

TITLE:

How can speech technologies support learners to improve their skills of speaking, listening, conversation and more?

LECTURER:

Prof. and Dr. Eng. Nobuaki MINEMATSU (The University of Tokyo)

ABSTRACT:

In the globalization era, not only students but also immigrant workers have to learn new languages for smooth oral communication in those languages. In this talk, the lecturer illustrates how speech technologies, i.e. speech synthesis, speech recognition, voice conversion, etc can support learners to improve their skills of speaking, listening, conversation, and more. Text does not show any prosodic structure explicitly and native speakers use their implicit knowledge on prosodic control to read aloud that text naturally. Implicit knowledge is difficult for teachers to explain explicitly and therefore prosody training is rare in classrooms. Text-to-speech systems often use a text-based prosody prediction module and this module is used effectively to teach prosodic control required to read given texts aloud explicitly to learners. In High Variability Phonetic Training (HVPT), teachers use speech stimuli with different ages, genders, accents, background noises, etc. Being exposed to those variabilities, learners can obtain robust listening skills. However, teachers prepare those stimuli manually. By introducing speech analysis and voice conversion techniques, those variabilities are easily enhanced. In the talk, an interesting example of adversarial training, which was originally used for machine learners and is newly introduced to human learners, and its effectiveness for acquiring robust listening skills are explained. Further, use of speech recognition technologies for shadowing assessment to improve parallel processing skills for conversation is described. In the lecturer's laboratory, a new project has started to realize a novel speech assessment framework, where not native-likeness but comprehensibility of learners' speech is mainly focused on for assessment. The lecturer shows recently obtained results of objective measurement of comprehensibility of learners' speech.

BIOGRAPHY:

Nobuaki MINEMATSU was born in Hyogo, Japan in 1966. He earned the doctor of Engineering in 1995 from The University of Tokyo and now is a full professor at Graduate School of Engineering there. When he was a high school boy, he wanted to become a language teacher, but finally became a professor of Engineering. He has a very wide interest in speech communication covering the areas of speech science and speech engineering, especially he has a good and practical knowledge on Computer-Aided Language Learning (CALL) and has applied a variety of speech technologies for CALL. He received paper awards from RISP, JSAI, ICIST, O-COCOSDA, IEICE in 2005, 2007, 2011, 2014, and 2016 and received an encouragement award from PSJ in 2014. He gave tutorial talks on CALL at APSIPA2011, INTERSPEECH2012, and CASTEL/J2017. He was a distinguished lecturer of APSIPA from 2015 to 2016. He has made remarkable contributions to technical and scientific societies. He served as editorial chair of IEICE from 2014 to 2016 and chair of SIG-SLP of IPSJ from 2016 to 2017, and has been serving as member of the PSJ council from 2016. He also served as secretary of Speech Prosody 2004, secretary of INTERSPEECH2010, co-organizer of SLaTE2010 (L2 workshop 2010), and program chair of O-CCOSDA2018. He will serve as general chair of Speech Prosody 2020. He is a member of IEEE, ISCA, SLaTE, IPA, APSIPA, IEICE, IPSJ, ASJ, PSJ, JSAI, LET, etc.