Development of automatic speech recognition and synthesis technologies to support language learning - the CUHK experience

Helen Meng

Director of Big Data Decision Analytics Research Centre, CUHK

This talk presents an ongoing research initiative in the development of speech technologies that strives to raise the efficacy of computer-aided pronunciation training, especially for Chinese learners of English. Our approach is grounded on the theory of language transfer and involves a systematic phonological comparison between the primary language (L1 being Chinese) and secondary language (L2 being English) to predict possible segmental and suprasegmental realizations that constitute mispronunciations in L2 English. The predictions are validated based on a specially designed corpus that consists of several hundred hours of L2 English speech. The speech data supports the development of automatic speech recognition technologies that can detect and diagnose mispronunciations. The diagnosis aims to support the design of pedagogical and remedial instructions, which involves text-to-speech synthesis technologies for corrective feedback generation in audiovisual forms. This talk offers an overview of the technologies, related experimental results and ongoing work as well as future plans.