ECE 598: The Speech Chain

Lecture 9: Consonants
Today

- International Phonetic Alphabet
  - History
  - SAMPA – an IPA for ASCII
- Sounds with a Side Branch
  - Nasal Consonants
  - Reminder: Impedance of a uniform tube
  - Liquids: /l/, /r/
- Events in the Release of a Syllable-Initial Stop Consonant
  - Transient and Frication Sources
  - Aspiration
  - Formant Transitions
Topic #1: International Phonetic Alphabet
International Phonetic Alphabet: Purpose and Brief History

- **Purpose of the alphabet:** to provide a universal notation for the sounds of the world’s languages
  - “Universal” = If any language on Earth distinguishes two phonemes, IPA must also distinguish them
  - “Distinguish” = Meaning of a word changes when the phoneme changes, e.g. “cat” vs. “bat.”

- **Very Brief History:**
  - 1446: King Sejong of Chosen publishes a distinctive-feature based phonetic notation for Korean.
  - 1867: Alexander Melville Bell publishes a distinctive-feature-based universal phonetic notation in “Visible Speech: The Science of the Universal Alphabetic.” His notation is rejected as being too expensive to print.
  - 1886: International Phonetic Association founded in Paris by phoneticians from across Europe; begins developing IPA notation.
*SAMPA: An international standard for the ASCII transcription of the IPA phonemes. Maps most IPA phones to the ASCII printable characters.*
Vowels: Sample Words

From [http://www.phon.ucl.ac.uk/home/sampa/american.htm](http://www.phon.ucl.ac.uk/home/sampa/american.htm) With phonological feature annotation. All are [+syllabic].

- **Lax Vowels** [-reduced,+lax,-blade]: lax: like IPA “central,” but also short duration
  - I pit pIt [+high,+front] high: like IPA “close.”
  - E pet pEt [-high,+front]
  - V cut kVt [-high,-front]
  - U put pUt [+high,-front]
- **Tense Vowels**: [-reduced,-lax]
  - i ease iz [-low,+high,+front] low: like IPA “open.” Specified if [-high,-lax]
  - e raise rez [-low,-high,+front]
  - { pat p{t [+low,+front]
  - u lose luz [-low,+high,-front,+round] [+round]:
  - o nose noz [-low,-high,-front,+round] possible if [-lax,-front]
  - O cause kOz [+low,-front,+round]
  - A pot fAD’ [+low,-front,-round]
- **Classic Diphthongs**:
  - aI rise raIz
  - OI noise nOIz
  - aU rouse raUz
- **Schwa and Syllabic /r/**: 
  - 3’ furs f3’z [-reduced,+lax,+blade,-anterior,-distributed]
  - @ allow @laU [+reduced,-blade]
  - @’ corner kOrn@’ [+reduced,+blade,-anterior,-distributed]
### IPA: Regular Consonants

#### THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

<table>
<thead>
<tr>
<th>CONSONANTS (PULMONIC)</th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Postalveolar</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plosive</strong></td>
<td><strong>p</strong></td>
<td><strong>b</strong></td>
<td><strong>t</strong></td>
<td><strong>d</strong></td>
<td><strong>t</strong></td>
<td><strong>d</strong></td>
<td><strong>k</strong></td>
<td><strong>q</strong></td>
<td><strong>g</strong></td>
<td></td>
<td><strong>?</strong></td>
</tr>
<tr>
<td><strong>Nasal</strong></td>
<td><strong>m</strong></td>
<td><strong>n</strong></td>
<td><strong>n</strong></td>
<td><strong>n</strong></td>
<td><strong>N</strong></td>
<td><strong>N</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trill</strong></td>
<td><strong>B</strong></td>
<td></td>
<td><strong>r</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>R</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tap or Flap</strong></td>
<td></td>
<td></td>
<td><strong>4r</strong></td>
<td></td>
<td><strong>t</strong></td>
<td><strong>t</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td><strong>φ</strong></td>
<td><strong>θ</strong></td>
<td><strong>θ</strong></td>
<td><strong>s</strong></td>
<td><strong>s</strong></td>
<td><strong>z</strong></td>
<td><strong>ç</strong></td>
<td><strong>x</strong></td>
<td><strong>y</strong></td>
<td><strong>χ</strong></td>
<td><strong>h</strong></td>
</tr>
<tr>
<td><strong>Lateral fricative</strong></td>
<td></td>
<td></td>
<td><strong>D</strong></td>
<td><strong>s</strong></td>
<td><strong>s</strong></td>
<td><strong>z</strong></td>
<td><strong>ç</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>h</strong></td>
</tr>
<tr>
<td><strong>Approximant</strong></td>
<td><strong>voy</strong></td>
<td><strong>r</strong></td>
<td><strong>l</strong></td>
<td><strong>j</strong></td>
<td><strong>j</strong></td>
<td><strong>w</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lateral approximant</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>l</strong></td>
<td><strong>L</strong></td>
<td><strong>L</strong></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 areas denote articulations judged impossible.
Obstruent Consonants: Sample Words

From http://www.phon.ucl.ac.uk/home/sampa/american.htm

With phonological feature annotation. All are [-syllabic].

- **Stops** [-sonorant,-continuant]:
  - p pin pIn [+lips,-round,-voice]
  - b bin bIn [+lips,-round,+voice]
  - t tin tIn [+blade,+anterior,-distributed,-voice] anterior: of the alveolar ridge
  - d din dIn [+blade,+anterior,-distributed,+voice] distributed: tongue tip flat
  - k kin kIn [+body,+high,-voice]
  - g give gIv [+body,+high,+voice]

- **Affricates** [-sonorant,-continuant]:
  - tS chin tSIn [+blade,-anterior,+distributed,-voice]
  - dZ gin dZIn [+blade,-anterior,+distributed,+voice]

- **Fricatives** [-sonorant,+continuant]:
  - f fin fIn [+lips,-round,-voice]
  - v vim vIm [+lips,-round,+voice]
  - T thin TIn [+blade,+anterior,+distributed,-voice]
  - D this DIs [+blade,+anterior,+distributed,+voice]
  - s sin sIn [+blade,+anterior,-distributed,-voice]
  - z zin zIn [+blade,+anterior,-distributed,+voice]
  - S shin SIn [+blade,-anterior,+distributed,-voice]
  - Z measure mEZ@’ [+blade,-anterior,+distributed,+voice]
  - h hit hIt [+glottal] --- voicing is not specified, depends on context
Doubly-Articulated Consonants

**SAMPA**

- `M` Voiceless labial-velar fricative
- `W` Voiced labial-velar approximant
- `H` Voiced labial-palatal approximant
- `H` Voiceless epiglottal fricative
- `]>` Voiced epiglottal fricative
- `?` Epiglottal plosive

**OTHER SYMBOLS**

- `?)` Alveolo-palatal fricatives
- `J` Alveolar lateral flap
- `f` Simultaneous `J` and `X`

Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary.

\[ kp \quad ts \]
Sonorant Consonants: Sample Words

Mostly from [http://www.phon.ucl.ac.uk/home/sampa/american.htm](http://www.phon.ucl.ac.uk/home/sampa/american.htm)

With phonological feature annotation

- **Flap** [+sonorant, -continuant, -nasal, -syllabic]:
  - 4 butter bV43’ [+blade, +anterior, -distributed]

- **Nasals** [+sonorant, -continuant, +nasal]:
  - m mock mAk [+lips, -round, -syllabic]
  - n knock nAk [+blade, +anterior, -distributed, -syllabic]
  - =n button bV4=n [+blade, +anterior, -distributed, +syllabic]
  - N thing TIN [+body, +high, -syllabic]

- **Liquids** [+sonorant, +continuant, -syllabic]:
  - r wrong rON [+blade, -anterior, -distributed]
  - l long lON [+blade, +anterior, -distributed]

- **Glides** [+sonorant, +continuant, -syllabic]:
  - w wasp wAsp [+lips, +round]
  - j yacht jAt [+blade, -anterior, -distributed]
1. You wish to know all about my grandfather. Well, he is nearly ninety-three years old, but he still thinks as swiftly as ever.

2. ju wIS tu no Al @baUt maI gr}ndfaD@r. wEl, hi iz nirli naInti-Tri yirz old, bVt hi stIl Tinks }z swIftlI }z Ev@r.
### Example: Phonological Feature Matrix

<table>
<thead>
<tr>
<th>Feature</th>
<th>@</th>
<th>k</th>
<th>u</th>
<th>s</th>
<th>t</th>
<th>I</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>[syllabic]</td>
<td></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>[sonorant]</td>
<td></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>[continuant]</td>
<td></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[reduced]</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>[lax]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>[lips]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[round]</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[blade]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>[anterior]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>[distributed]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[body]</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>[high]</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[pharyngeal/low]</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[front]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>[nasal]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>[glottal]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>[voice]</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
## Non-Pulmonic Consonants

<table>
<thead>
<tr>
<th>CONSONANTS (NON-PULMONIC)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clicks</strong></td>
<td><strong>Voiced implosives</strong></td>
<td><strong>Ejectives</strong></td>
</tr>
<tr>
<td>Bilabial</td>
<td>Bilabial</td>
<td>' as in:</td>
</tr>
<tr>
<td>Dental</td>
<td>Dental/alveolar</td>
<td>p' Bilabial</td>
</tr>
<tr>
<td>(Post)alveolar</td>
<td>Palatal</td>
<td>t' Dental/alveolar</td>
</tr>
<tr>
<td>Palatoalveolar</td>
<td>Velar</td>
<td>k' Velar</td>
</tr>
<tr>
<td>Alveolar lateral</td>
<td>Uvular</td>
<td>s' Alveolar fricative</td>
</tr>
</tbody>
</table>
Topic #2:
Sounds with Side
Branches: Nasal
Consonants, /l/, /r/
Nasal Murmur

“the mug”

“the nut”

“sing a song”

Observations:

- Low-frequency resonance (about 300Hz) always present
- Low-frequency resonance has wide bandwidth (about 150Hz)
- Energy of low-frequency resonance is very constant
- Most high-frequency resonances cancelled by zeros
- Different places of articulation have different high frequency spectra
- High-frequency spectrum is talker-dependent and variable
Acoustic Description of Vocal Tract with a Side Branch

Continuity Equations at the Juncture (x=0):
1. Conservation of mass
   \[ u_P(0,\omega) + u_N(0,\omega) + u_M(0,\omega) = 0 \]
2. Continuity of pressure
   \[ p_P(0,\omega) = p_N(0,\omega) = p_M(0,\omega) \]
Reminder: How to Calculate Impedance of a Uniform Tube

1. Express \( u(x, \omega) \), \( p(x, \omega) \) in terms of \( p_+ \) and \( p_- \)
   \[
   p(x, \omega) = p_+e^{-j\omega x/c} + p_-e^{j\omega x/c}
   \]
   \[
   u(x, \omega) = A \]
   \[v(x, \omega) = (A/\rho c)(p_+e^{-j\omega x/c} - p_-e^{j\omega x/c})
   \]

2. Impose one boundary condition
   1. \( u(L, \omega) = 0 \): \( p_- = p_+e^{2j\omega L/c} \)
   2. ... or ...
   3. \( p(L, \omega) = 0 \): \( p_- = -p_+e^{2j\omega L/c} \)

3. Calculate impedance at the other boundary
   1. \( z(0, \omega) = p(0, \omega)/v(0, \omega) = (\rho c)(p_+ + p_-)/(p_+ - p_-) \)
   2. ... or ...
   3. \( z(0, \omega) = p(0, \omega)/u(0, \omega) = (\rho c/A)(p_+ + p_-)/(p_+ - p_-) \)

4. Result (for a uniform tube, using \( z = p/u \)):
   1. \( u(L, \omega) = 0 \): \( z(0, \omega) = j(\rho c/A)\cot(\omega L/c) \)
   2. ... or ...
   3. \( p(L, \omega) = 0 \): \( z(0, \omega) = -j(\rho c/A)\tan(\omega L/c) \)
Resonant Frequencies ("Formants") of a Nasal Consonant

1. Air Flow from Pharynx must equal Air Flow to Nose and Mouth
   \[ u_p + u_N + u_M = 0 \]

2. Therefore, the admittances sum to zero
   \[ \frac{1}{z_p} + \frac{1}{z_N} + \frac{1}{z_M} = 0 \]

3. Plug in the true admittances
   \[ -j \left( \frac{A_p}{\rho c} \tan(\omega L_p/c) \right) + j \left( \frac{A_N}{\rho c} \cot(\omega L_N/c) \right) - j \left( \frac{A_M}{\rho c} \tan(\omega L_M/c) \right) = 0 \]

4. For most resonant frequencies: Eq. (3) must be solved numerically on a computer (or graphically, as in Fujimura, JASA 1962)
First Nasal Formant

1. **The true resonance equation:**
   \[-jA_p \tan(\omega L_p/c) + jA_N \cot(\omega L_N/c) - jA_M \tan(\omega L_M/c) = 0\]

2. **Low-frequency approximation:**
   \[-A_p (\omega L_p/c) + A_N / (\omega L_N/c) - A_M (\omega L_M/c) = 0\]
   \[-(A_M L_M + A_p L_p) (\omega/c)^2 + A_N / L_N = 0\]
   \[\omega^2 = c^2 (A_N / L_N) / (A_M L_M + A_p L_p)\]

3. **Typical example, nasal F₁:**
   \[A_p L_p = 60\text{cm}^3, A_M L_M = 40\text{cm}^3,\]
   \[A_N / L_N = (3.5\text{cm}^2 / 9\text{cm}) = 2/5\text{cm}\]
   \[F_1 = (c/2\pi)(1/250)^{1/2} \approx 350\text{Hz}\]
Nasal Consonant Formants = Resonances of the Oral-Nasal-Pharyngeal Combined System

- First Nasal Resonance ≈ 350 Hz, with broad bandwidth $B \approx 300$ Hz
- Second Nasal Resonance ≈ 1400 Hz
- Third Nasal Resonance ≈ 1800 Hz
- Fourth Nasal Resonance ≈ 2400 Hz
Nasal Consonant Anti-Resonances

1. Anti-Resonances (zeros) occur because of energy lost into the side-branch.
2. Continuity of pressure at the juncture:
   \[ p_P(0, \omega) = p_N(0, \omega) = p_M(0, \omega) \]
3. If it turns out that \( p_N(0, \omega) = 0 \), that’s OK; that’s just part of the normal standing wave pattern in the nostrils.
4. If it turns out that \( p_M(0, \omega) = 0 \), that’s extra: it means that the side branch (into the mouth) is draining away all of the energy that would otherwise go out through the nostrils.
Nasal Consonant Anti-Resonances

1. Anti-Resonances (zeros) occur at any frequency such that, regardless of $u_M(0,\omega)$, $p_M(0,\omega)$ is required to be zero.

2. In other words: $z_M(0,\omega)=p_M/u_M=0$

3. Uniform tube: $z_M(0,\omega)=\left(\rho c/A_M\right)\cot(\omega L_M/c)$

4. Therefore $F_{zm}=(mc/2L_M)-(c/4L_M)$

5. Anti-resonances of the English nasal consonants:

   1. /m/: $L_M \approx 8$cm, $F_{zm} \approx 1100, 3300, 5500$Hz
   2. /n/: $L_M \approx 5$cm, $F_{zm} \approx 1700, 5100, 8500$Hz
   3. /ng/: $L_M \approx 2$cm, $F_{zm} \approx 4400$Hz, ...
Liquids

/l/:  
- Main airway: around the tongue (on both sides, so there may be zeros because of added transfer functions)
- Side branch: above the tongue
  - $L = 6\text{cm}$
  - $F_Z = \frac{35400}{4L} = 1500\text{Hz}$, between F2 and F3, pushes F2 down

/r/:  
- Main airway: as in a vowel
- Side branch: under the tongue
  - $L = 3.5\text{cm}$
  - $F_Z = 2500\text{Hz}$, between F3 and F4, pushes F3 down
Topic #3:
Events in the Release of a Stop (Plosive) Consonant
“Burst” = transient + frication (the part of the spectrogram whose transfer function has poles only at the front cavity resonance frequencies, not at the back cavity resonances).
Events in the Release of a Stop

Unaspirated (/b/)  
Transient  
Frication  
Aspiration  
Voicing  
Aspirated (/t/)
Pre-voicing during Closure

To make a voiced stop in most European languages:

- Tongue root is relaxed, allowing it to expand so that vocal folds can continue to vibrating for a little while after oral closure.

- Result is a low-frequency “voice bar” that may continue well into closure.

- In English, closure voicing is typical of read speech, but not casual speech.

“the bug”
Transient: The Release of Pressure

Initial Condition: Constriction of length $L_c$ Located at $x = \xi$:

$$A(\xi, t) = 0, \quad t < 0$$

Front Cavity: $p(x, t) = 0 = p_+(t - x/c) + p_-(t + x/c)$

Back Cavity: $p(x, t) = P_0$

Immediately After Release:

$$- \left. \frac{\partial p}{\partial x} \right|_{x=\xi} = \frac{P_{\text{back}} - P_{\text{front}}}{L_c} = \frac{P_0}{L_c}$$

Newton’s Law:

$$- \left. \frac{\partial p}{\partial x} \right|_{x=\xi} = \rho_0 \frac{\partial u}{\partial t}$$

$$u(\xi, t) = \frac{P_0 t}{\rho_0 L_c}$$

Solve for the Forward-Going Wave:

$$u(\xi, t) = \frac{1}{\rho_0 c} (p_+(t - \xi/c) - p_-(t + \xi/c))$$

For Small $t$, $p_-(t + \xi/c) = 0$

$$p_+(t - \xi/c) = \begin{cases} 
0 & t < 0 \\
\frac{P_0 t c}{L_c} & 0 < t < L_c/c \\
0 & t > L_c/c
\end{cases}$$
Transfer Function During Transient and Friction: Poles

Turbulence striking an obstacle makes noise

Front cavity resonance frequency:

\[ F_R = \frac{c}{4L_f} \]

Example: \[ FR = 2250 = \frac{c}{4L_f} \]

\[ L_f = 4\text{cm} \]

"the shutter"
Turbulence Striking an Obstacle Creates a Pressure Source at Position $\xi$:

$$P(\xi, j\Omega) = P_s(j\Omega)$$

Resulting Velocity Depends on Impedance:

$$U(\xi, j\Omega) = -\frac{P_s(j\Omega)}{Z_b(\xi, j\Omega)} = \frac{P_s(j\Omega)}{Z_f(L - \xi, j\Omega)}$$

where $Z_b(\xi, j\Omega)$ is impedance of cavity behind $\xi$, $Z_f(\xi, j\Omega)$ is impedance of cavity in front of $\xi$.

Cavity Behind the Source is Small

$$Z_b(\xi, j\Omega) = -\frac{j\rho c}{A_0} \cot(\Omega\xi/c) \approx -\frac{j\rho c^2}{\Omega A_0 \xi}$$

So $\Omega = 0$ is a Zero of the Transfer Function:

$$U(\xi, 0) = -\frac{P_s(0)}{Z_b(\xi, 0)} = 0$$
Transfer Function During Frication: An Important Zero
Transfer Function During Aspiration

Vowel Transfer Function

Zero in the Aspiration TF

Total Aspiration Transfer Function (Zero Reduces F1)

Frequency (Hz)
Formant Transitions: A Perceptual Study

Perception of Formant Transitions

Fig. 1. Synthetic spectrograms showing second-formant transitions that produce the voiced stops before various vowels.
Summary

- **International Phonetic Alphabet**
  - “Distinct” = distinguishes words
  - “Universal” = distinct any language → distinct in IPA

- **Sounds with a Side Branch**
  - Poles=poles of entire system (oral-nasal-pharyngeal)
  - Zeros=resonances of side branch (P=0 at juncture)

- **Events in the Release of a Syllable-Initial Stop Consonant**
  - Transient “pop,” triangular shape, about 0.5ms long
  - Friction = turbulence at the constriction, about 5ms
  - Aspiration = turbulence at glottis, 0-70ms
  - Formant Transitions may start during aspiration (in the case of an unvoiced stop release)