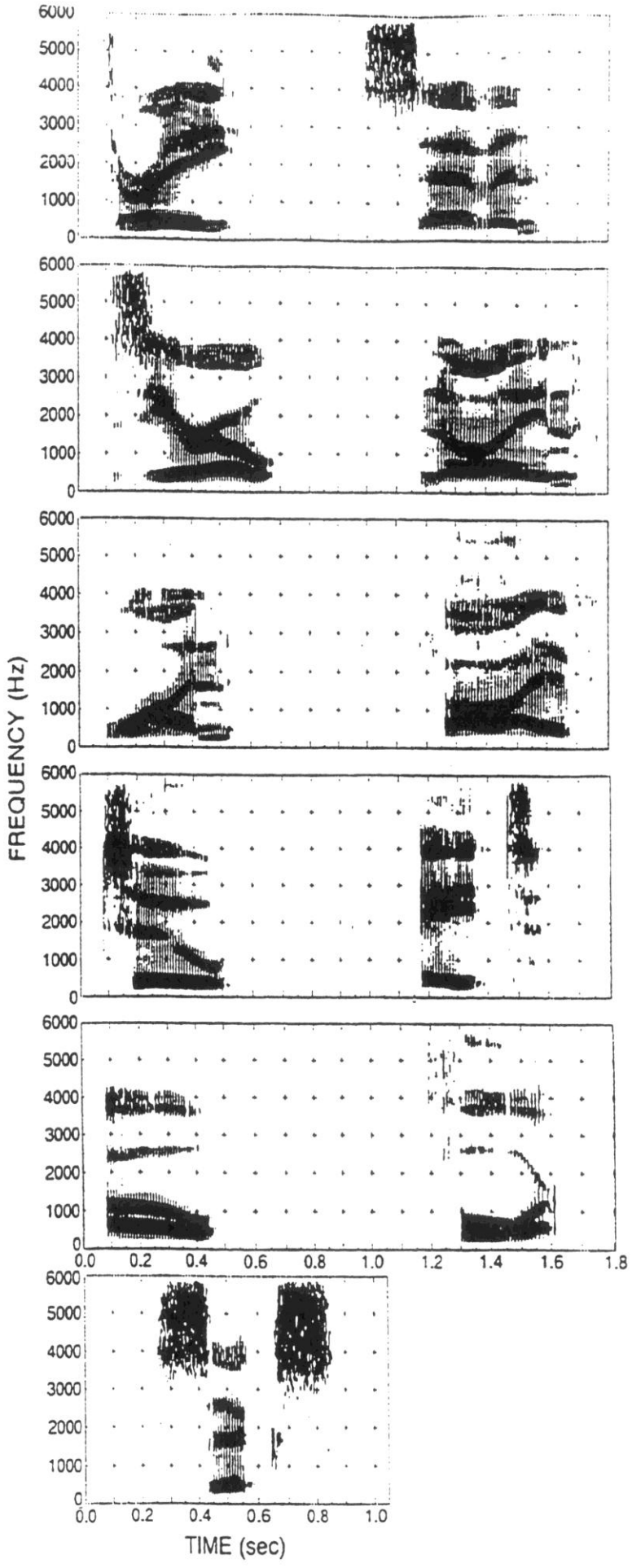
Homework 2 (ECE6255 Spring 2010) Grade=3/100

The following exercises are for *DSP Fundamentals* and *reading spectrograms.*

1) Work out Problem 2.13 in Quatieri.
2) A speech signal is sampled at a rate of 20000 samples per sec (20 KHz). A segment of length 1024 samples is selected and the 1024-point DFT is computed.
   a) What is the time duration of the selected speech segment?
   b) What is the frequency resolution (spacing in Hz) between the DFT values?
   c) How do your answers in Parts (a) and (b) change if we compute the 1024-point DFT with only 512 samples of the speech signal. (The 512-sample signal would be augmented with 512 zero samples before DFT is computed).
3) Write a MATLAB program to change the sampling rate of a speech file from 16 kHz to both 10 kHz and 8 kHz. Listen to all three speech files to hear any noticible differences, and describe them. (Use the file 1.wav and 5.wav to test your program).
4) Record your own voice in your own selected utterances. Watch for harmonic lines generated by your pitch over time in vowel regions on a narrowband spectrogram. Estimate them and give your reasons. Compare with the results from wavesurfer.
5) A lexicon for the digits in American English is given below. From the displayed sound spectrograms in the next page, identify each of the 11 digits (two in the top five panels and one in the bottom). This could be a tough exercise and you may want to use both forward analysis, i.e., finding the sound classes and matching to individual digits to these sound classes, as well as elimination analysis, i.e., removing certain words not matching the sound classes. In any case, it is intended to show how difficult it is to 'read spectrograms', even for the simple task of recognizing isolated English digits given by a single speaker. In general, human beings are capable of recognizing 10 digits spoken by most speakers in the English speaking population.

TABLE 2.3. Sound Lexicon of Digits

Word	Sounds	ARPABET
Zero	/z I r o/	Z–IH–R–OW
One	/w ʌ n/	W–AH–N
Two	/t u/	T–UW
Three	/θ r i/	TH–R–IY
Four	/f o r/	F–OW–R
Five	/f a^y v/	F–AY–V
Six	/s I k s/	S–IH–K–S
Seven	/s ɛ v ə n/	S–EH–V–AX–N
Eight	/e^y t/	EY–T
Nine	/n a^y n/	N–AY–N
Oh	/o/	OW


FREQUENCY (Hz)
6000
5000
4000
3000
2000
1000
0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
TIME (sec)