−1−

Perturbation theory

2/19/04

Figure 3.18 Illustrating a perturbation ΔA in the area of an acoustic tube at a short segment of length $\Delta\ell$ centered at point $x = x_1$.

Figure 3.19 Curves showing the relative magnitude and direction of the shift ΔF_n in formant frequency F_n for a uniform tube when the cross-sectional area is decreased at some point along the length of the tube. The abscissa represents the point at which the area perturbation is made. The minus sign represents a decrease in formant frequency and the plus sign an increase.

Effect of vocal tract walls

(a)   (b)
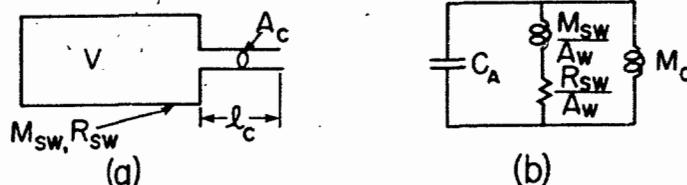
Figure 3.26 (a) Model for a constricted vocal tract configuration with yielding walls. (b) Low-frequency equivalent circuit for the model in (a).
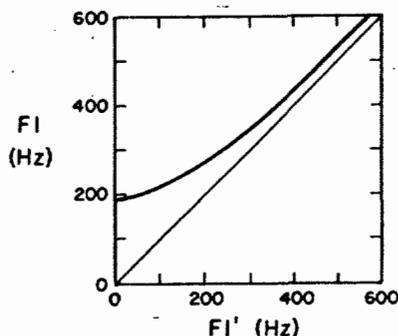
Figure 3.27 Natural frequency F1 for configuration in figure 3.26, with yielding walls, as a function of natural frequency F1' computed on the assumption of hard walls (i.e., $M_w = \infty$ in figure 3.26). Deviation of the curve from the diagonal line is a measure of the effect of the walls.

19 Feb/04
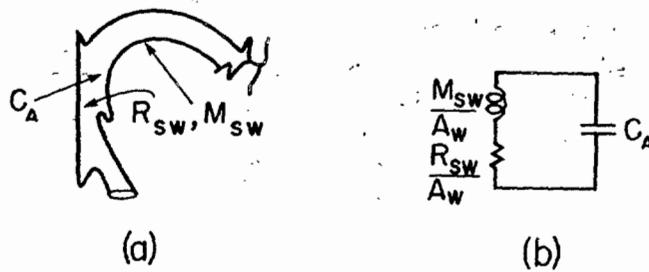
Figure 3.25 (a) Midsagittal section for a vocal tract configuration with closure at the lips. The resistance and mass of the walls are shown, together with the acoustic compliance of the vocal tract volume. (b) Low-frequency equivalent circuit for the configuration in (a) with closed glottis. A_w is the surface area of the vocal tract walls, and M_{sw} and R_{sw} are mass and resistance of walls per unit area.

Table 3.1 Calculation of contributions of radiation (B_r), vocal tract walls (B_w), viscosity (B_v), and heat conduction (B_h) to the formant bandwidths for two different vocal tract configurations

a. Uniform tube, length 15 cm, cross-sectional area 3 cm²

	Formant frequency (Hz)	B_r (Hz)	B_w (Hz)	B_v (Hz)	B_h (Hz)	Total B (Hz)
First formant	592	3	8	6	3	20
Second formant	1682	24	1	10	4	39
Third formant	2804	67	0	12	5	84
Fourth formant	3927	131	0	15	6	152

b. Resonator with dimensions in figure 3.28a, with area of opening equal to 0.32 cm²

	Formant frequency (Hz)	B_r (Hz)	B_w (Hz)	B_v (Hz)	B_h (Hz)	Total B (Hz)
First formant	300	0	28	12	2	42
Second formant	1475	0	1	8	3	12
Third formant	2950	0	0	11	5	16

Excitation of resonators

Figure 3.31 (a) Plot of magnitude of transfer function $T(f) = U_o/U_s$, expressed in decibels for an ideal uniform, lossless acoustic tube, shown in figure 3.8. (b) Magnitude of transfer function $T(f)$ for an ideal uniform tube of length 15 cm with losses similar to those occurring in the vocal tract.

Figure 3.30 The lower panel shows the distribution of amplitude of sound pressure p and volume velocity U for the second natural frequency of a uniform tube, shown in the upper panel. At points 1 and 3 a volume velocity source gives maximum excitation of this mode, whereas at points 2 and 4 a sound pressure source gives maximum excitation.
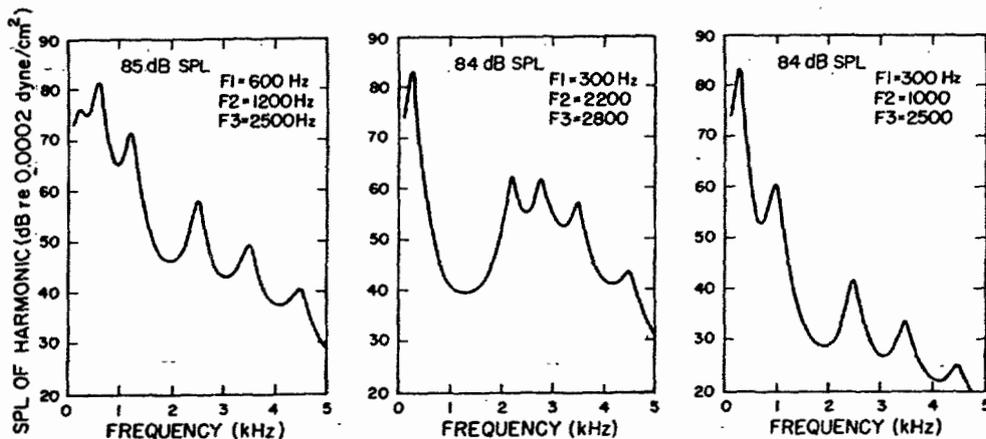
Figure 3.32 Computed spectrum envelopes approximating the vowels /ɑ/ (left), /i/ (middle), and /u/ (right). The formant frequencies are indicated in each panel, and formant bandwidths are selected to approximate those observed in natural utterances. The ordinate is the calculated sound pressure level for each harmonic at a distance of 15 cm from the lips, assuming a fundamental frequency of 125 Hz. A smooth curve is drawn through the amplitudes of the individual harmonics. The spectrum of the glottal source is that for a male voice, from figure 2.10. The calculated overall sound pressure levels are shown in each panel.

Constructing a transfer function for vowels
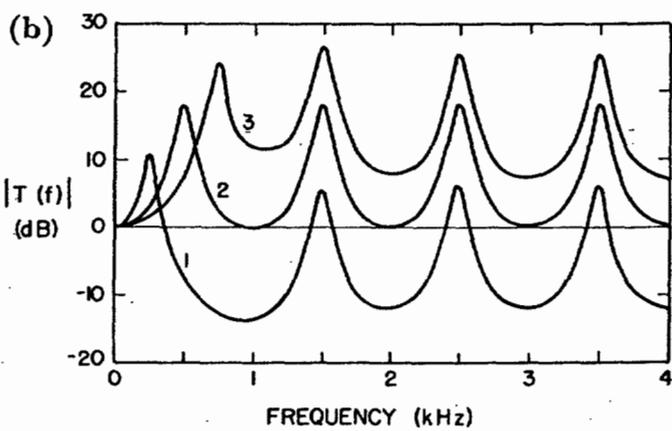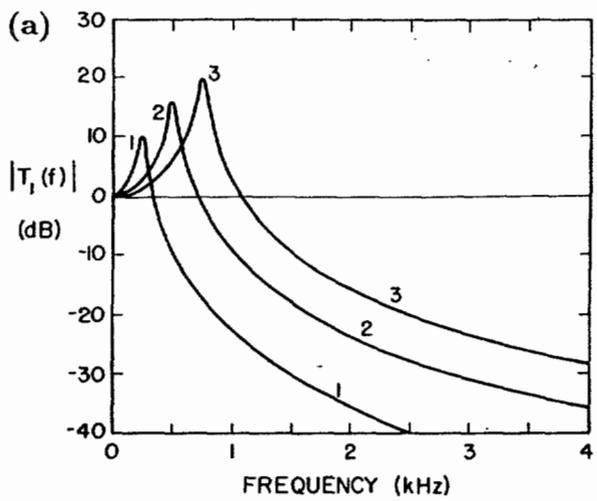
Figure 3.4 (a) The component of the vocal tract transfer function (in decibels) corresponding to the first formant for three different values of F1. Note the change in amplitude of the peak and the shift in level at higher frequencies. (b) The effect of a change in F1 on the overall transfer function, assuming formants above F1 remain fixed. The labels 1, 2, and 3 identify low, medium, and high values of F1.
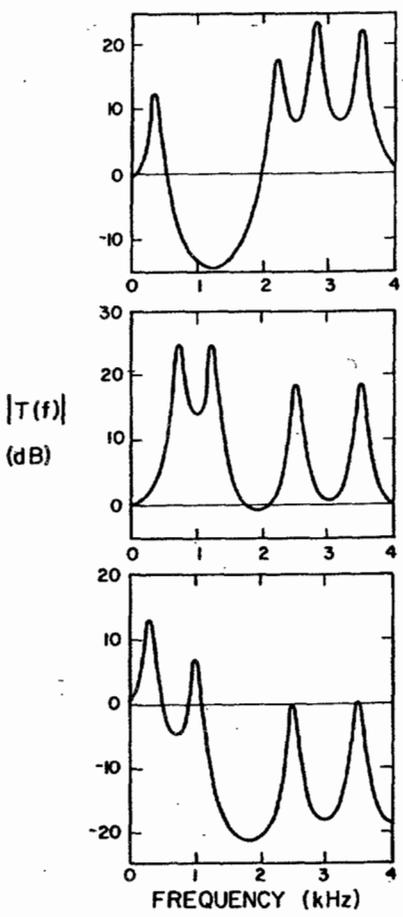
Figure 3.5 Computed transfer functions for three different configurations of formant frequencies, illustrating changes in relative amplitudes of peaks and valleys in the transfer function. Bandwidths of all resonances are fixed at 80 Hz.

Constructing Transfer functions
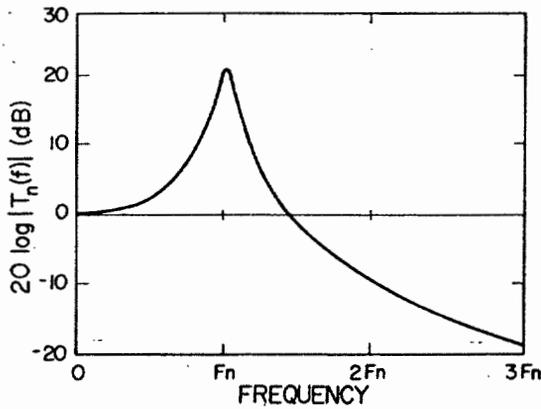
Figure 3.2 A plot of one of the terms of equation (3.9), that is, the component of the transfer function $T(s)$ associated with one conjugate pair of poles. The equation for this component is

$$T_n(s) = \frac{s_n s_n^*}{(s - s_n)(s - s_n^*)}$$

where $s = j2\pi f$, s_n is complex frequency of pole, and $s_n = \sigma_n + j2\pi F_n$. Ordinate represents magnitude of $T_n(s)$ on a decibel scale. Abscissa is frequency f. The bandwidth of the pole for this example is approximately $F_n/12$, so that $\sigma_n = F_n/12\pi$.
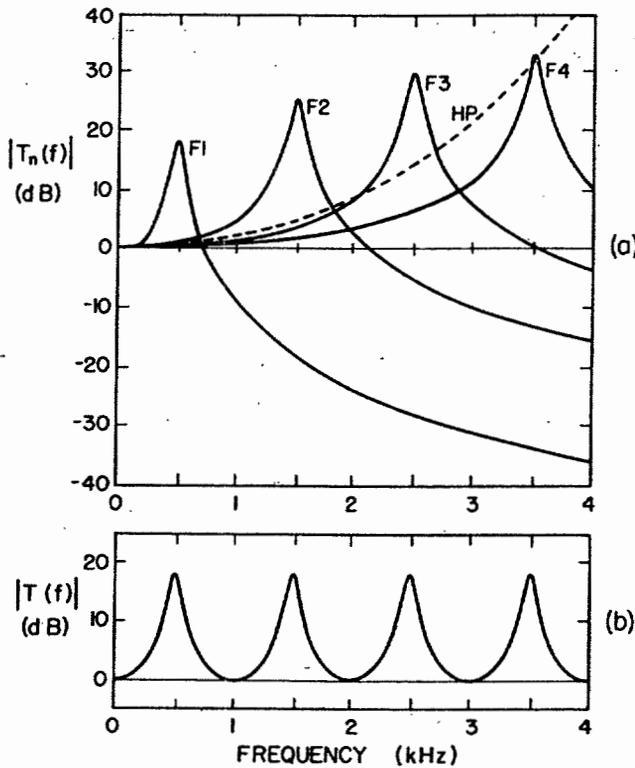
Figure 3.3 (a) The components of the vocal tract transfer function corresponding to four formants $F1$, $F2$, $F3$, and $F4$, together with the effect of higher formants (dashed curve, labeled HP). The sum of all these curves (in decibels), yielding the overall transfer function, is shown in (b).

19 Feb /04.

Some measurements of formant bandwidths
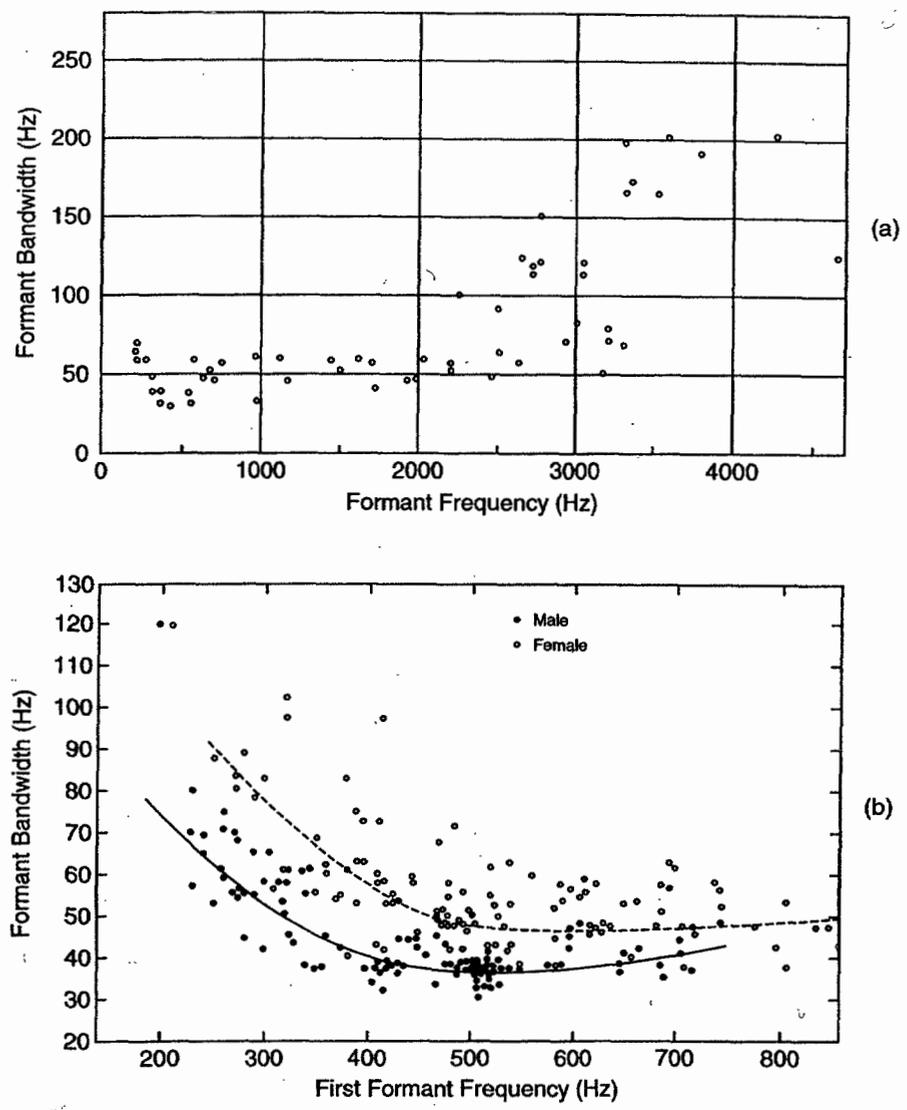
Figure 6.1 Measurements of formant bandwidths for a variety of vowels with a closed-glottis condition. The data in (a) were obtained using a sweep-tone method (Fant, 1962), and cover a range of vowel formants. The first-formant bandwidths in (b) were obtained by Fujimura and Lindqvist (1971), also using a sweep-tone method. Average curves are given for male and female speakers.

19 Feb/04
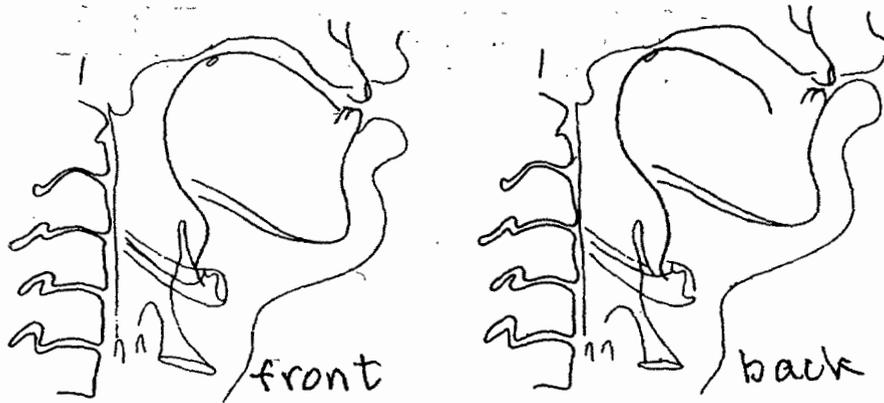
High Vowels

Figure 6.2 Midsagittal vocal tract configurations for the high vowels /i/ (left) and /u/ (right). Adult male speaker of English. (From Perkell, 1969.)

Low vowels

front /æ/   back /a/
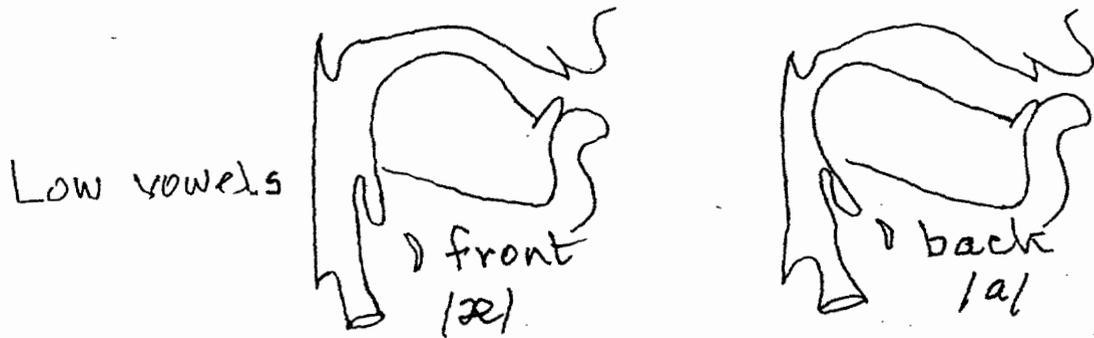
Figure 6.7 Midsagittal vocal tract configurations for the non-low, non-high vowels /e/ (left) and /o/ (right). Adult male speaker of French. (Adapted from Bothorel et al., 1986.)
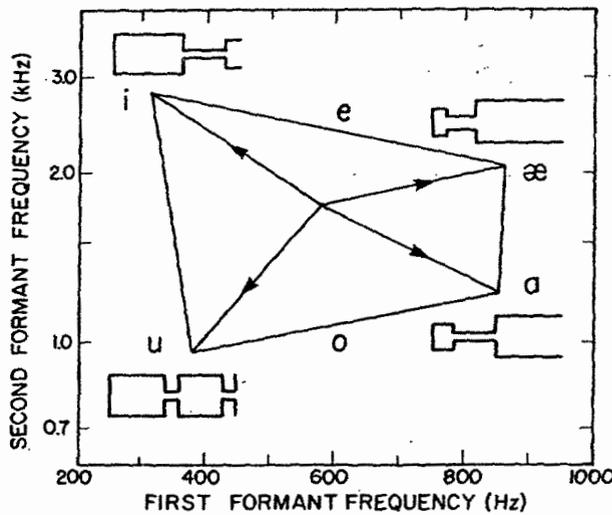
Figure 6.16 Plot of F2 vs. F1 showing how formants shift when the shape of an acoustic tube is perturbed in different ways. The midpoint represents equally spaced formants for a uniform tube of length 15.4 cm. The lines with arrows indicate how the formant frequencies change when the tube is modified as shown by the tube shapes. The corners of the diagram are labeled with vowel symbols corresponding roughly to the tube shapes. Approximate locations for the vowels /e/ and /o/ are also shown. Dimensions are selected to approximate the vocal tract size of an adult female speaker.

19 Feb/04
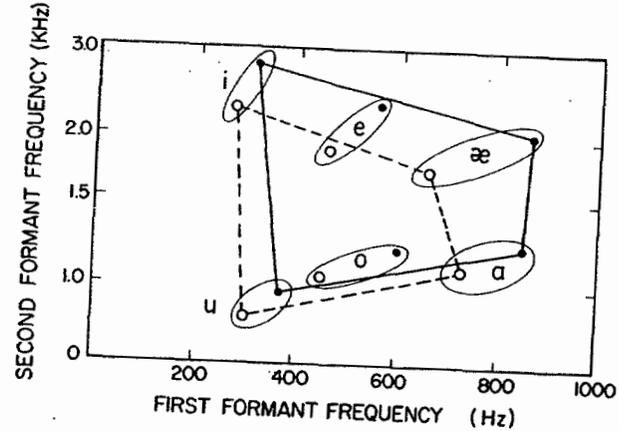
Vowels for English

Figure 6.17 Plots of F2 vs. F1 for several vowels of American English. Open circles (joined by dashed lines) are data for adult male speakers and filled circles (solid lines) are for adult female speakers. The data for the vowels /i æ ɑ u/ are averages from Peterson and Barney (1952). Data for /e o/ are averages for two male and two female speakers.

Table 6.2 Average values of the first three formant frequencies and the fundamental frequency for six basic vowels of American English produced by adult male and female speakers

Vowel	F1 Hz	F2 Hz	F3 Hz	F0 Hz	B0 Bark	B1 Bark	B2 Bark	B3 Bark	B2 − B1 Bark	B3 − B2 Bark	B1 − B0 Bark
i (female)	310	2790	3310	235	2.8	3.6	15.7	16.8	12.1	1.1	0.8
i (male)	270	2290	3010	136	1.9	3.2	14.4	16.2	11.2	1.8	1.3
e (female)	560	2320	2950	223	2.7	5.9	14.5	16.1	8.6	1.6	3.2
e (male)	460	1890	2670	130	1.8	5.1	13.1	15.4	8.0	2.3	3.3
æ (female)	860	2050	2850	220	2.6	8.1	13.7	15.9	5.6	2.2	5.5
æ (male)	660	1720	2410	127	1.8	6.7	12.5	14.7	5.8	2.2	4.9
ɑ (female)	850	1220	2810	212	2.6	8.1	10.2	15.8	2.1	5.6	5.5
ɑ (male)	730	1090	2440	124	1.7	7.2	9.5	14.8	2.3	5.3	5.5
o (female)	600	1200	2540	220	2.7	6.2	10.2	15.1	4.0	4.9	3.5
o (male)	450	1050	2610	130	1.8	5.0	9.3	15.2	4.3	5.9	3.2
u (female)	370	950	2670	232	2.8	4.2	8.7	15.4	4.5	6.7	1.4
u (male)	300	870	2240	137	1.9	3.5	8.2	14.3	4.7	6.1	1.6

Note: Frequencies are given in hertz and in bark, and bark differences are also tabulated. Data for the vowels /i æ ɑ u/ are taken from Peterson and Barney (1952). Data for /e o/ are from a separate study with two female and two male speakers.
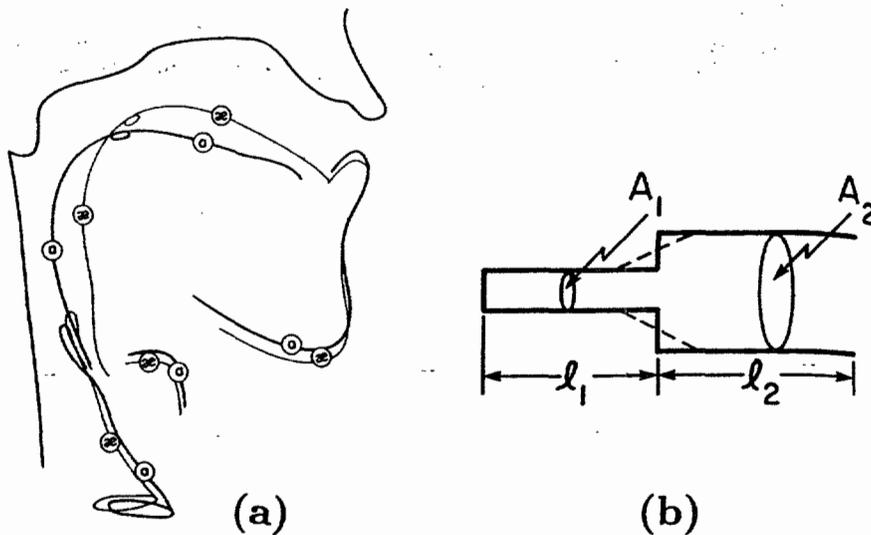
Distinction between /a/ (back) and /æ/ (front)

Figure 6.8 (a) Superimposed midsagittal configurations for the low vowels /æ/ and /ɑ/. (From Perkell, 1971.) (b) Model of low vowel vocal tract shape as a concatenation of two tubes. The dashed line indicates a tapered transition between the tubes.
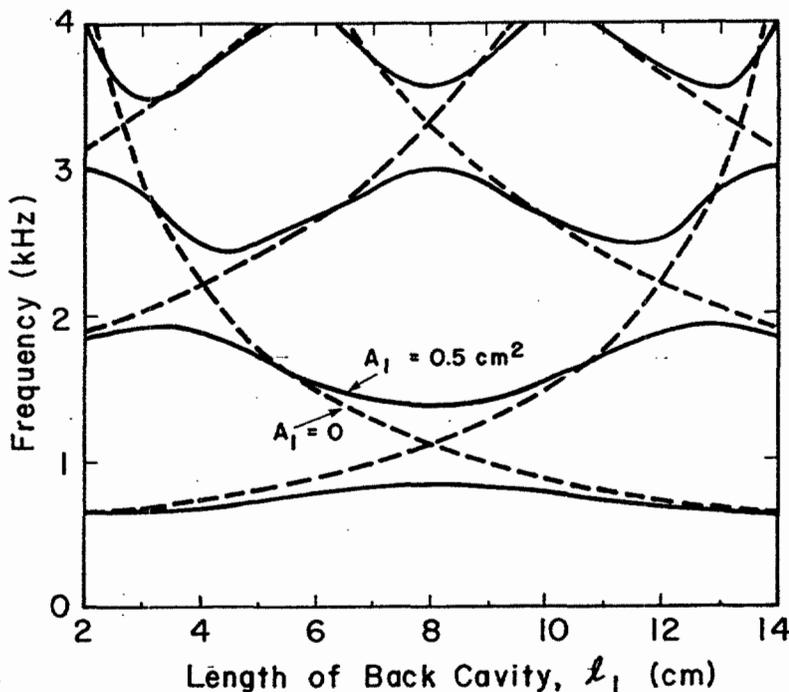
Figure 6.9 Frequencies of the first four natural frequencies for the nontapered configuration of figure 6.8, as the length ℓ_1 of the back cavity is manipulated. The total length $\ell_1 + \ell_2 = 16$ cm, and the cross-sectional area $A_2 = 3$ cm². The dashed line corresponds to the case where $A_1 \ll A_2$, and the solid line is for $A_1 = 0.5$ cm². The radiation impedance is assumed to be zero. (From K. N. Stevens, 1989.)