24.941J / 6.543J / 9.587J / HST.727J The Lexicon and Its Features
Spring 2007

THE LEXICON AND ITS FEATURES
6.976-9.912-24.921 FALL 2004

Professor Kenneth N. Stevens

Lecture 2 *09/16/04*

SOME ARTICULATOR-BOUND FEATURES

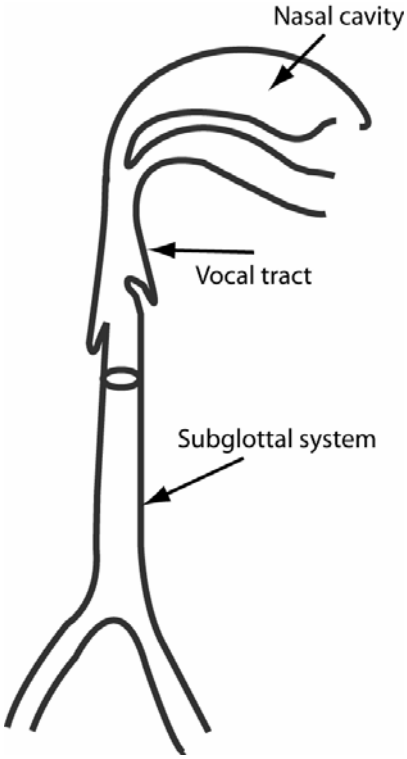
Features for vowels

back
high
low
round
tense
nasal

Dispersion theory (Liljencrants and Lindblom)

Quantal theory

Three parts to speech production system



Some midsagittal vocal tract shapes for vowels

Source: Perkell, J. S. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study. Research monograph No. 53*. Cambridge, MA: MIT Press, 1969.

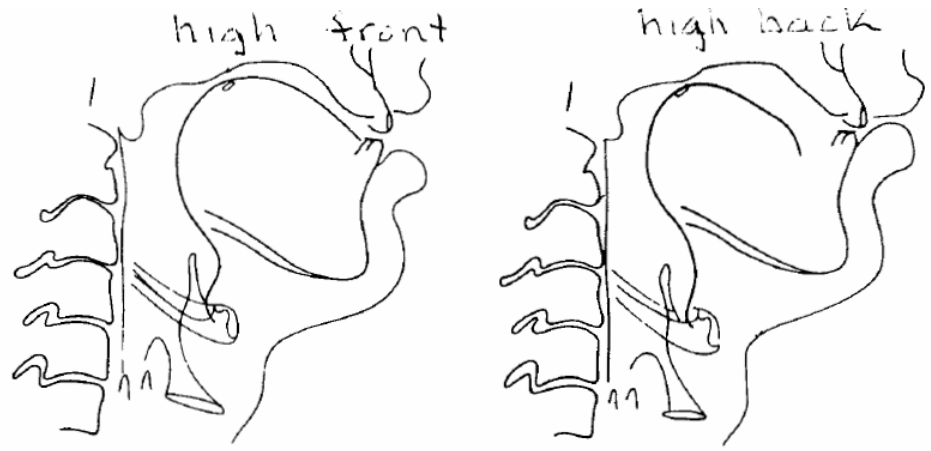
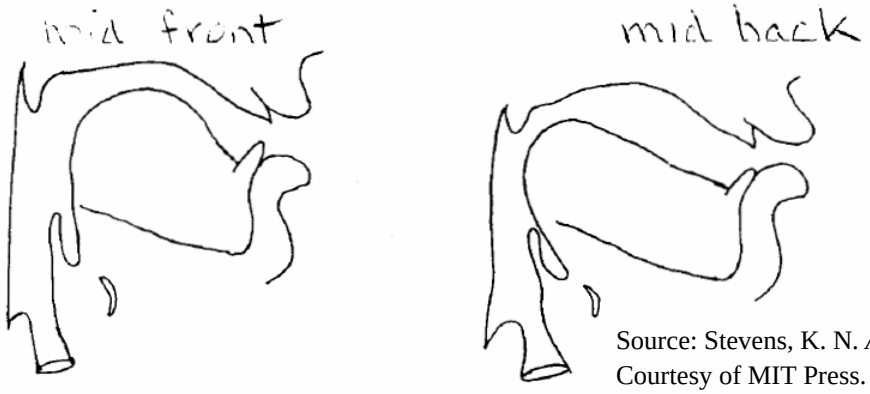


Figure 6.2 Midsagittal vocal tract configurations for the high vowels /i/ (left) and /u/ (right). Adult male speaker of English. (From Perkell, 1969.)



Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.

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.)

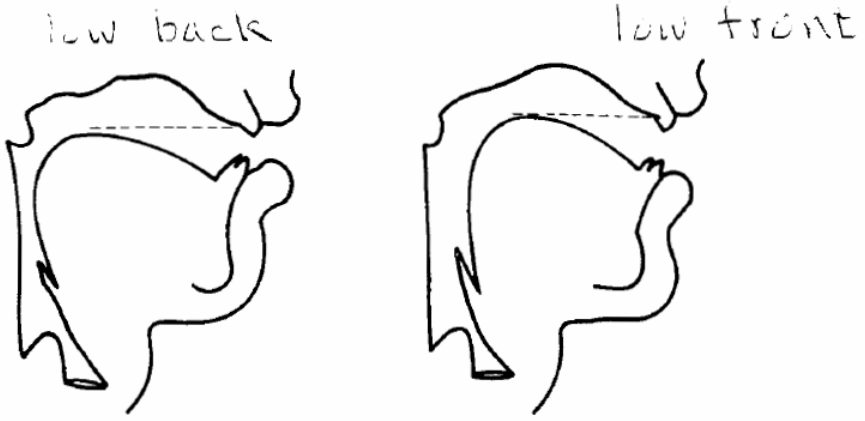


Figure 6.6 Midsagittal vocal tract configurations for the low vowels /ɑ/ (left) and /æ/ (right). The horizontal lines show the approximate location of the plane of the surfaces of the upper teeth. Adult male speaker of English. (From Perkell, 1969.)

Source: Perkell, J. S. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study. Research monograph No. 53*. Cambridge, MA: MIT Press, 1969.

Some vowel spectra illustrating variations in amplitudes of formant peaks as well as differences in frequencies

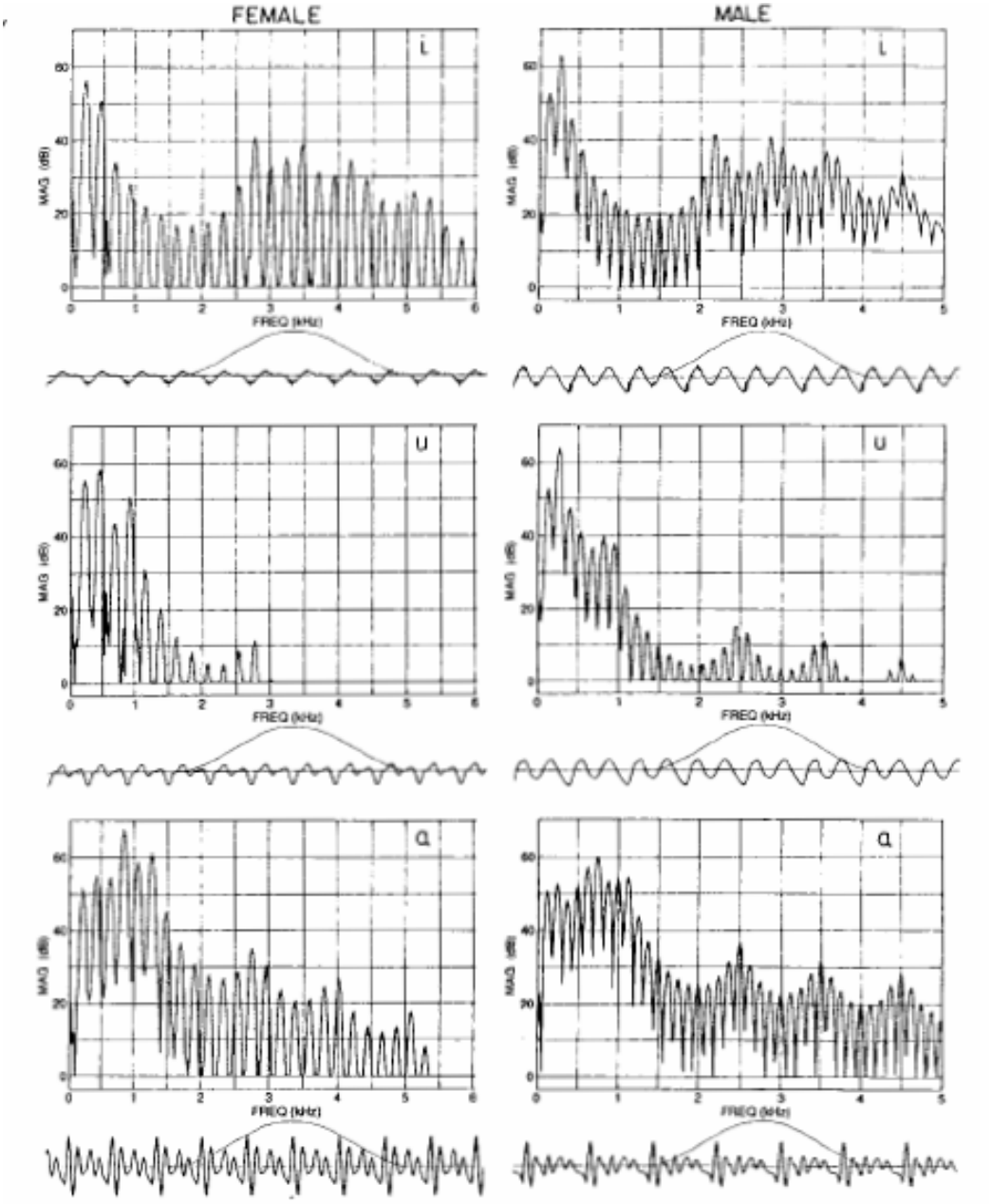
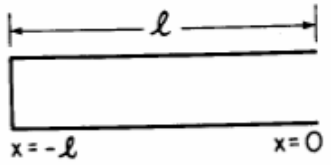


Figure 6.4 Spectra of the two high vowels /i/ and /u/ (top and middle panels) and the low vowel /a/ (bottom panel), with formant frequencies and fundamental frequencies appropriate for an adult female speaker (left column) and an adult male speaker (right column). The waveform is shown below each panel. The window used to calculate the spectrum is shown on each waveform. The vowels were synthesized using a Klatt synthesizer (Klatt and Klatt, 1990).

Some simple resonator shapes

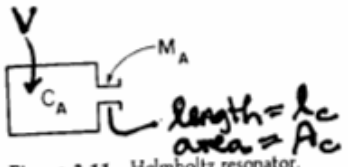


F1 = c/4ℓ
F2 = 3c/4ℓ
F3 = 5c/4ℓ
c = velocity of sound

Figure 3.8 Uniform tube of length ℓ, closed at one end and open at the other.

F1 = 0
F2 = c/2ℓ
F3 = c/ℓ

Figure 3.10 (a) Uniform tube open at both ends. (b) Uniform tube closed at both ends.



$$F1 = \frac{c}{2\pi}\sqrt{\frac{A_c}{V\ell_c}}$$

Figure 3.11 Helmholtz resonator.

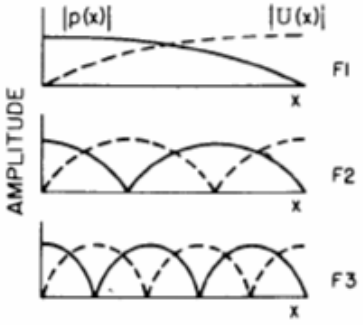


Figure 3.9 Distribution of sound pressure amplitude $|p(x)|$ and volume velocity amplitude $|U(x)|$ in a uniform tube for the first three natural frequencies $F1$, $F2$, and $F3$. Tube is closed at the left-hand end and open at the right-hand end.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Perturbation theory

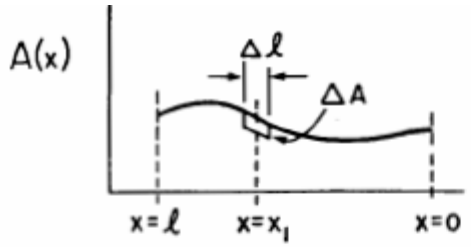
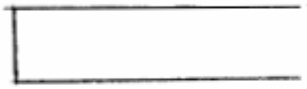


Figure 3.18 Illustrating a perturbation ΔA in the area of an acoustic tube at a short segment of length $\Delta\ell$ centered at point $x = x_1$.



Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

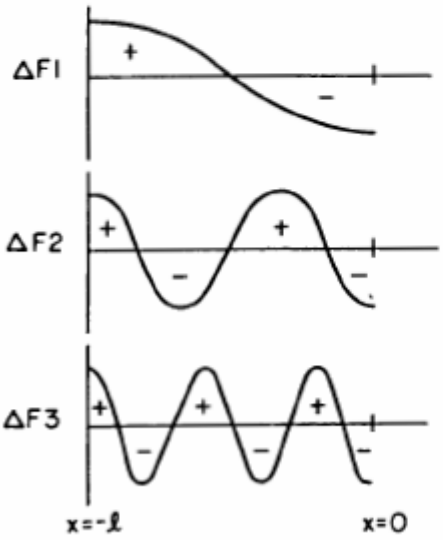


Figure 3.19 Curves showing the relative magnitude and direction of the shift ΔFn in formant frequency Fn for a uniform tube when the cross-sectional area is decreased at some point along the length of the tube. The abscissa represents the point at which the area perturbation is made. The minus sign represents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

Shifting position of a constriction

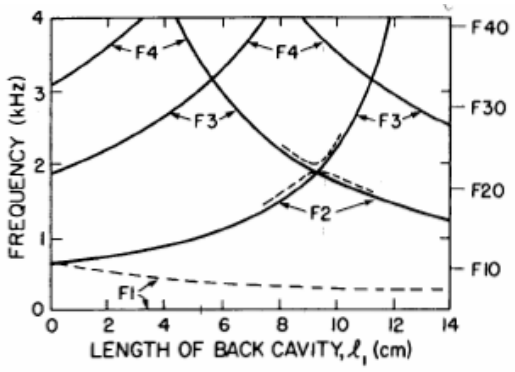


Figure 3.15 Uniform tube with a narrow constriction, used to approximate area functions for consonants.

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.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

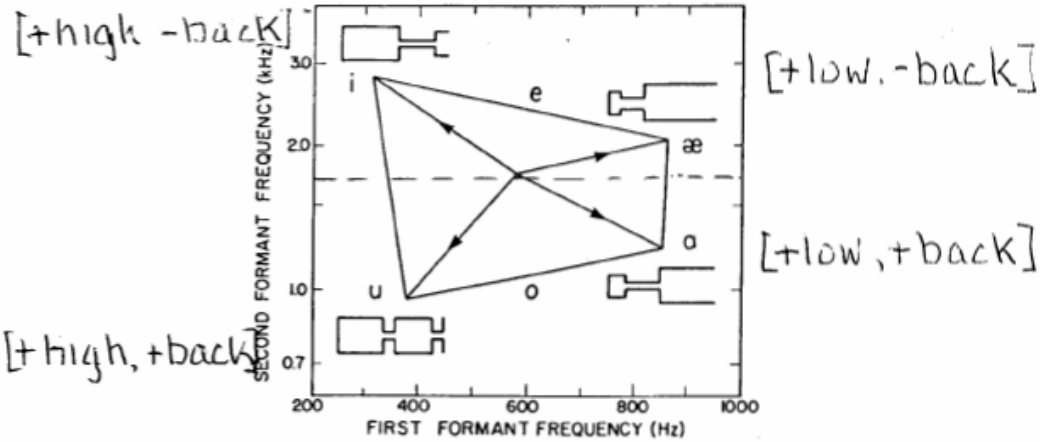


Figure 6.16 Plot of $F2$ vs. $F1$ showing how formants shift when the shape of an acoustic tube is perturbed in different ways. The midpoint represents equally spaced formants for a uniform tube of length 15.4 cm. The lines with arrows indicate how the formant frequencies change when the tube is modified 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.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

spectra

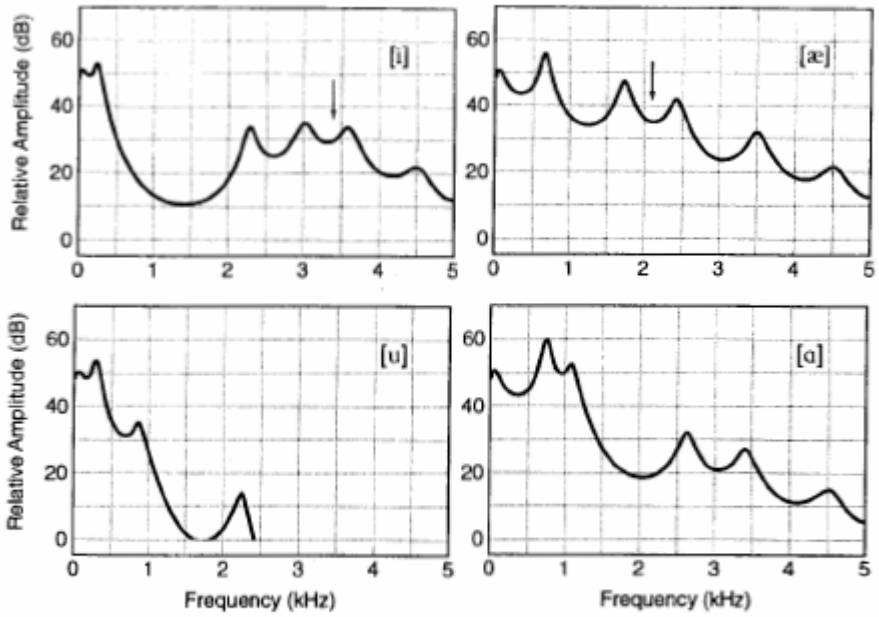


Figure 6.18 Calculated spectrum envelopes for the four vowels corresponding to the four corners of the quadilateral in figure 6.17, for an adult male speaker. The figure shows the relative amplitudes of the spectral prominences for the different vowels. The arrows in the two upper panels are estimates of where listeners would place an equivalent second-formant frequency *F2'* for a two-formant matching vowel. The values of *F2'* for front vowels are estimated from the data of Carlson et al. (1970), shown in figure 4.34.

male vs. female

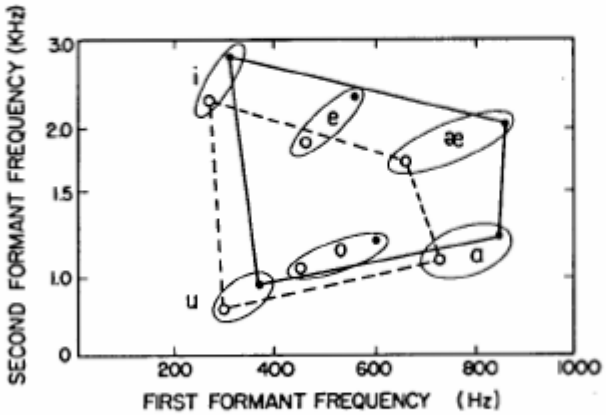


Figure 6.17 Plots of *F2* vs. *F1* for several vowels of American English. Open circles (joined by dashed lines) are data for adult male speakers and filled circles (solid lines) are for adult female speakers. The data for the vowels /i æ ɑ u/ are averages from Peterson and Barney (1952). Data for /e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

Effect of coupling to trachea

Source: Stevens, K. N. "Acoustic and Perceptual Evidence for Universal Phonological Features." Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, pp. 33-38, August 3-9, 2003.

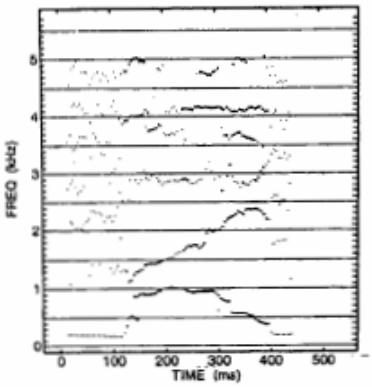
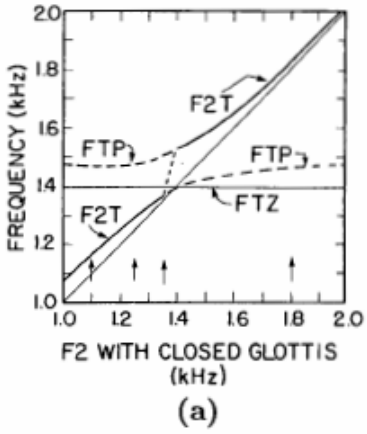
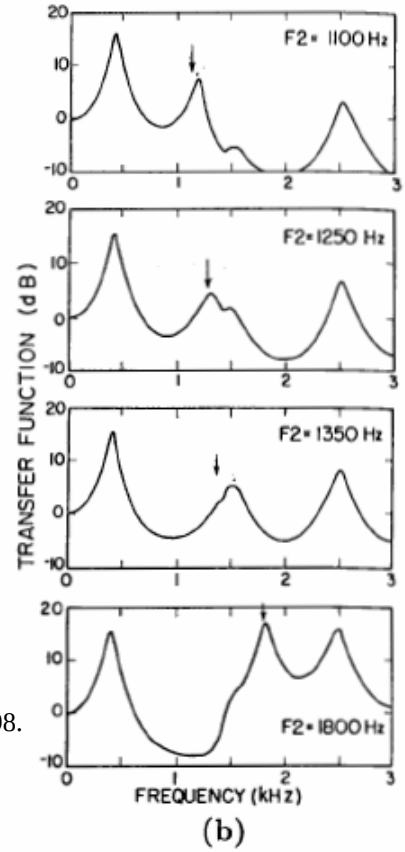


Figure 1: Trajectories produced by an LPC-based formant tracker for the word bide produced by a female speaker. There are discontinuous jumps as the formants pass through tracheal resonances for this speaker.





Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

Figure 6.27 (a) Calculations of the frequencies of the two poles and zero in the vicinity of *F2* when there is coupling to the trachea through a partially open glottis. The abscissa *F2* is the formant frequency that would exist if the glottis were closed and there were no acoustic coupling to the trachea. *F1* and *F3* are held constant at 400 and 2500 Hz. The frequency of the tracheal zero (*FTZ*) is assumed to be fixed at 1400 Hz. *F2T* represents the pole corresponding to *F2*, shifted by the influence of the tracheal system. *FTP* is the tracheal pole. The solid lines indicate the frequency of the most prominent spectral peak, which shows an abrupt jump in frequency (short-dashed) when *F2* is just below 1400 Hz. The dashed lines represent the poles that are less prominent in the spectrum. (b) Calculated transfer function for four different values of *F2*, as indicated by the arrows in (a) and by the values of *F2* given on each panel. The arrows on the transfer functions give the values of *F2*, and the deviation of the major spectral prominence from *F2* can be seen. The cross-sectional area of the glottal opening is assumed to be 0.05 cm^2.

Effect of subglottal resonance on vowel spectrum

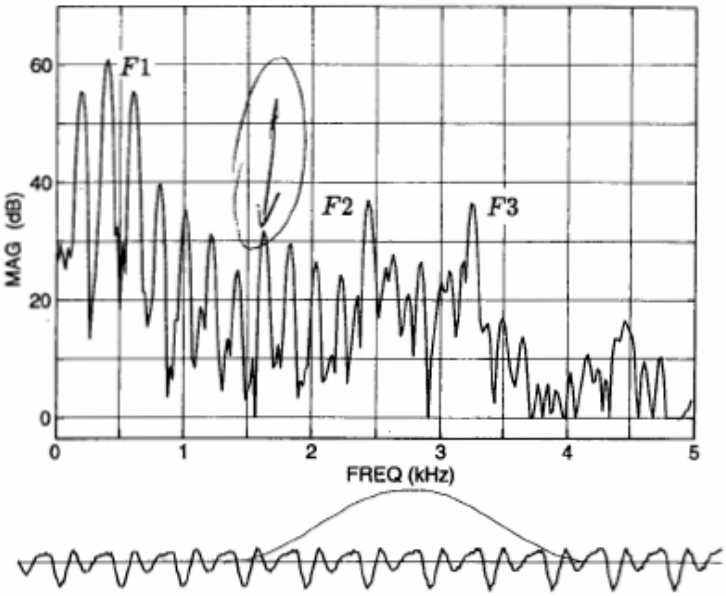
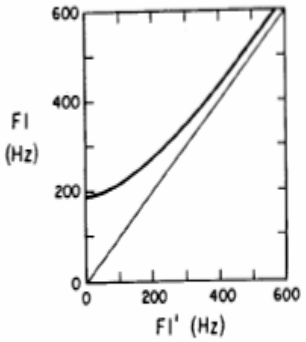


Figure 6.26 Spectrum of the vowel /ɪ/ sampled in the word *bit*, produced by an adult female speaker. The first three formant frequencies are labeled. The small prominence around 1600 Hz is assumed to be the effect of the second subglottal resonance. The waveform and the time window for calculating the spectrum are shown below the spectrum.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Effect of vocal tract walls on F1



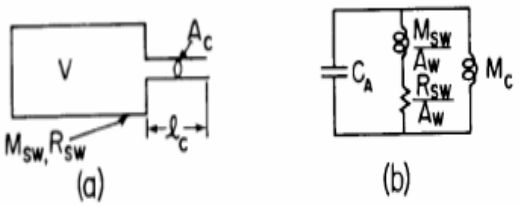
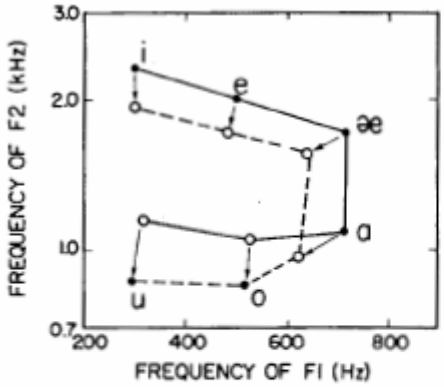


Figure 3.26 (a) Model for a constricted vocal tract configuration with yielding walls. (b) Low-frequency equivalent circuit for the model in (a).

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.

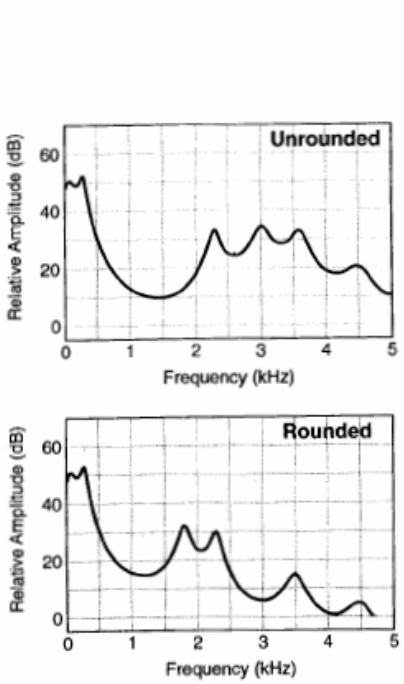
Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Rounding



Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

Figure 6.22 Comparison of the values of *F*1 and *F*2 for a set of six unrounded vowels (points joined by solid lines) and six rounded vowels (points joined by dashed lines). The labeled vowels (solid circles) are approximations to the six basic extreme vowels discussed in the text. The arrows indicate how the formants change when rounding occurs. The frequencies are in the range corresponding to adult male speakers.



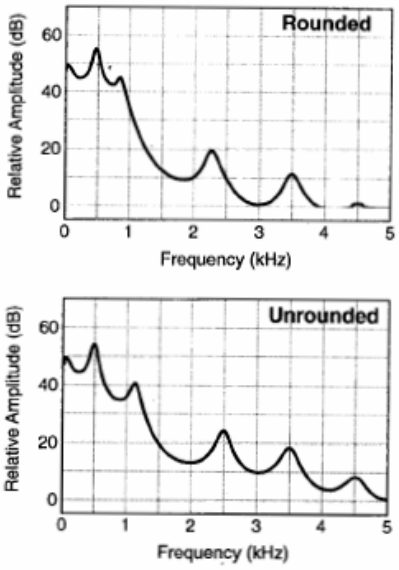


Figure 6.19 Calculated spectrum envelopes for two back vowels, illustrating the contrast between a rounded (top) and an unrounded (bottom) back vowel. The formants *F*1 and *F*2 are closer and higher-frequency peaks are weaker for the rounded vowel.

Figure 6.21 Calculated spectrum envelopes for an unrounded high front vowel (top) and a contrasting rounded front vowel (bottom).

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

The tense-lax distinction (in English)

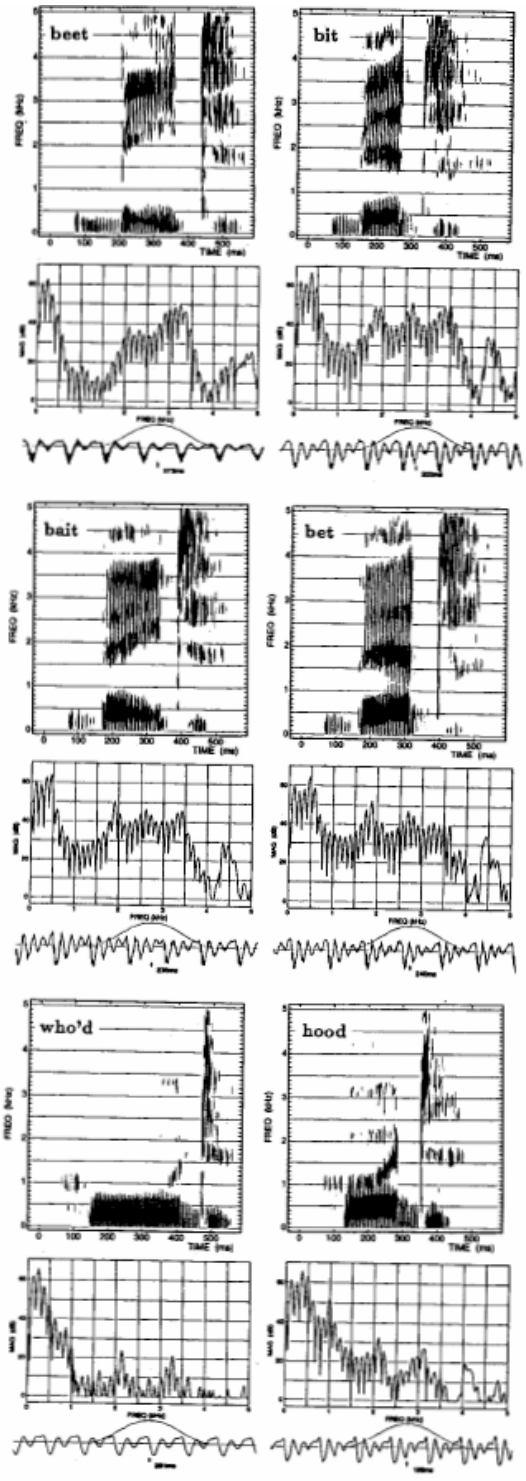


Figure 6.25 Spectrograms and spectra of utterances illustrating the tense-lax pairs of vowels in American English produced by a male speaker. Tense vowels are on the left and lax vowels on the right. The spectrum under each spectrogram is sampled at a point about one-third of the vowel length from the onset. Waveforms are displayed below each spectrum. The spectra show the locations of the formant peaks and the different relative amplitudes of these peaks.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

Figure 6.23 Comparing midsagittal vocal tract configurations for tense and lax high vowels for a speaker of American English. (Data from Perkell, 1969.)

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

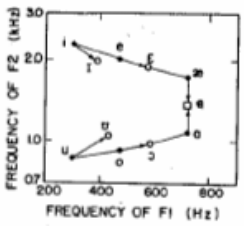
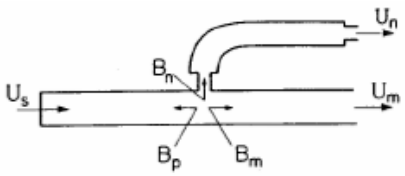


Figure 6.24 Comparison of the values of *F1* and *F2* for the basic set of six tense vowels (closed circles) and the four non-low lax vowels /ɪ ɛ ɔ ʊ/ (open circles). Arrows point from tense vowels to corresponding lax vowels. Data are appropriate for adult male speakers. For the high vowels, data are from American English (Peterson and Barney, 1952). For the vowels /e ɛ ɔ o/, data are estimates from measurements of vowels in English, Italian, and Korean, and for any one language the pair /e ɛ/ or /o ɔ/ may shift somewhat relative to the position shown in the figure. The point for the "lax" low vowel /a/ is arbitrarily placed between /ɑ/ and /æ/.

The feature [nasal] for vowels



Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

Figure 6.30 Schematization of the shapes of the vocal and nasal tracts for a nasal vowel. The volume velocities U_s, U_n, and U_m at the glottis, nostrils, and mouth are shown, as are the acoustic susceptances at the coupling point looking into the pharynx (B_p), the nasal cavity (B_n), and

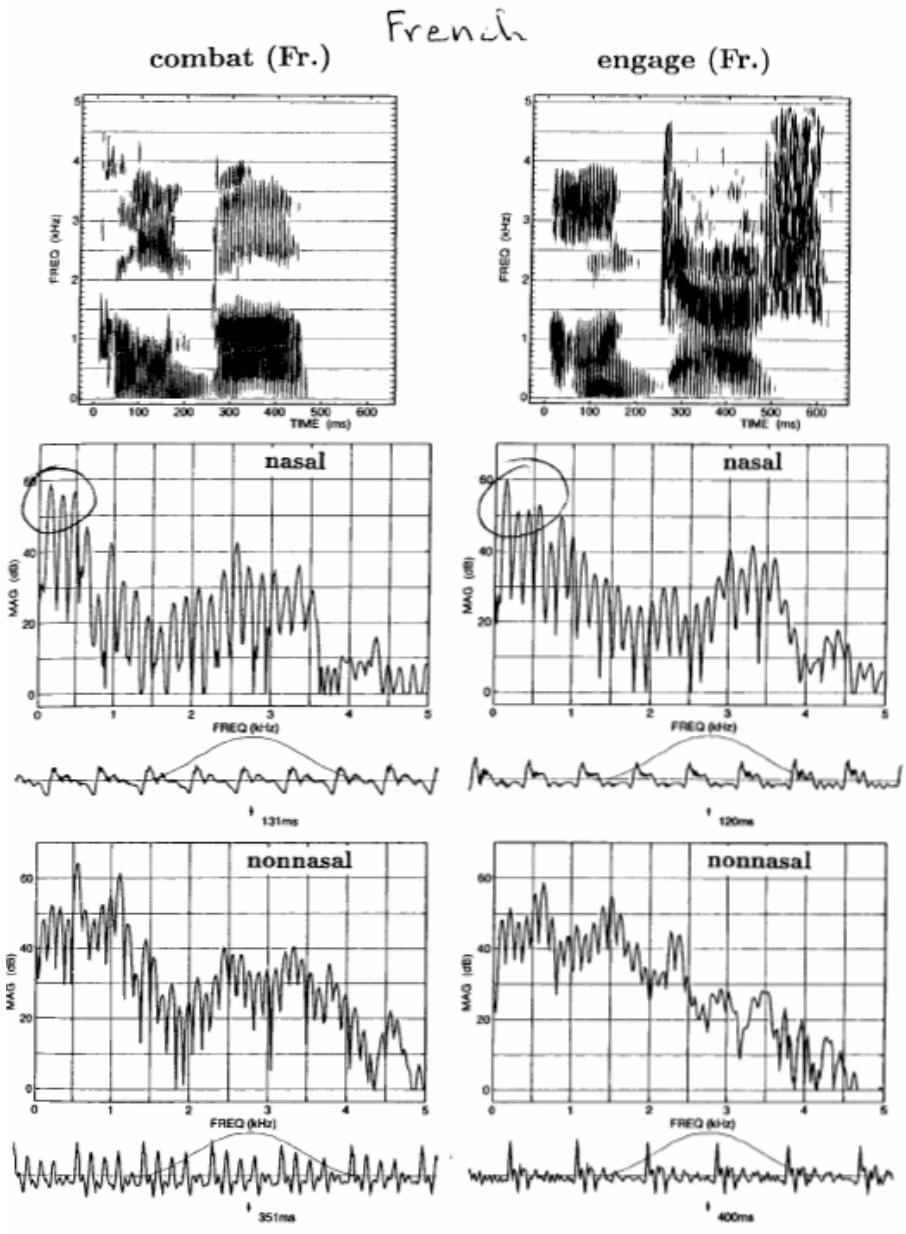


Figure 6.41 Spectrograms and spectra illustrating nasal and non-nasal vowels in French. The left column shows a spectrogram of the word *combat*, and the two spectra below are sampled in the nasal vowel and in the non-nasal vowel, respectively. Similar displays in the right column are for the word *engage*.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.

Figure 6.29 Midsagittal section of the vocal tract for the nasal vowel /ɑ̃/ produced by a speaker of French. (From Bothorel et al., 1986.)

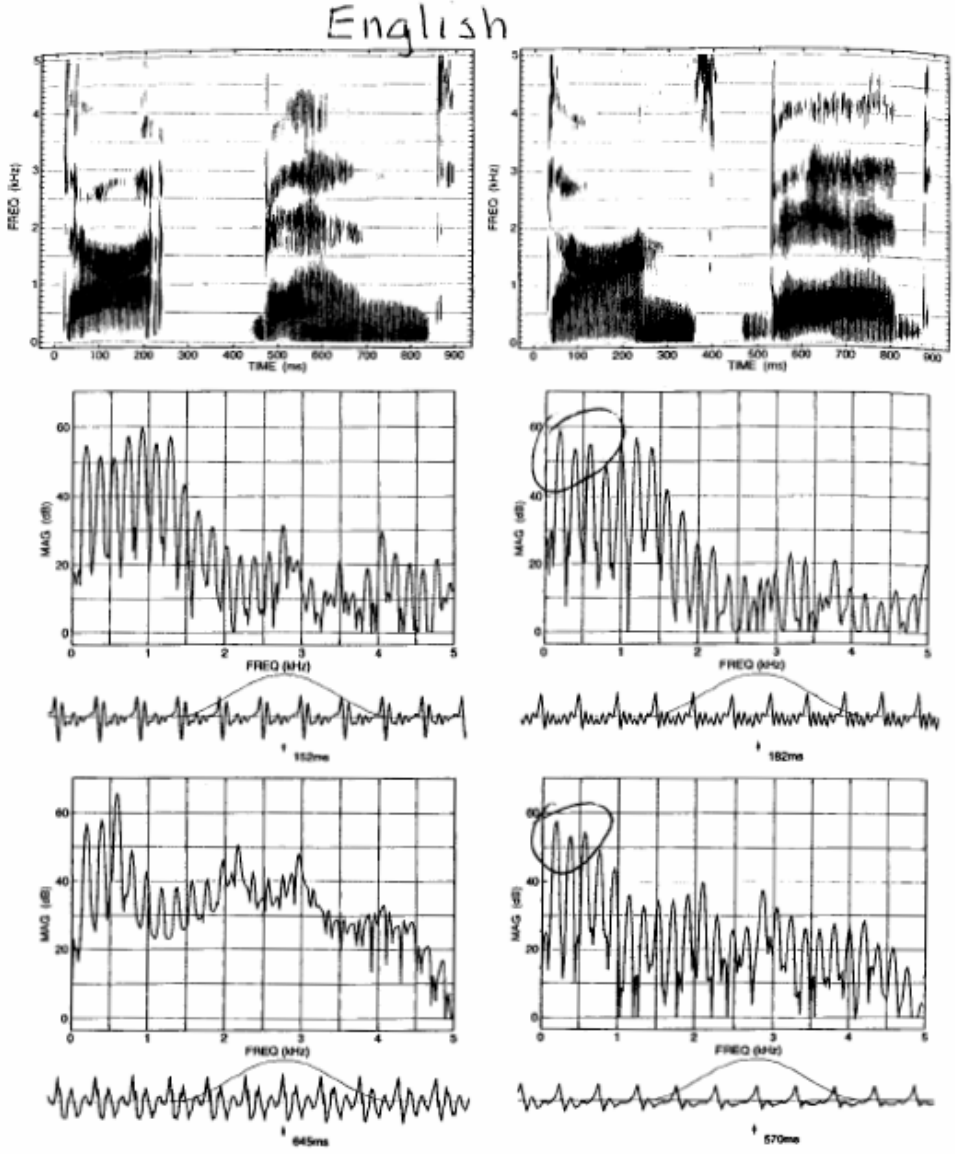


Figure 6.42 Spectrograms and spectra illustrating nasalized vowels and their non-nasal cognates in English for a female speaker. The spectrograms at the top are the utterances *dot bend* (left) and *don's bed* (right). The spectra immediately below the spectrograms are sampled in the non-nasal /ɑ/ in *dot* at the left and the nasalized /ɑ/ in *don* at the right. The spectra at the bottom are sampled in the non-nasal /ɛ/ in *bed* (left) and the nasalized /ɛ/ in *bend* (right).

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998. Courtesy of MIT Press. Used with permission.