Eighty challenges facing speech input/output technologies. Victor Zue, MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA, USA. [Full Paper Available on CD]

During the past three decades, we have witnessed remarkable progress in the development of speech input/output technologies. Despite these successes, we are far from reaching human capabilities of recognizing nearly perfectly the speech spoken by many speakers, under varying acoustic environments, with essentially unrestricted vocabulary. Synthetic speech still sounds stilted and robot-like, lacking in real personality and emotion. There are many challenges that will remain unmet unless we can advance our fundamental understanding of human communication – how speech is produced and perceived, utilizing our innate linguistic competence. This paper outlines some of these challenges, ranging from signal presentation and lexical access to language understanding and multimodal integration, and speculates on how these challenges could be met.