CS630: Speech Technology
LAB-3: Nature of Speech Signal

OBJECTIVE:

To obtain the pitch and formant frequencies from the nature of the speech signal in the time and frequency domains respectively.

Tasks

(a) Record speech for three vowels in the context of h*ee*d (for vowel /i/), h*e*ad (for vowel /e/) and h*o*d (for vowel /a/).

(b) Note the average pitch using time domain representation.

(c) Note average formant frequencies (F_1, F_2 and F_3) from 10^{th} order LPC spectrum.

(d) Repeat for one more set of vowels.

(e) Compare the readings from the two sets (variability within a given speaker).

(f) Compare the average pitch and formant frequency values of one speaker with values from other speakers (variability among speakers).

(g) Write a brief note on the observations.

1 PROCEDURE

1. **Recording Speech Signal**

   - Record a speech signal for heed, head, hod using 'brec' command as
     brec -s 8000 -b 16 -t 2 -w samplefile.wav
     where "brec" is the linux command
     "s" is the sampling frequency in Hz
     "b" is the number of bits/sample used for quantization
     "t" is the time interval used to record the speech signal
     "w" format (wav) in which the speech signal to be represented

   - Play the signal back using 'bplay' command
     bplay -s 8000 -b 16 -w samplefile.wav

2. To obtain Average Pitch

   - Display the recorded speech waveform using the utility "wavesurfer".

     wavesurfer filename.wav

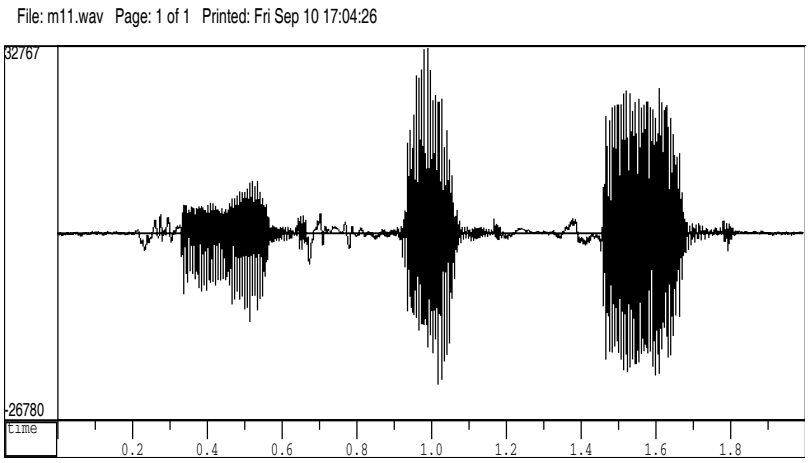

Figure 1: Speech waveform for the utterance "heed- head - hod"

- Zoom the vowel segment using 'zoom in' command from menu options 'view'. Asegment of the vowel /i/ is shown in the Figure 2.

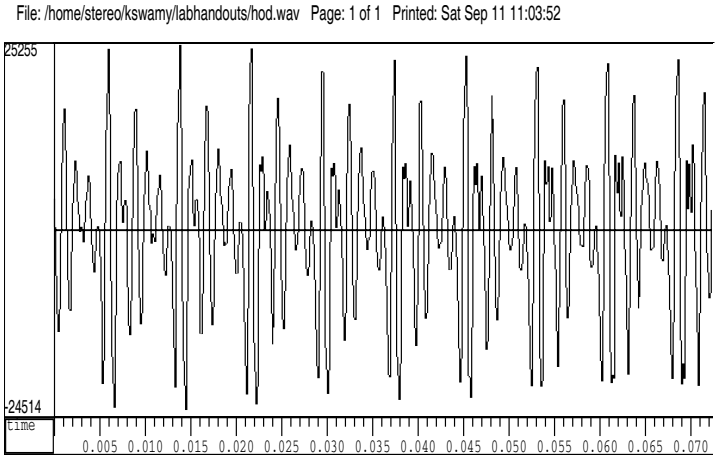

Figure 2: A part of the vowel /a/ in the speech segment of 'hod' obtained using 'zoom in' command in menu options

- Note the time in msec of the period for three or four cycles (say) and measure the average pitch period, preferably in the steady portion as illustrated in the Figure 3.
- As pitch period is not constant over the entire vowel segment, make four or five pitch period measurements.
- Mean of these values gives average pitch.

3. To obtain Average Formant Frequencies

   - Display voiced segment of the recorded speech segment /hod/ using wavesurfer.
   - Select the spectrum section in the menu popped up by right click of the mouse. Choose LPC section and set order to 10 and pre-emphasis to 1.0.This will display LPC spectrum as shown in the Figure 4. The peaks in the spectrum represents formant frequencies and are noted down in the Table **??**.

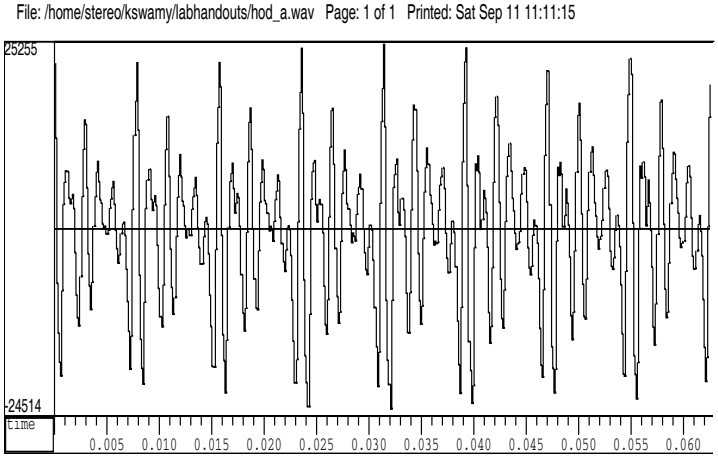

Figure 3: A part of the vowel /a/ in the speech segment of 'hod'. As shown in the Figure 3, pitch is measured as the interval between any two peaks

- The mean of each formants is calculated to obtain corresponding average values.
- Repeat the previous step for two or three short segments of the voiced region.
- Repeat the above step for recorded speech of /heed/ and /head/.
- Repeat the above procedure for recorded speech of different speakers.

4. OBSERVATIONS:

5. Conclusions

- The pitch for the same speaker varies with respect to time as well as varies across speakers. This clearly illustrates the variability of speech signal.
- For given sound-unit, the formant frequency values varies slightly for the same speaker and also across speakers. This again illustrates the variability of the speech in frequency domain for the same speaker.

Table 1:

Speaker	Vowel	Formants (Hz)						Pitch(msec)
		F1		F2		F3		
/i/								
/e/								
/a/								

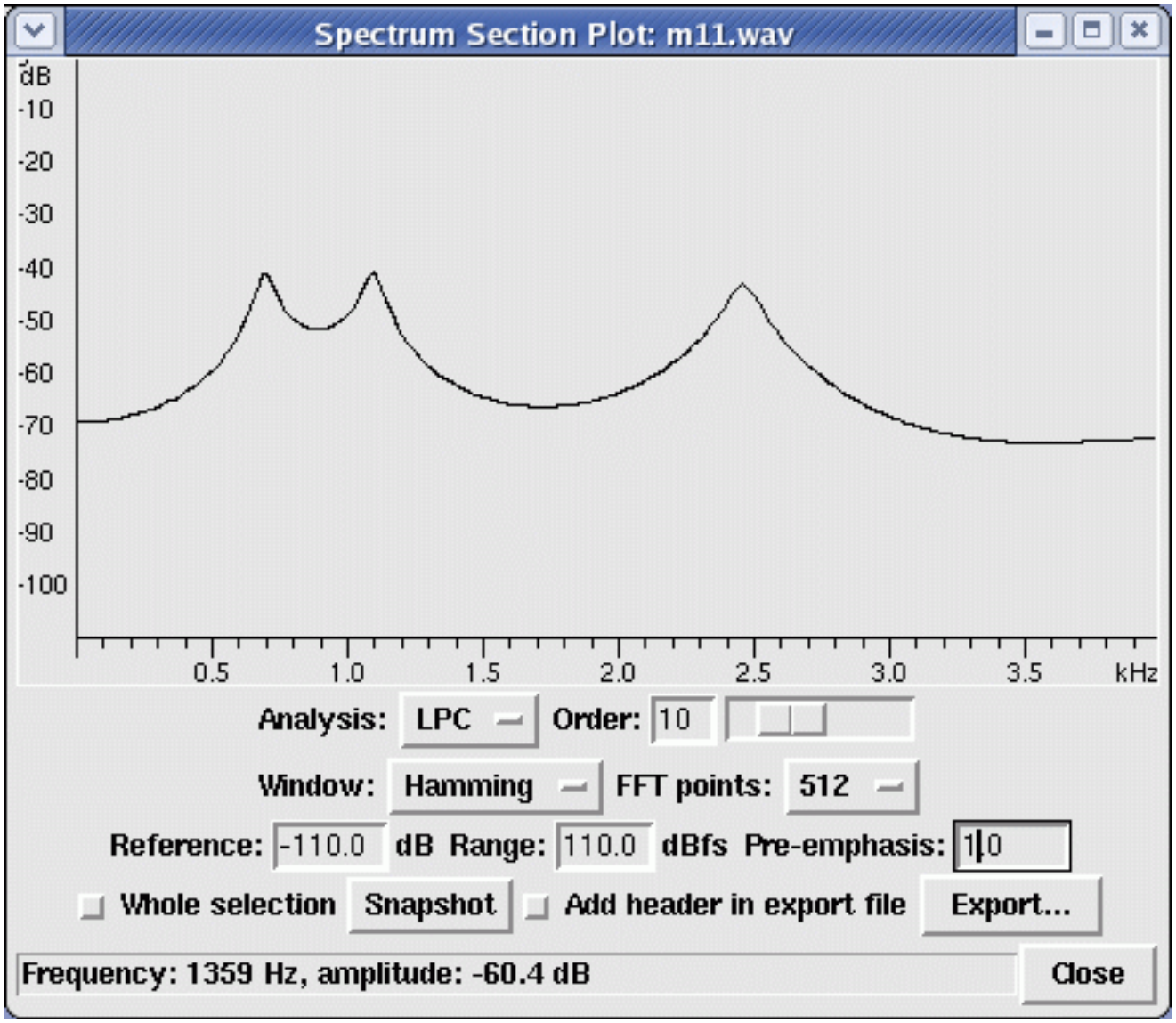

Figure 4: LPC spectrum of a short segment of vowel $/a/$ in the utterance of 'hod'. The spectrum is obtained using 10^{th} order LP analysis. As shown in the figure, the peaks of the LP spectrum represent the formant frequencies.