Mini Project

Time and Frequency Domain Analysis of Your Recorded Voice Signal

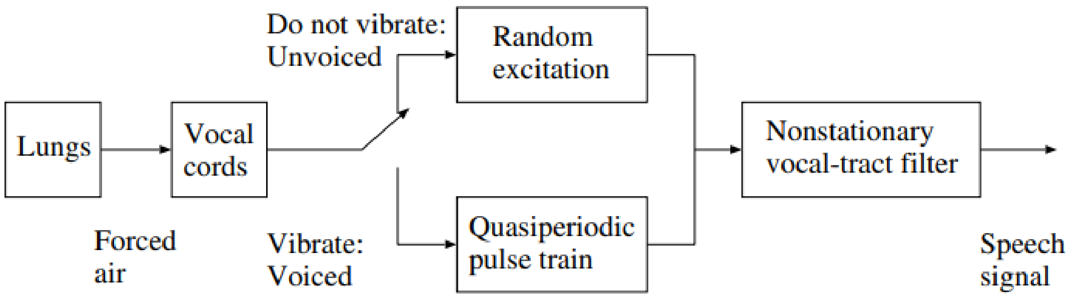

Figure 1.49 Schematic representation of the production of voiced and unvoiced speech.

quasiperiodic pulses of air which is passed through the vocal tract; see Figure 1.49. The input to the vocal tract may be treated as a train of impulses or pulses that is almost periodic. The vocal tract acts as a filter: Upon convolution with the impulse response of the vocal tract, which is held steady in a certain configuration for the duration of the voiced sound desired, a quasiperiodic signal is produced with a characteristic waveshape that is repeated. All vowels are voiced sounds. Figure 1.50 shows the speech signal of the word “safety” spoken by a male. Figure 1.51 shows, in the upper trace, a portion of the signal corresponding to the phoneme /E/ (the letter “a” in the word). The quasiperiodic nature of the signal is evident. Features of interest in voiced signals are the pitch (average interval between the repetitions of the vocal-tract impulse response or basic wavelet) and the resonance or formant frequencies of the vocal-tract system.

An unvoiced sound (or fricative) is produced by forcing a steady stream of air through a narrow opening or constriction formed at a specific position along the vocal tract. The vocal cords do not vibrate for such sounds; see Figure 1.49. The result is a turbulent signal that appears like random noise. The input to the vocal tract is a broadband random signal, which is filtered by the vocal tract to yield the desired sound. Fricatives are unvoiced sounds, as they do not involve any activity (vibration) of the vocal cords. The phonemes /S/, /SH/, /Z/, and /F/ are examples of fricatives. The lower trace in Figure 1.51 shows a portion of the signal corresponding to the phoneme /S/ in the word "safety." The signal has no identifiable structure and appears to be random (see also Figures 3.1, 3.3, and 3.4, as well as Section 3.2.4). The transfer function of the vocal tract, as evidenced by the spectrum of the signal itself, would be of interest in analyzing a fricative.

Plosives, also known as stops, involve complete closure of the vocal tract, followed by an abrupt release of built-up pressure. The phonemes /P/, /T/, /K/, and /D/ are examples of plosives. The sudden burst of activity at about $1.1\ s$ in Figure 1.50 illustrates the plosive nature of /T/. Plosives are difficult to characterize as they are transients; their properties are affected by the preceding phoneme as well. For more details on the speech signal, see Rabiner and Schafer [79].

Parkinson's disease, which causes tremor, rigidity, and loss of muscle control, is also known to affect speech. The changes in speech caused by the disease include reduced loudness, increased vocal tremor, and breath-related noise. Vocal impairment caused by the disease are labeled as dysphonia, which refers to the inability to pro-

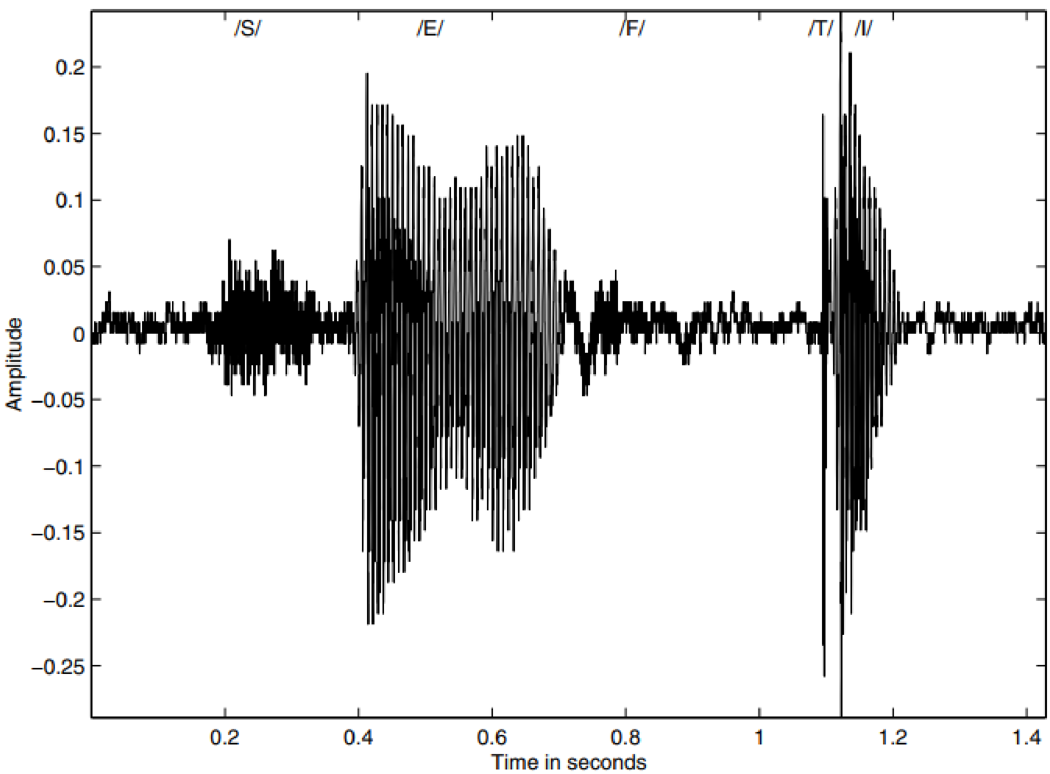

Figure 1.50 Speech signal of the word "safety" uttered by a male speaker. Approximate time intervals of the various phonemes in the word are /S/: $0.2 - 0.35\ s$; /E/: $0.4 - 0.7\ s$; /F/: $0.75 - 0.95\ s$; /T/: transient at $1.1\ s$; /I/: $1.1 - 1.2\ s$. Background noise is also seen in the signal before the beginning and after the termination of the speech segment, as well as during the stop interval before the plosive /T/.

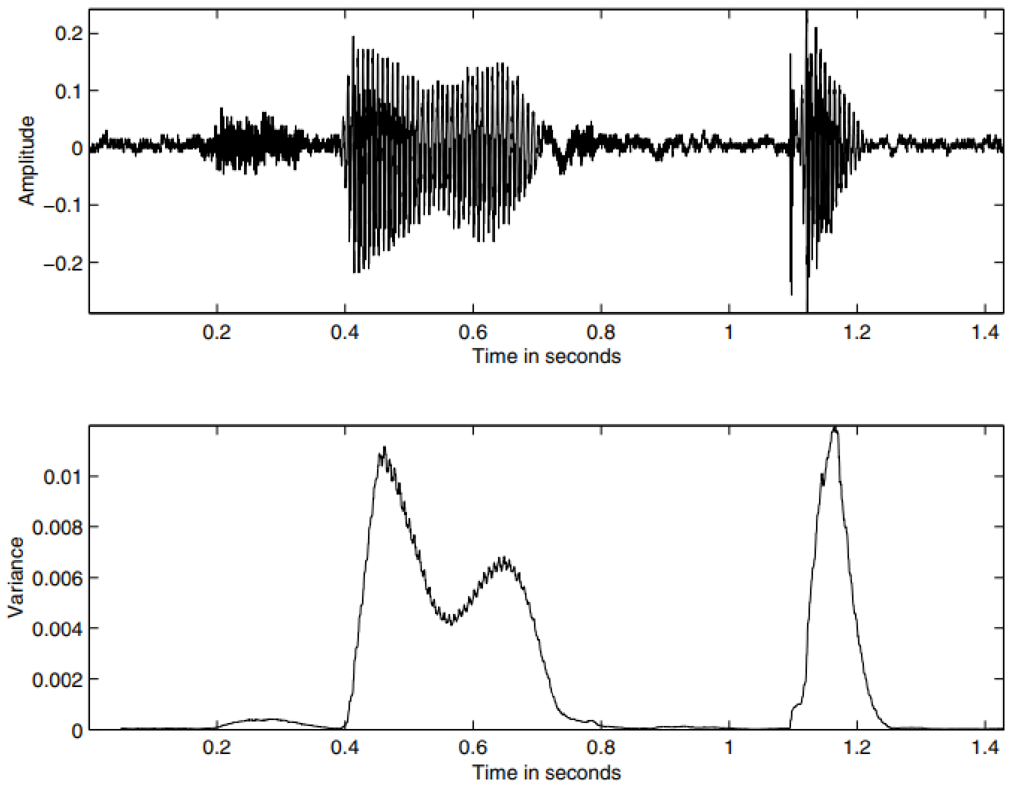

Figure 3.3 Top: Speech signal of the word "safety" uttered by a male speaker. Bottom: Variance computed in a moving window of 50 ms (400 samples with $f_s = 8\ kHz$).

- Capture one word in your voice
- Capture 10-15 second of your voice.

Time Domain Analysis:

Objective 1: Visual Analysis of phonetics

Objective 2: Effect of LPF filtering in Time Domain

Objective 3: Effect of HPF filtering in Time Domain

Objective 3: Segregate Voice and Un-voice part

Frequency Domain Analysis:

Objective 1: Visual Analysis about the frequency Details

Objective 2: Effect of LPF filtering in Frequency Domain

Objective 3: Effect of HPF filtering in Frequency Domain

Objective 3: Segregate Voice and Un-voice part

Problem statement: Study of time and frequency domain analysis of your recorded voice signal.

Procedure:

1. Capture 10-15 sec of your voice signal.
2. Convert the recorded signal to .wav type if applicable, and save the recorded file in a folder.
3. In order to visually analyse the phonetics of your voice signal, use ''audioread'' command in MATLAB; for example, [y]=audioread('airtel_flute.wav');
4. Analyse the time domain representation, and frequency domain representation of audio vector ''y'' using DTFT.
5. Study the effect of ''low-pass'', and ''high-pass'' filtering of the audioread file using MATLAB.
6. Explore the phonetic variation by using time, and frequency domain representation.