110
OS4.6 Synthetic Voice Detection and Audio Splicing Detection using SE-Res2Net-Conformer Architecture <i>Lei Wang, Benedict Yeoh and Jun Wah Ng</i>	115
OS4.7 Audio Splicing Localization: Can We Accurately Locate the Splicing Tampering? <i>Zhiping Zeng and Zhizheng Wu</i>	120
Oral 5: Speech Enhancement and Separation	
OS5.1 Masking-based Neural Beamformer for Multichannel Speech Enhancement <i>Shuai Nie, Shan Liang, Zhanlei Yang, Longshuai Xiao, Wenju Liu and Jianhua Tao</i>	125
OS5.2 Deep Multi-task Cascaded Acoustic Echo Cancellation and Noise Suppression <i>Junjie Li, Meng Ge, Longbiao Wang and Jianwu Dang</i>	130
OS5.3 Boosting the Performance of SpEx+ by Attention and Contextual Mechanism <i>Chenyi Li, Zhiyong Wu, Wei Rao, Yannan Wang and Helen Meng</i>	135
OS5.4 Assessing the Effect of Temporal Misalignment between the Probe and Processed Speech Signals on Objective Speech Quality Evaluation <i>Shangdi Liao and Fei Chen</i>	140
OS5.5 Speech-enhanced and Noise-aware Networks for Robust Speech Recognition <i>Hung-Shin Lee, Pin-Yuan Chen, Yao-Fei Cheng, Yu Tsao and Hsin-Min Wang</i>	145
OS5.6 Separate-to-Recognize: Joint Multi-target Speech Separation and Speech Recognition for Speaker-attributed ASR <i>Yuxiao Lin, Zhihao Du, Shiliang Zhang, Fan Yu, Zhou Zhao and Fei Wu</i>	150
OS5.7 Speech Enhancement Based on CycleGAN with Noise-informed Training <i>Wen-Yuan Ting, Syu-Siang Wang, Hsin-Li Chang, Borching Su and Yu Tsao</i>	155
Oral 6: Speech Recognition II	
OS6.1 Incorporating VAD into ASR System by Multi-task Learning <i>Meng Li, Yan Xia and Feng Lin</i>	160
OS6.2 Improving ASR in Reverberant Environments <i>Yen-Lun Liao, Chi-Han Lin, Ren-Yuan Lyu and Jyh-Shing Roger Jang</i>	165

OS6.3 3M: Multi-loss, Multi-path and Multi-level Neural Networks for speech recognition <i>Zhao You, Shulin Feng, Dan Su and Dong Yu</i>	170
OS6.4 Multi-Level Modeling Units for End-to-End Mandarin Speech Recognition <i>Yuting Yang, Binbin Du and Yuke Li</i>	175
OS6.5 Exploiting Single-Channel Speech for Multi-Channel End-to-End Speech Recognition: A Comparative Study <i>Keyu An, Ji Xiao and Zhijian Ou</i>	180
OS6.6 Ensemble and Re-ranking based on Language Models to Improve ASR <i>Shu-Fen Tsai, Shih-Chan Kuo, Ren-Yuan Lyu and Jyh-Shing Roger Jang</i>	185
Oral 7: Speech Production and Perception II	
OS7.1 Acoustic and Perceptual Study of Tones in Jin Chinese (Togtoh variety) <i>Yue Wang and Wen Liu</i>	190
OS7.2 Acoustic-perceptual correlates of whispered Mandarin consonants <i>Min Xu, Jing Shao, Hongwei Ding and Lan Wang</i>	195
OS7.3 Bilingual Advantage? Perception of the Japanese Consonant Length Contrast by Monolingual vs Bilingual Speakers of Mongolian <i>Kimiko Tsukada, Yurong and Badmaavanchin Munguntsetseg</i>	200
OS7.4 Multichannel Emotional Perception in Chinese Female: Faces, Voices and Bodies <i>Ruiqi Ge and Xiyu Wu</i>	205
OS7.5 Coda Nasal Perception in Wenzhou Wu and Rugao Mandarin by Native Speakers of Standard Mandarin <i>Yanyang Chen, Xinya Zhang, Ying Chen and Jiazheng Wang</i>	210
OS7.6 Objective Hand Complexity Comparison between Two Mandarin Chinese Cued Speech Systems <i>Li Liu, Gang Feng, Xiaoxi Ren and Xianping Ma</i>	215
Oral 8: Speech Synthesis & Speaker Embedding	
OS8.1 Rhythm-controllable Attention with High Robustness for Long Sentence Speech Synthesis <i>Dengfeng Ke, Yayue Deng, Yukang Jia, Jinlong Xue, Qi Luo, Ya Li, Jianqing Sun, Jiaen Liang and Binghuai Lin</i>	220
OS8.2 AdaptiveFormer : A Few-shot Speaker Adaptive Speech Synthesis Model based on FastSpeech2 <i>Dengfeng Ke, Ruixin Hu, Qi Luo, Liangjie Huang, WenHan Yao, Wentao Shu, Jinsong Zhang and Yanlu Xie</i>	225
OS8.3 ECAPA-TDNN for Multi-speaker Text-to-speech Synthesis <i>Jinlong Xue, Yayue Deng, Yichen Han, Ya Li, Jianqing Sun and Jiaen Liang</i>	230
OS8.4 Low-Resource Speech Synthesis with Speaker-Aware Embedding <i>Li-Jen Yang, I-Ping Yeh and Jen-Tzung Chien</i>	235

OS8.5 A Phone-Level Speaker Embedding Extraction Framework with Multi-Gate Mixture-of-Experts Based Multi-Task Learning <i>Zhijunyi Yang, Mengjie Du, Rongfeng Su, Xiaokang Liu, Nan Yan and Lan Wang</i>	240
---	-----

OS8.6 Shuffle is What You Need <i>Wan Lin, Lantian Li and Dong Wang</i>	245
---	-----

Oral 9: Multimodality

OS9.1 Deep Learning Based Audio-Visual Multi-Speaker DOA Estimation Using Permutation-Free Loss Function <i>Qing Wang, Hang Chen, Ya Jiang, Zhe Wang, Yuyang Wang, Jun Du and Chin-Hui Lee</i>	250
--	-----

OS9.2 Multi-Task Joint Learning for Embedding Aware Audio-Visual Speech Enhancement <i>Chenxi Wang, Hang Chen, Jun Du, Baocai Yin and Jia Pan</i>	255
---	-----

OS9.3 Multimodal automatic speech fluency evaluation method for Putonghua Proficiency Test propositional speaking section <i>Jiajun Liu, Huazhen Meng, Yunfei Shen, Linna Zheng and Aishan Wumaier</i>	260
--	-----

OS9.4 Cantonese neural speech synthesis from found newscasting video data and its speaker adaptation <i>Raymond Chung</i>	265
---	-----

OS9.5 A Preliminary Study on Taiwanese OCR for Assisting Textual Database Construction from Historical Documents <i>Yuan-Fu Liao, Yu-Hsuan Huang, Matus Pleva, Daniel Hladek and Ming-Hsiang Su</i>	270
--	-----

OS9.6 Reconstruction of speech spectrogram based on non-invasive EEG signal <i>Di Zhou, Masashi Unoki, Gaoyan Zhang and Jianwu Dang</i>	275
---	-----

Oral 10: Speech Prosody

OS10.1 J-TranPSP: A Joint Transition-based Model for Prosodic Structure Prediction, Word Segmentation and PoS Tagging <i>Binbin Shen, Jian Luan, Shengyan Zhang, Quanbo Shen and Yujun Wang</i>	280
---	-----

OS10.2 A Mandarin Prosodic Boundary Prediction Model Based on Multi-Source Semi-Supervision <i>Peiyang Shi, Zengqiang Shang and Pengyuan Zhang</i>	285
--	-----

OS10.3 English lexical stresses in non-native speech under adverse conditions <i>Mosi He, Ting Zhang, Bin Li and Kin Cheung</i>	290
---	-----

OS10.4 Stress Gravity of Neutral Tone Words in Different Information Structures <i>Jingwen Huang and Aijun Li</i>	295
---	-----

OS10.5 Prosodic Encoding of Mandarin Chinese Intonation by Uygur Speakers in Declarative and Interrogative Sentences <i>Tong Li, Hui Feng and Yuan Jia</i>	300
--	-----

OS10.6 In-group Advantage for Chinese and English Emotional Prosody in Quiet and Noise Conditions <i>Yuhan Yan, Shanpeng Li and Ying Chen</i>	305
---	-----

Oral 11: Lightweight Model & Knowledge Distillation

OS11.1 Multi-Resolution Stacked 1D-CNN for Small-Footprint keyword Spotting with Two-Stage Detection <i>Jian Tang and Shaofei Xue</i>	310
---	-----

OS11.2 Lightweight End-to-End Deep Learning Model for Music Source Separation <i>Yao-Ting Wang, Yi-Xing Lin, Kai-Wen Liang, Tzu-Chiang Tai and Jia-Ching Wang</i>	315
---	-----

OS11.3 AdaVITS: Tiny VITS for Low Computing Resource Speaker Adaptation <i>Kun Song, Heyang Xue, Xinsheng Wang, Jian Cong, Yongmao Zhang, Lei Xie, Bing Yang, Xiong Zhang and Dan Su</i>	319
--	-----

OS11.4 Label-free Knowledge Distillation with Contrastive Loss for Light-weight Speaker Recognition <i>Zhiyuan Peng, Xuanji He, Ke Ding, Tan Lee and Guanglu Wan</i>	324
--	-----

OS11.5 Improving Speech Separation with Knowledge Distilled from Self-supervised Pre-trained Models <i>Bowen Qu, Chenda Li, Jinfeng Bai and Yanmin Qian</i>	329
---	-----

OS11.6 Text-Informed Knowledge Distillation for Robust Speech Enhancement and Recognition <i>Wei Wang, Wangyou Zhang, Shaoxiong Lin and Yanmin Qian</i>	334
---	-----

Oral 12: Speech Technology for Health

OS12.1 Prediction of Depression Severity Based on Transformer Encoder and CNN Model <i>Jiahao Lu, Bin Liu, Zheng Lian, Cong Cai, Jianhua Tao and Ziping Zhao</i>	339
--	-----

OS12.2 Depressive Tendency Recognition by Fusing Speech and Text Features: A Comparative Analysis <i>Yimin He, Xiaoyong Lu, Jingyi Yuan, Tao Pan and Yafan Wang</i>	344
---	-----

OS12.3 Medical Difficult Airway Detection using Speech Technology <i>Zhikai Zhou, Shuang Cao, Zhengyang Chen, Bei Liu, Ming Xia, Hong Jiang and Yanmin Qian</i>	349
---	-----

OS12.4 CUEMPATY: A Counseling Speech Dataset for Psychotherapy Research <i>Dehua Tao, Harold Chui, Sarah Luk and Tan Lee</i>	354
--	-----

OS12.5 Aphasia Detection for Cantonese-Speaking and Mandarin-Speaking Patients Using Pre-Trained Language Models <i>Ying Qin, Tan Lee, Anthony Pak Hin Kong and Feng Lin</i>	359
--	-----

OS12.6 Respiratory and laryngeal influences on voice in post-stroke dysarthria: a pilot study <i>Tinghao Zhao, Xiaoxia Du, Juan Liu, Rongfeng Su, Nan Yan and Lan Wang</i>	364
--	-----

Oral 13: Listening Comprehension of Machines and Humans

OS13.1 End-to-end speech topic classification based on pre-trained model Wavlm <i>Tengfei Cao, Liang He and Fangjing Niu</i>	369
OS13.2 BERT-based Chinese Medicine Named Entity Recognition Model Applied to Medication Reminder Dialogue System <i>Tsung-Hsien Yang, Matus Pleva, Daniel Hládek and Ming-Hsiang Su</i>	374
OS13.3 Dialogue scenario classification based on social factors <i>Yuning Liu, Di Zhou, Masashi Unoki, Jianwu Dang and Aijun Li</i>	379
OS13.4 BERT-LID: Leveraging BERT to Improve Spoken Language Identification <i>Yuting Nie, Junhong Zhao, Wei-Qiang Zhang and Jinfeng Bai</i>	384
OS13.5 An Exploratory Study for Quantifying the Contextual Information for Successful Chinese L2 Speech Comprehension <i>Rian Bao, Linkai Peng, Yuchen Yan and Jinsong Zhang</i>	389
OS13.6 The Contribution of Phonological and Fluency Factors to Chinese L2 Comprehensibility Ratings: A Case Study of Urdu-speaking Learners <i>Rian Bao, Linkai Peng, Yingming Gao and Jinsong Zhang</i>	394

Oral 14: Acoustic Phonetics & Prosody

OS14.1 An Acoustic Study on Fricative Vowel [iz] in Zhongwei Chinese <i>Xinyi Zhang and Wen Liu</i>	399
OS14.2 Acoustic Features of Consonants of Standard Chinese and English by Uyghur Native Speakers <i>Yuan Jia and Xintong Zuo</i>	404
OS14.3 A Study on Mandarin Chinese ""Bu"" Tone Sandhi Followed by English Words <i>Kaige Gao and Xiyu Wu</i>	409
OS14.4 An Entropy-based Study on the Acquisition of Mandarin Initial Consonants by Korean Learners <i>Xiaoli Feng, Yingming Gao, Jinsong Zhang and Yanchun Cao</i>	414
OS14.5 Impacts of aging on suprasegmental and segmental encoding of vocally-expressed confidence in Wuxi dialect <i>Yujie Ji, Qiqi Sun, Zhikang Peng and Xiaoming Jiang</i>	419
OS14.6 Acceptance of tonal and segmental variability correlates to inventory size in Mandarin Chinese <i>Julie Siying Chen and Stephen Politzer-Ahles</i>	424

Special Session 1: Data Augmentation in Speech Technologies

SS1.1 Dynamic Thresholding on FixMatch with Weak and Strong Data Augmentations for Sound Event Detection <i>Tanmay Khandelwal and Rohan Kumar Das</i>	428
--	-----

SS1.2 Data Augmentation for Infant Cry Classification
Aastha Kachhi, Shreya Chaturvedi, Hemant A. Patil and Dipesh Kumar Singh 433

SS1.3 Low Pass Filtering and Bandwidth Extension for Robust Anti-spoofing
 Countermeasure Against Codec Variabilities
Yikang Wang, Xingming Wang, Hiromitsu Nishizaki and Ming Li 438

SS1.4 Improving Speech Recognition with Augmented Synthesized Data and Conditional
 Model Training
Shaofei Xue, Jian Tang and Yazhu Liu 443

SS1.5 Speaking style compensation on synthetic audio for robust keyword spotting
Houjun Huang and Yanmin Qian 448

SS1.6 A Study on Joint Modeling and Data Augmentation of Multi-Modalities for Audio-
 Visual Scene Classification
*Qing Wang, Jun Du, Siyuan Zheng, Yunqing Li, Yajian Wang, Yuzhong Wu, Hu Hu, Chao-
 Han Huck Yang, Sabato Marco Siniscalchi, Yannan Wang and Chin-Hui Lee* 453

Special Session 2: Deep Noise Reduction

SS2.1 On the Use of Absolute Threshold of Hearing-based Loss for Full-band Speech
 Enhancement
Rohith Mars and Rohan Kumar Das 458

SS2.2 RAT: RNN-Attention Transformer for Speech Enhancement
Tailong Zhang, Shulin He, Hao Li and Xueliang Zhang 463

SS2.3 A Speech-Noise-Equilibrium Loss Function for Deep Learning-Based Speech
 Enhancement
Weitong Zhao, Fushi Xie, Kang Ouyang and Nengheng Zheng 468

SS2.4 Speakerfilter-Pro: an improved target speaker extractor combines the time domain
 and frequency domain
Shulin He, Hao Li and Xueliang Zhang 473

SS2.5 Two-Branch Network with Selective Kernel Convolution for Time-Domain Speech
 Enhancement
Hui Li, Zhihua Huang and Chuangjian Guo 478

SS2.6 Optimizing Shoulder to Shoulder: A Coordinated Sub-Band Fusion Model for Full-
 Band Speech Enhancement
Guochen Yu, Andong Li, Wenzhe Liu, Chengshi Zheng, Yutian Wang and Hui Wang 483

Challenge 1: Conversational Short-Phrase Speaker Diarization Challenge (CSSD)

GC1.1 The Conversational Short-phrase Speaker Diarization (CSSD) Task: Dataset,
 Evaluation Metric and Baselines
*Gaofeng Cheng, Yifan Chen, Runyan Yang, Qingxuan Li, Zehui Yang, Lingxuan Ye,
 Pengyuan Zhang, Qingqing Zhang, Lei Xie, Yanmin Qian, Kong Aik Lee and Yonghong Yan*
 488

GC1.2 Spectral Clustering Based EEND-vector Clustering: A Robust System Fine-tuned on Simulated Conversations
Kai Li 493

GC1.3 The X-Lance Speaker Diarization System for the Conversational Short-phrase Speaker Diarization Challenge 2022
Tao Liu, Xu Xiang, Zhengyang Chen, Bing Han, Kai Yu and Yanmin Qian 498

GC1.4 TSUP Speaker Diarization System for Conversational Short-phrase Speaker Diarization Challenge
Bowen Pang, Huan Zhao, Gaosheng Zhang, Xiaoyue Yang, Yang Sun, Li Zhang, Qing Wang and Lei Xie 502

Challenge 2: Intelligent Cockpit Speech Recognition Challenge (ICSRC)

GC2.1 The ISCSLP 2022 Intelligent Cockpit Speech Recognition Challenge (ICSRC): Dataset, Tracks, Baseline and Results
Ao Zhang, Fan Yu, Kaixun Huang, Lei Xie, Longbiao Wang, Eng Siong Chng, Hui Bu, Binbin Zhang, Wei Chen and Xin Xu 507

GC2.2 The FawAI ASR System for the ISCSLP 2022 Intelligent Cockpit Speech Recognition Challenge
Yujia Sun, Bing Ge, Bo Chen, Zhen Fu, Jinxin He, Hongwei Gao and Xue Wang 512

GC2.3 LeVoice ASR Systems for the ISCSLP 2022 Intelligent Cockpit Speech Recognition Challenge
Yan Jia, Mi Hong, Jingyu Hou, Kailong Ren, Sifan Ma, Jin Wang, Yinglin Ji, Fangzhen Peng, Lin Yang and Junjie Wang 517

GC2.4 Efficient Conformer-Based CTC Model for Intelligent Cockpit Speech Recognition
Hanzhi Guo, Yunshu Chen, Xukang Xie, Gaopeng Xu and Wei Guo 522

Challenge 3: Chinese-English Code-Switching Automatic Speech Recognition (CSASR)

GC3.1 Summary on the ISCSLP 2022 Chinese-English Code-switching ASR Challenge
Shuhao Deng, Chengfei Li, Jinfeng Bai, Qingqing Zhang, Wei-Qiang Zhang, Runyan Yang, Gaofeng Cheng, Pengyuan Zhang and Yonghong Yan 527

GC3.2 The NPU-ASLP System for The ISCSLP 2022 Magichub Code-Switching ASR Challenge
Yuhao Liang, Peikun Chen, Fan Yu, Xinfu Zhu, Tianyi Xu, Yingying Gao and Lei Xie 532

GC3.3 Hybrid CTC Language Identification Structure for Mandarin-English Code-Switching ASR
Hengxin Yin, Guangyu Hu, Fei Wang and Pengfei Ren 537