ACOUSTICS2008/94
The role of the cochlear processing in human speech recognition

Jont Allen[a], Marion Regnier[b], Sandeep Phatak[c] and Feipeng Li[a]
[a]University of IL, 405 N. Mathews, Room 2061 Beckman Inst. (MC 251), Urbana, IL 61801, USA
[b]208 S. 3rd St. Apt 5A, Brooklyn, NY 11211, USA
[c]Walter Reed Hospital, Silver Springs, MD 20901, USA

Little is know about how the auditory system decodes speech. We may think of speech communication re Shannon's source-channel model, thus viewed, the most complex part of the speech communication channel is the auditory system (the receiver). In my speech-perception research, I have attempted to limit the assumptions, and have thus fallen back on Shannon's basic source-channel model. The basic tool is the confusion matrix (CM) for isolated natural consonant and vowels (CV), as a function of the speech to noise ratio (SNR), with several types of masking noise. We have used large numbers of talkers and listeners (i.e., 20). In a second experiment we selectively remove islands of speech in time-frequency, and then correlate the resulting modified speech against subject scores. Our most important conclusions are: 1) The across-frequency onset transient portion of the signal is typically the most important. 2) The spectral regions of these transient are used to code different consonants. 3) While the frequency regions for a given consonant are correlated to the following vowel, this may not be important for perception. 4) Compact spectral-temporal amplitude modulations components (e.g., a 10 Hz modulation) do not seem to play a significant role, at least above 1-2 kHz.