Homework 1

- Available on website and Canvas (we will release homeworks on website only moving forward).
  - Due on Monday 4/11 11:59pm Pacific
- Colab and written sections. Today’s lecture will help with phonetic transcription!
Week 1

- Course introduction
- Course Logistics
- Course topics overview
- Articulatory Phonetics
- ARPAbet transcription
Phonetics

• **ARPAbet**
  • An alphabet for transcribing American English phonetic sounds.

• **Articulatory Phonetics**
  • How speech sounds are made by articulators (moving organs) in mouth.

• **Acoustic Phonetics**
  • Acoustic properties of speech sounds
Phonetics

- Modern systems are less reliant on encoding phonetic domain knowledge directly.
- Basic understanding helps with describing and debugging spoken language systems
  - E.g. how does an accent change the sound of pronunciations?
- Phonetic categories derived from how humans produce speech
# International Phonetic Alphabet (IPA)

## CONSONANTS (PULMONIC)

<table>
<thead>
<tr>
<th></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>Plosive</td>
<td>p b</td>
<td>t d</td>
<td>t d</td>
<td>c j</td>
<td>k g</td>
<td>q g</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m nŋ</td>
<td>n</td>
<td>nŋ</td>
<td>nŋ</td>
<td>n</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>B</td>
<td>r</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap or Flap</td>
<td>v̑</td>
<td>ȓ</td>
<td>ɾ̑</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>f v̑</td>
<td>θ ō̃</td>
<td>s z̃</td>
<td>s z̃</td>
<td>ç j</td>
<td>x y</td>
<td>χ Ø̃</td>
<td>h i</td>
<td>h i</td>
<td></td>
</tr>
<tr>
<td>Lateral fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>u j̃</td>
<td>j l̃</td>
<td>j l̃</td>
<td>j̃</td>
<td>w̃</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

## CONSONANTS (NON-PULMONIC)

<table>
<thead>
<tr>
<th></th>
<th>Voiced implosives</th>
<th>Ejectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
<td>B</td>
<td>’</td>
</tr>
<tr>
<td>Dental</td>
<td>D</td>
<td>p’</td>
</tr>
<tr>
<td>(Post)alveolar</td>
<td>F</td>
<td>t’</td>
</tr>
<tr>
<td>Palatoalveolar</td>
<td>G</td>
<td>k’</td>
</tr>
<tr>
<td>Alveolar lateral</td>
<td>G’</td>
<td>s’</td>
</tr>
</tbody>
</table>

## VOWELS

Front | Central | Back
---|---------|-------
Close | i • y   | i • u |
Close-mid | ø        | ø     |
Open-mid | ø        | ø     |
Open    | ø        | ø     |

OTHER SYMBOLS

- Voiceless labial-velar fricative
- Alveolo-palatal fricatives

(Phonetics IPA with sounds)
Articulatory parameters for English consonants (in ARPAbet)

<table>
<thead>
<tr>
<th>PLACE OF ARTICULATION</th>
<th>bilabial</th>
<th>labio-dental</th>
<th>inter-dental</th>
<th>alveolar</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>p</td>
<td>b</td>
<td></td>
<td>t</td>
<td>d</td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td>fric.</td>
<td></td>
<td>f</td>
<td>v</td>
<td>th</td>
<td>dh</td>
<td>s</td>
<td>z</td>
</tr>
<tr>
<td>affric.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ch</td>
</tr>
<tr>
<td>nasal</td>
<td></td>
<td>m</td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approx</td>
<td></td>
<td>w</td>
<td></td>
<td></td>
<td>l/r</td>
<td>y</td>
<td></td>
</tr>
<tr>
<td>flap</td>
<td></td>
<td></td>
<td></td>
<td>dx</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table from Jennifer Venditti

**VOICING:**
- voiceless
- voiced
### ARPAbet Vowels

<table>
<thead>
<tr>
<th></th>
<th>b_d</th>
<th>ARPA</th>
<th></th>
<th>b_d</th>
<th>ARPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>bead</td>
<td>iy</td>
<td>9</td>
<td>bode</td>
<td>ow</td>
</tr>
<tr>
<td>2</td>
<td>bid</td>
<td>ih</td>
<td>10</td>
<td>booed</td>
<td>uw</td>
</tr>
<tr>
<td>3</td>
<td>bayed</td>
<td>ey</td>
<td>11</td>
<td>bud</td>
<td>ah</td>
</tr>
<tr>
<td>4</td>
<td>bed</td>
<td>eh</td>
<td>12</td>
<td>bird</td>
<td>er</td>
</tr>
<tr>
<td>5</td>
<td>bad</td>
<td>ae</td>
<td>13</td>
<td>bide</td>
<td>ay</td>
</tr>
<tr>
<td>6</td>
<td>bod(y)</td>
<td>aa</td>
<td>14</td>
<td>bowed</td>
<td>aw</td>
</tr>
<tr>
<td>7</td>
<td>bawd</td>
<td>ao</td>
<td>15</td>
<td>Boyd</td>
<td>oy</td>
</tr>
<tr>
<td>8</td>
<td>Budd(hist)</td>
<td>uh</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Many speakers pronounce Buddhist with the vowel uw as in booed, so for them [uh] is instead the vowel in “put” or “book”

[https://corpus.linguistics.berkeley.edu/acip/](https://corpus.linguistics.berkeley.edu/acip/)
Speech Production: Flow, Resonance, & Articulation

- **Flow**
  - We (normally) speak while breathing out. Respiration provides airflow. “Pulmonic egressive airstream”
  - Airstream sets vocal folds in motion. Vibration of vocal folds produces sounds. Sound is then modulated by:

- **Resonance:** Shape of vocal tract causing harmonics

- **Articulation:** Manipulation of airflow
  - Oral tract: uvula, soft palate (velum), hard palate, tongue, lips, teeth
  - Nasal tract

Text adopted from Sharon Rose
Sagittal section of the vocal tract (Techmer 1880)

- Nasal Cavity
- Pharynx
- Vocal Folds (within the Larynx)
- Trachea
- Lungs

Text copyright J. J. Ohala, Sept 2001, from Sharon Rose slide
USC’s SAIL Lab
Shri Narayanan
Tamil
Larynx and Vocal Folds

- The Larynx (voice box)
  - A structure made of cartilage and muscle
  - Located above the trachea (windpipe) and below the pharynx (throat)
  - Contains the vocal folds
  - (adjective for larynx: laryngeal)

- Vocal Folds (older term: vocal cords)
  - Two bands of muscle and tissue in the larynx
  - Can be set in motion to produce sound (voicing)
Voicing:

- Air comes up from lungs
- Forces its way through vocal cords, pushing open (2,3,4)
- This causes air pressure in glottis to fall, since:
  - when gas runs through constricted passage, its velocity increases (Venturi tube effect)
  - this increase in velocity results in a drop in pressure (Bernoulli principle)
- Because of drop in pressure, vocal cords snap together again (6-10)
- Single cycle: ~1/100 of a second.

Figure & text from John Coleman’s web site
Vocal Fold Vibration
Voicelessness

- When vocal cords are open, air passes through unobstructed
- Voiceless sounds: p/t/k/s/f/sh/th/ch
- If the air moves very quickly, the turbulence causes a different kind of phonation: whisper
Consonants and Vowels

- **Consonants**: phonetically, sounds with audible noise produced by a constriction
- **Vowels**: phonetically, sounds with no audible noise produced by a constriction

- (it’s more complicated than this, since we have to consider syllabic function, but this will do for now)
Place of Articulation

- Consonants are classified according to the location where the airflow is most constricted.
- This is called **place of articulation**
- Three major kinds of place articulation:
  - **Labial** (with lips)
  - **Coronal** (using tip or blade of tongue)
  - **Dorsal** (using back of tongue)
Places of articulation

- labial
- dental
- alveolar
- post-alveolar/palatal
- velar
- uvular
- pharyngeal
- laryngeal/glottal

Figure thanks to Jennifer Venditti
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

Thanks to Jong-bok Kim for this figure!
More on Manner of articulation of consonants

- **Fricatives**
  - Close approximation of two articulators, resulting in turbulent airflow between them, producing a hissing sound.
    - f, v, s, z, th, dh

- **Approximant**
  - Not quite-so-close approximation of two articulators, so no turbulence
    - y, r

- **Lateral approximant**
  - Obstruction of airstream along center of oral tract, with opening around sides of tongue.
    - l

Text from Ladefoged “A Course in Phonetics”
Tongue position for vowels
[iy] (bead) vs. [uw] (booed)

Figure from Jennifer Venditti, from a lecture given by Rochelle Newman
Articulatory parameters for English consonants (in ARPAbet)

<table>
<thead>
<tr>
<th>MANNER OF ARTICULATION</th>
<th>PLACE OF ARTICULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bilabial</td>
</tr>
<tr>
<td>stop</td>
<td>p</td>
</tr>
<tr>
<td>fric.</td>
<td>f</td>
</tr>
<tr>
<td>affric.</td>
<td>m</td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
</tr>
<tr>
<td>approx</td>
<td>w</td>
</tr>
<tr>
<td>flap</td>
<td>dx</td>
</tr>
</tbody>
</table>

VOICING: voiceless voiced

Table from Jennifer Venditti
The art of language invention

- Fun, informative book on phonetics and phonotactics across languages.

Great audio book!

Talk video
Phonetics

- ARPAbet
  - An alphabet for transcribing American English phonetic sounds.

- Articulatory Phonetics
  - How speech sounds are made by articulators (moving organs) in mouth.

- **Acoustic Phonetics**
  - Acoustic properties of speech sounds
Sound waves are longitudinal waves
particle displacement

longitudinal (left-right) particle displacement

pressure

Dan Russell Figure
Back to waves:
Fundamental frequency

- Waveform of the vowel [iy]

- Frequency: 10 repetitions / 0.03875 seconds = 258 Hz
- This is speed that vocal folds move, hence voicing
- Each peak corresponds to an opening of the vocal folds
- The low frequency of the complex wave is called the fundamental frequency of the wave or F0
She just had a baby

- Note that vowels all have regular amplitude peaks
- Stop consonant
  - Closure followed by release
  - Notice the silence followed by slight bursts of emphasis: very clear for [b] of “baby”
- Fricative: noisy. [sh] of “she” at beginning
Spectrogram: spectrum + time dimension
Source filter model of vowels

- Any body of air will vibrate in a way that depends on its size and shape.
- Vocal tract as "amplifier"; amplifies certain harmonics
- Formants are result of different shapes of vocal tract.
Source-filter model of speech production

Input → Filter → Output

Glottal spectrum → Vocal tract frequency response function

Source and filter are independent, so:
Different vowels can have same pitch
The same vowel can have different pitch

Figures and text from Ratree Wayland slide from his website
Resonances of the vocal tract

- The human vocal tract as an open tube

Figure from Ladefoged(1996) p 117
From Mark Liberman’s Web site
Defining Intonation

- “The use of suprasegmental phonetic features
  Suprasegmental = above & beyond the segment/phone
  - F0
  - Intensity (energy)
  - Duration
- to convey sentence-level pragmatic meanings”
  - I.e. meanings that apply to phrases or utterances as a whole, not lexical stress, not lexical tone.
Pitch track

Time (s)
Pitch is not Frequency

- Pitch is the mental sensation or perceptual correlated of F0
- Relationship between pitch and F0 is not linear;
  - human pitch perception is most accurate between 100Hz and 1000Hz.
    - Linear in this range
    - Logarithmic above 1000Hz
- Mel scale is one model of this F0-pitch mapping
  - A mel is a unit of pitch defined so that pairs of sounds which are perceptually equidistant in pitch are separated by an equal number of mels
  - Frequency in mels = 1127 ln (1 + f/700)
Plot of Intensity

is  it  a  long  movie?

Time (s)
Three aspects of prosody

- **Prominence**: some syllables/words are more prominent than others
- **Structure/boundaries**: sentences have prosodic structure
  - Some words group naturally together
  - Others have a noticeable break or disjuncture between them
- **Tune**: the intonational melody of an utterance.

From Ladd (1996)
Prosodic Boundaries

- I met Mary and Elena’s mother at the mall yesterday.
- French [bread and cheese]
- [French bread] and [cheese]
Appendix
The Speech Chain (Denes and Pinson)
ARPAbet

- http://www.stanford.edu/class/cs224s/arpabet.html

- The CMU Pronouncing Dictionary
  - http://www.speech.cs.cmu.edu/cgi-bin/cmudict

- International Phonetic Alphabet:
Vocal folds open during breathing

From Mark Liberman’s web site, from Ultimate Visual Dictionary
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

From Ladefoged “A Course in Phonetics”
She came back and started again

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

From Ladefoged “A Course in Phonetics”
More on manner of articulation of consonants

- Tap or flap
  - Tongue makes a single tap against the alveolar ridge
    - $dx$ in “butter”

- Affricate
  - Stop immediately followed by a fricative
    - $ch$, $jh$
Vowels

Fig. from Eric Keller
The oral cavity amplifies some harmonics

$F_0 = 150 \text{ Hz}$

10th harmonic = 1,500 Hz