Student: Minshu Zhan

Term: 2012 Summer

Faculty Supervisor: Stefanie Shattuck-Hufnagel

Date: June 11, 2012

**Automatic Generation of Phonetic Landmarks**

The overall goal of the project is to model the process that human listeners recognize and understand spoken language. We assume that this process involves the recognition of sound segments and their features using acoustic cues. One particular kind of acoustic cues is the robust spectral changes, called landmarks. Under the assumption that these robustly implemented cues play a particularly important role in human’s sound perception, the project requires developing programs that processes of spoken utterances automatically.

My job in this project includes the following main tasks: (a). predict landmarks from word sequences using existing rules; (b). detects the mutations, deletions, and insertions in hand landmark labels compared with the landmark predicted ones; (c) extract contextual information, e.g. syllable stress and position, associated with each landmark; (d). based on the contexts, generate a proper decision tree that models the differences between predicted and actual landmarks. I have completed three python scripts for the first task in the spring term. During 2012 summer, I will: (i). modify our current landmark comparison code and ensure its performance and correctness on various input files; (ii) implement the data representation of landmarks’ context information, including stresses and positions of relevant syllables, words, and phrases, as well as word frequencies; (iv) add context extraction functions into the landmark prediction script; (iv). apply the data to C4.5, an existing decision tree module. Upon completion of these tasks, I will start implementing similar programs for another type of acoustic cues is related to the distinctive features that describe speech characteristics, such as voicing.

My motivation for joining this project is to acquire a better understanding of human’s acoustic perception process. I have always been interested in the field of automatic sound perception and wish to be more familiar to the basic techniques and models in sound processing through this project. The project is a very good practice in programming and artificial intelligence for me.