

SPEECH COMMUNICATION
6.541J-24.968J-HST710J
Spring 2004

Lab 1

Assigned: 02/24/04

Due: 03/02/04

In this laboratory exercise, we will examine several different ways of doing spectral analysis of speech, and we will collect some acoustic data from a few vowels and consonants. These data will be compared with data reported in the literature and will be interpreted in terms of speech-production theories described in class.

Each lab group has a directory on the LINUX workstations. Stored in this directory are two utterances of the same sentence, spoken by an adult female and by an adult male speaker. The sentence is

“I have seen Pete at Todd's car.”

These utterances have already been low-pass filtered at 7.5 kHz, and digitized with a sampling rate of 16 kHz.

If you are unable to complete the lab in the 2-hour time period, you can also use the server.

- (a) Print out a spectrogram of each of the two sentences. Estimate the average speaking rate for each in syllables per second.
- (b) Examine the waveform and spectrum of the vowel in the middle of the word “Todd” (either the male or the female version), as follows:
 - (i) Look at the expanded waveform in the vowel. Measure the fundamental frequency. Estimate roughly the first formant frequency from the waveform by looking at the frequency of the decaying oscillation in each glottal period.
 - (ii) Look at the spectrum in the middle of the vowel. Four types of spectra should be examined. Record the fundamental frequency and first three formant frequencies for the vowel. Print out each of the four spectra described below.
 - (1) The discrete Fourier transform (dft) with a time window of about 26 ms. Also displayed with this spectrum is the smoothed spectrum envelope, obtained by smoothing the dft with a window of about 300 Hz.
 - (2) The dft with a time window of about 6 ms. In this case the window should be centered on the first part of the glottal period.

- (3) An average of a series of 6-ms dft's over a time interval of 15 ms.
- (4) An lpc spectrum superimposed on the 26-ms dft. The lpc spectrum can be considered as a “best” all-pole fit to the dft spectrum.

Discuss some possible advantages and disadvantages of using each type of spectrum for determining vowel formant frequencies and fundamental frequency.

- (c) Measure the first three formant frequencies and the fundamental frequency in the middle of the word “Todd” for the other speaker. Measure the first three formant frequencies and the fundamental frequency in the middle of the word “Pete” for both speakers. Compare your results with data published by Peterson and Barney (1952), which are attached to this description.
- (d) Repeat (c) for the unstressed vowel in the word “at.” What can you say about the vocal-tract shape for this vowel compared with /a/ and /i/.
- (e) Measure the fundamental frequency in the middle of the remaining vowels in one of the sentences. Which vowel has the highest F0? Which has the lowest F0?
- (f) For file “SEENPETEHH”, examine the spectrum (25 ms window) in the middle of “seen” and compare it with that in “Pete”. Can you observe the effect of nasalization in the vowel in “seen”? (See pages 303-322 of **Acoustic Phonetics**.)
- (g) Using a 6-ms window, examine the spectrum of the noise burst at the release of [t] in “Todd” and at the release of [k] in “car”. Print these spectra. Discuss reasons for the differences in these spectra. (See pages 357-375 of **Acoustic Phonetics**.)
- (h) If time permits, have one (or more) speaker(s) record the same sentence into the computer and make some measurements similar to those described above. Ask Professor Stevens or Xuemin for assistance with this.

You should write up a brief report of the lab (suggested length 3-4 pages) giving the results of your measurements and some discussion. Each member of the group should prepare a separate report.