It was emphasized that fine structure and transitory characteristics of both the source and system, such as formant transitions, voice onset time, and degree and timing of consonantal aspiration, are aspects of a sound that are important for its perceptual identity, and thus must be considered in the development of signal processing algorithms. These features were shown to form the smallest elements of a language, being combined to form its phonemes, and hypotheses were presented that the auditory system is predisposed to these kinds of particular acoustic and articulatory features. The property of coarticulation, which is the influence of the articulation of one sound on the articulation of other sounds in the same utterance, and the importance of prosodics were also briefly discussed. In the next chapter, we develop a more quantitative description of the acoustics of speech production, showing how the heuristics of this chapter are approximately supported with mathematical models. We also will predict other acoustic effects not seen by a qualitative treatment.

EXERCISES

3.1 Consider the following two-pole model for the glottal pulse:

$$G(z) = \frac{1}{(1 - \alpha z^{-1})(1 - \beta z^{-1})} \tag{3.7}$$

with α and β both real, positive, and less than one, and where the region of convergence includes the unit circle. In the time domain, $g[n]$ can be expressed as the convolution of two decaying exponentials:

$$g[n] = (\alpha^n u[n]) * (\beta^n u[n]). \tag{3.8}$$

(a) Sketch $g[n]$ and $|G(\omega)|$. Assume α and β are close to unity, say, about 0.95. Why are Equations (3.7) and (3.8) a reasonable model for the spectral magnitude, but not for the shape of the glottal pulse?

(b) Explain why an improved model for the glottal pulse is given by

$$\tilde{g}[n] = g[-n].$$

Derive the z-transform of $\tilde{g}[n]$. Where are the poles of $\tilde{g}[n]$ in relation to those of $g[n]$?

(c) Consider the periodic glottal flow waveform $u[n] = \tilde{g}[n] * \sum_{k=-\infty}^{\infty} \delta[n - kP]$ where P is the pitch period. Assuming the glottal pulse is two convolved decaying exponentials as in part (b), sketch the Fourier transform magnitude of the windowed glottal flow waveform $u[n]$ for a rectangular window with length equal to P and also $2P$. Which window length would be used in the calculation of a narrowband spectrogram of the glottal flow waveform?

3.2 Suppose we ask a speaker to hold a vowel steady. Ideally, the speaker would give a perfectly periodic waveform with constant pitch and loudness. In practice, however, these characteristics change in time for a "steady-state" vowel. For example, small alternating deviations can result in hoarseness in a voice. In this exercise, you explore the consequences of small changes in pitch and loudness on the speech spectrum.