# Speech Processing

0.5 x 10k = 5k

The following phonemes from the available digitized (fs = 10 kHz) speech records will be of interest: (1) /i/ ''eve'', (2) /u/ "boot", and (3) / η / "rung". Repeat the following for each of the three:

A. Conventional Spectral Analysis (3 plots/ phoneme)

1. Plot the N=256 point time series : n=1024, ..., 1279

2. Compute the corresponding N-point FFT and plot the spectrum (dB)

3. Take just the first 64-points of the block under consideration, zero fill out to N-1, FFT, and plot the spectrum (dB).

B. Autocorrelation Method of Linear Prediction

Process both the 256-point (A1) and 64-point (A3) data blocks and make estimates of the inverse filter.

1. Determine $E_p$ for inverse filter orders p = 2, 4, 6, 8, 10, 12, and 14 and plot $E_p$ vs. p.

2. For just order p = 14, zero fill the filter coefficient sequence out to N-1, compute the N = 256-point FFT, and plot $10 \log\left(\frac{1}{|\hat{A}(k)|^2}\right)$, $k = 0, \ldots, \frac{N}{2} - 1$.

C. Comment on your results.

Note: Make sure that the time series is windowed prior to computation of the autocorrelation function used in estimating the inverse filter.

In your frequency domain plots, the horizontal axis should cover 0 to fs/2 Hz.

Reference:

[1] J. Markel, "Digital Inverse Filtering - A New Tool for Formant Trajectory Estimation", IEEE Trans. on Audio and Electroacoustics, AU-20: 129-127 (1972)