VOICE Research Lab 聲線研究所

Beyond the "Rainbow Passage": Evaluating the Phonetic Balance of Large-Scale AI-Generated Chinese Texts via Low-Code Platforms

Professor Feifan Wang is an Associate Professor at the School of Humanities and Deputy Director of the Key Laboratory of Linguistic *Sciences* at Shanghai Normal University. He received his Ph.D. in Speech and Hearing Sciences from the University of Hong Kong. Professor Wang's research focuses on experimental phonetics, speech-language pathology, and voice science, with publications in the *Journal of Speech, Language, and Hearing Research* and *Journal of Voice*. Currently, he serves as the Principal Investigator for a youth project funded by the National Social Science Fund of China and serves as a cultural researcher for the "Shanghai Sound" project at the West Bund Museum.

Speaker: Prof. Feifan Wang
School of Humanities, Shanghai Normal University

ABSTRACT

Standardized reading passages are fundamental tools in speech-language pathology and phonetic research for ensuring controlled elicitation of speech. While classic passages provide consistency, they often suffer from limited ecological validity and a lack of variety. The emergence of Generative AI (GenAI) offers a potential solution by enabling the rapid production of infinite reading materials. However, it remains unclear whether these machine-generated texts statistically reflect the natural phonological distribution of spoken Chinese.

This study introduces a novel workflow utilizing a low-code platform to automate the generation of large-scale short Chinese passages via Large Language Models (LLMs). We analyze the phonetic coverage and balance of these AI-generated texts. Specifically, we compare the distribution of phonemes and tones in the generated corpus against reference distributions derived from current spoken Chinese corpora. This presentation outlines the automated generation pipeline and discusses the preliminary findings regarding the "phonetic naturalness" of AI-written content. The study aims to determine if GenAI can serve as a reliable engine for creating phonetically balanced, customized assessment materials for clinical and pedagogical applications.

Jan 27, 2026 (Tuesday)

12:00 - 13:00

Room 754, Meng Wah Complex, HKU

All are Welcome