6.541 HST710 4 Mar 04

Voiceless aspirated stop consonants

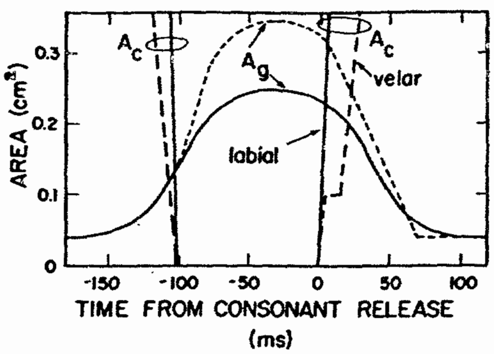


Figure 8.58 The various curves represent the assumed time variation of the cross-sectional area of the glottal opening A_g and the supraglottal constriction A_c when a voiceless aspirated stop consonant is produced in intervocalic position. In the case of the curves for A_g, the solid curve represents the trajectory A_g would have if there were no pressure increase in the mouth; it is the same A_g that was assumed for /h/, in figure 8.34. The short-dashed curve represents the actual A_g that would occur as a consequence of the abducting forces due to the increased intraoral pressure for the stop consonant. The sloping straight line following the consonant release (from $t = 0$ to $t = 70$ ms) is an approximation to $A_g(t)$ during this time interval, and was used to determine the calculated curves in figure 8.60. The supraglottal curves $A_c(t)$ are intended to approximate the changes in cross-sectional area for a velar consonant (dashed lines) and for a labial consonant (solid lines).

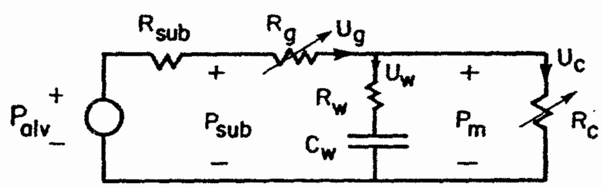


Figure 8.59 Equivalent circuit used to calculate the flows and pressures at the release of a voiceless aspirated stop consonant. Values of the elements are discussed in the text and in the legend to figure 7.4. $R_{sub} = 1.5$ acoustic ohms.

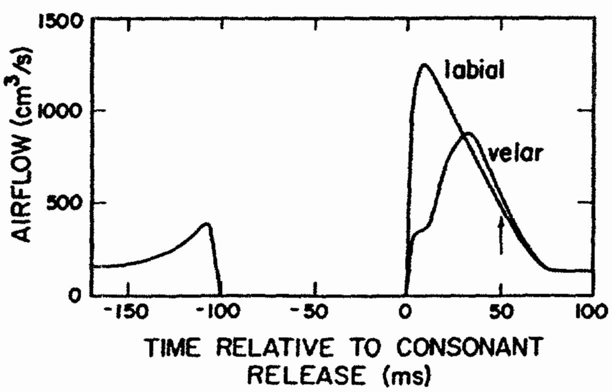
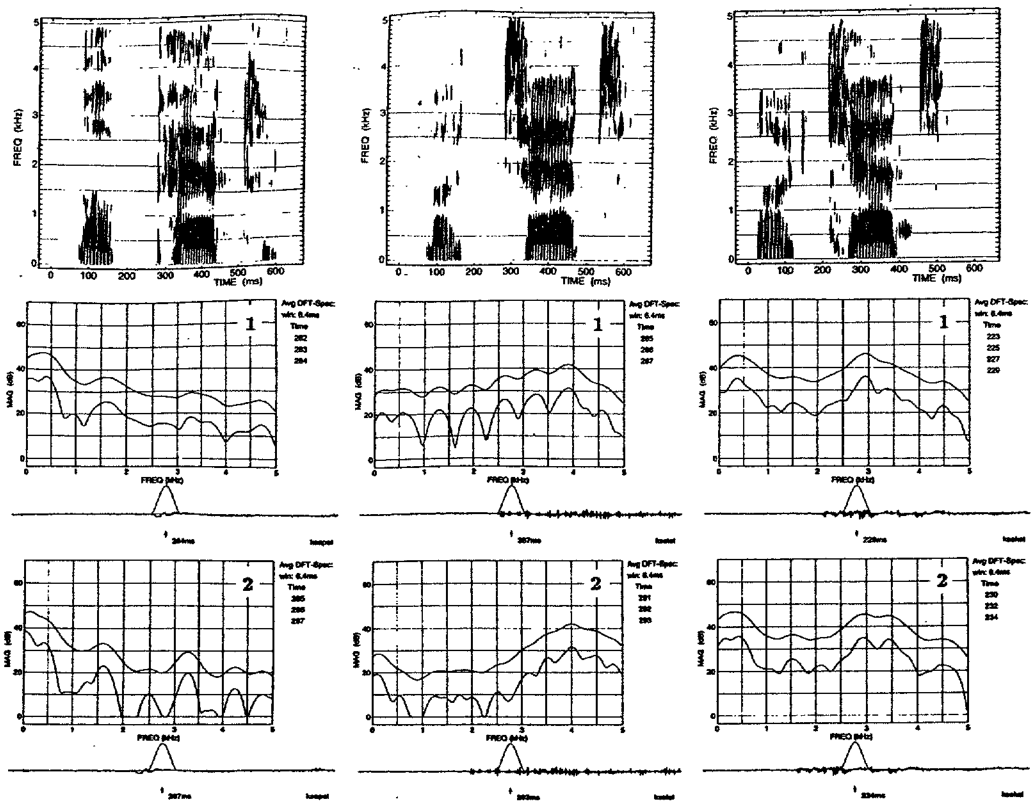
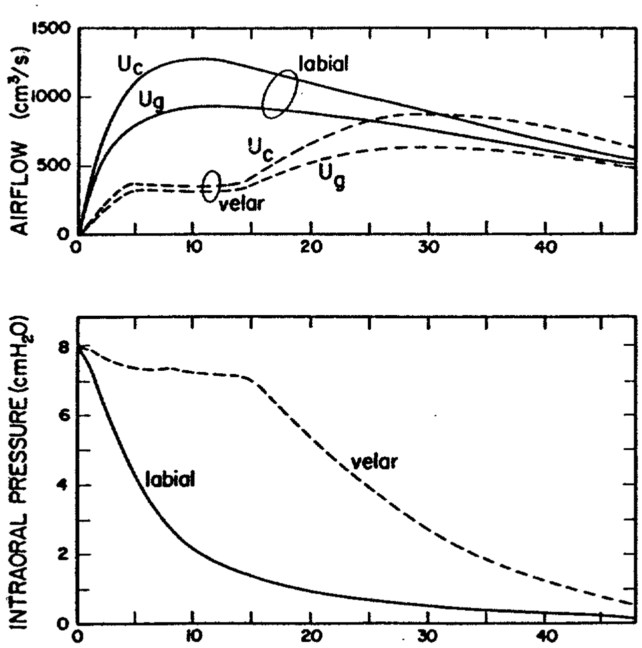


Figure 8.60 Calculated values of airflow versus time for intervocalic voiceless aspirated stop consonants, corresponding to the time variation of glottal and supraglottal areas given in figure 8.58. Calculations are based on the equivalent circuit in figure 8.59, with a subglottal pressure of 8 cm H_2O. Element values for C_w and R_w are given in the legend for figure 7.4. The estimated time of onset of vocal fold vibration is indicated by the arrow at 50 ms.

/ə'pɛt/ /ə'tɛt/ /ə'kɛt/





calculated airflows & pressures

Figure 8.61 Calculations of airflow (top) and intraoral pressure (bottom) at the release of a voiceless aspirated stop consonant. Curves are shown for airflow through the supraglottal constriction (U_c) and through the glottis (U_g). Dashed lines represent estimates for velar release, and solid lines for labial release. See text.

Voiced stop consonants

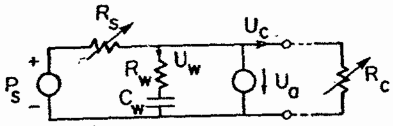


Figure 8.68 Low-frequency equivalent circuit for estimating flows and pressures in the vocal tract for voiced obstruent consonants. The volume velocity source U_a accounts for the increase in vocal tract volume due to active expansion of this volume. The resistance R_c is connected to the circuit during fricatives and after release of the consonants.

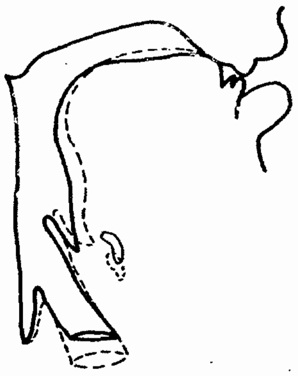
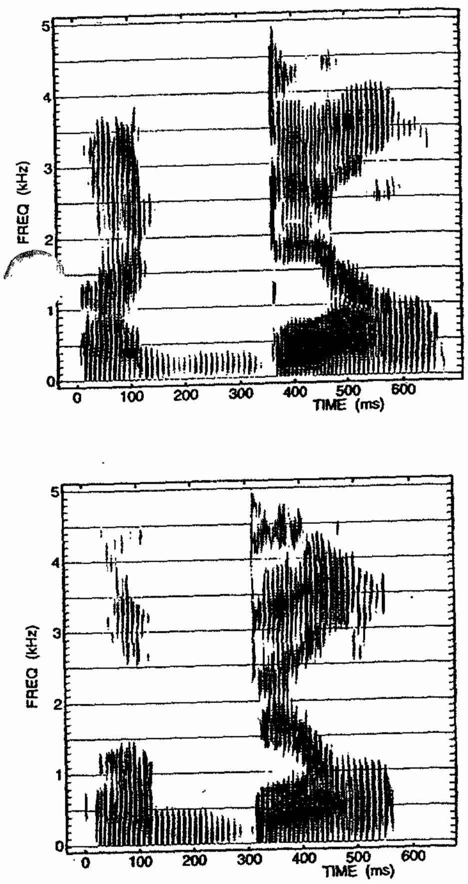


Figure 8.69 Schematic representation of midsagittal section of the vocal tract at two points in time during production of the voiced stop consonant /d/ in intervocalic position: at the time the closure for the consonant is produced (solid contour), and immediately before the consonant is released (dashed contour). The changing midsagittal contour is intended to illustrate the expansion of the vocal tract volume during the closure interval for the consonant.



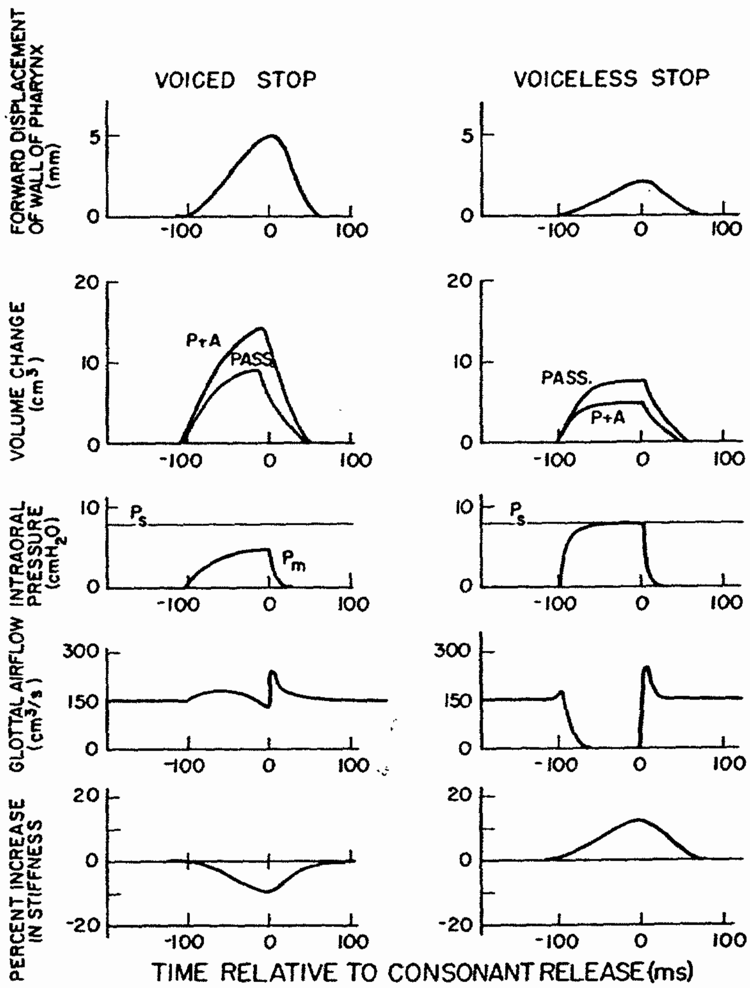


Figure 8.70 Schematized representation of the time course of various articulatory and aerodynamic parameters during the production of an intervocalic voiced stop consonant (left) and voiceless stop consonant (right), both consonants being unaspirated. The implosion of the consonant occurs at time −100 ms on the abscissa and the consonant release occurs at time zero. An intraoral pressure of 8 cm H_2O is assumed. The top panel in each column gives the forward displacement of the pharyngeal wall (data estimated from Perkell, 1969). The next panel shows the estimated change in volume of the vocal tract during the closure interval. The curves labeled PASS. give the estimated passive volume change as a consequence of the increased intraoral pressure; the curves P + A are the estimated combined volume increase due to active volume change as well as passive change. The third panel gives the intraoral pressure P_m, with the subglottal pressure indicated by the horizontal line. The fourth panel shows the estimated glottal airflow. The bottom panel gives estimates of the percent change in vocal fold stiffness, relative to the stiffness in the vowels, that occurs during the closure interval and immediately following the release.

Fricative consonants

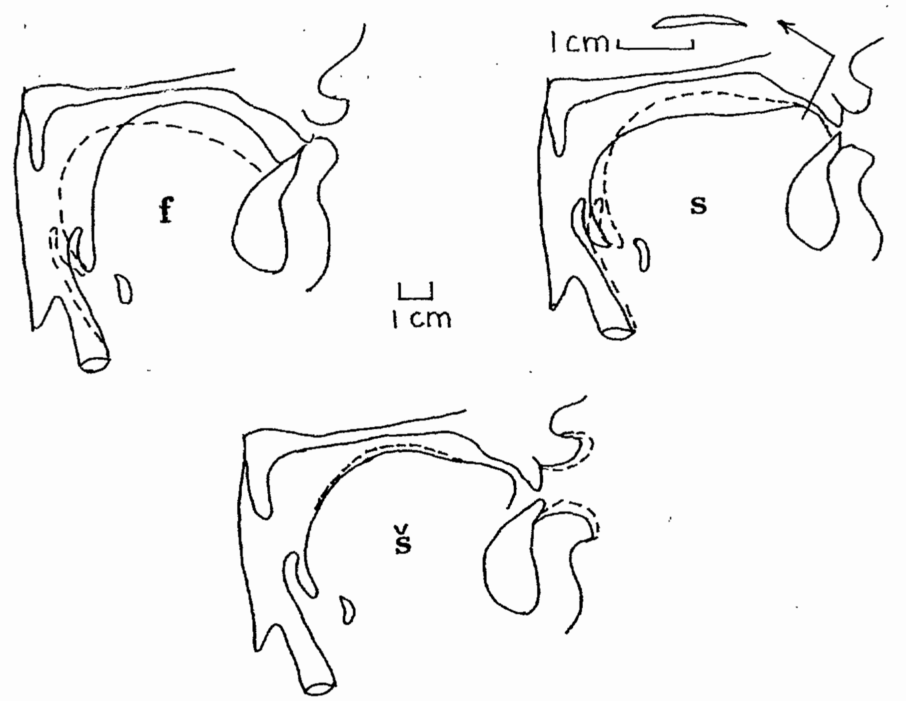
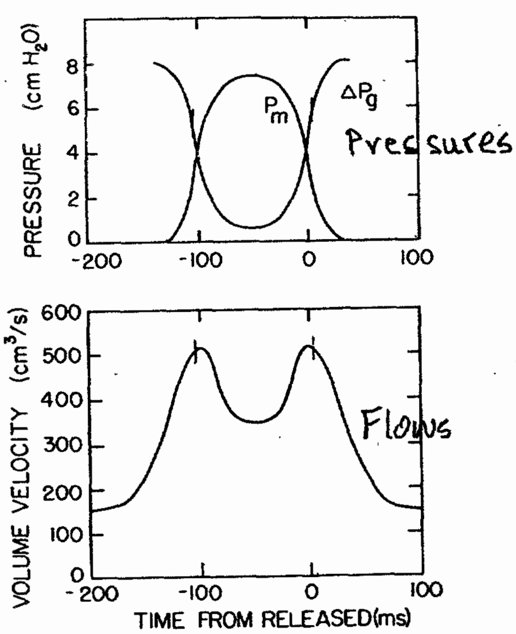


Figure 8.1 Midsagittal sections obtained from cineradiographs during the production of the three fricative consonants /f/, /s/, and /š/, as indicated. The subject is an adult female speaker of French. The approximate scale is given between the /f/ and /s/ configurations. In the case of the /s/ configuration, an estimate of the cross-sectional shape at the constriction is given. (This shape is drawn with an enlarged scale, as shown.) For each panel, two midsagittal sections for the tongue (and, in the case of /š/, for the lips) are shown, representing the fricative in two different phonetic environments. For /f/, the following vowel contexts are the high front rounded vowel /y/ (solid line) and /ɔ/ (dashed line); for /s/, /a/ (solid line) and /i/ (dashed line); and for /š/, /a/ (solid line) and /u/ (dashed line). (After Bothorel et al., 1986.)



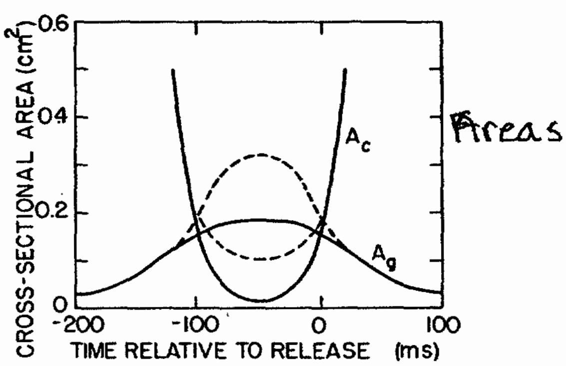


Figure 8.2 Estimates of the time course of the area of the glottal opening (A_g) and the area of the supraglottal constriction (A_c) when a fricative consonant is produced in intervocalic position. The solid lines indicate the area trajectories that would occur in the absence of a subglottal pressure, and the dashed lines indicate the modified areas that are calculated when forces on the structures due to the increased intraoral pressure are taken into account. The subglottal pressure is assumed to be 8 cm H_2O.

Figure 8.3 Top: Time course of intraoral pressure (P_m) and transglottal pressure (ΔP_g) corresponding to the area trajectories for an intervocalic fricative consonant, in figure 8.2. The vertical marks on the ΔP_g curve indicate times when vocal fold vibration is expected to stop and to start, that is, at about −105 ms and 5 ms. Bottom: Calculated airflow vs. time for the intervocalic fricative consonant. Only the average flow is shown in the figure. There are fluctuations in the flow when the vocal folds are vibrating. Offset and onset of vocal fold vibration are expected at times indicated by the vertical marks.

/afa/

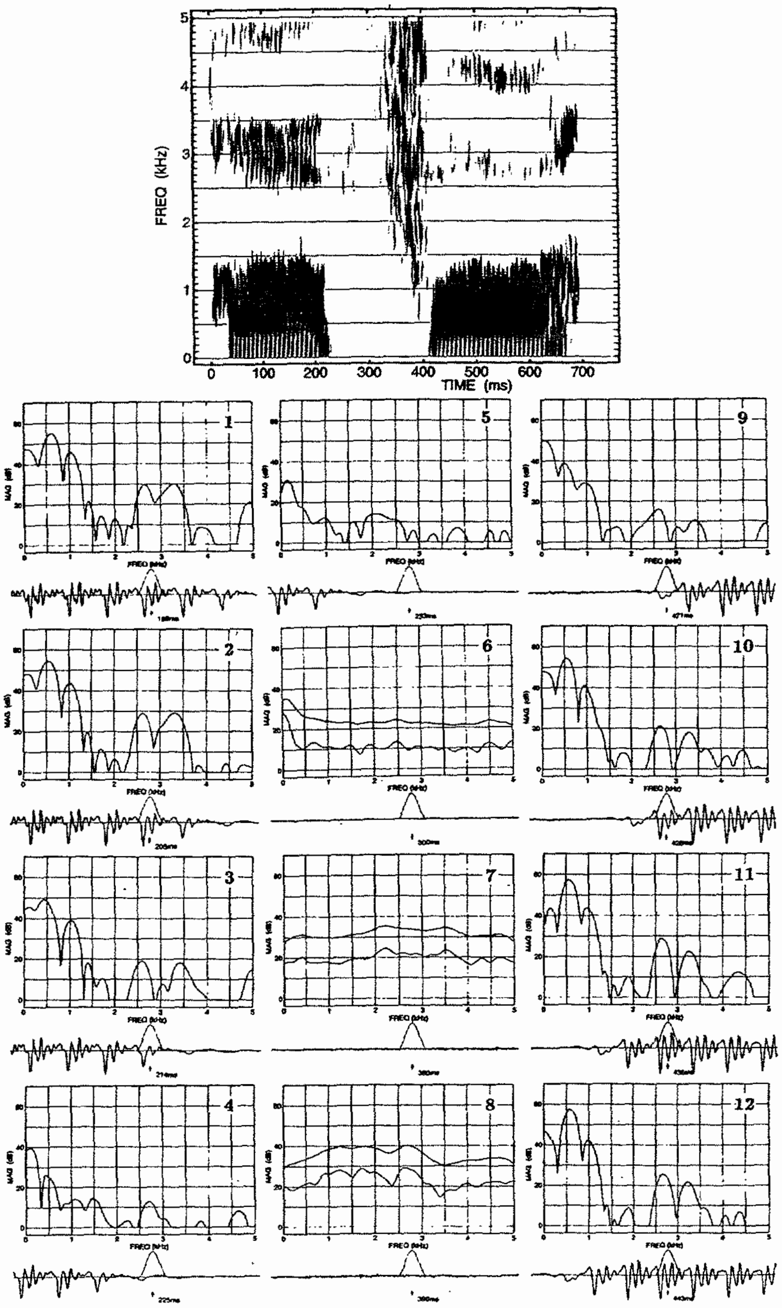


Figure 8.12 A spectrogram of the utterance /afa/ is shown at the top. The spectra in the three columns immediately below the spectrogram are sampled at several points in three regions of the utterance: in the vowel immediately preceding the consonant (left, panels 1 to 4); within the fricative region (middle, panels 5 to 8); and in the vowel immediately following the consonant (right, panels 9 to 12). These spectra are obtained with a 6.4-ms Hamming window. In the vowel regions, spectra are sampled at four adjacent glottal periods, whereas in the consonant the spectra are sampled more sparsely. In panels 6 to 8, where there is no evidence of glottal vibration, the lower spectra are averages of a series of spectra obtained with a 6.4-ms window over a 20-ms interval. The upper curves in these panels are smoothed versions of the average spectra. Waveforms and time windows, together with the times at which the spectra are sampled, are shown below each spectrum panel.

/asa/

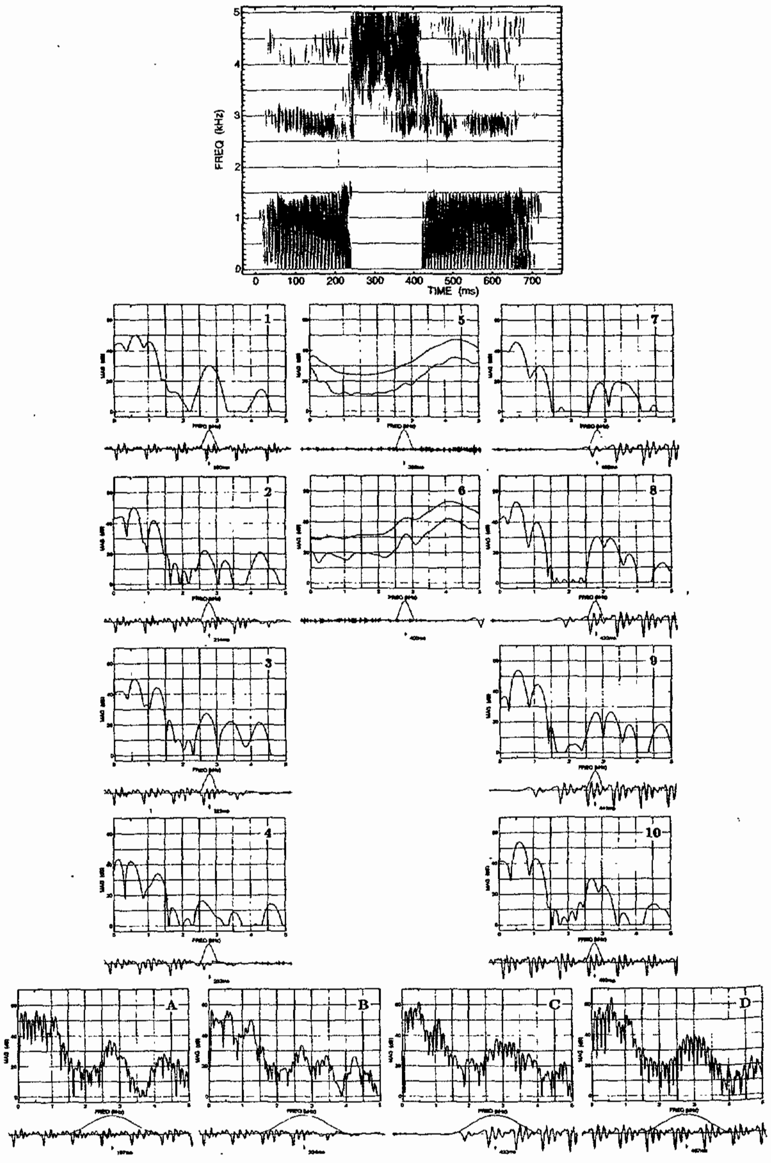
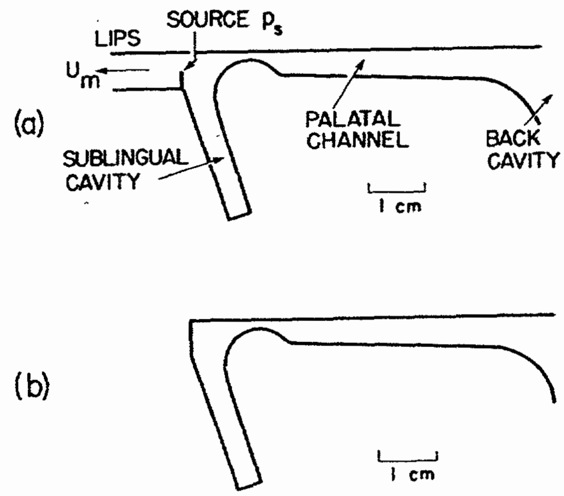


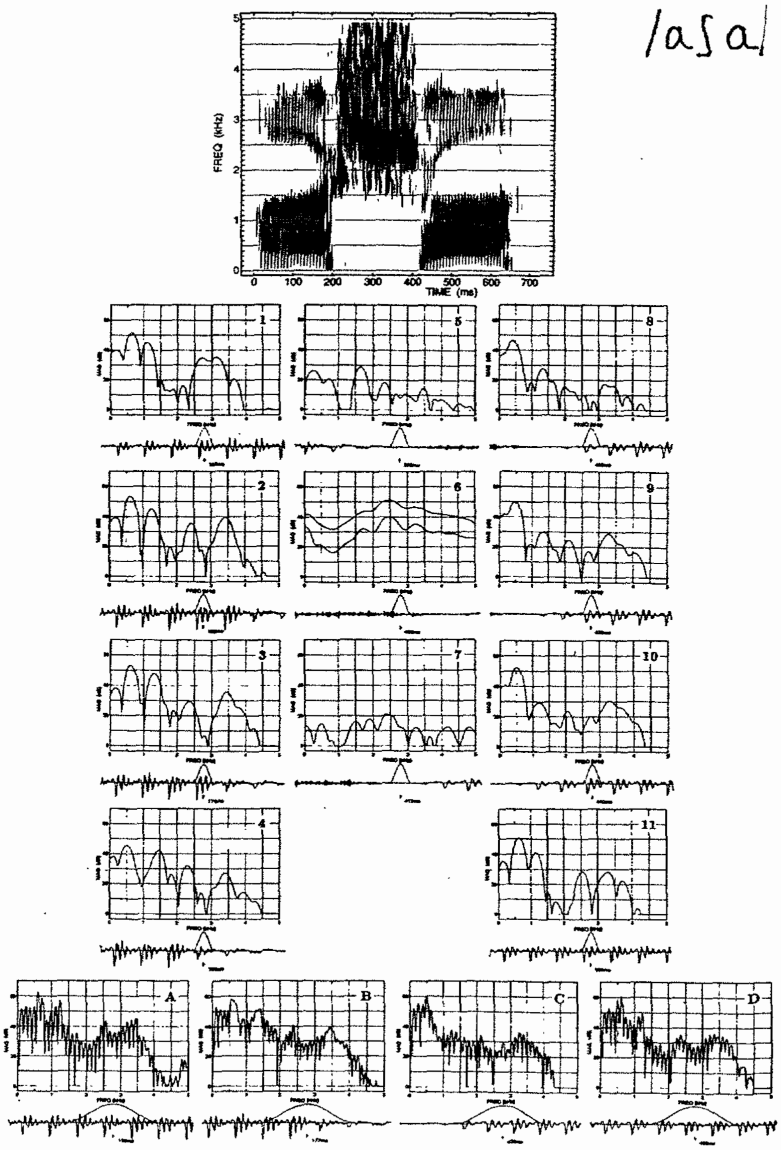
Figure 8.16 Same as figure 8.12, except that the utterance is /asɑ/. Panels 1 to 4 represent spectra (6.4-ms window) sampled at several times preceding the VC boundary, and panels 7 to 10 are sampled following the CV boundary. Panels 5 and 6 are average spectra calculated over the first and second halves of the frication region, respectively. The spectra in panels A, B, C, and D were obtained with a longer time window. See text.



Model for /ʃ/

Figure 8.19 (a) Schematized representation of the detailed cavity shapes in the anterior portion of the vocal tract for /š/. The source of sound is assumed to be located at the lower incisors, as indicated. (From Halle and Stevens, 1991b.) (b) The zeros in the transfer function from source to mouth output in (a) are the natural frequencies of this configuration.

/aʃa/



Voiced fricative

a zip

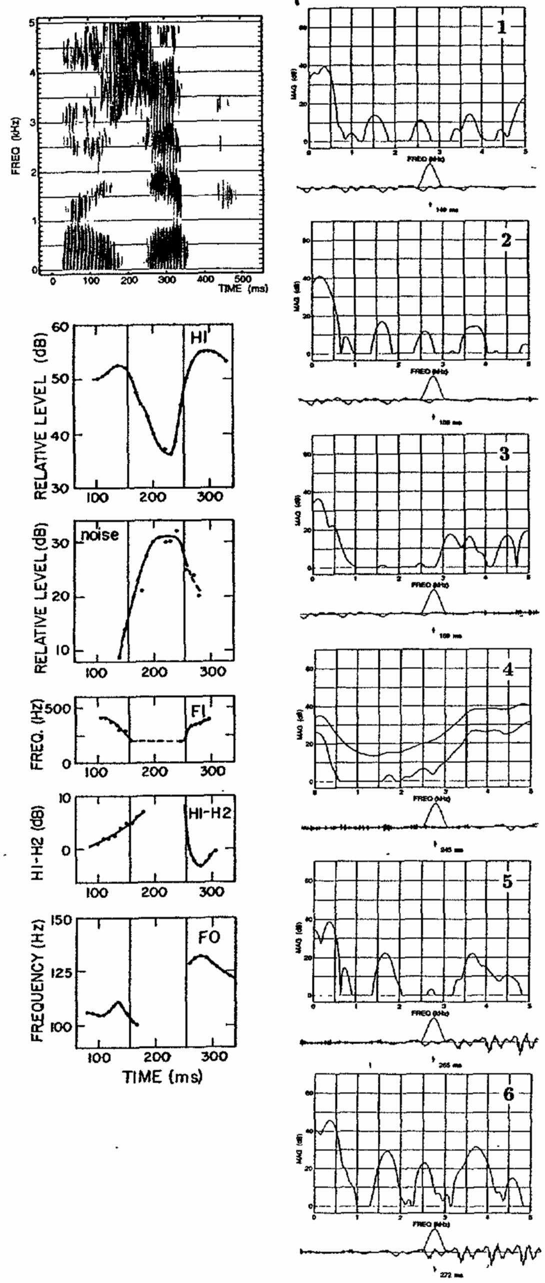


Figure 8.76 Samples of acoustic data for the intervocalic voiced fricative consonant /z/ in the utterance *a zip*. The spectrogram of the utterance is shown at the top left, and below this are: *H1*, the amplitude of the first harmonic, measured as in figure 8.71; "noise," the amplitude of the frication noise, taken to be the peak spectrum amplitude (6.4-ms window) in the frequency range of 3.5 to 5 kHz; *F1*, the first formant frequency, measured as in figure 8.71; the difference $H1 - H2$ between the amplitudes (in decibels) of the first and second harmonics; and *F0*, the fundamental frequency. The vertical lines are estimates of the points where intraoral pressure begins to build up and where the pressure begins to decrease, based on examination of the acoustic record. Spectra at selected points near the left and right boundaries of the fricative are displayed at the right, together with waveforms and time windows (6.4-ms duration). The noise spectrum (panel 4) was obtained by averaging spectra over a 55-ms time interval. All other spectra were calculated with a single time window.

Affricates

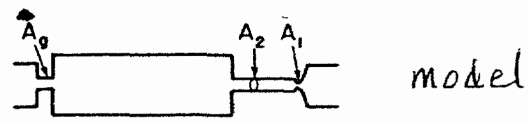


Figure 8.25 Schematized model of the vocal tract for an affricate consonant. Closure is formed by setting the area A_1 to zero at the front of the constriction. The channel behind the constriction, with area A_2, forms the fricative portion of the affricate after A_1 is released.

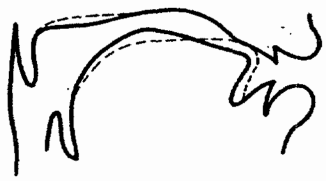
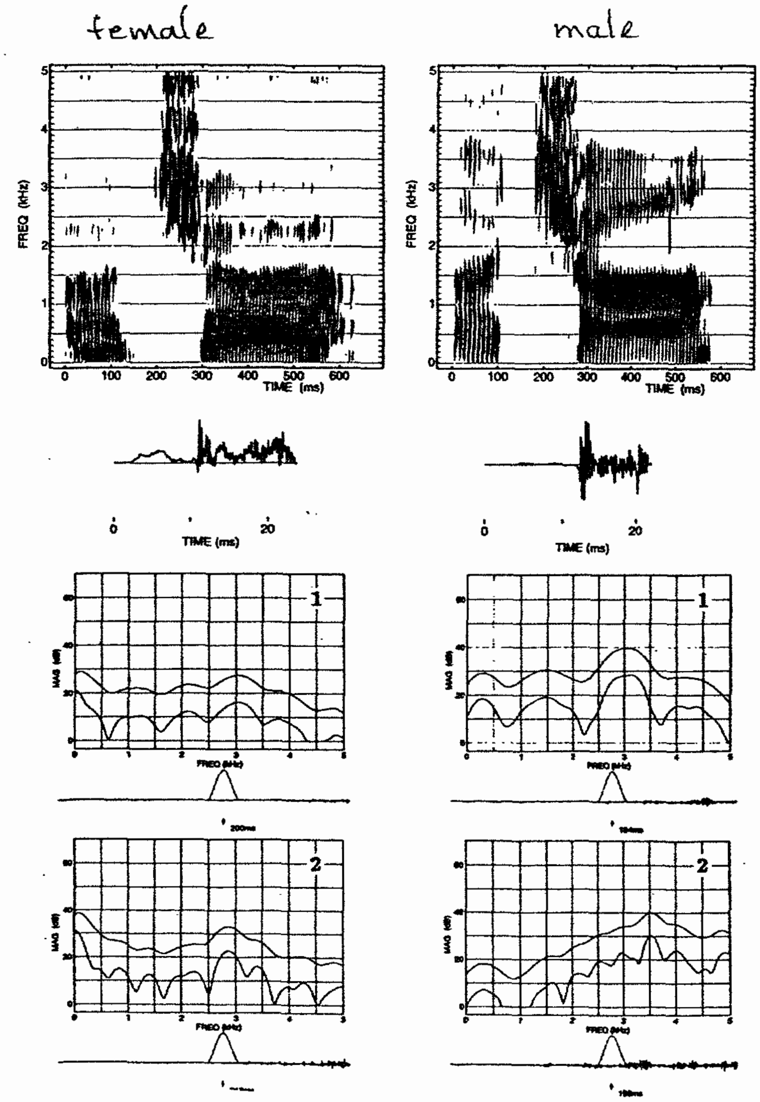


Figure 8.26 Conception of the midsagittal section of the anterior part of the vocal tract just prior to the initial release (dashed line) and 20 to 30 ms following the release (solid line) of a palatoalveolar affricate.



/h/ and aspiration
noise
and breathy voicing

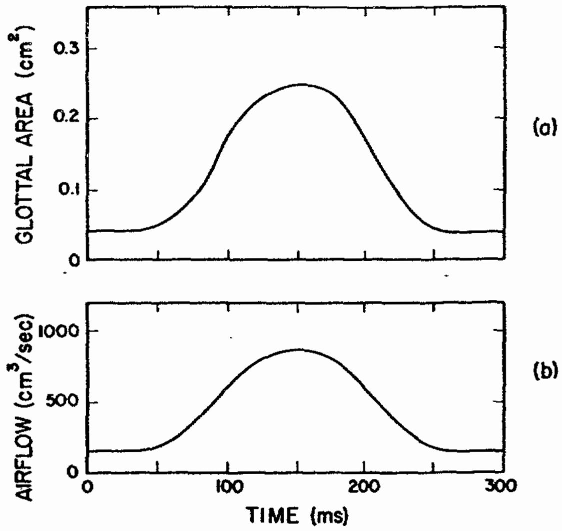


Figure 8.34 (a) Estimate of time course of glottal cross-sectional area during the production of /h/ in intervocalic position. (b) Calculated airflow corresponding to area variation in (a). A lung pressure of 8 cm H_2O is assumed. Resistance of subglottal airways is taken to be 1.5 acoustic ohms.

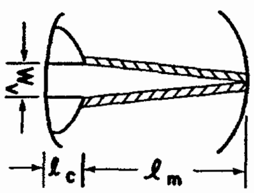


Figure 8.33 Schematization of glottal opening for an abducted glottal configuration. The glottal width at the level of the arytenoid cartilages is w_v; ℓ_m and ℓ_c are the lengths of the membranous and cartilaginous portions of the glottis.

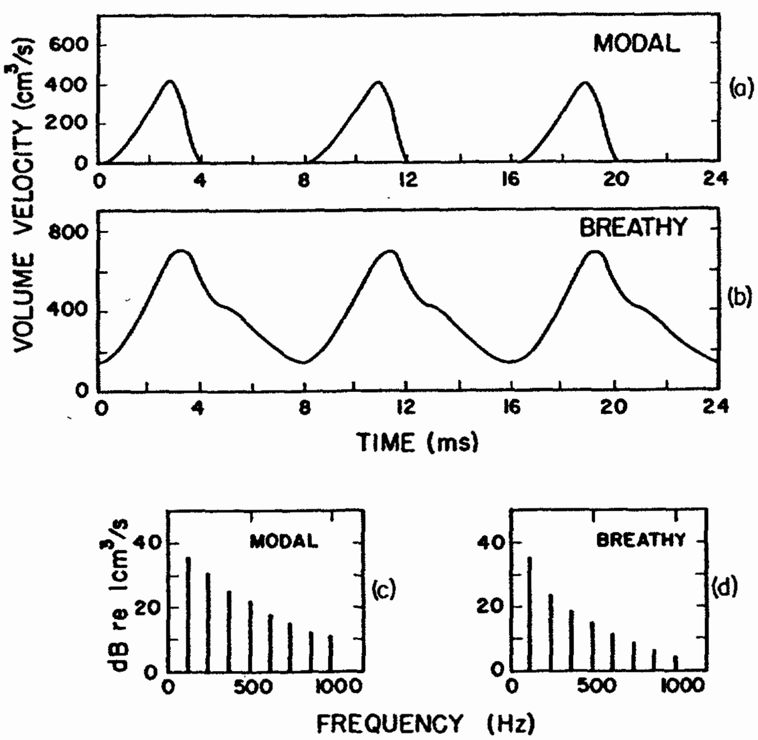


Figure 8.35 (a) and (b) Schematic representation of the waveforms and (c) and (d) spectra for modal voicing and for breathy voicing. For breathy voicing, the amplitude of the fundamental component is enhanced relative to the amplitudes of higher-frequency components. A fundamental frequency of 125 Hz is assumed. The amplitudes of the waveforms and of the spectral components are selected to be in the range observed for male voices.

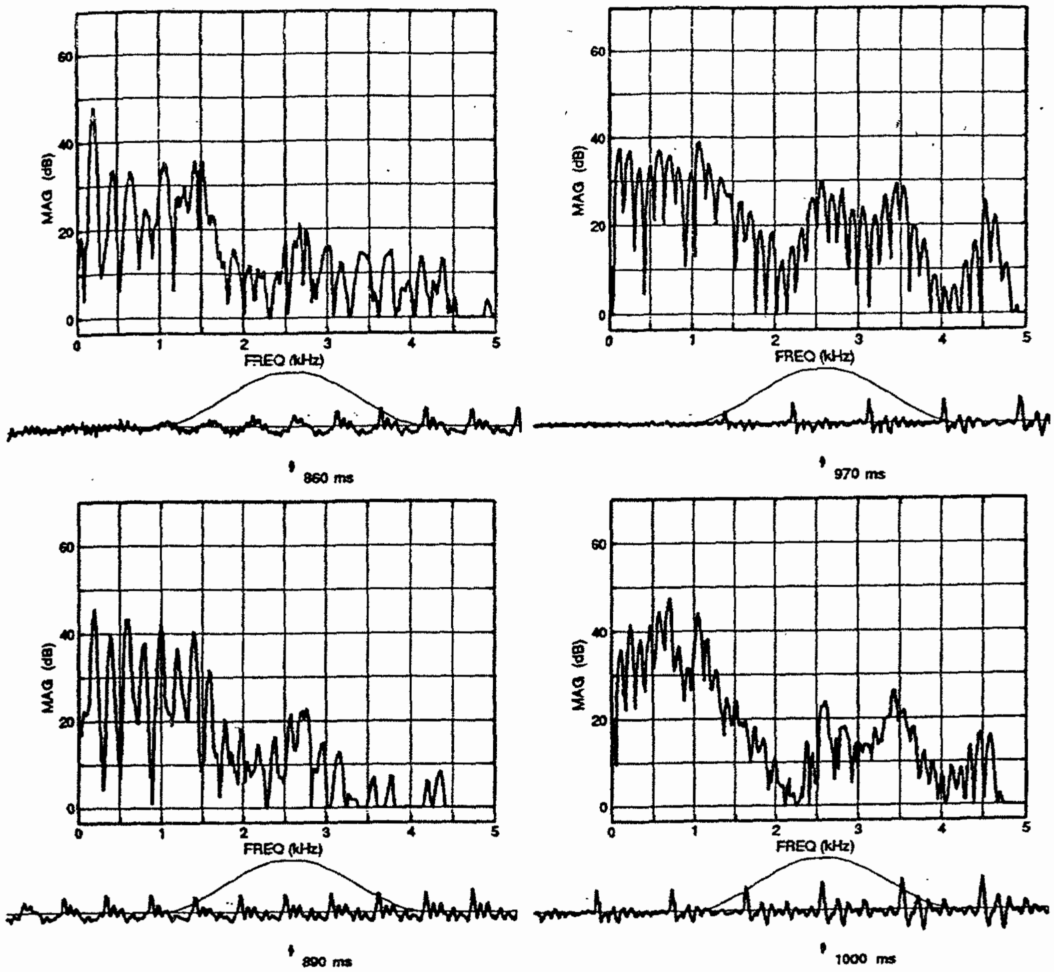
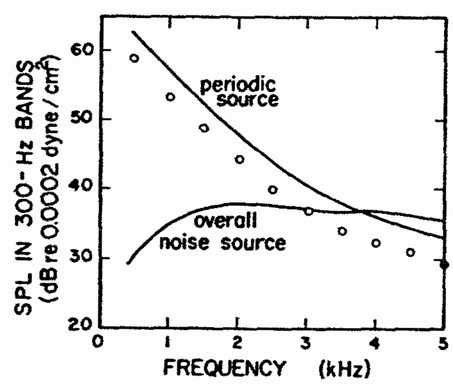


Figure 8.36 Spectra sampled near the onset of the vowel (upper panels) and about 30 ms later (lower panels) in the syllable /hɑ/ produced by two speakers: a female (left) and a male (right). Waveforms are shown below each spectrum, together with the time window over which the discrete Fourier transform is calculated.



comparing periodic and noise source at glottis

Figure 8.41 Calculated spectra of radiated sound (before adding the effect of an all-pole transfer function) due to sources of turbulence noise in the laryngeal region during /h/ and to the periodic glottal source at a time when the glottis is in a configuration for modal voicing in an adjacent vowel. For the periodic source, the solid line gives the levels in 300-Hz bands, whereas the open circles give the levels of every fourth harmonic, assuming a fundamental frequency of 125 Hz. The spectrum of the noise is in 300-Hz bands. The distance from the mouth is 20 cm. To obtain the spectra for a vowel, the all-pole component of the vocal tract transfer function must be added to these spectra. See text.

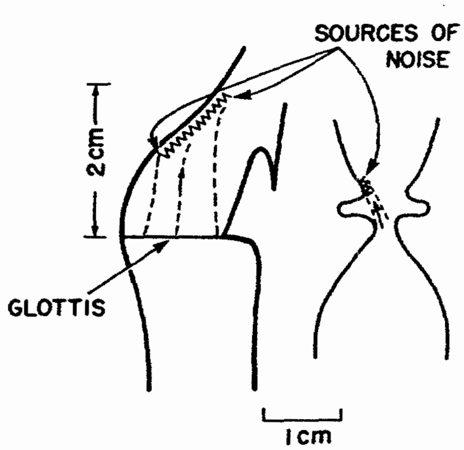


Figure 8.38 Schematized views of the laryngeal region in the sagittal plane (left) and in coronal section (right), indicating how airflow through the glottis might impinge on the surface of the epiglottis (left) or on the ventricular folds (right) to produce turbulence noise that is represented as a source of sound pressure.

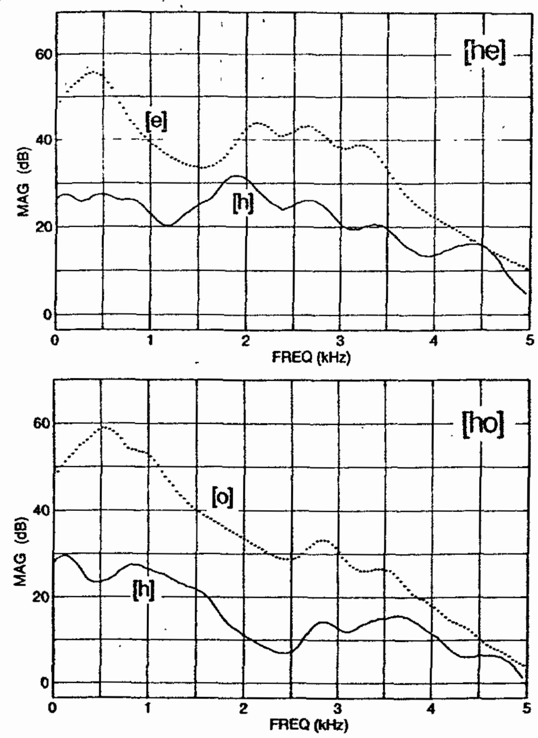


Figure 8.44 Examples of smoothed spectra sampled in /h/ (solid lines) and in the follow ing vowel (dotted lines) for the utterances /he/ (male speaker) and /ho/ (female speaker) a indicated. Spectra were obtained by calculating a discrete Fourier transform (time window c 30 ms) and then smoothing this spectrum with a weighted frequency window of width abou 400 Hz.

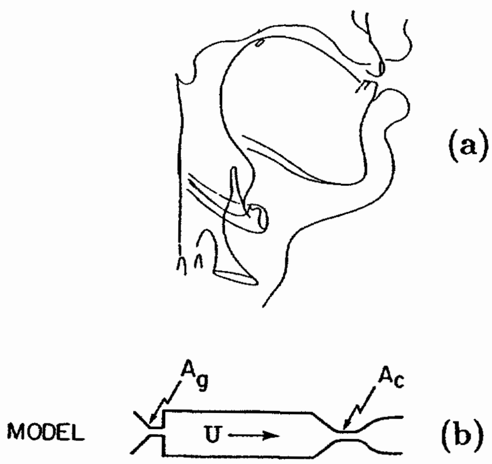


Figure 8.50 (a) Midsagittal section of vocal tract for vowel /i/, showing constriction in the palatal region. (From Perkell, 1969.) (b) Schematization of the airway with two constrictions—one at the glottis, with cross-sectional area A_g, and the other in the supraglottal airway, with cross-sectional area A_c.

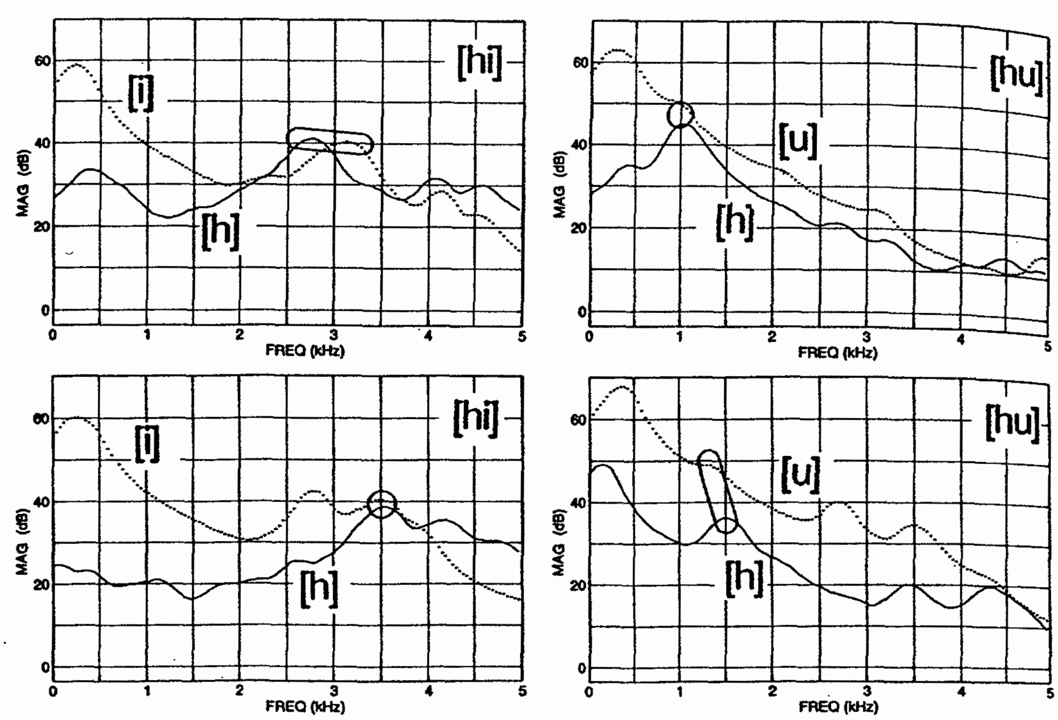


Figure 8.54 Some smoothed spectra sampled in the syllables /hi/ and /hu/. For each utterance, two spectra are shown: a spectrum sampled within /h/ (solid line) and a spectrum sampled in the following vowel (dotted line). The prominence of the *F2* or *F3* peak in the noise spectrum is indicated in each panel. The upper row represents a male speaker, and the lower row a female speaker.

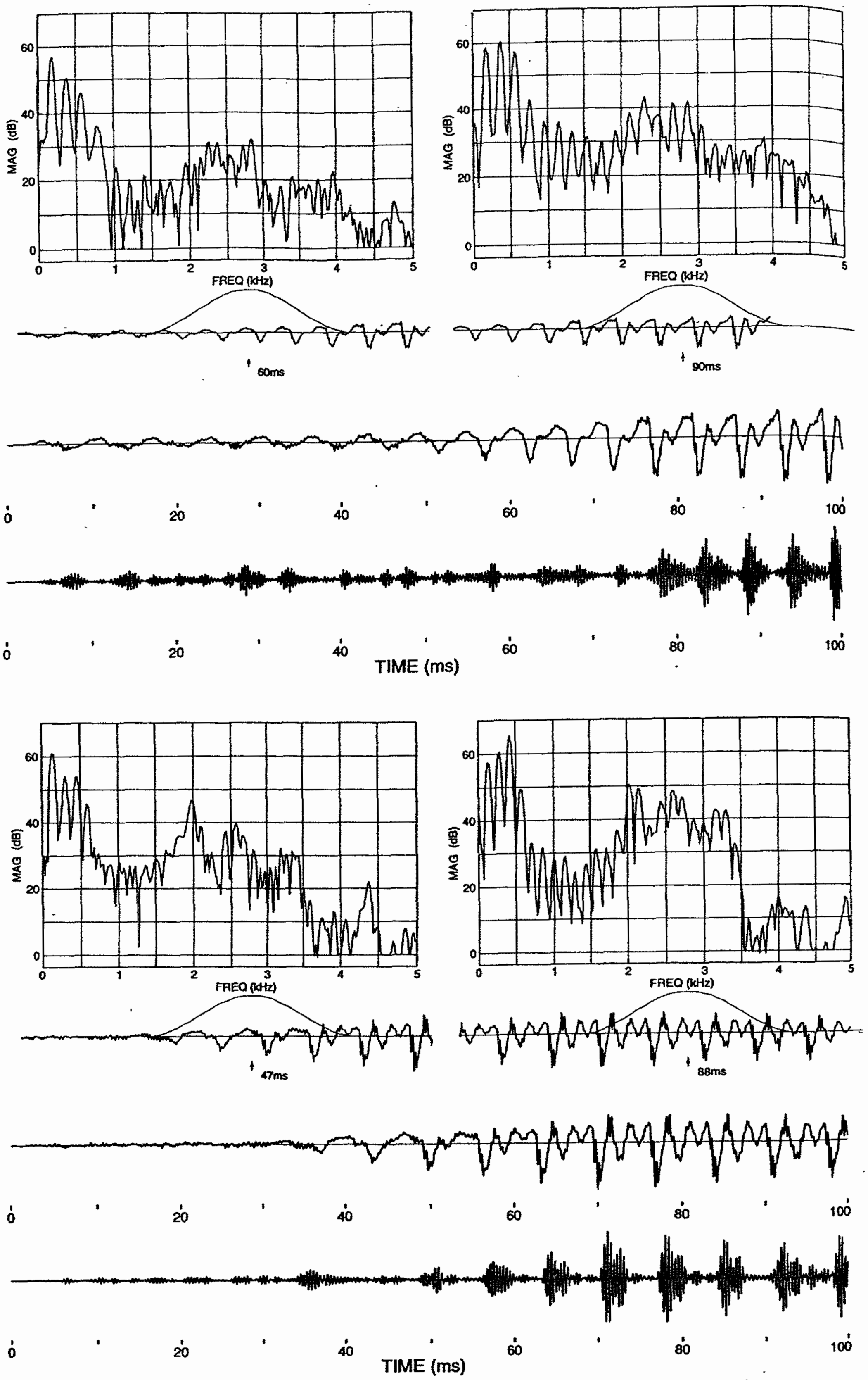


Figure 8.57 Displays of several aspects of the acoustic signal in the vicinity of voicing onset in the syllable /he/ produced by a female speaker (top) and a male speaker (bottom). The upper panel of each section gives the spectrum (30-ms window) sampled near the onset of glottal vibration (left) and 30 to 40 ms later (right), as indicated by the waveforms immediately below the spectra. Below these waveforms is the complete waveform for about 100 ms near voicing onset. The bottom waveform in each section was obtained by passing the signal through a bandpass filter (bandwidth = 600 Hz) centered on the third formant. This waveform is less regular near voicing onset than it is later, indicating the dominance of the noise component in the glottal source.