Liquids

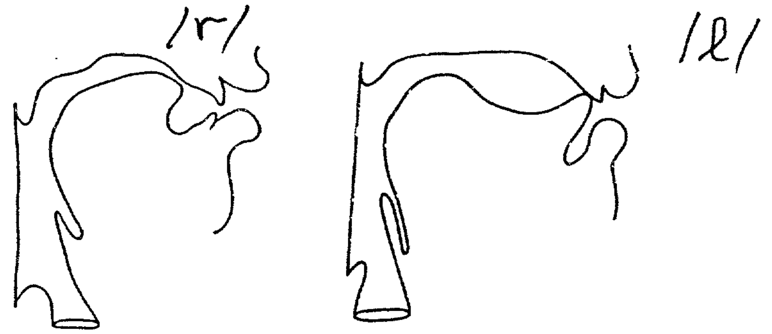


Figure 9.34 Midsagittal sections of the vocal tract for the liquid consonants /r/ (left) and /l/ (right). The midsagittal sections do not reflect the fact that in each case there is an acoustic path around the lateral edges of the tongue blade. The representation of /l/ is adapted from Narayanan et al. (1995b).

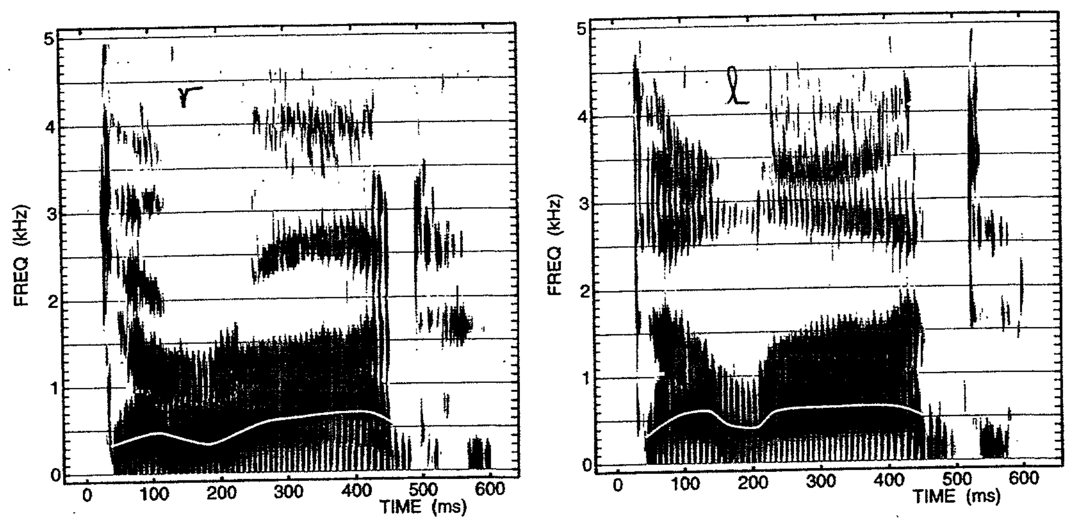


Figure 9.35 Spectrograms of the utterances /də'rɑd/ (left) and /də'lɑd/ (right) produced by a male speaker. The time course of the first-formant frequency during the liquid consonant is overlaid on the spectrogram.

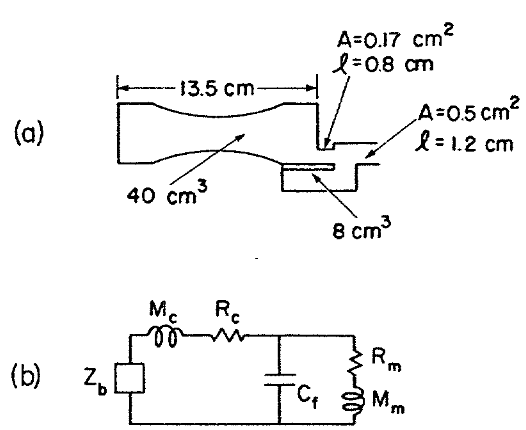


Figure 9.36 (a) Schematized model of the vocal tract used for estimating the acoustic characteristics of a retroflex consonant. The side branch representing the space under the tongue is shown, together with the constrictions formed by the tongue blade and the lips. (b) Equivalent circuit used to obtain the approximate frequency of the front-cavity resonance of the configuration in (a). The volume of the front cavity is represented by C_f, and the constrictions posterior and anterior to the cavity are represented by R_c, M_c, R_m, and M_m. Z_b is the impedance of the cavity behind the constriction.

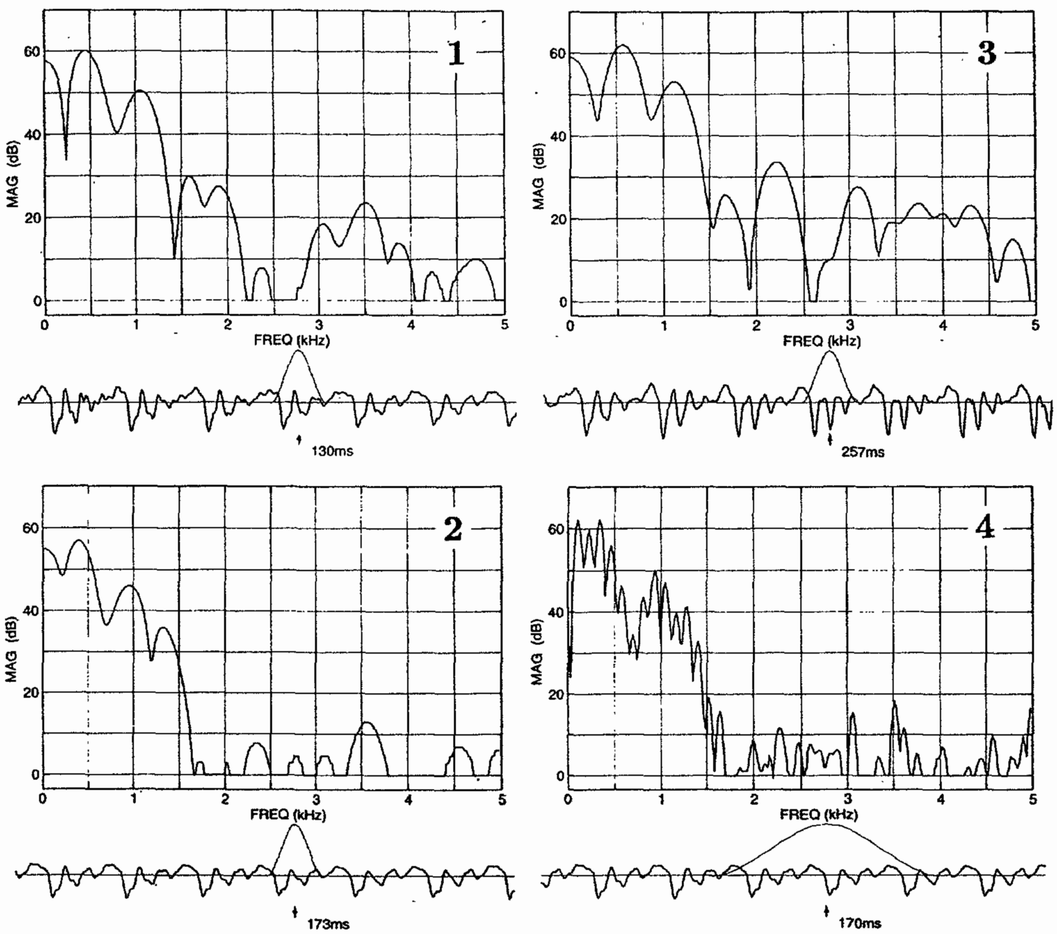
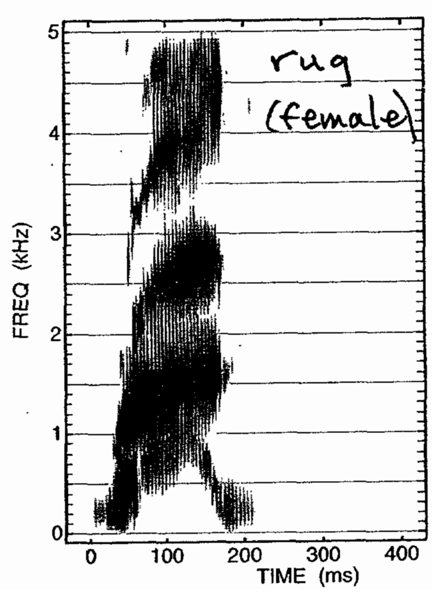


Figure 9.40 Spectra sampled at several points in the utterance /də'rad/, for which a spectrogram is shown in figure 9.35. Spectra 1, 2, and 3 are calculated using a short (6.4 ms) time window, whereas spectrum 4 uses a longer (30 ms)time window. Spectra 2 and 4 are sampled near the most constricted part of the consonant /r/, and show the close proximity of $F2$ and F_R, at about 950 and 1350 Hz, with evidence of a weak $F3$ peak at 2300 Hz and a zero just below 2000 Hz. Spectra 1 and 3 are sampled during the initial movement from the initial vowel and the final movement toward the final vowel, respectively. These spectra show a double peak corresponding to F_R and $F3$ in the frequency range 1.5 to 2.5 kHz.



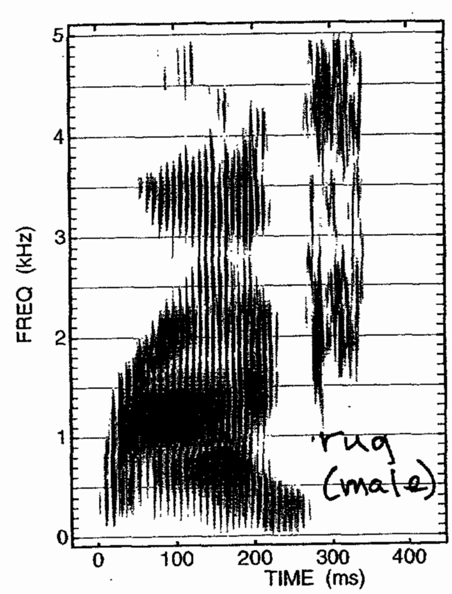


Table 9.2 Ranges of average values of frequencies (in hertz) of first three resonances of prevocalic /r/ as reported by Espy-Wilson (1987), Hagiwara (1995), and Westbury (1995)

	$F1$	$F2$	F_R
Female speakers	360–480	1030–1240	1800–2050
Male speakers	330–430	880–1200	1380–1610

Note: The different studies used different numbers of speakers and contexts.

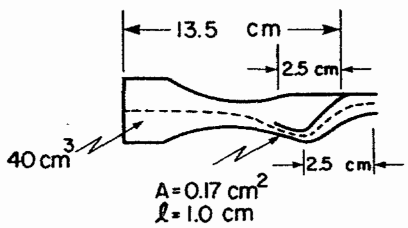


Figure 9.43 Schematization of the vocal tract shape for purposes of calculating the acoustic characteristics of a lateral consonant. Approximate dimensions are shown for the back cavity, the side branch formed by the tongue blade, and the cavity in front of the constriction. The dashed line indicates the midline of the acoustic path from glottis to lips.

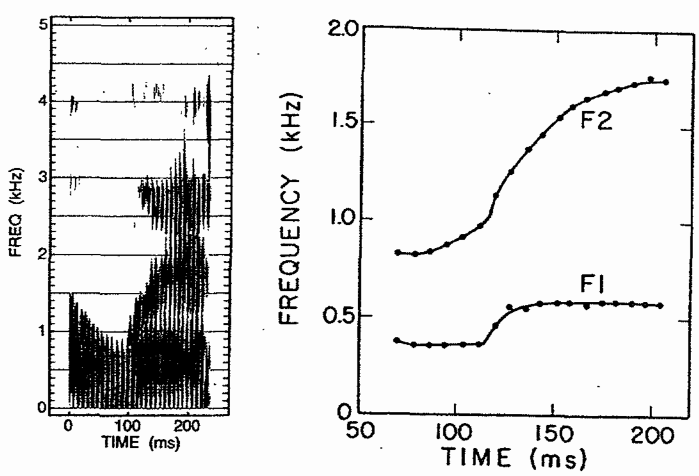


Figure 9.48 Left, Spectrogram of the utterance *a let*. Right, Frequencies of first two formants preceding and following the release of /l/, with points given at each glottal period.

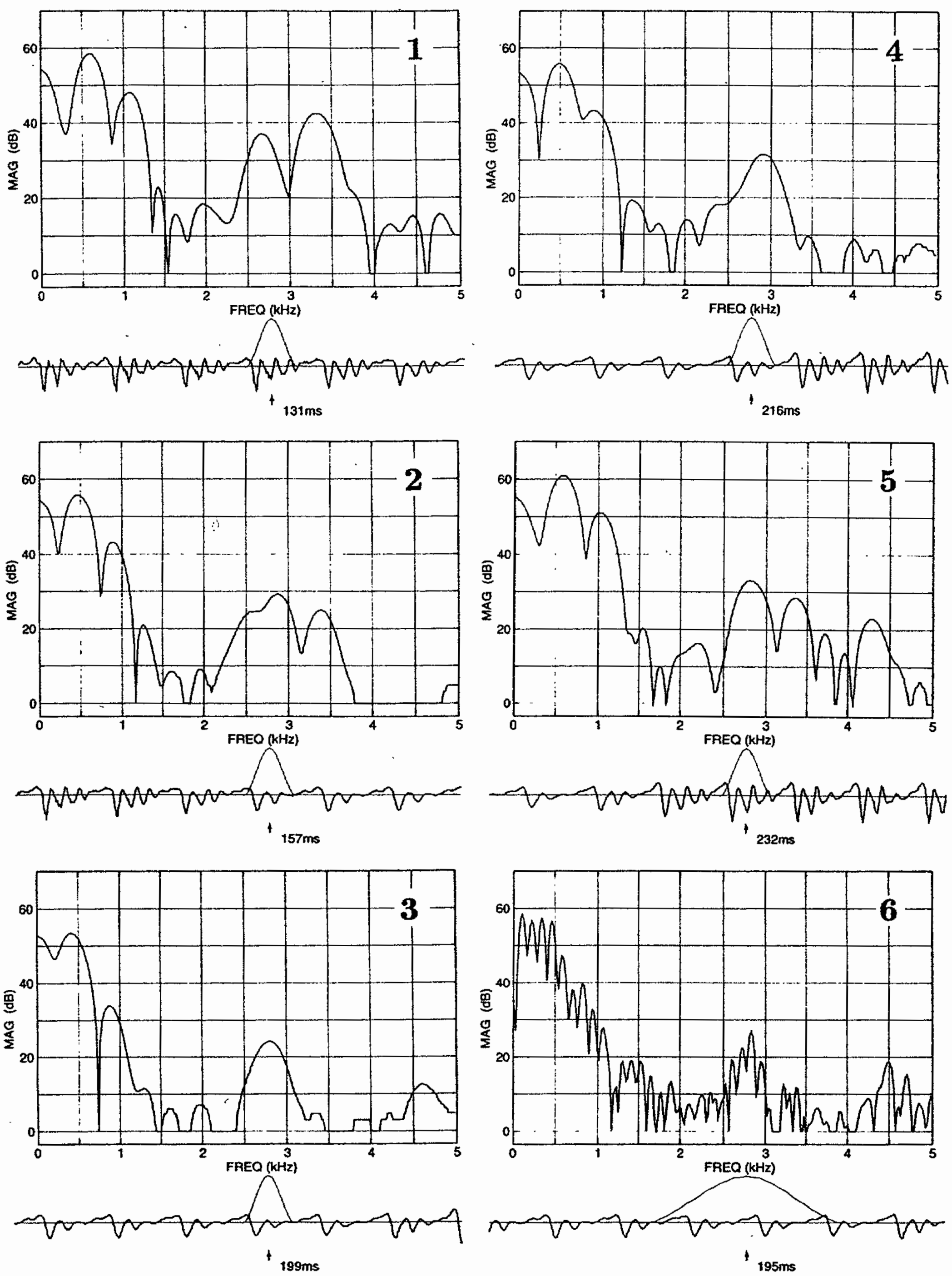
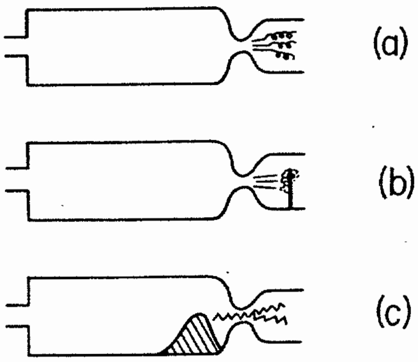


Figure 9.46 Spectra sampled at several points within the utterance /də'lɑd/, a spectrogram of which is shown in figure 9.35. Spectra 1 to 5 are obtained with a time window of 6.4 ms, centered on individual glottal pulses. At time 157 ms (spectrum 2), there is evidence of a double peak in the *F3* range, apparently resulting from two slightly asymmetrical acoustic paths around the tongue blade. The increase in spectrum amplitude of the *F4* peak (around 3400 Hz) is especially abrupt from 217 to 233 ms (spectra 4 and 5). Spectrum 6 is obtained with a 30-ms time window centered at a time immediately prior to the release (about the same time as spectrum 3). The peak at 1400 Hz is assumed to be a tracheal resonance.

Noise due to turbulence in flow



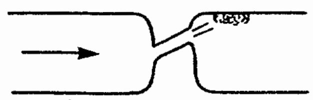


Figure 2.31 Airflow impinges on the wall of a tube downstream from a constriction, to generate turbulence noise at the wall.

Figure 2.30 Schematic representation of three types of acoustic sources due to turbulence in the airstream near a constriction: (a) no obstacle; (b) obstacle in the airstream downstream from the constriction; (c) fluctuation in flow through the constriction.

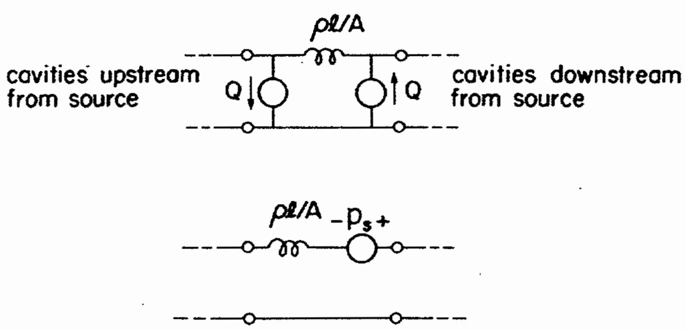


Figure 2.32 Equivalent circuit representation of a turbulence noise source as a one-dimensional dipole with two equal and opposite volume velocity sources Q (top) and as a sound pressure source (bottom).

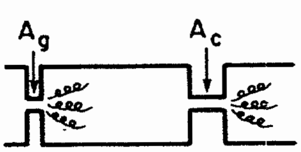


Figure 2.36 Tube with two constrictions, with turbulence noise generated at both constrictions. The cross-sectional areas of the constrictions are A_g and A_c, as indicated, representing the glottal and supraglottal constrictions.

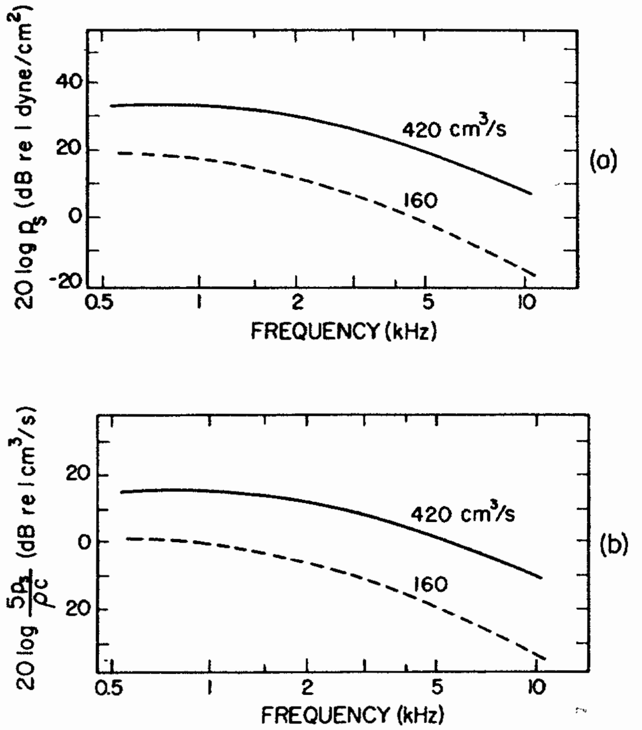


Figure 2.33 Spectrum of sound pressure source p_s for a configuration similar to that in figure 2.30b, for two values of airflow. Distance from the constriction to the obstacle is 3 cm, and the diameter of the (circular) constriction is 0.32 cm. The cross-sectional area of the tube is 5.0 cm^2. The spectra are in 300-Hz bands of frequency. (a) The ordinate is the absolute level of the sound pressure source, in decibels re 1 dyne/cm^2. (b) The same curves are given, but with an ordinate that is in units of volume velocity, to permit comparison with other volume velocity sources in the vocal tract. The sound pressure source p_s is normalized by dividing by the characteristic impedance $\rho c/5$. For 0 dB on the ordinate, $5p_s/(\rho c) = 1$ cm^3/s. When the area A_t of the tube downstream from the constriction is different from 5 cm^2, the curves in (b) should be scaled up by $20 \log (A_t/5)$ dB. For the upper curve in both panels, the value of $20 \log U^3 A^{-2.5} = 212$ dB, where U = volume velocity in cubic centimeters per second and A = area of constriction in square centimeters. When $20 \log U^3 A^{-2.5}$ is different from 212 dB, the upper curve should be scaled up by $20 \log U^3 A^{-2.5} - 212$ dB. Curves are based on experimental data of Shadle (1985a).

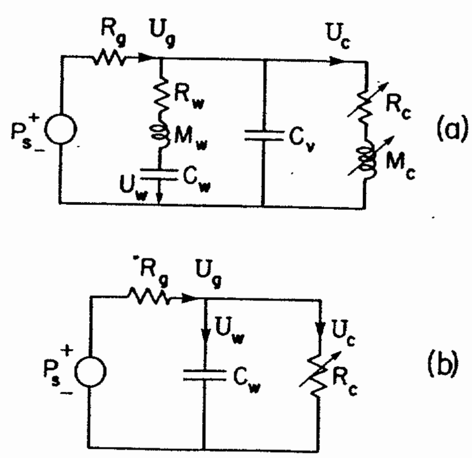


Figure 2.42 (a) Equivalent circuit of the vocal tract that can be used to estimate the time course of the turbulence noise source at the time-varying supraglottal constriction following the release of a stop consonant. (b) Equivalent circuit containing the principal elements that determine the turbulence noise source.

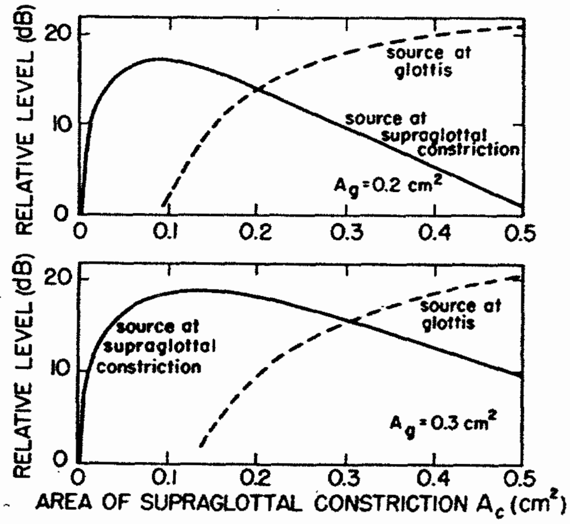


Figure 2.37 Relative level of noise source at each constriction of figure 2.36, as a function of the cross-sectional area of the supraglottal constriction. The cross-sectional area A_g is fixed at a different value for each panel. The constant K in equation (2.21) is assumed to be the same for the noise generated at each constriction.

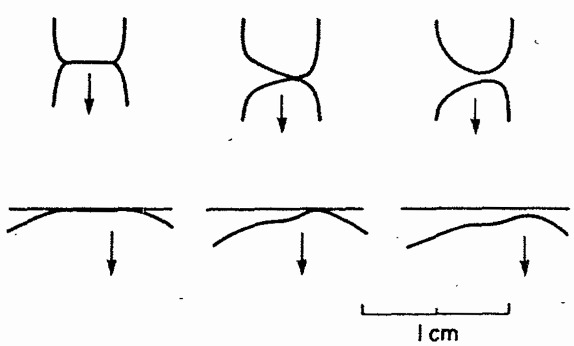


Figure 2.46 Schematic representation of the shape of the lips (top) and tongue surface (bottom) as a labial or a velar stop consonant is released. At the instant of release, the length of the constriction is relatively short, due to the distortion of the surfaces by the action of the intraoral pressure.

Stop consonants

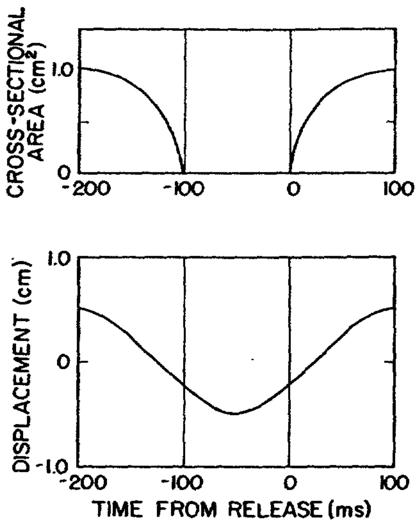


Figure 7.1 Schematic representation of articulatory movements when a stop consonant is produced in intervocalic position. Time from release of the closure is plotted on the abscissa. The upper panel is a representation of the cross-sectional area at a point in the vocal tract where the constriction is formed. The lower panel indicates the type of trajectory that might be followed by a secondary articulatory structure such as the mandible or the tongue body as the intervocalic stop consonant is produced. The durations, rates, and extents of movement are selected to be within the ranges observed in actual utterances, but do not necessarily correspond to a particular stop-consonant trajectory. The time course of the cross-sectional area may be modified somewhat from the smooth trajectory shown above as a consequence of forces due to the intraoral pressure.

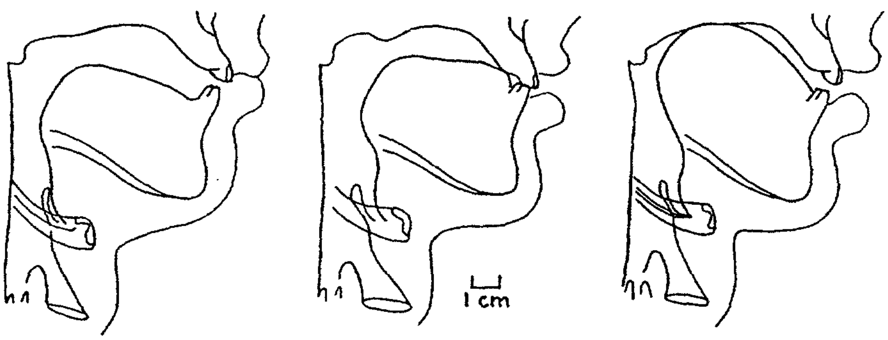


Figure 7.2 Midsagittal sections through the vocal tract during the production of stop consonants produced by closing the lips (left), raising the tongue tip to the alveolar ridge (middle), and raising the tongue body (right). Adult male speaker. (From Perkell, 1969.)

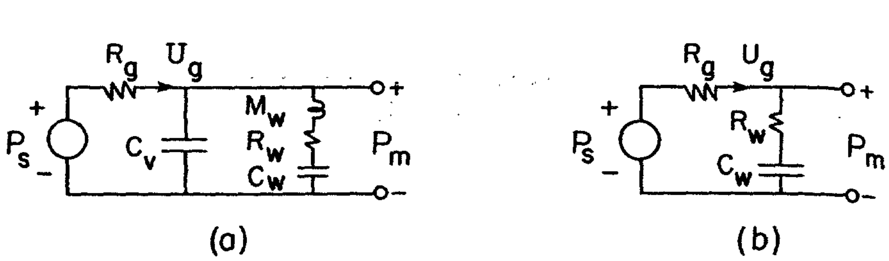
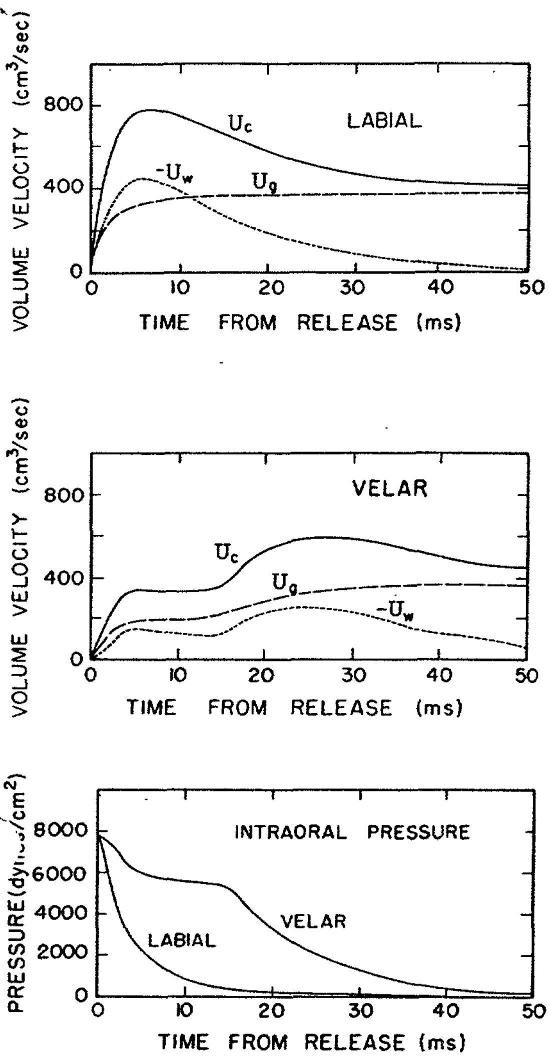


Figure 7.3 (a) Low-frequency equivalent circuit of the vocal tract when there is a complete closure in the oral cavity. (b) Approximate equivalent circuit that is used for calculating intraoral pressure P_m and glottal airflow U_g immediately following closure of a stop consonant. See text.



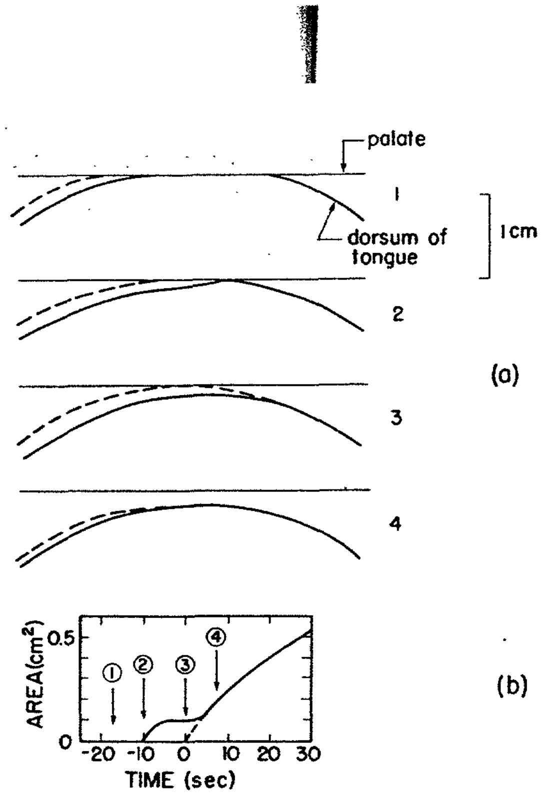


Figure 7.6 (a) Schematic representation of the surfaces of the tongue and the palate at several times during the release of a stop consonant produced by making a closure with the tongue body. The dashed lines indicate the tongue contour if there were no intraoral pressure, and the solid lines show the contour as modified by forces due to the intraoral pressure. (b) Estimated cross-sectional area vs. time at the constriction without the influence of intraoral pressure (dashed line) and with intraoral pressure (solid line). The labeled arrows indicate the points in time corresponding to the panels in (a).

Figure 7.8 Calculated airflows (top two panels) and intraoral pressures (bottom panel) as a function of time following the release of voiceless unaspirated stop consonants. The constriction area follows a trajectory of the form $A_c = A_{max}(1 - e^{-t/\tau})$, where A_{max} is taken to be 1.0 cm^2, and $1/\tau = 100\ s^{-1}$ for labials releases and 25 s^{-1} for velars. The area for the velar release is further modified as in figure 7.6. The glottal area is assumed to be constant at 0.1 cm^2. Initial intraoral pressure is 8 cm H_2O, and other parameters are as given in figure 7.4. Negative U_w indicates that the walls of the vocal tract displace inward following the release. Note that $U_g = U_c + U_w$.

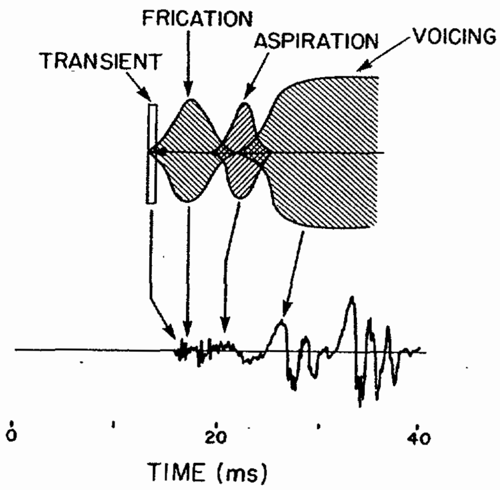


Figure 7.17 Schematic representation of sequence of events at the release of a voiceless unaspirated stop consonant. (From K. N. Stevens, 1993b.)

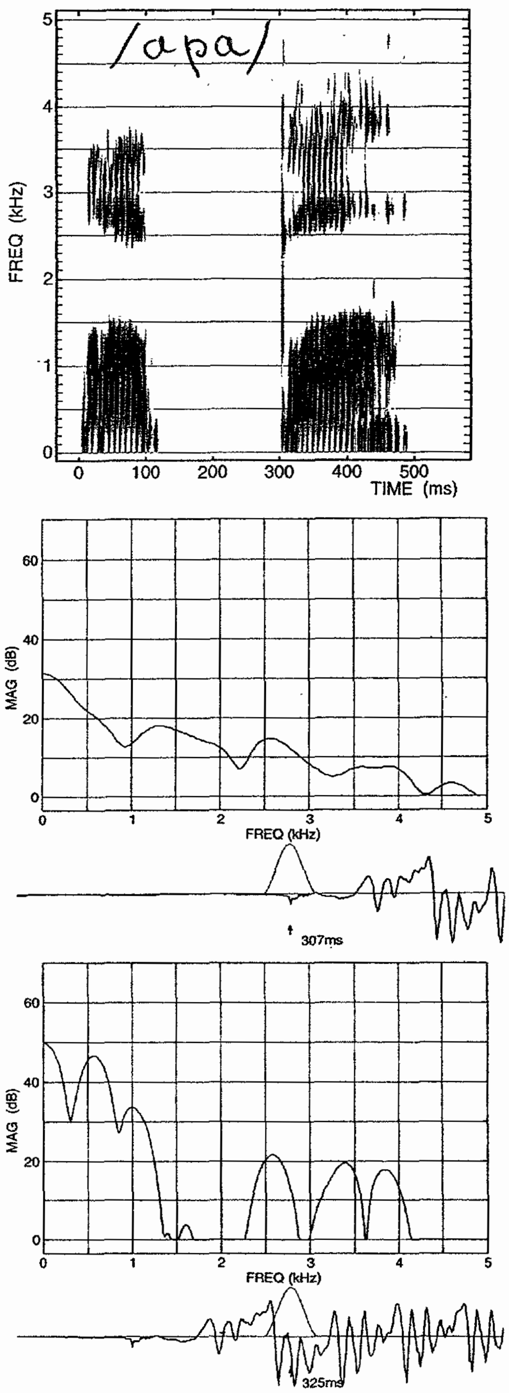


Figure 7.20 Acoustic data obtained from an utterance of a voiceless unaspirated stop consonant /p/ in the intervocalic environment /ɑpɑ/ (the same utterance as that displayed in figure 7.12a). A spectrogram of the utterance is shown at the top, and spectra and waveforms at the burst and 18 ms later are given below. The spectra are discrete Fourier transforms calculated with a Hamming window of 6.4-ms duration.

/ata/

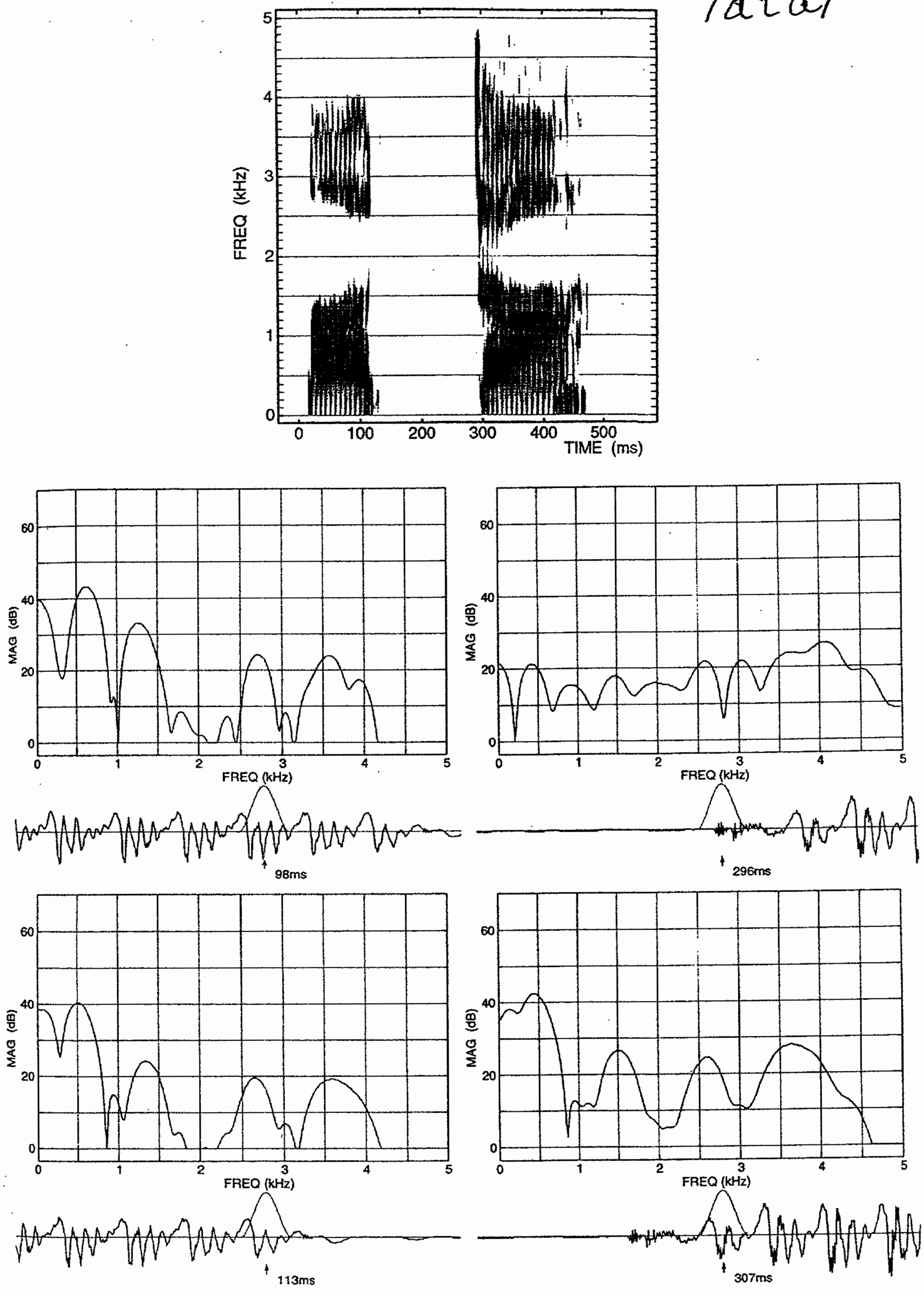
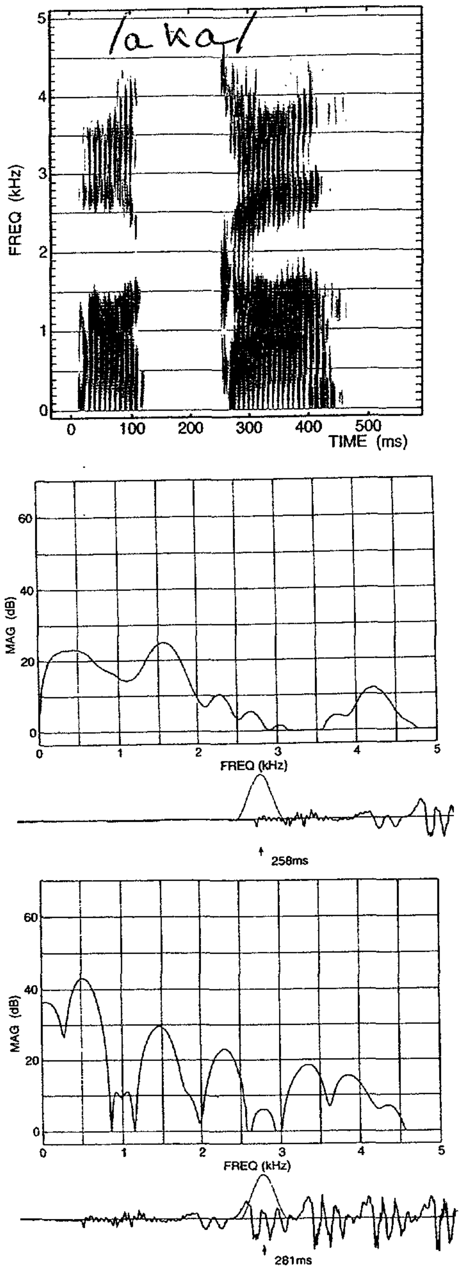


Figure 7.27 Acoustic data obtained from an utterance of a voiceless unaspirated stop consonant /t/ in the intervocalic environment /ata/. The four spectra below the spectrogram are sampled at selected points before the consonant closure and after the release, as indicated by the waveforms and time windows below the spectra. See legend for figure 7.20.



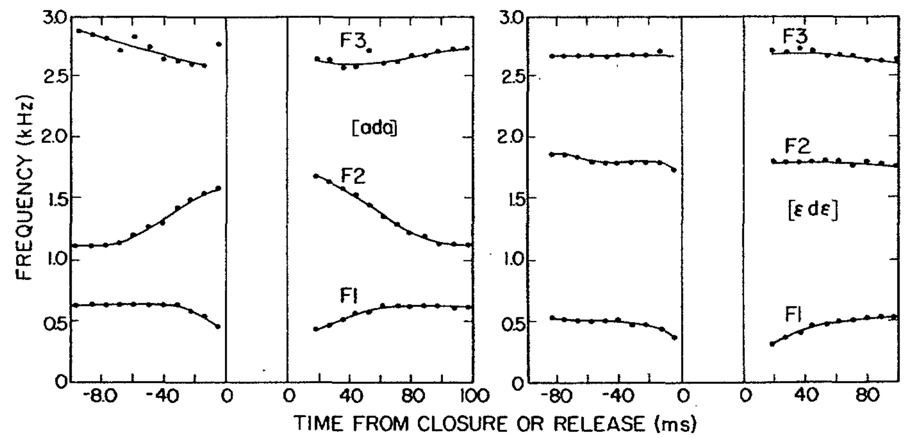


Figure 7.23 Measured frequencies of the first three formants for the utterances /ada/ (left) and /ɛdɛ/ (right) produced by a male speaker. See legend for figure 7.15.

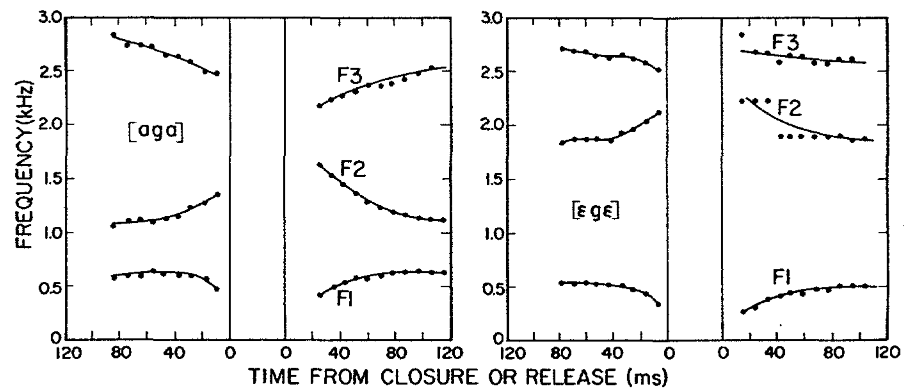


Figure 7.31 Measured frequencies of the first three formants for the utterances /aga/ (left) and /ɛgɛ/ (right) produced by a male speaker. See legend for figure 7.15. The frequency of the major spectral peak in the vicinity of $F2$ for /ɛgɛ/ shows a discontinuity at about 40 ms. This discontinuity is presumably a result of interference by the third subglottal resonance for this utterance. The $F2$ curve is an estimate of the $F2$ movement if this interference were not present.

Figure 7.35 Acoustic data obtained from an utterance of a voiceless unaspirated stop consonant /k/ in the intervocalic environment /aka/. See legend for figure 7.20.

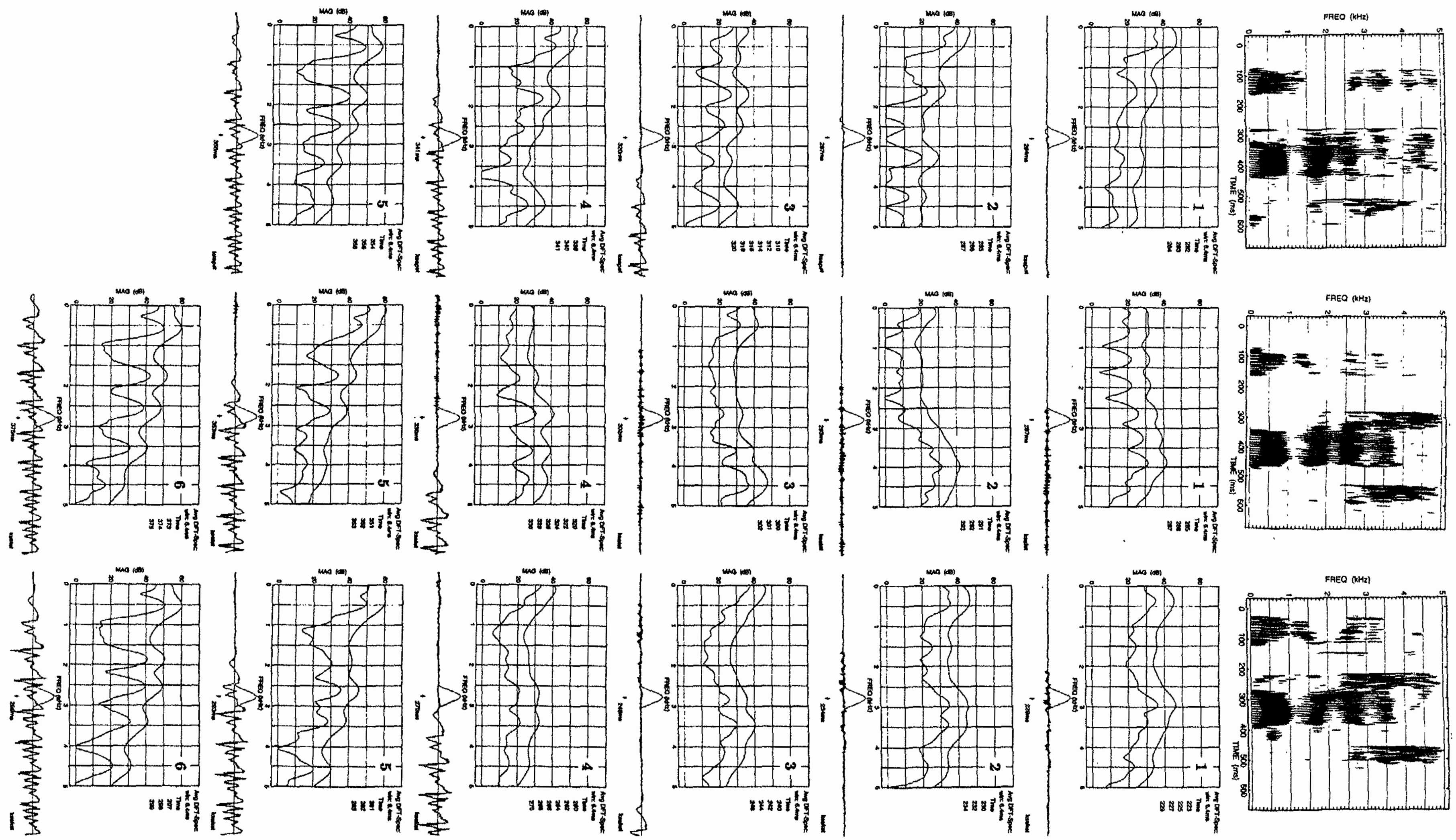


Figure 8.67 Spectrograms and spectra for the utterances /ə'pʰɛt/ (left), /ə'tʰɛt/ (middle), and /ə'kʰɛt/ (right). In each case, the spectrogram is shown at the top, and spectra (time window of 6.4 ms) sampled at several times following the release are given below the spectrogram. The spectra are obtained by averaging spectra sampled at several adjacent times in the burst, aspiration, or vowel onset. The times at which these spectra were sampled are shown at the right of the displays. Two curves are shown on each spectrum display: the lower curve is the discrete Fourier transform (DFT) and the upper curve is a smoothed version of the DFT, with a smoothing window of about 400 Hz. Waveforms are shown below each spectrum. See text.