SOUNDS AND LANGUAGE

Phonology, phonetics, & spectrograms
About this class

1. **Phonology**: vowels and consonants
2. How sound works
3. Spectrograms! A tool for looking at sounds
4. **Phonetics**: vowels and consonants
About this class

1. Phonology: vowels and consonants
2. How sound works
3. Spectrograms! A tool for looking at sounds
4. Phonetics: vowels and consonants

Note:

- Phonics: A method for teaching people how to read
- Phonology: The systematic categorization of sounds.
- Phonetics: The study of how we make and hear sounds.
About this class

1. **Phonology**: vowels and consonants
2. How sound works
3. Spectrograms! A tool for looking at sounds
4. **Phonetics**: vowels and consonants

**Note:**

- **Phonics**: A method for teaching people how to read
- **Phonology**: The systematic categorization of sounds.
- **Phonetics**: The study of how we make and hear sounds.
PHONOLOGY!
A phoneme is a sound used by languages in their words.
- Hawaiian: 10 vowels, 8 consonants
  - (a, e, i, o, u, ā, ē, ī, ō, ū)
  - (h, k, l, m, n, p, w, `)
- English: 14 vowels, 24 consonants

The International Phonetic Alphabet (IPA) is what we use to write down phonemes.

Let’s look in more detail at:
- 1: Vowels
- 2: Consonants
PHONOLOGY: VOWELS

- **Vowels**
- Vowel sounds produced in *voice box* or *larynx* (try it!)
- Different sounds made by different tongue positions (try it!)

- /u/ back closed
- /α/ back open
- /i/ front closed
PHONOLOGY: VOWELS

- **Vowels**
- Vowel sounds produced in *voice box* or *larynx* (try it!)
- Different sounds made by different tongue positions (try it!)
- Human mouth: specialized for speech?
PHONOLOGY: VOWELS

- **Vowels**
- Vowel sounds produced in *voice box* or *larynx* (try it!)
- Different sounds made by different tongue positions (try it!)
- Human mouth: specialized for speech?
- **Axes**: front-to-back, open-to-closed, unrounded-to-rounded

(of English)
PHONOLOGY: VOWELS

- **Vowels**
- Vowel sounds produced in *voice box* or *larynx* (try it!)
- Different sounds made by different tongue positions (try it!)
- Human mouth: specialized for speech?
- Axes: front-to-back, open-to-closed, unrounded-to-rounded
- **Diphthongs**: when you have two or more vowels combined together into a sound.
  - “oy”
  - “ow”
  - “ey”
  - “ai”
Consonants

What’s the difference between /s/ and /z/?
Consonants

What’s the difference between /s/ and /z/? Voicing.

- /s/ is voiceless. /z/ is voiced. Try it!
- Consider the words: author, father. Which /th/ is voiced?
Consonants

What’s the difference between /s/ and /z/? Voicing.

- /s/ is voiceless. /z/ is voiced.
- Consider the words: author, father. Which /th/ is voiced?
  - /th/ as in author is voiceless. IPA uses /θ/. Old English uses /þ/.
  - /th/ as in father is voiced. IPA uses /ð/, and so does Old English.
**PHONOLOGY: CONSONANTS**

- **Consonants**
- **What’s the difference between /s/ and /z/? Voicing.**
  - /s/ is *voiceless*. /z/ is *voiced*.
  - Consider the words: *author, father*. Which /th/ is voiced?
    - /th/ as in *author* is *voiceless*. IPA uses /θ/. Old English uses /þ/.
    - /th/ as in *father* is *voiced*. IPA uses /ð/, and so does Old English.

---

*Beowulf in Old English*

_Hwæt! Wé Gárdena in gêardagum_  
þéodcyninga þrym gefrúnon·  
hú ðá æþelingas ellen fremedon.

_Oft Scyld Scéfing sceapena þréatum_  
monegum maégbum meodosetla oftéah  
egsode Eorle yððan aérest wearð  
féasceaf funden hé þæs frófre gebád  
wéox under wolcnum · weordmyndum þáh  
oð þæt him aéghwylc þára ymsittendra  
ofer hronráde hýran scolde,  
gomban gyldan · þæt wæs gód cyning!_
Consonants

What’s the difference between /s/ and /z/? Voicing.

- /s/ is voiceless. /z/ is voiced.
- Consider the words: author, father. Which /th/ is voiced?
  - /th/ as in author is voiceless. IPA uses /θ/. Old English uses /þ/.
  - /th/ as in father is voiced. IPA uses /ð/, and so does Old English.

What’s the difference between /s/ and /θ/?
Consonants

What’s the difference between /s/ and /z/? Voicing.

- /s/ is voiceless. /z/ is voiced.
- Consider the words: author, father. Which /th/ is voiced?
  - /th/ as in author is voiceless. IPA uses /θ/. Old English uses /þ/.
  - /th/ as in father is voiced. IPA uses /ð/, and so does Old English.

What’s the difference between /s/ and /θ/?

- Place of articulation. /s/ is alveolar. /θ/ is dental.

What’s the difference between /s/ and /t/?
Consonants

What’s the difference between /s/ and /z/? Voicing.

- /s/ is voiceless. /z/ is voiced.
- Consider the words: author, father. Which /th/ is voiced?
  - /th/ as in author is voiceless. IPA uses /θ/. Old English uses /þ/.
  - /th/ as in father is voiced. IPA uses /ð/, and so does Old English.

What’s the difference between /s/ and /θ/?

- Place of articulation. /s/ is alveolar. /θ/ is dental.

What’s the difference between /s/ and /t/?

- Frication. /s/ is a fricative. /t/ is a stop.

What in the world does “voiced alveolar fricative” mean?
Consonants

What’s the difference between /s/ and /z/? **Voicing.**
- /s/ is *voiceless*. /z/ is *voiced*.
- Consider the words: *author, father*. Which /th/ is voiced?
  - /th/ as in *author* is *voiceless*. IPA uses /\θ/. Old English uses /\p/.
  - /th/ as in *father* is *voiced*. IPA uses /\ð/, and so does Old English.

What’s the difference between /s/ and /\θ/?
- **Place of articulation.** /s/ is *alveolar*. /\θ/ is *dental*.

What’s the difference between /s/ and /t/?
- **Frication.** /s/ is a *fricative*. /t/ is a *stop*.

What in the world does “voiced alveolar fricative” mean?
- /z/.

Let’s fill out the rest of the chart!
## PHONOLOGY: CONSONANTS

- **Consonants**
- **Some of the consonants of English**

### CONSONANTS (PULMONIC)

<table>
<thead>
<tr>
<th>Consonant Type</th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Epiglottal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal</td>
<td>m</td>
<td>m</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>N</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Plosive</td>
<td>p b</td>
<td>d</td>
<td>t d</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>
</tr>
<tr>
<td>Fricative</td>
<td>f v</td>
<td>s z</td>
<td>s z</td>
<td>s z</td>
<td>ç j</td>
<td>x y</td>
<td>x k</td>
<td>h</td>
<td>h</td>
<td>H</td>
<td>H</td>
<td>Z</td>
</tr>
<tr>
<td>Approximant</td>
<td>v</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>J</td>
<td>J</td>
<td>J</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>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Tap, Flap</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
<td>n f</td>
</tr>
<tr>
<td>Lateral fricative</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
<td>l k</td>
</tr>
<tr>
<td>Lateral approximant</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
<td>j l</td>
</tr>
<tr>
<td>Lateral flap</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
</tbody>
</table>

Where symbols appear in pairs, the one to the right represents a modally voiced consonant, except for murmured ŋ. Shaded areas denote articulations judged to be impossible. Light grey letters are unofficial extensions of the IPA.
SOUND!
HOW SOUND WORKS
“YOUR EARS CAN DO MATH”
“YOUR EARS CAN DO MATH”

Cochlea

(flexible) Low pitch: 20 Hz

Neurons

Vibration

(stiff) High pitch: 20,000 Hz
YOUR EARS CAN DO MATH

sounds

calculating device

frequencies (neurons)
YOUR EARS CAN DO MATH

Fourier transform!

sounds

calculating device

frequencies (graph)
YOUR EARS CAN DO MATH

Fourier transform!
SOUND: NOT SO SIMPLE

- Frequency
  - 5000 Hz
  - 440 Hz
  - 0 Hz
Overtones determine timbre, which is like texture for sound.
SOUNDS CHANGE OVER TIME

Fourier transform!

Spectrogram!
SOUNDS CHANGE OVER TIME

Fourier transform!

Spectrogram!

Time

Frequencies

Visible part 2.000000 seconds
SOUNDS CHANGE OVER TIME

Spectrogram

5000 Hz

0 Hz

frequency
time
SPECTROGRAMS!
NARROWBAND VERSUS BROADBAND

- Determined by mathematical parameters.

**Narrowband**: This is what we’ve seen so far.
Better frequency resolution.
Worse time resolution.

**Broadband**: Almost everything else for the rest of the class.
Better time resolution.
Worse frequency resolution.
A SMALL PUZZLE

- Two _______band spectrograms. Which is from a male talker, which is from a female talker?
Two _______band spectrograms. Which is from a male talker, which is from a female talker?
Two narrowband spectrograms. Which is from a male talker, which is from a female talker?
Two narrowband spectrograms. Which is from a male talker, which is from a female talker?
A SMALL PUZZLE

- Two narrowband spectrograms.
- What do these look like in broadband instead? This:
Let’s look at some spectrograms.
VOWELS!
Fig. 7.5: The position of the vocal organs (based on data from X-ray photographs of the author) and the spectra of the vowel sounds in the middle of the words head, bad, head, bad, had, moved, head, rode'd, in the author's speech.
FORMANTS: YOU TRY!

/a/ “ahh”
/i/ “eeh”
/u/ “ooh”
Spectrograms for five words are given in scrambled order here. Match them!

- spooky
- maki
- kiwi
- pie
CONSONANTS
# Consonants

- Some of the consonants of English

## CONSONANTS (PULMONIC)

<table>
<thead>
<tr>
<th>Consonant Type</th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Epiglottal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>n</td>
<td>η</td>
<td>η</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<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>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>φ β</td>
<td>θ ð s z</td>
<td>θ ð s z</td>
<td>s z</td>
<td>ç j</td>
<td>x y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>v</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>w l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>β</td>
<td>r</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap, Flap</td>
<td>v</td>
<td>θ</td>
<td>θ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral fricative</td>
<td>l h</td>
<td>l</td>
<td>l</td>
<td></td>
<td></td>
<td></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>
<td></td>
</tr>
<tr>
<td>Lateral flap</td>
<td></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>

Where symbols appear in pairs, the one to the right represents a modally voiced consonant, except for murmured h. Shaded areas denote articulations judged to be impossible. Light grey letters are unofficial extensions of the IPA.
SPEECH SYNTHESIS
Speech recognition is **hard**!

**Segmentation problem**
- How many words are in that → spectrogram?
- Where does one word end and the other begin?
Speech recognition is *hard*!

**Segmentation problem**
- How many words are in that spectrogram?
- Where does one word end and the other begin?

**Variation problem**
- Different speech from different people varies!
- We’re good at normalizing for what one person’s voice sounds like, but computers aren’t.
Speech recognition is hard!

Segmentation problem
- How many words are in that spectrogram?
- Where does one word end and the other begin?

Variation problem
- Different speech from different people varies!
- We’re good at normalizing for what one person’s voice sounds like, but computers aren’t.

Syllable stress and tones

诗 shī 石 shí 始 shǐ 室 shì
Speech recognition is hard!

Segmentation problem
- How many words are in that spectrogram?
- Where does one word end and the other begin?

Variation problem
- Different speech from different people varies!
- We’re good at normalizing for what one person’s voice sounds like, but computers aren’t.

Syllable stress and tones

But we’re working on it!
- (Siri, anyone?)

诗 shī 石 shí 始 shǐ 室 shì
If you think that this is all really cool:

- **Download Praat** to your computer at home and have fun using it to experiment! (It’s *really fun.*) [http://www.fon.hum.uva.nl/praat/](http://www.fon.hum.uva.nl/praat/)

- **Download RTgram**, which can make spectrograms in real-time: [http://www.phon.ucl.ac.uk/resource/sfs/rtgram/](http://www.phon.ucl.ac.uk/resource/sfs/rtgram/)

- **Come to my Linguistics Problem Solving walk-in activity!**
  Tomorrow (Sunday), 1:30pm-3:30pm, in Lobby 13.

- **Consider trying the North American Computational Linguistics Olympiad (NACLO)!** [http://naclo.cs.cmu.edu/](http://naclo.cs.cmu.edu/)
  - The first round is **January 30, 2014**.
  - You can sign up to take it at MIT or at many other universities, or ask a high school teacher to proctor the exam.

- **Browse Wikipedia to learn more about phonology!**