



CASE

COLLEGE OF ARTS AND SCIENCES

Department of Communication Sciences, CASE Speech Production Lab

SEMINAR SERIES IN COMMUNICATION SCIENCES

Fall 2005

**INTERDISCIPLINARY APPROACHES TO
SPEECH PRODUCTION AND PERCEPTION RESEARCH**

PRELIMINARY LECTURE

Monday, November 28th – 12:00 pm

Nord Hall, Room 310B

Refreshments will be available

Prof. D. Wang

Department of Computer Science and Engineering

The Ohio State University

**“An Auditory Scene Analysis Approach to
Speech Segregation and Restoration”**

Abstract

The acoustic environment is typically composed of multiple simultaneous events. A remarkable achievement of the auditory system is its ability to disentangle the acoustic mixture and group the sound energy that originates from the same event or source. This process of auditory organization is referred to as auditory scene analysis. The cocktail party problem, or segregation of speech from interfering sounds, has proven to be very challenging computationally. In this talk I describe an auditory scene analysis approach to the cocktail party problem. Our model performs segmentation and grouping in a two-dimensional time-frequency representation that encodes proximity in frequency and time, periodicity, and amplitude modulation. In addition to feature-based organization, I also describe a schema-based model for the well-known perceptual phenomenon of phonemic restoration, referring to perceptual synthesis of phonemes that are masked by acoustic interference based on linguistic context. This model employs a speech recognition system to decode speech based on unmasked portions and activates templates corresponding to the words containing the masked phonemes. Our computational effort has significantly advanced the state-of-the-art performance in speech segregation. Potential applications include hearing prosthesis and automatic speech recognition in noise.