

# Applying Physiologically-Motivated Models of Auditory Processing to Automatic Speech Recognition: Promise and Progress

Richard Stern<sup>1,a)</sup>

**Abstract:** The human auditory system has been an inspiration for developers of automatic speech recognition systems for many years because of its ability to interpret speech accurately in a wide variety of difficult acoustical environments. This talk will discuss the application of physiologically-motivated and psychophysically-motivated approaches to signal processing that facilitates robust automatic speech recognition. The talk will begin by reviewing selected aspects of auditory processing that are believed to be especially relevant to speech perception, and that had been components of signal processing schemes that were proposed in the 1980s. We will review and discuss the motivation for, and the structure of, classical and contemporary computational models of auditory processing that have been applied to speech recognition, and we will evaluate and compare their impact on improving speech recognition accuracy. We will discuss some of the general observations and results that have been obtained during the renaissance of activity in auditory-based features over the past 15 years. Finally, we will identify certain attributes of auditory processing that we believe to be generally helpful, and share insights that we have gleaned about auditory processing from recent work at Carnegie Mellon.

**Biography:** Richard M. Stern received the S.B. degree from the Massachusetts Institute of Technology in 1970, the M.S. from the University of California, Berkeley, in 1972, and the Ph.D. from MIT in 1977, all in electrical engineering. He has been on the faculty of Carnegie Mellon University since 1977, where he is currently a Professor in the Department of Electrical and Computer Engineering, the Department of Computer Science, and the Language Technologies Institute, and a Lecturer in the School of Music. Much of Dr. Stern's current research is in spoken language systems, where he is particularly concerned with the development of techniques with which automatic speech recognition can be made more robust with respect to changes in environment and acoustical ambience. In addition to his work in speech recognition, Dr. Stern has worked extensively in psychoacoustics, where he is best known for theoretical work in binaural perception. Dr. Stern is a Fellow of the IEEE, the Acoustical Society of America, and the International Speech Communication Association (ISCA). He was the ISCA 2008-2009 Distinguished Lecturer, a recipient of the Allen Newell Award for Research Excellence in 1992, and he served as the General Chair of Interspeech 2006.

---

<sup>1</sup> Department of Electrical and Computer Engineering and Language Technologies Institute Carnegie Mellon University Pittsburgh, Pennsylvania 15213 USA

<sup>a)</sup> rms@cs.cmu.edu; <http://www.ece.cmu.edu/~rms>