6.551J/HST714J ACOUSTICS OF SPEECH AND HEARING

Kenneth N. Stevens

Lecture 22 *11/30/04*

Glides

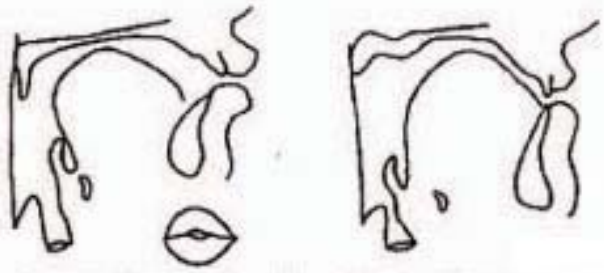


Figure 9.19 Midsagittal sections when the vocal tract reaches its most constricted configuration for the glides /w/ (left) and /j/ (right). In the case of /w/, the frontal lip contour is also shown, to illustrate the lip rounding. (Adapted from Bothorel et al., 1986.)

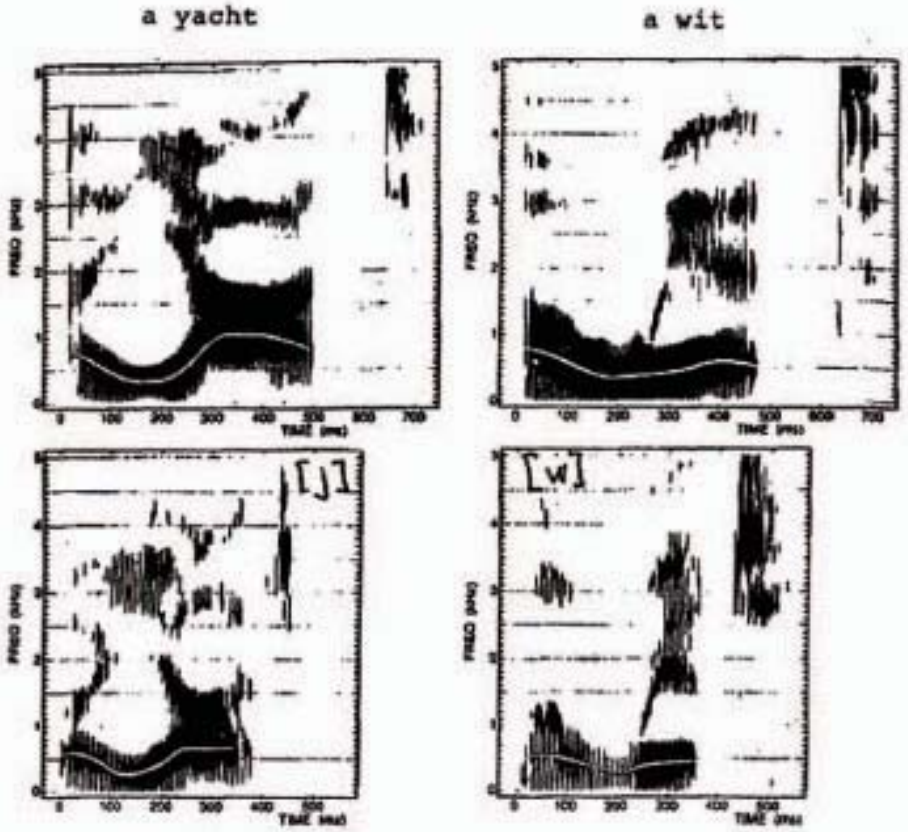


Figure 9.24 Spectrograms of the utterances /ə'jɑt/ and /ə'wɪt/ produced by a female speaker (top) and a male speaker (bottom). Measured F1 trajectories are superimposed on the spectrograms.

preceding vowel

following vowel

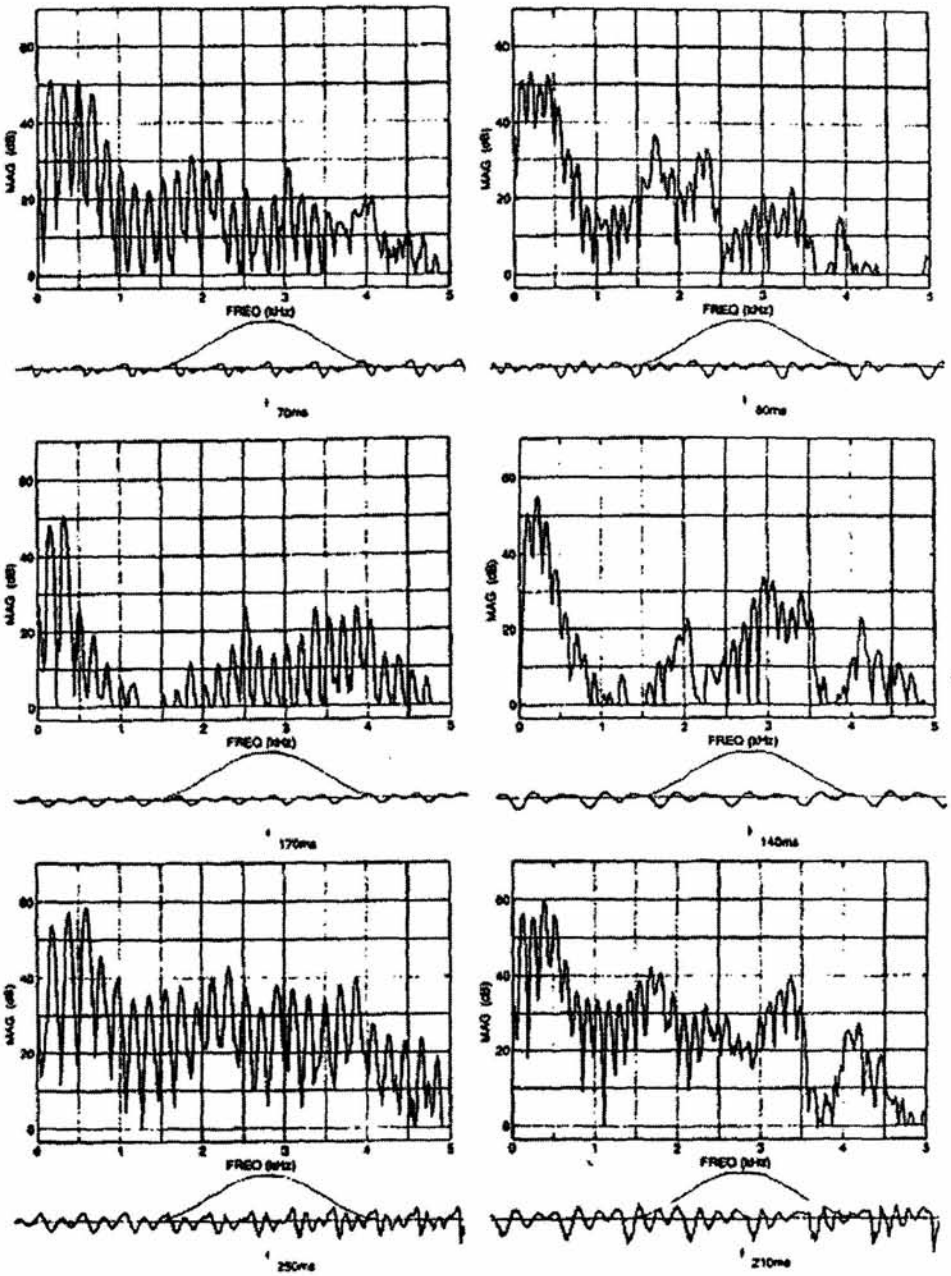


Figure 9.33 Same as figure 9.27, except that the utterance is /s'jot/. The female speaker is represented in the left panels and the male speaker in the right panels.

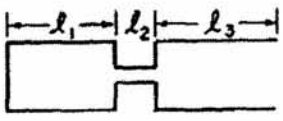


Figure 3.15 Uniform tube with a narrow constriction used to approximate area functions for consonants.

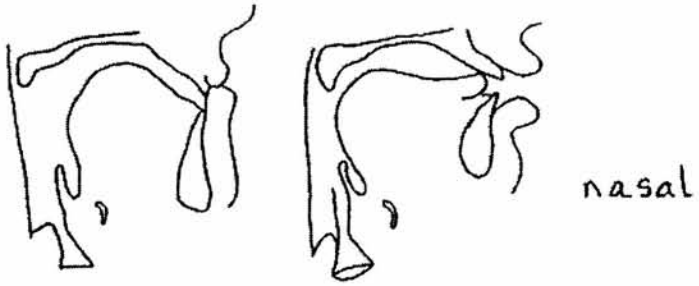
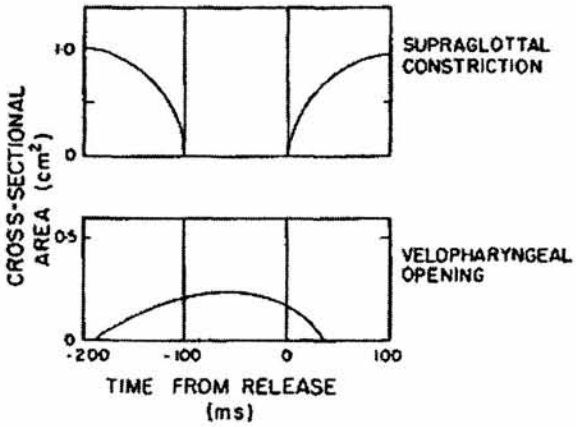


Figure 9.1 Midsagittal sections through the vocal tract during the production of nasal consonants produced by closing the lips (left) and raising the tongue tip to the alveolar ridge (right). For the labial consonant, the context is "Mets tes beaux habits," and for the alveolar consonant the context is "Une réponse ...." (Adapted from Bothorel et al., 1986.)



see next page

Figure 9.2 Schematization of the time course of the cross-sectional area of the oral constriction (upper panel) and of the velopharyngeal opening (lower panel) during production of a nasal consonant in intervocalic position. The detailed shapes of these trajectories depend on the place of articulation for the consonant and for the adjacent sounds.

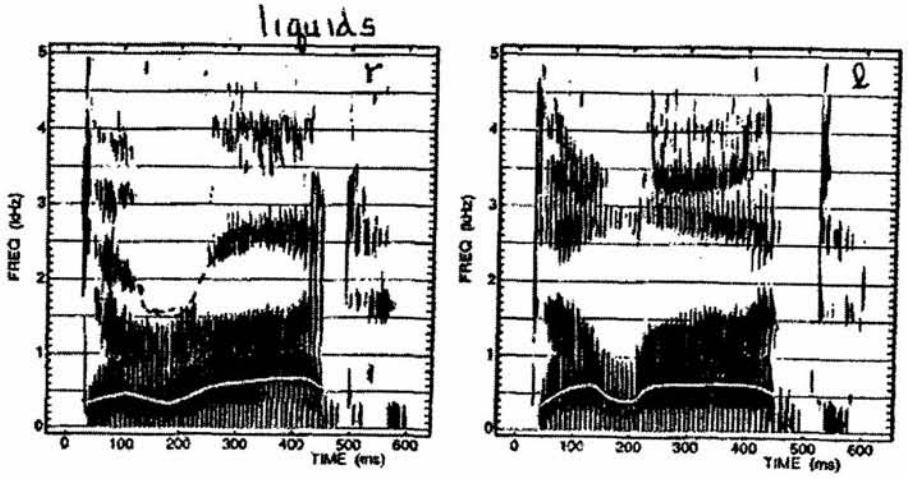


Figure 9.35 Spectrograms of the utterances /da'rod/ (left) and /da'lod/ (right) produced by a male speaker. The time course of the first-formant frequency during the liquid consonant is overlaid on the spectrogram.

nasal consonant

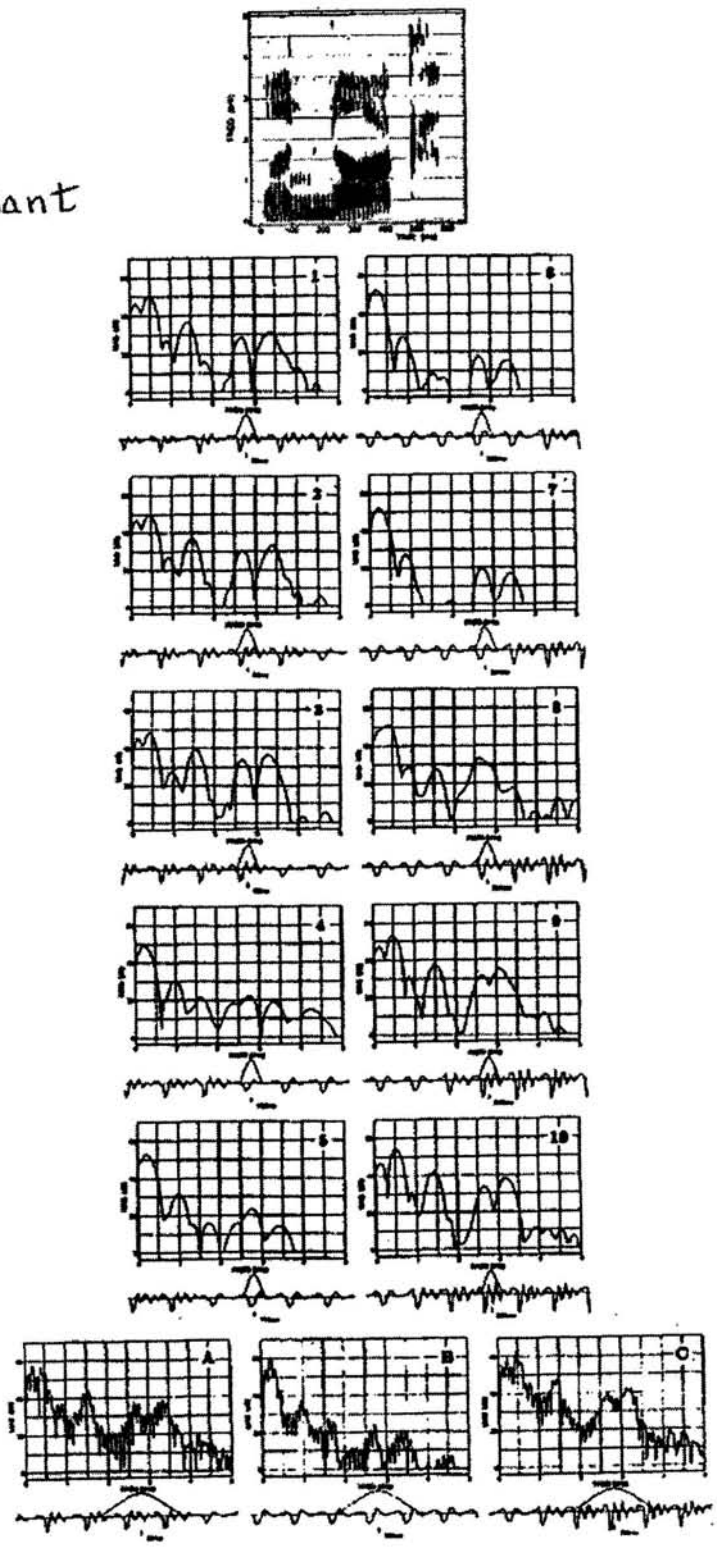


Figure 9.11 Acoustic data obtained from an alveolar nasal consonant /n/ in the utterance *a knock*. A spectrogram of the utterance is shown at the top, and the panels below the spectrogram are spectra sampled at several points and with different time windows. The sampling times and windows are as described in the legend for figure 9.7.

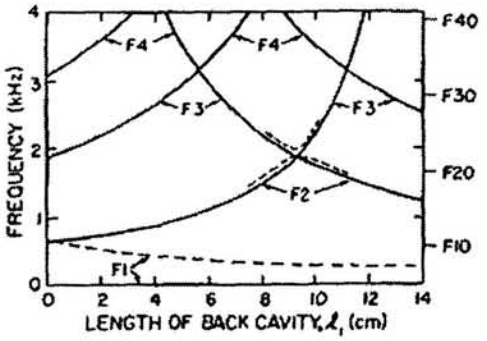


Figure 3.16 Relations between natural frequencies and the position of the constriction for the configuration shown in figure 3.15. The overall length of the tube is 16 cm and the length of the constriction is 2 cm. The lines sloping up to the right represent the resonances of the front cavity (anterior to the constriction); the lines sloping down to the right represent the resonances of the back cavity. The solid lines are the resonances if there is no coupling between front and back cavities. The dashed lines near the point of coincidence of two resonances at $\ell_1 = 9.3$ cm illustrate the shift in the resonant frequencies for the case where there is a small amount of coupling between front and back cavities. When the area of the constriction is very small, $F1 = 0$ as shown. When the constriction is larger, F1 increases as shown by the dashed line. The resonances of a 16-cm tube with no constriction are shown by the labeled marks at the right. The curves are labeled with the appropriate formant numbers.

From *Acoustic Phonetics* by Kenneth Stevens. (c) 1998, MIT Press. Used with permission.

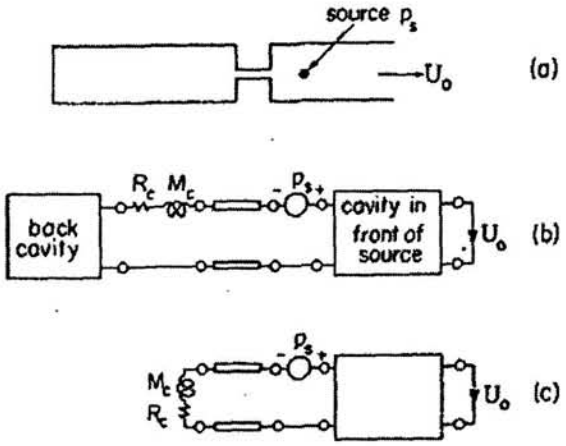


Figure 3.36 (a) Model of acoustic resonator with noise source in front of the constriction. (b) Equivalent circuit of configuration in (a). (c) Equivalent circuit neglecting the effect of the back cavity. From *Acoustic Phonetics* by Kenneth Stevens. (c) 1998, MIT Press. Used with permission.

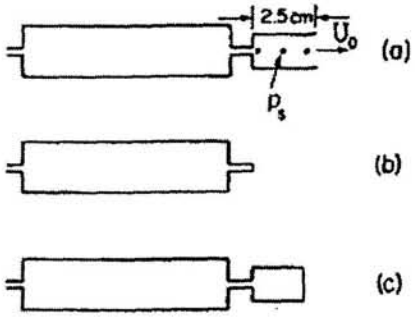


Figure 3.37 (a) Model for illustrating calculation of radiated sound pressure when a noise source is located in front of a constriction in the vocal tract. Three different source locations are indicated. (b) The natural frequencies of this configuration are the zeros of the transfer function U_r/p_s of the configuration of (a), with the leftmost source location. (c) Same as (b) but with the rightmost source location.

From *Acoustic Phonetics* by Kenneth Stevens. (c) 1998, MIT Press. Used with permission.

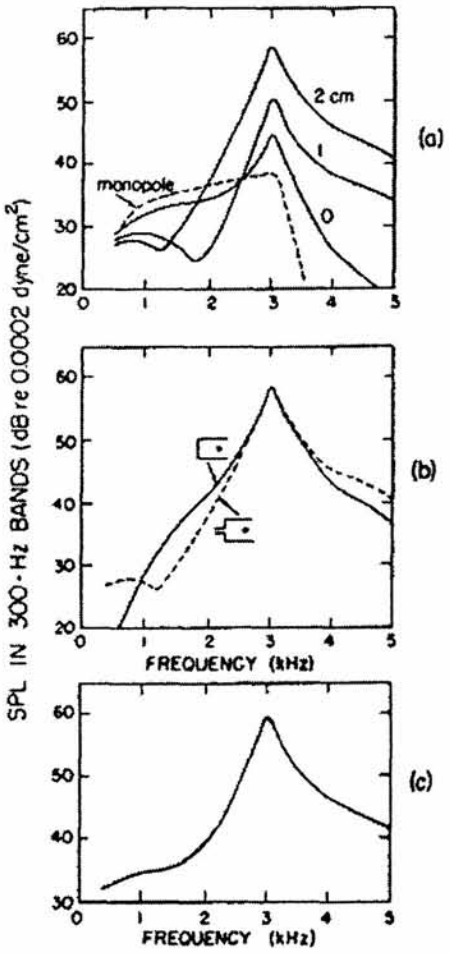


Figure 3.38 (a) Calculated spectrum of radiated sound pressure (at a distance of 20 cm) for the three source locations in the model in figure 3.37a. See text. (b) Comparison of calculated spectrum of radiated sound pressure for the source located 2 cm from the constriction when the constriction is modeled by an infinite impedance (solid line) and by an acoustic resistance and mass (dashed line). (c) Spectrum obtained by adding the contributions of the three source locations in (a), approximating the sound pressure for a distributed source.

stop consonants

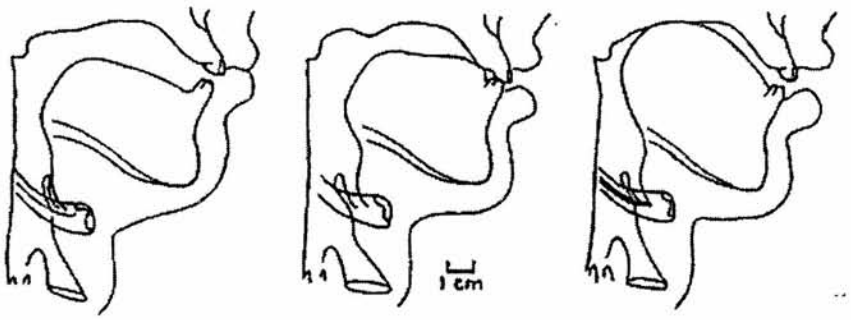


Figure 7.2 Midsagittal sections through the vocal tract during the production of stop consonants produced by closing the lips (left), raising the tongue tip to the alveolar ridge (middle), and raising the tongue body (right). Adult male speaker. (From Perkell, 1969.)

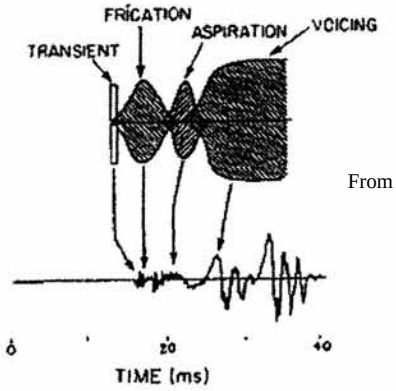
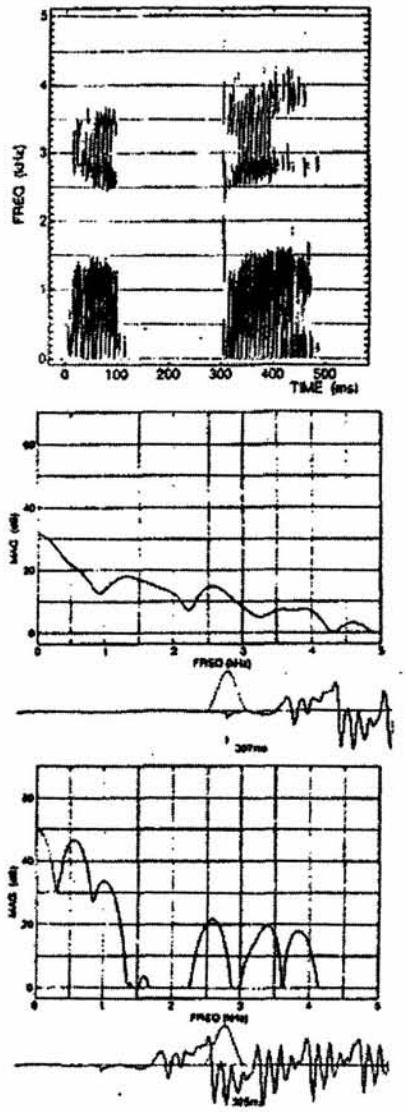


Figure 7.17 Schematic representation of sequence of events at the release of a voiceless unaspirated stop consonant. (From K. N. Stevens. 1993b.)



unaspirated [p]

Figure 7.20 Acoustic data obtained from an utterance of a voiceless unaspirated stop consonant /p/ in the intervocalic environment /opo/ (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.

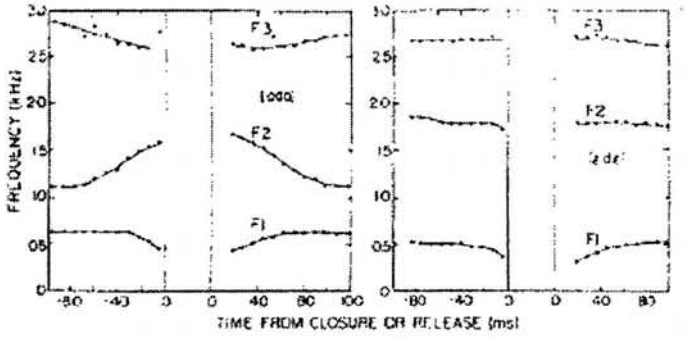
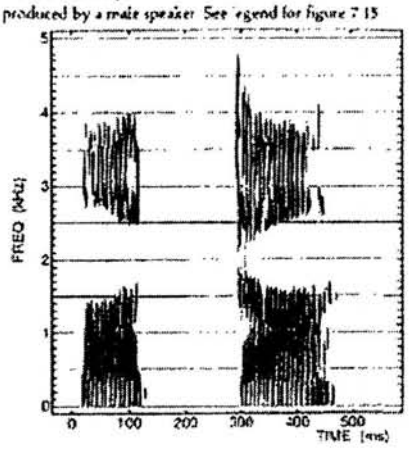


Figure 7.23 Measured frequencies of the first three formants for the utterances /ada/ (left) and /ede/ (right) produced by a male speaker. See legend for figure 7.15.



[t]

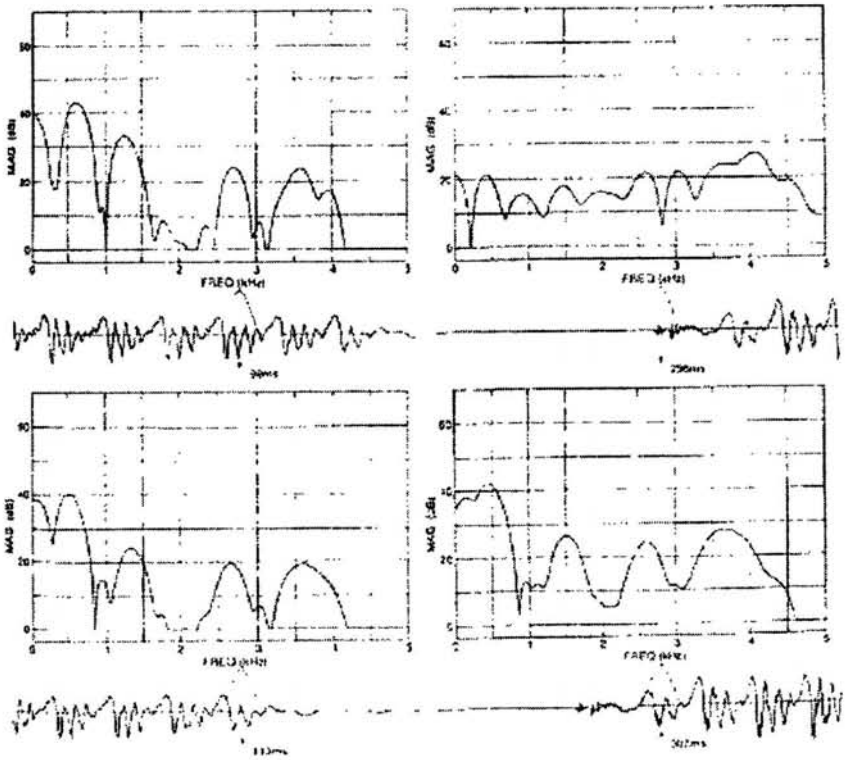
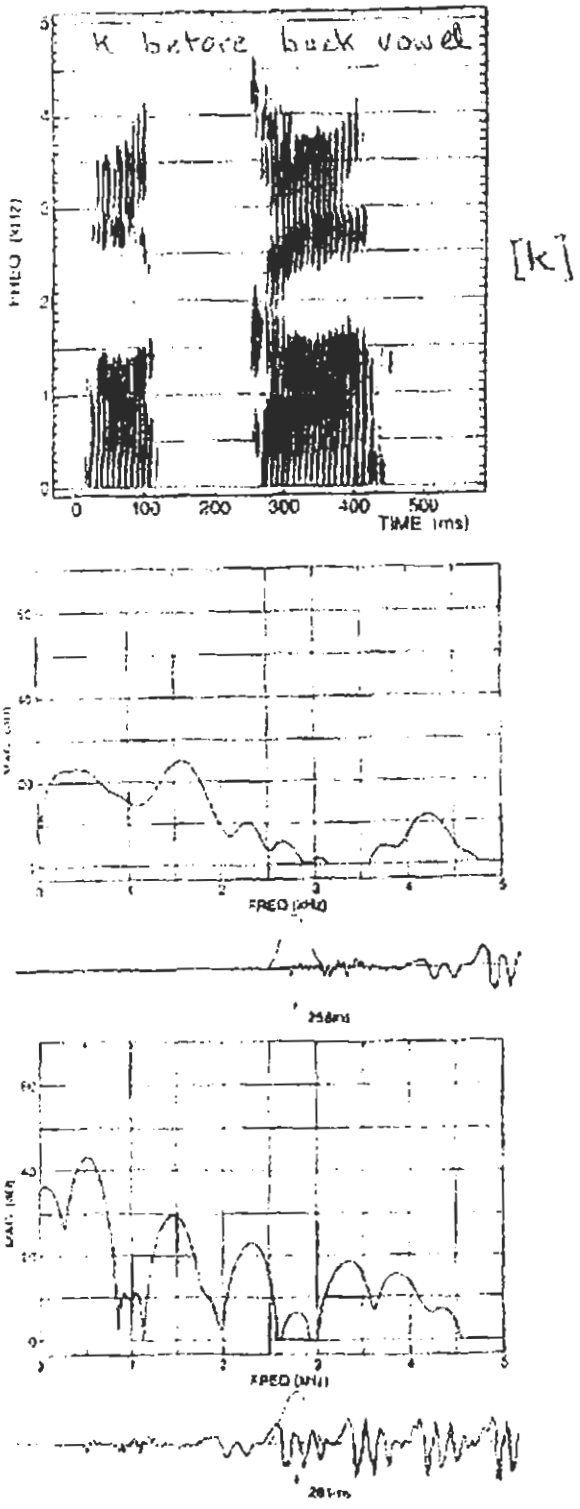
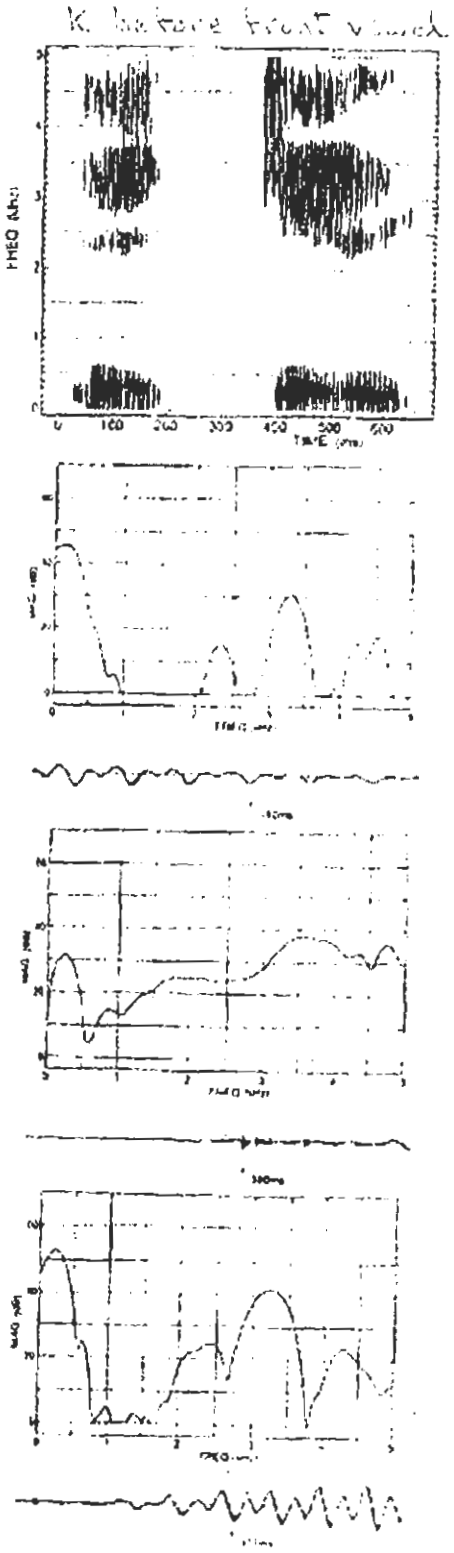


Figure 7.27 Acoustic data obtained from an utterance of a voiceless unaspirated stop consonant /t/ in the intervocalic environment /ata/. The four spectra below the spectrogram are sampled at selected points before the consonant closure and after the release, as indicated by the waveforms and time windows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


k before back vowel
FREQ (kHz)
0
100
200
300
400
500
TIME (ms)
[k]


k before front vowel
FREQ (kHz)
TIME (ms)

(a)

voiced stop

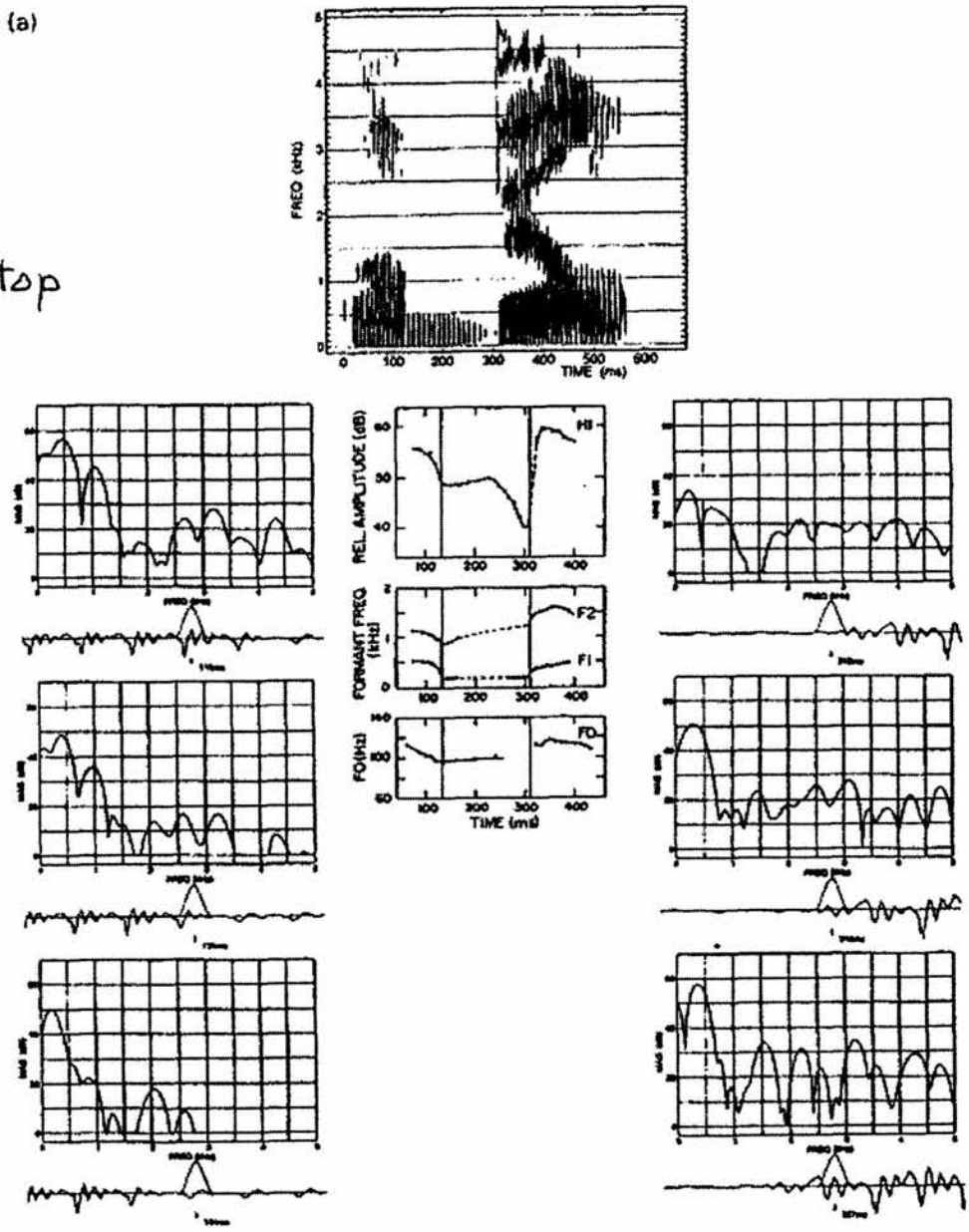
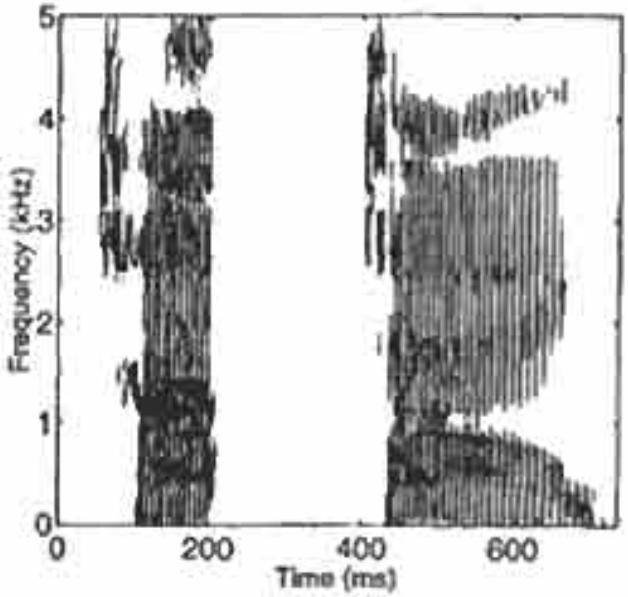


Figure 8.71a and b Samples of acoustic data for two voiced stop consonants in intervocalic position. The utterances are *a bill* (a) ~~and *a dill*~~ (b). The spectrogram of the utterance is displayed at the top. In the center below the spectrogram are plots of the time course of several parameters: the amplitude of the first harmonic, obtained from a discrete Fourier transform at 10-ms intervals; the frequencies of the first and second formants measured from the peaks in a 6.4-ms discrete Fourier transform centered on the initial part of each glottal period, with interpolations in the closure interval shown as dashed lines, and the fundamental frequency. The vertical lines indicate the instants of consonant closure and release, as estimated from the acoustic signal. The spectra at the left and right are sampled at times near the implosion and release respectively; they are discrete Fourier transforms with a 6.4-ms Hamming window. The waveform and the time window are shown below each spectrum.

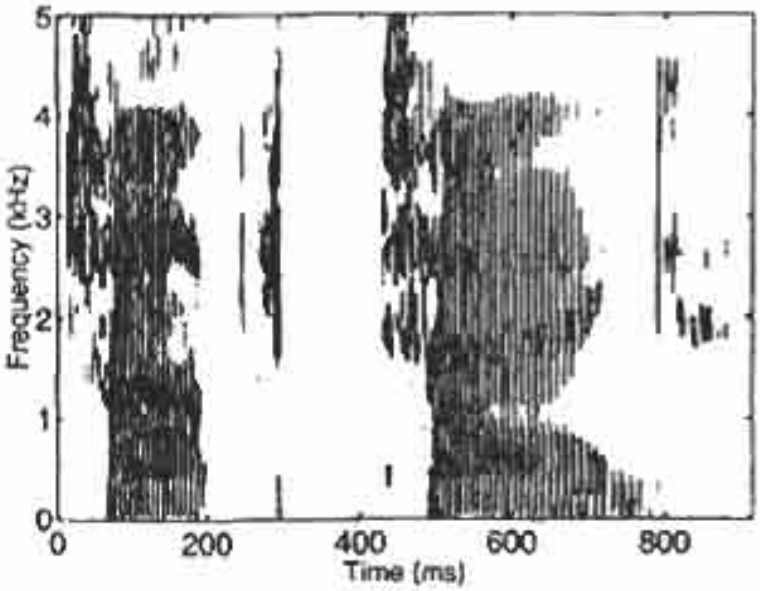
472



top tag

aspirated

stop consonants



top tag

with /p/ release

fricatives

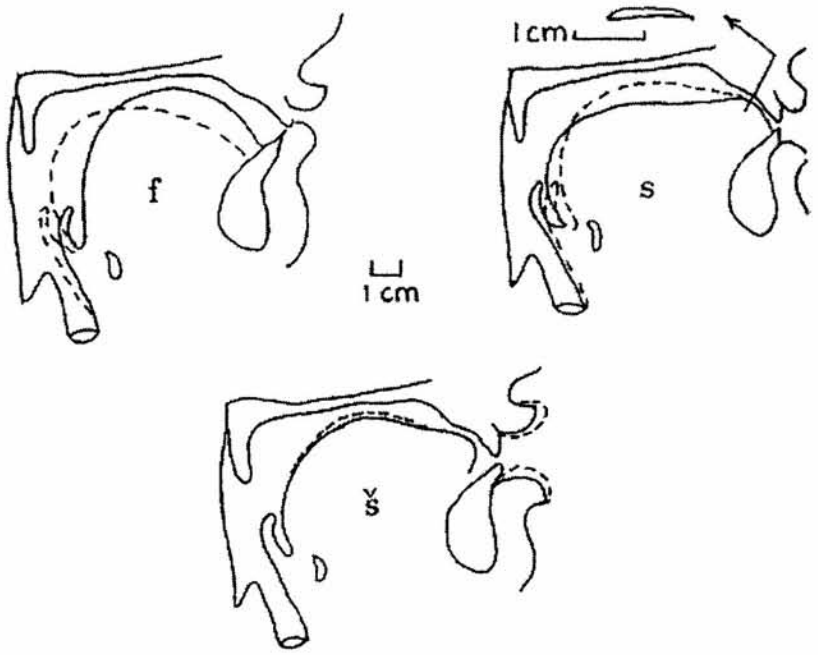
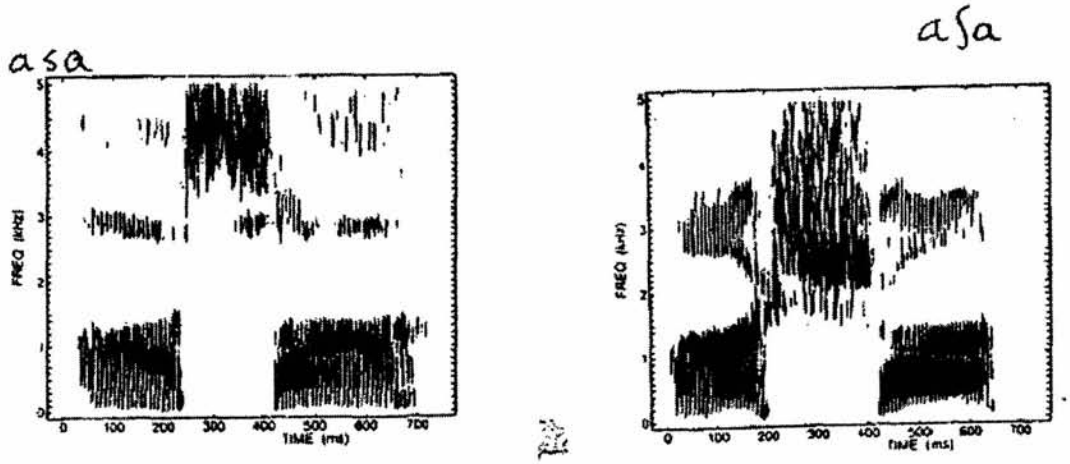


Figure 8.1 Midsagittal sections obtained from cineradiographs during the production of the three fricative consonants /f/, /s/, and /ʃ/, as indicated. The subject is an adult female speaker of French. The approximate scale is given between the /f/ and /s/ configurations. In the case of the /s/ configuration, 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 /ʃ/, 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 /ʃ/, /a/ (solid line) and /u/ (dashed line). (After Bothorel et al., 1986.)

From *Acoustic Phonetics* by Kenneth Stevens. (c) 1998, MIT Press. Used with permission.



voiced fricative

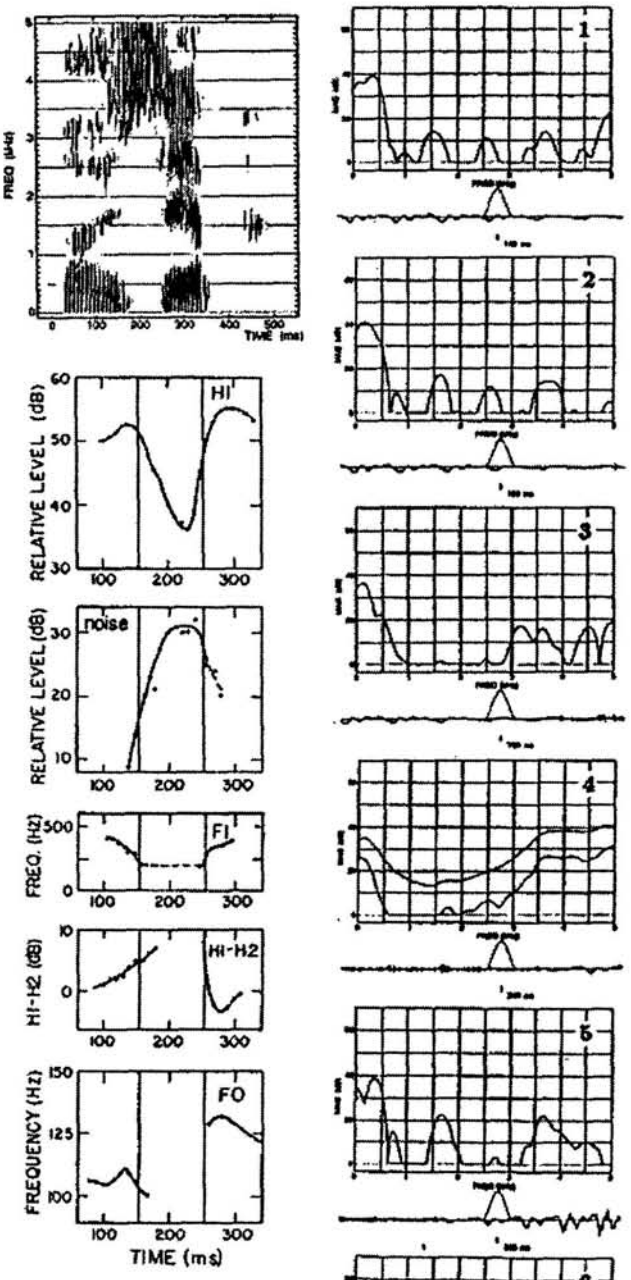


Figure 8.76 Samples of acoustic data for the intervocalic voiced fricative consonant /z/ in the utterance *a zip*. The spectrogram of the utterance is shown at the top left, and below this are: *H1*, the amplitude of the first harmonic, measured as in figure 8.71; "noise," the amplitude of the frication 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.