[Virtual oral session 12] Speech recognition and synthesis and spoken language understanding Document classification, information retrieval and cross-lingual retrieval Dialogue and Conversational Interaction	
Date	Day 1. Wednesday 12 Oct.
Time	18:25–19:15 KST
	11:25-12:15 CEST
	05:25-06:15 EDT
Place	101-102
Session Chair	Björn Schuller

Incorporating Instructional Prompts into A Unified Generative Framework for Joint Multiple Intent Detection and Slot Filling

Yangjun Wu, Han Wang, Dongxiang Zhang, Gang Chen and Hao Zhang

Unsupervised Multi-scale Expressive Speaking Style Modeling with Hierarchical Context Information for Audiobook Speech Synthesis

Xueyuan Chen, Shun Lei, Zhiyong Wu, Dong Xu, Weifeng Zhao and Helen Meng

BERT-Flow-VAE: A Weakly-supervised Model for Multi-Label Text Classification Ziwen Liu, Josep Grau-Bove and Scott Allan Orr

Recycle Your Wav2Vec2 Codebook: a Speech Perceiver for Keyword Spotting Guillermo Cámbara, Jordi Luque and Mireia Farrús

A Transformer-based Threshold-Free Framework for Multi-Intent NLU Lisung Chen, Nuo Chen, Yuexian Zou, Yong Wang and Xinzhong Sun