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
• 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

Spectrum of the sentence "Joe took father's shoe bench out." Note that the pauses (the absence of energy) are not between words. (from Tannen, V. C. [1986], Language Process, NY: Holt, Rinehart and Winston, Figure 7.1 page 210.)
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

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>t (tin)</td>
<td>d (din)</td>
<td>k (kon)</td>
<td>g (gin)</td>
<td></td>
</tr>
<tr>
<td>Nasal stop voiced</td>
<td>m (map)</td>
<td>n (nap)</td>
<td></td>
<td></td>
<td>g (ing)</td>
<td></td>
</tr>
<tr>
<td>Friction voiced</td>
<td>f (fin)</td>
<td>θ (thin)</td>
<td>s (sin)</td>
<td>l (lin)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate voiced</td>
<td>j (jan)</td>
<td>j (jan)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide voiced</td>
<td>j (jen)</td>
<td>w (wen)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11/4/2008
International Phonetic Alphabet (consonants)

<table>
<thead>
<tr>
<th>CONSONANTS</th>
<th>Nasal</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Postalveolar</th>
<th>Velar</th>
<th>Uvular</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless</td>
<td>p b</td>
<td>t d</td>
<td>l q</td>
<td>c j</td>
<td>k g</td>
<td>q g</td>
<td>g ?</td>
<td>g ?</td>
<td>g ?</td>
</tr>
<tr>
<td>Nasal</td>
<td>m n</td>
<td>n η</td>
<td>n η</td>
<td>n η</td>
<td>n η</td>
<td>n η</td>
<td>n η</td>
<td>n η</td>
<td>n η</td>
</tr>
<tr>
<td>Trill</td>
<td>B r</td>
<td>r</td>
<td>r</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Tap or Flap</td>
<td>r 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</td>
<td>v θ δ</td>
<td>s z</td>
<td>s z</td>
<td>s z</td>
<td>s z</td>
<td>s z</td>
<td>s z</td>
<td>s z</td>
</tr>
<tr>
<td>Lateral Fricative</td>
<td>f v</td>
<td>η h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>v j</td>
<td>η j</td>
<td>η j</td>
<td>j w</td>
<td>j w</td>
<td>j w</td>
<td>j w</td>
<td>j w</td>
<td>j w</td>
</tr>
<tr>
<td>Lateral Approximant</td>
<td>l l</td>
<td>l s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ejective stop</td>
<td>p' t'</td>
<td>η' η'</td>
<td>η' η'</td>
<td>k' q'</td>
<td>k' q'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implosive</td>
<td>b b</td>
<td>f f</td>
<td>η f</td>
<td>η f</td>
<td>η f</td>
<td>η f</td>
<td>η f</td>
<td>η f</td>
<td>η f</td>
</tr>
</tbody>
</table>

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

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/