24.914
Speech acoustics
Introduction to acoustics

• Sound consists of pressure fluctuations in a medium (usually air).

• Movements at a source produce a sound wave in the medium which carries energy to the perceiver.

www.acs.psu.edu/drussell/Demos/waves/wavemotion.html

Image courtesy of Dan Russell. Used with permission.
Speech acoustics

- Pressure fluctuations move through space, but each air particle moves only a small distance.

www.acs.psu.edu/drussell/Demos/waves/wavemotion.html

Image courtesy of Dan Russell. Used with permission.
Representing sound waves

Sound is a Pressure Wave

C = Compression
R = Rarefaction

Image by MIT OCW.
Adapted from The Physics Classroom Tutorial.
Periodic sounds

- A waveform is periodic if it repeats at regular intervals.
- Frequency of a wave is the number of cycles occurring per unit of time.
  - Units: 1 Hertz (Hz) is 1 cycle/second

![Diagram of a periodic waveform with amplitude and time axes labeled.](Image by MIT OCW.)
Periodic sounds

- Voiced sounds have complex (quasi-)periodic wave forms.
- The perceived pitch of a sound depends on its frequency.
Aperiodic sounds

- Aperiodic sounds have waveforms that do not repeat.
- Fricative noise is aperiodic.

Segment of [s]
Spectrums and spectrograms

• The spectrum of a sound plays a central role in determining its quality or timbre.
Spectral representation

- Any complex wave can be analyzed as the combination of a number of sinusoidal waves of different frequencies and intensities (Fourier theorem).
- In the case of a periodic sound like a vowel these will be
  - the fundamental frequency
  - multiples of the fundamental frequency (harmonics)
- The quality of a periodic sound depends on the relative amplitude of its harmonics.
Spectral representation
Fundamental frequency
2nd harmonic
Spectral representation

- Phase differences are relatively unimportant to sound quality, so key properties of a complex wave can be specified in terms of the frequencies and amplitudes of its sinusoidal components.

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>200</td>
<td>0.6</td>
</tr>
<tr>
<td>300</td>
<td>0.45</td>
</tr>
<tr>
<td>400</td>
<td>0.3</td>
</tr>
<tr>
<td>500</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Power spectrum
Idealized vowel spectrum

Intensity level (decibels)

Frequency (Hertz)
vowel spectrum

Frequency (Hz)

[æ]
Vowel quality

- The quality of a vowel depends on the shape of its spectrum.
- The shape of the spectrum depends on the shape of the vocal tract.
Source-Filter Model of Speech Production

Vowel quality

- The peaks in the spectrum of a vowel are called **formants**.
- Perceived vowel quality depends primarily on the frequencies of the first three formants.
Spectrograms
Spectrograms

Image by MIT OCW.
narrow band
(long window)

broad band
(short window)
Dialect differences in vowels

- Californian speakers (M open, F closed) (Hagiwara 1997).

- N. Midwest speakers (M open, F closed) (Hillenbrand et al 1995).