Speech in a Slide

- Frequency gives pitch; amplitude gives volume
- Frequencies at each time slice processed into observation vectors

Articulatory System

- Nasal cavity
- Oral cavity
- Pharynx
- Vocal folds (in the larynx)
- Trachea
- Lungs

Places of Articulation

- Bilabial: p, b, m
- Labiodental: f, v
- Alveolar: t/d/s/z/l/n
- Velar: k/g/ɑ
- Labial: l/ɹ/j
- Coronal: f, v, θ/dh

Labial place

Coronal place
Manner of Articulation

- In addition to varying by place, sounds vary by manner.
- Stop: complete closure of articulators, no air escapes via mouth.
  - Oral stop: palate is raised (p, t, k, b, d, g).
  - Nasal stop: oral closure, but palate is lowered (m, n, ng).
- Fricatives: substantial closure, turbulent: (f, v, s, z).
- Approximants: slight closure, sonorant: (l, r, w).
- Vowels: no closure, sonorant: (i, e, a).

Vowel Space

- What can we learn from a wavefile?
  - No gaps between words (!)
  - Vowels are voiced, long, loud.
  - Length in time = length in space in waveform picture.
  - Voicing: regular peaks in amplitude.
  - When stops closed: no peaks, silence.
  - Peaks in voicing: .46 to .58 (vowel [i]), from second .65 to .74 (vowel [ax]) and so on.
  - Silences of stop closure (1.06 to 1.08 for first [b], or 1.26 to 1.28 for second [b]).
  - Fricatives like [sh]: intense irregular pattern; see .33 to .46.
Non-Local Cues

Example from Ladefoged

Simple Periodic Waves of Sound

- Y axis: Amplitude = amount of air pressure at that point in time
- Zero is normal air pressure, negative is rarefaction
- X axis: Time.
- Frequency = number of cycles per second.
- 20 cycles in .02 seconds = 1000 cycles/second = 1000 Hz

Complex Waves: 100Hz+1000Hz

Spectrum

Frequency components (100 and 1000 Hz) on x-axis

Spectrum of an Actual Soundwave

Part of [ae] waveform from “had”

- Note complex wave repeating nine times in figure
- Plus smaller waves which repeats 4 times for every large pattern
- Large wave has frequency of 250 Hz (9 times in .036 seconds)
- Small wave roughly 4 times this, or roughly 1000 Hz
- Two little tiny waves on top of peak of 1000 Hz waves
Back to Spectra

- Spectrum represents these freq components
- Computed by Fourier transform, algorithm which separates out each frequency component of wave.

- x-axis shows frequency, y-axis shows magnitude (in decibels, a log measure of amplitude)
- Peaks at 930 Hz, 1860 Hz, and 3020 Hz.

Why these Peaks?

- Articulator process:
  - The vocal cord vibrations create harmonics
  - The mouth is an amplifier
  - Depending on shape of mouth, some harmonics are amplified more than others

Vowel [i] sung at successively higher pitches

- Figures from Rattee Wayland

Resonances of the Vocal Tract

- The human vocal tract as an open tube:
  - Air in a tube of a given length will tend to vibrate at resonance frequency of tube.
  - Constraint: Pressure differential should be maximal at (closed) glottal end and minimal at (open) lip end.

- Length 17.5 cm.

- From Sundberg

Computing the 3 Formants of Schwa

- Let the length of the tube be L
  - \( F_1 = \frac{c}{4L} = \frac{35,000}{4 \times 17.5} = 500 \text{Hz} \)
  - \( F_2 = \frac{3c}{4L} = \frac{3 \times 35,000}{4 \times 17.5} = 1500 \text{Hz} \)
  - \( F_3 = \frac{5c}{4L} = \frac{5 \times 35,000}{4 \times 17.5} = 2500 \text{Hz} \)

- So we expect a neutral vowel to have 3 resonances at 500, 1500, and 2500 Hz
- These vowel resonances are called formants
American English Vowel Space

How to Read Spectrograms

- bab: closure of lips lowers all formants: so rapid increase in all formants at beginning of "bab"
- dad: first formant increases, but F2 and F3 slight fall
- gag: F2 and F3 come together: this is a characteristic of velars. Formant transitions take longer in velars than in alveolars or labials

"She came back and started again"

- lots of high-freq energy
- closure for k
- burst of aspiration for k
- ey vowel; faint 1100 Hz formant is nasalization
- bilabial nasal
- short b closure, voicing barely visible.
- ae; note upward transitions after bilabial stop at beginning
- note F2 and F3 coming together for "k"

From Mark Liberman's Web site

From Jennifer Venditti, H. T. Bunnell

From Ladefoged "A Course in Phonetics"