



There is a big difference between hearing and listening. Listening requires complex auditory processing, which facilitates learning. It's a skill humans use automatically in order to filter out background noise to understand someone's speech; remember a previously heard tune and hum along; or recognize the difference between a ringing phone and ringing alarm and understand what an appropriate response to those sounds would be.

Human listeners are able to handle such mixed signals, but machines—such as automatic speech recognizers—are vulnerable to added interference, even at levels that listeners barely notice. Consider the implications of machines that could respond when called, technology that could classify and retrieve videos by their sound tracks, or applications that could automatically search for audio data the same way we do now for text data.

To make these advances possible, it is important to understand how perceptual systems manage to make precise judgments in noisy and uncertain circumstances. This understanding can then be applied to extracting information from sound commonly encountered in daily life, identifying characteristics of the sounds, classifying them, and matching the sounds to appropriate responses.

Daniel P. Ellis is working on such advances. He is the founder and principal investigator at the Laboratory for Recognition and Organization of Speech and Audio (LabROSA) at Columbia Engineering. This lab is the only one in the nation to combine research in speech recognition, music processing, signal separation, and content-based retrieval in order to implement sound processing in machines.

His chief focus is to develop and apply signal processing and machine learning techniques to extract high-level, perceptually relevant information from sound. His intention is to test theories about how human auditory perception works and enable the creation of machines that can make use of sound information in the same way humans do.

Ellis' work in soundtrack classification pioneered the idea of using statistical classification of audio data for general classification of videos by their soundtracks. Current projects in the research group include speech processing and recognition; source separation and organization; music audio information extraction; personal audio organization; and marine mammal sound recognition.

He is a member of the Audio Engineering Society, International Speech Communications Association, Institute of Electrical and Electronics Engineers, and the Acoustical Society of America.

B.A., University of Cambridge, 1987; M.S., Massachusetts Institute of Technology, 1992; Ph.D., MIT, 1996

Delving into the Science of Listening

DANIEL P. ELLIS

Associate Professor of Electrical Engineering