MAS.632 Conversational Computer Systems
Fall 2008

Problem Set 3

1. Text-to-speech synthesis can be broken into two stages. The first converts text to the appropriate phonemes, and the second realizes the phonemes as an audio signal. Considering only phoneme production aspect, i.e., the acoustic complexity of the various speech sounds we use as phonemes, which are more difficult to synthesize, vowels or consonants? Why?

   [Note: you don't really need to know more about speech synthesis than you have read so far to answer this question.]

2. Is F1 more similar to F0 or F2. Why?