The Central Auditory System

- There are many parallel pathways in the auditory brainstem.
- The *binaural* system receives input from both ears.
- The *monaural* system receives input from one ear only.

Each set of auditory pathways has a specialized function.
Speech Production:
  Respiration (lungs)
  Phonation (vocal cords/folds)
  Articulation (vocal tract)

Vocal Cords/Folds
The Vocal Tract:

- The airway above the larynx used for production of speech. Includes the oral tract and nasal tract
  - Humans are capable of producing lots of different speech sounds
  - 5000 languages spoken today, utilizing over 850 different speech sounds
  - Flexibility of vocal tract: Important in speech production

Articulators

- Vocal organs that produce sounds are called articulators
Articulators + Source Energy = Speech Sounds

• Speech:
  – Vocal cords open slowly and close quickly
  – Airflow pulses to produce a buzz (a waveform with a characteristic period and, hence, frequency)
  – The complex buzz can be decomposed into constituent sinusoidal frequencies (each an integer multiple of the original—fundamental—frequency)
  – Original source spectrum filtered by vocal tract; precise effect depends on articulator position
How do we look at speech?

- Fourier transformations create spectrograms
Spectrogram

• A landscape overview of the spectral characteristics and spectral changes in the time domain
• A 3-D display with y-axis representing frequency, x-axis representing time, and the shade of darkness representing amplitude
Audio Spectrum of the Song of the Prothonotary Warbler
Speech Basics

• Key entity is the phoneme
  – minimal distinguishing sounds (consonants and vowels)
  – change one phoneme in a word and you change the word (*heed vs. head vs. had vs. hid*)

Vowel Acoustics

• **Articulatory:**
  – articulation involves no significant constriction of articulators in vocal tract (air flows with little to no obstruction)

• **Acoustic:**
  – spectrograms demonstrate characteristic steady formant patterns

• **Perceptual:**
  – most prominent in a syllable
Vowel Classification

A vowel chart showing the relative vowel qualities represented by some of the symbols used in transcribing English. The symbols [e, a, o] occur as the first elements of diphthongs.

Looking at Formants in Spectrograms

- Depending on tract shape, some frequency components get through better than others
  - Formants: frequency regions where transmission is good, thanks to resonance in vocal tract
  - F1: lowest frequency formant; F2: next up, etc.
  - Vowels are distinguishable by values of F1, F2, etc.
Consonant Acoustics

- **Articulatory:**
  - significant constriction of articulators in the vocal tract
  - air stream is partially or totally obstructed in the vocal tract
- **Acoustic:**
  - classified according to the nature of constriction in articulation
- **Perceptual:**
  - generally not the most prominent segment or nucleus of a syllable
## Consonant Classification

<table>
<thead>
<tr>
<th>Place of articulation</th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Interdental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral stop voiced</td>
<td>p (pin)</td>
<td>b (ben)</td>
<td>1 (ihn)</td>
<td>d (dnt)</td>
<td>t (tin)</td>
<td>k (kin)</td>
</tr>
<tr>
<td>nasal stop voiced</td>
<td>m (nap)</td>
<td>n (nap)</td>
<td>ɑ (ahn)</td>
<td>ɑ (ahn)</td>
<td>ɑ (ahn)</td>
<td>ɒ (ong)</td>
</tr>
<tr>
<td>Fricative voiced</td>
<td>f (fin)</td>
<td>v (van)</td>
<td>ɹ (thr)</td>
<td>s (sin)</td>
<td>ʃ (shin)</td>
<td>ʒ (zhon)</td>
</tr>
<tr>
<td>Affricate voiced</td>
<td>tʃ (chin)</td>
<td>dʒ (dzhin)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid voiced</td>
<td>l (law)</td>
<td>r (raw)</td>
<td>ɪ (in)</td>
<td></td>
<td>w (wet)</td>
<td></td>
</tr>
<tr>
<td>Glides voiced</td>
<td>j (yen)</td>
<td>w (wet)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
International Phonetic Alphabet (consonants)

<table>
<thead>
<tr>
<th>CONSONANTS</th>
<th>Nasal</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Anterior</th>
<th>Posterior</th>
<th>Velar</th>
<th>Uvular</th>
<th>Under</th>
<th>Pharyngeal</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice</td>
<td>p b</td>
<td>t d</td>
<td>t d c j k g q g ?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m n</td>
<td>n n n j n N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>B R</td>
<td>r r r r r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap or Flag</td>
<td>f t</td>
<td>f t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>φ β f v θ ð s z s z s z s z j x y χ k h h h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral Ejective</td>
<td>p' t' t' c' k' q'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implosive</td>
<td>b b</td>
<td>l d c f k g q q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded rows denote articulations judged impossible.*
International Phonetic Alphabet (vowels)

![Vowel Chart]

Where symbols appear in pairs, the one to the right represents a rounded vowel.

Peter Ladefoged 1926-2006
http://www.linguistics.ucla.edu/people/ladefoge/