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

Places of articulation

Labial place
Labiodental: p, b, m
Bilabial: b, m

Coronal place
Dental: th/dh
Alveolar: t/d/s/z/l
Post: sh/zh/y

Text copyright J. J. Ohala, Sept 2001, from Sharon Rose slide
**Manner of Articulation**

- **Stop**: complete closure of articulators, so no air escapes through mouth
- **Oral stop**: palate is raised, no air escapes through nose. Air pressure builds up behind closure, explodes when released
  - p, t, k, b, d, g
- **Nasal stop**: oral closure, but palate is lowered, air escapes through nose.
  - m, n, ng

**Oral vs. Nasal Sounds**

**Vowels**

- **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 = voicing**: .46 to .58 (vowel [iy]), from second .65 to .74 (vowel [ax]) and so on
- **Silence 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

**Examples from Ladefoged**

- pad
- bad
- spat
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.
- Frequency = 1/Period
- 20 cycles in .02 seconds = 1000 cycles/second = 1000 Hz

Complex waves: 100Hz+1000Hz

- Spectrum
  - Frequency components (100 and 1000 Hz) on x-axis
  - Amplitude
  - Frequency in Hz
  - Peaks at 930 Hz, 1860 Hz, and 3020 Hz.

- 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

Resonances of the Vocal Tract

- The human vocal tract as an open tube

![Resonances of the Vocal Tract](image)

- 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.

Computing the 3 Formants of Schwa

- Let the length of the tube be \( L \)
  - \( F_1 = c/\lambda_1 = c/(4L) = 35,000/4*17.5 = 500 \text{ Hz} \)
  - \( F_2 = c/\lambda_2 = c/(4/3L) = 3c/4L = 3*35,000/4*17.5 = 1500 \text{ Hz} \)
  - \( F_3 = c/\lambda_3 = c/(4/5L) = 5c/4L = 5*35,000/4*17.5 = 2500 \text{ Hz} \)

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

Seeing formants: the spectrogram

![Seeing formants: the spectrogram](image)
**American English Vowel Space**

Vowel [i] sung at successively higher pitch.

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

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

From Ladefoged “A Course in Phonetics”