專 題 演 講

講 者：曹昱 博士（中央研究院資訊科技創新研究中心）

題 目： Enhancing Oral Communication with AI-Powered Assistive Technologies

大 綱：

This presentation is divided into three parts. Firstly, we will discuss our recent advancements in neural speech enhancement (SE), a critical element in various speech-related applications. The primary objective of SE is to enhance speech signals by mitigating distortions caused by additive and convoluted noises, thereby improving human-human and human-machine communication efficacy. We'll delve into the system architecture and fundamental theories behind neural SE approaches, as well as explore important directions aimed at achieving better performance. Moving on to the second part, we will focus on our recent progress in neural speech assessment (SA), which aims to effectively evaluate the quality and intelligibility of spoken audio—a crucial aspect in numerous speech-related applications. Traditionally, the evaluation process often relies on listening tests involving human participants, which can be both resource-intensive and impractical due to the need for a large number of listeners. To address this challenge, neural SA metrics have garnered notable attention. We will discuss the fundamental systems of neural SA, highlight several factors influencing performance, and explore emerging trends in this domain. Finally, we will present some applications of neural SE and SA in assistive oral communication technologies. These applications include impaired speech transformation and noise reduction for assistive hearing and speaking devices. Through these discussions, our aim is to illustrate the potential impact of neural-based approaches in improving communication accessibility for individuals with oral communication disorders.

簡 歷：

Yu Tsao (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1999 and 2001, respectively, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2008. From 2009 to 2011, he was a Researcher with the National Institute of Information and Communications Technology, Tokyo, Japan, where he engaged in research and product development in automatic speech recognition for multilingual speech-to-speech translation. He is currently a Research Fellow (Professor) and the Deputy Director with the Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan. He is also a Jointly Appointed Professor with the Department of Electrical Engineering, Chung Yuan Christian University, Taoyuan, Taiwan. His research interests include assistive oral communication technologies, audio coding, and bio-signal processing. He is currently an Associate Editor for the IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING and IEEE SIGNAL PROCESSING LETTERS. He was the recipient of the Academia Sinica Career Development Award in 2017, national innovation awards in 2018– 2021, Future Tech Breakthrough Award 2019, Outstanding Elite Award, Chung Hwa Rotary Educational Foundation 2019–2020, and NSTC FutureTech Award 2022. He is the corresponding author of a paper that received the 2021 IEEE Signal Processing Society (SPS), Young Author, Best Paper Award.