Quiz 8, Monday, November 6, 2006

ECE 598 AL
The Speech Chain

Problem 1 (10 points)

The glottal volume velocity is periodic, with a time-domain waveform that has the form

$$u_G(t) = H_1 e^{j\omega_0 t} + H_2 e^{2j\omega_0 t} + H_3 e^{3j\omega_0 t} + \ldots \tag{1}$$

For this problem, please assume an overly simplified form of the harmonic amplitudes, specifically

$$H_k = \frac{1}{k^2} \tag{2}$$

Assume that the vocal tract transfer function has just two formant frequencies, thus

$$T(\omega) = \left(\frac{(2\pi F_1)^2 + (\pi B_1)^2}{(j(\omega - 2\pi F_1) + \pi B_1)(j(\omega + 2\pi F_1) + \pi B_1)}\right)\left(\frac{(2\pi F_2)^2 + (\pi B_2)^2}{(j(\omega - 2\pi F_2) + \pi B_2)(j(\omega + 2\pi F_2) + \pi B_2)}\right) \tag{3}$$

Recall that the radiation characteristic is given by

$$R(\omega) = \frac{j\rho f}{r} \tag{4}$$

where you may assume MKS units for convenience (thus $\rho \approx 1\text{kg/m}^3$), and you may assume for convenience that the sound is recorded at a distance of $r = 1m$.

(a) Assume a pitch frequency of $F_0 = 200$Hz. Sketch $|U_G(\omega)|$ or $20\log_{10}|U_G(\omega)|$, the spectrum or log spectrum of the glottal volume velocity (whichever you like, but be sure to specify which one you are plotting), for $0 < f < 2000$Hz. You may label the abscissa in either Hertz or radians/second, whichever you find more convenient.

(b) Assume that the formant frequencies and bandwidths are given by $F_1 = 600$Hz, $B_1 = 100$Hz, $F_2 = 1200$Hz, $B_2 = 200$Hz. Sketch $|T(\omega)|$ or $20\log_{10}|T(\omega)|$ (whichever you like) for $0 < f < 2000$Hz. Label the approximate amplitudes at frequencies $f = 0$Hz, $f = F_1$, and $f = F_2$.

(c) Under the same assumptions as in part (b), plot $|U_L(\omega)|$ or $20\log_{10}|U_L(\omega)|$ for $0 < f < 2000$Hz. Recall that $U_L(\omega)$, the volume velocity at the lips, is given by $U_L(\omega) = T(\omega)U_G(\omega)$.

(d) Plot $|R(\omega)|$ or $20\log_{10}|R(\omega)|$.

(e) Under the same assumptions as in part (b), plot $|P_R(\omega)|$ or $20\log_{10}|P_R(\omega)|$ for $0 < f < 2000$Hz. Recall that $P_R(\omega)$, the recorded sound pressure, is given by $P_R(\omega) = R(\omega)T(\omega)U_G(\omega)$.