24.901 **Phonetics-1**

1. sound results from pressure fluctuations in a medium which displace the ear drum to stimulate the auditory nerve

- air is normal medium for speech
- air is elastic (cf. bicyle pump, plastic bag, etc.)
- pressure fluctuations originate at a source that produce a wave in the medium that carries energy to the ear
- air particles do not move from the source to the ear--rather the energy is passed through the medium
- oscillogram/waveform is a graphic (visual) representation of pressure fluctuations

2. sound waves

- **periodic**: wave repeats at regular intervals (figure 2.1, page 3)
- **frequency** is number of repetitions per unit of time: $f = 1/T$ Hertz (Hz) is 1 cycle per second
- the perceived pitch of a sound depends on its frequency; at frequencies above 1,000 Hz equal increases in frequency are not perceived as equal increases in pitch 1,000 – 2,000 Hz $\approx$ 2,000 – 4,000 Hz
- speed of sound in air is c. 340 m/s; wavelength is speed (m/s) * period (T)
- **amplitude** is maximal displacement of wave above zero line and corresponds to intensity of sound; relation between amplitude and perceived intensity is not linear; the relative difference in perceived intensity is approximated by a logrithmic scale whose unit is the **decibel** (dB); change of 1 dB is JND; 5 dB is about twice intensity

  [θ] = 13 dB, [s] = 17 dB, [i] = 34 dB, [ɑ] = 40 dB
- **aperiodic** sounds have waves that do not repeat
- fricatives like [s] are aperiodic while vowels like [a] are periodic

3. spectrum

- any complex wave can be analyzed as the combination of sinusoidal waves of different frequencies and intensities (Fourier Theorem); see fig. 4.1
- for a periodic sound like a vowel these are the **fundamental** frequency F0 and multiples of the fundamental known as **harmonics** or **overtones**
- the quality of a periodic sound depends on the relative amplitude of the harmonics
- these can be displayed in a **power spectrum** (fig. 4.2)
- examples from Ladefoged 1962: fig. 7.4 of vowel [ɔ]
- note different frequencies but same overall shape of spectrum
- differences in vowel quality result from different vocal tract shapes

- they give rise to different spectra (e.g. figure 7.5)
- the perceived quality of a vowel can be adequately described by the relative location of the peaks in the lower part of the spectrum, termed **formants**
- most vowels are adequately characterized by the first three formants: F1, F2, F3
- fig. 8.2 (CIP) shows formants for eight AE vowels
- **F1** primarily reflects vowel height in invrse fashion: greater F1 reflects lower vowel
- **F2** reflects vowel backness as well as lip rounding: lower F2 reflects greater backing or rounding
- the science of acoustic phonetics models speech as the behavior of waves in various types of tubes (Ken Stevens 1998 *Acoustic Phonetics*, 6.541)

4. **spectrogram**

- a graphic display of the components of a sound (e.g. figure 8.3)
- x-axis is time
- y-axis is frequency
- intensity of sound at a given frequency is indicated by gray scale: darker the wave the greater the intensity.
- **narrow-band** spectrograms give better resolution in the frequency dimension; striations are horizontal
- **wide-band** spectrograms give better resolution in the time dimension; striations are vertical
- **formant chart**
    - origin in NE corner
    - Bark scale*
    - lower part of spectrum has much more energy for vowel sounds

* The scale ranges from 1 to 24 and corresponds to the first 24 critical bands of hearing (in Hz): 20, 100,200,300,400,510, 630, 770, 920, 10801270,1480, 1720, 2000, 2320, 2700, 3150, 3700, 4400, 5300, 6400, 7700, 9500, 12000, 15500.

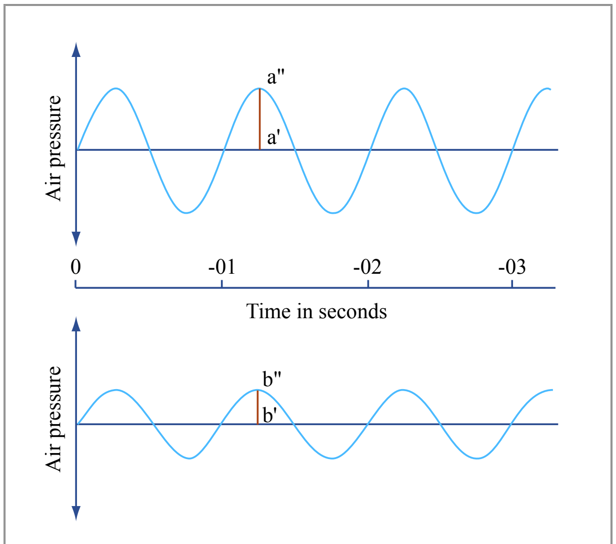

Fig. 2.1 Two sounds, one with twice the amplitude of the other.

Image by MIT OpenCourseWare.

Wave Analysis

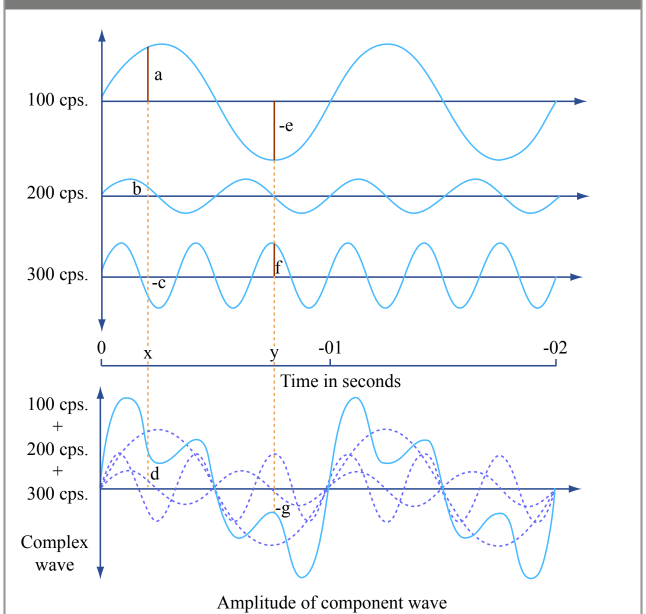

Fig. 4.1 A combination of 100 cps., 200 cps. and 300 cps waves forming a complex wave.

Image by MIT OpenCourseWare.

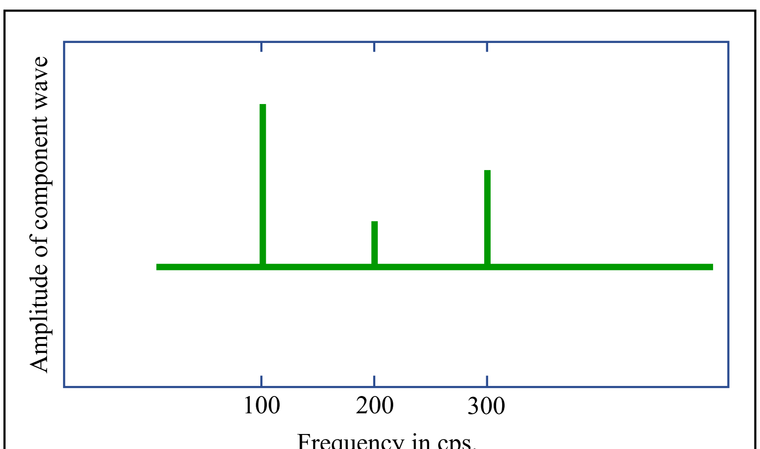

Fig. 4.2 The spectrum of the complex wave illustrated in fig. 4.1.

Image by MIT OpenCourseWare.

Elements of Acoustic Phonetics

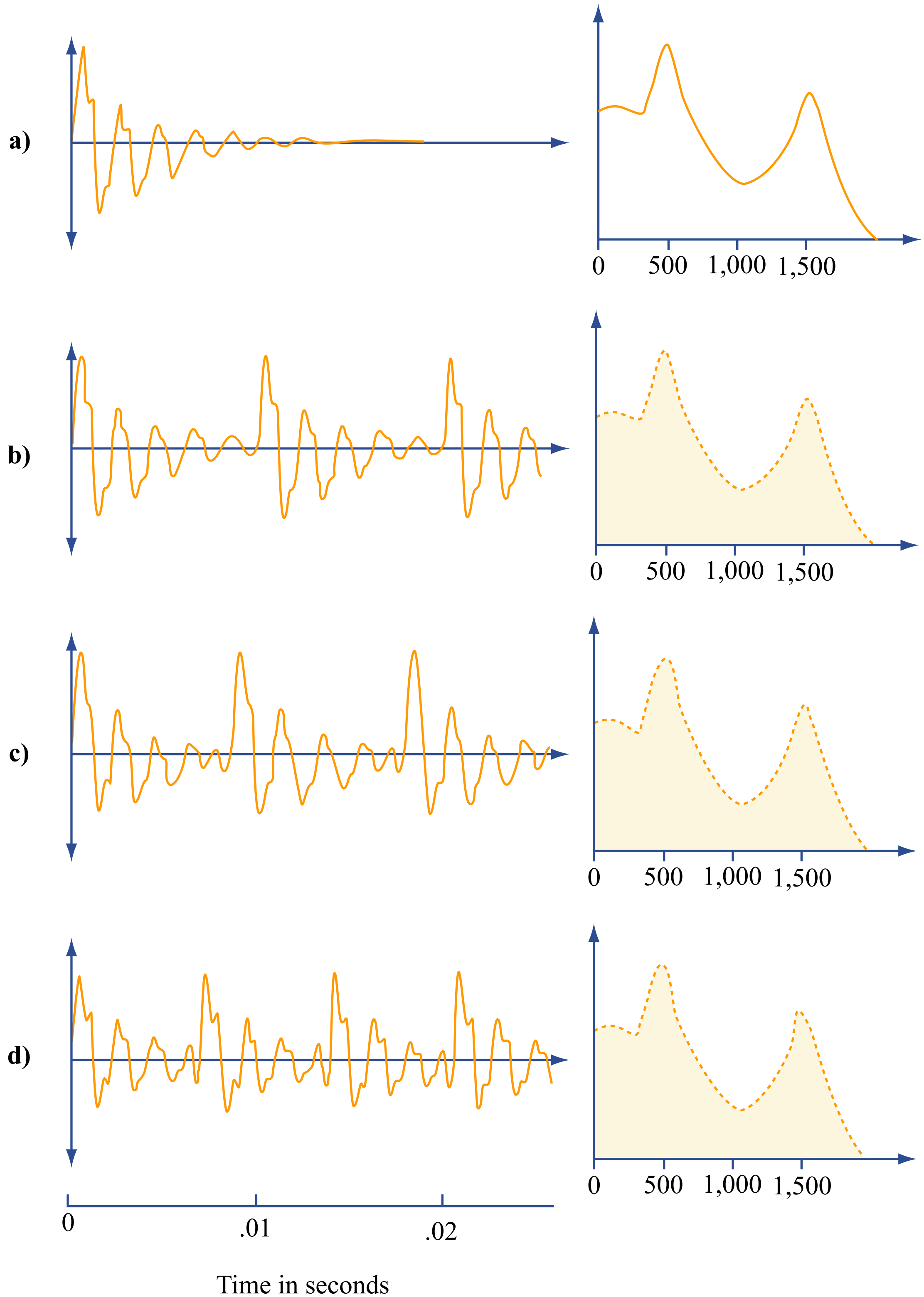

Fig. 7.4 Wave forms and spectra for a synthetic vowel similar to the vowel [ɔ] as in *caught*. (a) Effect of a single pulse on the resonating system; (b) pulses recurring at the rate of 100 a second; (c) 120 pulses a second; (d) 150 pulses a second.

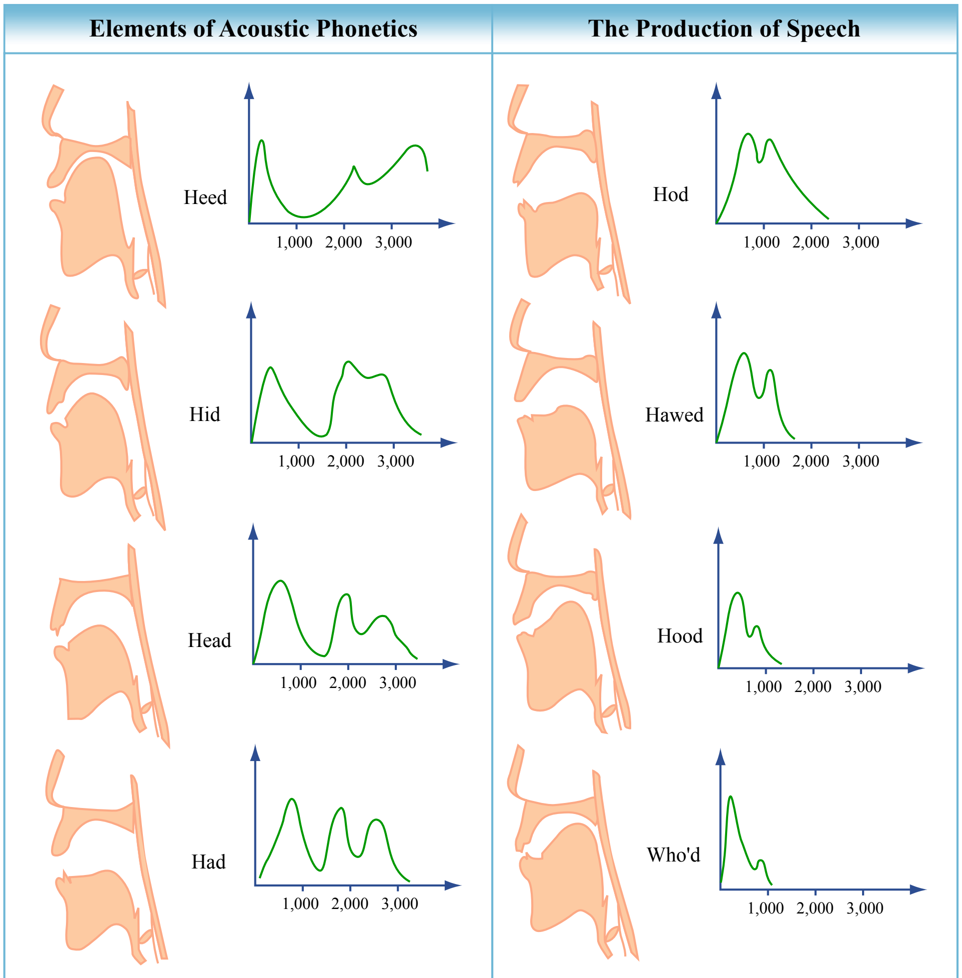

Fig. 7.5 The positions of the vocal organs (based on data from X-ray photographs of the author) and the spectra of the vowel sounds in the middle of the words heed, hid, head, had, hod, hawed, hood, who'd in the author's speech.

Image by MIT OpenCourseWare.

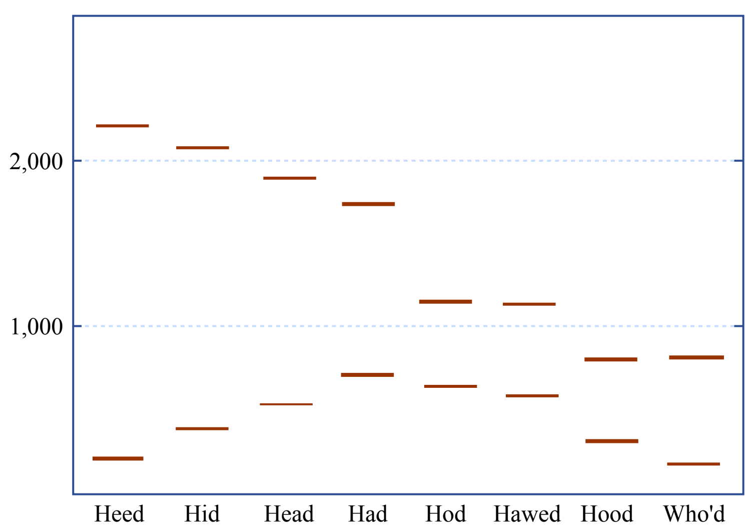

Fig. 7.9 A spectrogram showing the frequencies of the first and second formants of some of the English vowels as pronounced by the author.

Image by MIT OpenCourseWare.

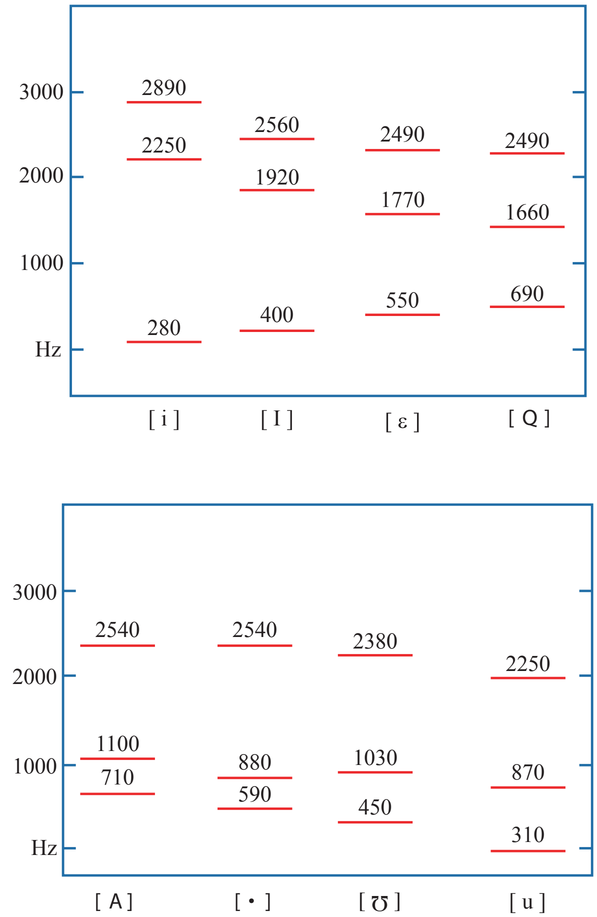

FIGURE 8.2: The Frequencies of the First Three Formants in Eight American English Vowel.

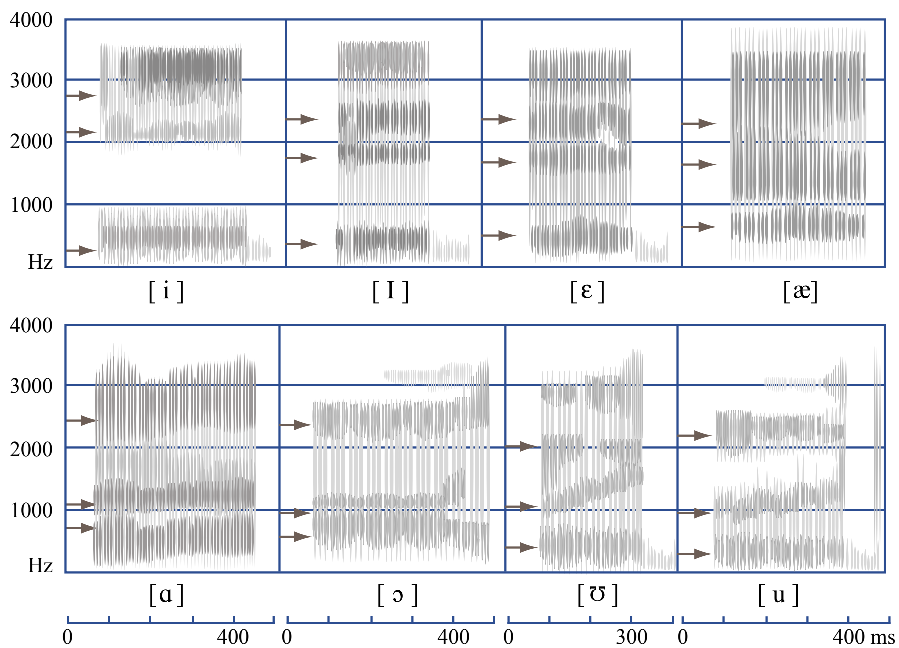

Figure 8.3 A spectrogram of the words heed, hid, head, had, hod, hawed, hood, who'd as spoken by a male speaker of American English. The locations of the first three formants are shown by arrows.

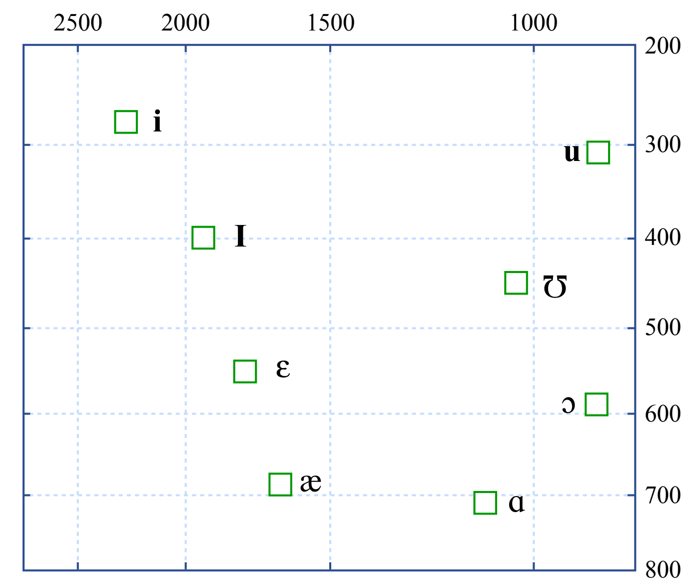

Figure 8.5 A formant chart showing the frequency of the first formant on the ordinate (the vertical axis) plotted against the second formant on the abscissa (the horizontal axis) for eight American English vowels. The scales are marked in Hz, arranged at Bark scale intervals.

Image by MIT OpenCourseWare.

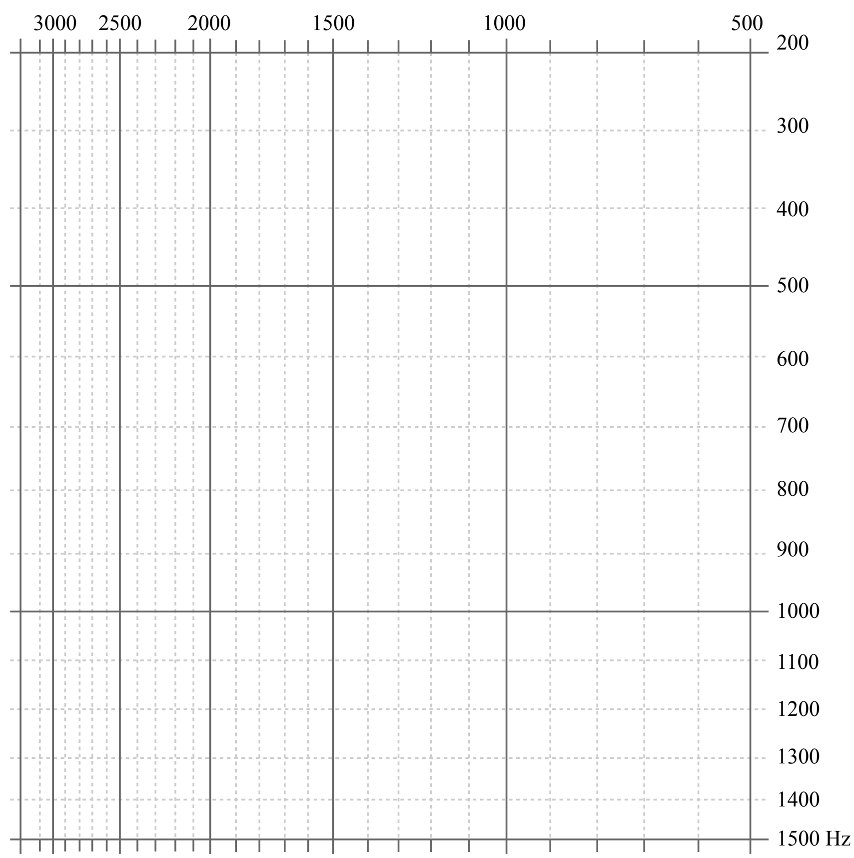

Figure 8.6 A blank formant chart for showing the relation between vowels. Using the information in Figures 8.3 and 8.4, plot the frequency of the first formant on the ordinate (the vertical axis) and the second formant on the abscissa (the horizontal axis).

Image by MIT OpenCourseWare.