Sounds
Syllabus & Intros

Name
Where you’re from
Year in school
Have you programmed?
What do you want to get out of this class?
Fun fact
International Phonetic Alphabet

Linguistics
What is IPA?

A set of symbols designed by the International Phonetic Association, used for the transcription of speech sounds in language-related fields.
Why do we need IPA?

It is the closest thing to a universal phonetic writing system.

It is orthographically transparent (one symbol for each sound)
Combinations of letters, diacritics, brackets

**Broad transcription** - conventional sounds of the language (or phonemes)

**Narrow transcription** - actual sounds (or phones)
Four Types of Symbols

1. Same as English
2. Same as English, but more specific
3. Looks like English, but different
4. Not in Modern English
Same as English

14 letters:
/b, d, f, h, k, l, m, n, p, s, t, v, w, z/
Same, but more specific

//i/ is the i in machine, not in pit or idol
//ɪ/ is the i in pit
//u/ is the u in hula, not in union or up
//g/ is the g in gift, not in gin
Looks like English, but not /j/ ("yod") is the y in yes
/æ/ ("ash") is the vowel in hat
/ɑ/ ("script a") is the first vowel in father
/ɛ/ ("epsilon") is the vowel in get
/ɔ/ ("open o") is the vowel in law
/ʊ/ ("upsilon") is the vowel in *book*
/
/ʌ/ ("caret") is the vowel in *up*
/ə/ ("schwa") is the first vowel in *above*

Don’t worry about the distinction between /ʌ/ and /ə/ for now
/ŋ/ ("engma") is the last sound in song
/θ/ ("theta") is the first sound in thin
/ð/ ("eth") is the first sound in then
/ʃ/ ("esh") is the first sound in she
/ʒ/ (“ezh”) is the second consonant in vision
/ɹ/ is the American r sound
/ʔ/ (“glottal stop”) is the sound in the middle of uh-oh
## 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><strong>Plosive</strong></td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td></td>
<td>t</td>
<td>q</td>
<td>c</td>
<td>j</td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td><strong>Nasal</strong></td>
<td>m</td>
<td>nj</td>
<td>n</td>
<td>η</td>
<td>η</td>
<td>η</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trill</strong></td>
<td>B</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tap or Flap</strong></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td>φ</td>
<td>β</td>
<td>f</td>
<td>v</td>
<td>θ</td>
<td>ð</td>
<td>s</td>
<td>z</td>
<td>s</td>
<td>z</td>
<td>j</td>
</tr>
<tr>
<td><strong>Lateral fricative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
<td>b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Approximant</strong></td>
<td></td>
<td>v</td>
<td></td>
<td>j</td>
<td></td>
<td></td>
<td>l</td>
<td>j</td>
<td></td>
<td>w</td>
<td></td>
</tr>
<tr>
<td><strong>Lateral approximant</strong></td>
<td></td>
<td>l</td>
<td></td>
<td>l</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 areas denote articulations judged impossible.
The Speech Mechanism
<table>
<thead>
<tr>
<th>Vowels</th>
<th>Front</th>
<th>Near front</th>
<th>Central</th>
<th>Near back</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>i</td>
<td>y</td>
<td>i</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>Near close</td>
<td>I · Y</td>
<td>· u</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Close mid</td>
<td>e</td>
<td>ø</td>
<td>e</td>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>ø</td>
<td>θ</td>
<td>y</td>
<td>o</td>
</tr>
<tr>
<td>Open mid</td>
<td>ε</td>
<td>øε</td>
<td>ða</td>
<td>ðæ</td>
<td>ða</td>
</tr>
<tr>
<td>Open</td>
<td>a</td>
<td>æ</td>
<td>æ−a</td>
<td>æ−a</td>
<td>æ−a</td>
</tr>
</tbody>
</table>

Vowels at right & left of bullets are rounded & unrounded.
Transcription Practice

cat
sing
build
thin
bees
think
this
tacks
strengths
enough
Transcription Practice

cat /kæt/
sing /sɪŋ/
build /bɪld/
thin /θɪn/
bees /biz/
think /θɪŋk/
this /ðɪs/
tacks /tækss/
strengths /streŋθs/
enough /ɪnʌf/
Lack of Invariance

- Adult males
- Adult females
- Young males
- Young females
Lack of Invariance
Lack of Invariance

Figure 1. Spectrographic patterns for the two two-formant synthetic syllables /di/ and /du/. Note the difference in formant transitions, marked by the dotted ovals.
Lack of Invariance

There is no 1:1 correlation between the physical properties of speech sounds and the perception of speech sounds.
Adaptation

Given that the speech signal is noisy and there is a lack of invariance, we have to adapt to the signal.

\[ P(\text{Message} | \text{Signal}) \propto P(\text{Signal} | \text{Message}) P(\text{Message}) \]

New beliefs                Likelihood       Prior Beliefs
PROLOG BREAK
Speech Recognition

- small vocab/many users vs. large vocab/few users
- discrete vs. continuous speech
Speech Recognition

How it works: Set-up (enrollment)

● set of words/phrases new users say before using the system
● how are these selected?
Speech Recognition

What information can a program use?
- Waveform of speech
- Decode sounds using formant information
Speech Recognition

Ghonim et al., 2007, 2013

https://www.llas.ac.uk/materialsbank(mb081/page_10.htm
Speech Recognition

Figure 5: Three long vowels in an /h_d/ context.

Speech Recognition

Lennes, M., 2005
Try Out Praat on Your Own

- Partner up
- Say vowels into Praat
- Compare yours to your partner’s
  - do they seem similar?
  - how are they different?
  - what differences do you see between different vowels for the same person?
Clustering
Machine Learning

An area at the intersection of mathematics, computer science, engineering and data science that focuses predominantly on one question:

Given some data, how can I learn a model to predict X.
k-Means Classifier

Initialize:
1. Pick a number $k$ and say that there are that many classes of thing.
2. Randomly assign all the data points to a class.
k-Means Classifier

Search:
1. Calculate the mean of each class.
2. Calculate the probability of each data point is from that class.
3. Reassign the data points according to which class is most likely for that data point.
4. Rinse and repeat until convergence.
Quick and Dirty k-Means

https://mollicaf.shinyapps.io/k-Means/