Independent Contributions of the Temporal Envelope and Fine Structure Neural Codes to Auditory Perception

Fan-Gang Zeng, Ph.D. (Telephone: 949-824-1539; Email: fzeng@uci.edu)

Departments of Anatomy and Neurobiology, Biomedical Engineering, Cognitive Sciences and Otolaryngology-Head and Neck Surgery, University of California, Irvine, CA 92697-1275

Different from traditional Fourier transform in the frequency domain, Hilbert transform decomposes a signal into a relatively slowing-varying envelope and fast-varying fine structure in the time domain. The temporal envelope and fine structure can be further reduced to amplitude and frequency modulations. Here I will discuss the computational relationship between Fourier and Hilbert transforms, the neural code for temporal envelope and fine structure, and their functional significance in auditory perception.

Motivated by how to deliver the temporal fine structure cue to cochlear implants, we have developed a signal processing strategy that independently extracts slowly-varying amplitude and frequency modulations. This novel strategy also provides a platform to test systematically the contribution of temporal envelope and fine structure to realistic listening tasks including speech recognition, speaker identification, sound localization, and music perception. Normal-hearing listeners were presented with original and processed sounds that contained amplitude modulation only or both amplitude and frequency modulations. The speech materials were vowels, consonants, and IEEE sentences presented in quiet, or with the speech-spectrum-shaped noise or a competing talker. The music materials included familiar melodies with and without the rhythm.

We found that amplitude modulation alone was sufficient for recognition of speech and rhythmic music materials with as few as four frequency bands, but produced essentially chance level for speech recognition in noise, speaker identification, and melody recognition without the rhythmic cue. Interestingly, the addition of slowly-varying (<400 Hz) frequency modulation could almost restore performance to a level comparable to the original sound. These data suggest that, while amplitude modulation provides essential information for speech recognition in quiet, frequency modulation is needed by the listener to identify the speaker and to separate signal from noise. While our finding has immediate impact on the design of auditory prostheses and audio encoders, how the neural system computes and encodes the temporal envelope and fine structure information is not clear and requires further physiological and modeling studies.