

# Prosody and Prosodic Models

Chilin Shih, Greg Kochanski

{cls,gpk}@bell-labs.com, +1-908-582-2926, +1-908-582-4706  
600 Mountain Ave, Room 2D-426, Murray Hill, NJ 07974, USA

## 1 Overview

The tutorial will summarize the state of the art of prosody research with an emphasis on intonation ( $f_0$ ). We will present it in a data and knowledge driven fashion, starting with physiology and speech data and building towards theoretical descriptions.

We will cover a wide range of intonation phenomenon, discussing the variations, the cause of variations, and their implications to prosodic modeling.

## 2 Outline of Syllabus

Major concepts that will be addressed are:

- The physical bases of prosody: what gestures and acoustic features carry information, and how the acoustics relate to physiology.
- Prosody as a communications channel. We will discuss what information is carried by prosody and show examples (from several languages) with variations. Linguistic theories will be discussed where applicable.
  - Prosody used to convey lexical meaning: Accents and tones.
  - Prosody used to convey non-lexical information: Intonation type (Question vs. declarative ...).
  - Prosody used to convey discourse functions: Focus, prominence, discourse segments, *etc.*
  - Prosody used to convey emotion.
  - Prosody tied to the physical system: declination.
- A survey of algorithms for modeling prosody. We will comment on the strengths and limitations of a variety of theories and modeling techniques.
  - Basic Questions: How is all this information carried on a shared channel? How is it encoded? How much of prosody can be predicted from text?

- Intonation Models. Some models describe prosody as a single composite signal, others as discrete signals.
  - Duration.
- The effect of physiological constraints on intonation generation.

### **3 Prerequisites**

The tutorial will be aimed at a broad audience of speech researchers, and will not assume specialized knowledge.

### **4 List of course materials**

- Course Notes, including a bibliography of current research.
- Selected reading materials will be provided.
- Speech examples in CD-ROM or web page.

### **5 Biographical Sketches**

- Dr. Chilin Shih works on prosodic modeling and text-to-speech systems at Bell Labs, and has built Text-to-Speech systems for many languages, including Chinese, Navajo, Romanian, Hindi, Greek, Spanish, Italian, and Russian. She has published in linguistics and speech research, including phonology, morphology, phonetics, and prosody.
- Dr. Greg Kochanski builds mathematical models of complex systems at Bell Laboratories, most recently focusing on prosody and speech synthesis. He has published in a variety of fields, including Astrophysics and experimental physics.
- Dr. Shih and Dr. Kochanski developed a multi-lingual prosodic modeling system Stem-ML (Soft Template Mark up Language) that is suitable for the modeling of natural intonation variations.