IAP Course: Technologies for Speech and Language Processing

Tuka Al-Hanai and Mandy Korpusik

1/8/16
Meet the Instructors

Tuka Al-Hanai  

Mandy Korpusik
Outline

I. Course overview
II. Today: acoustic-phonetics
III. Spectrograms
IV. Speech sounds
V. Reading spectrograms
VI. In-class activity: make a spectrogram
Outline

I. Course overview
II. Today: acoustic-phonetics
III. Spectrograms
IV. Speech sounds
V. Reading spectrograms
VI. In-class activity: make a spectrogram
Syllabus

1. Acoustic phonetics
2. Language modeling
3. Acoustic modeling
4. Building a recognizer
5. Bringing speech technologies to the world
Speech Recognition Overview

**Lexicon**: vocabulary of words in a language
Outline

I. Course overview
II. Today: acoustic-phonetics
III. Spectrograms
IV. Speech sounds
V. Reading spectrograms
VI. In-class activity: make a spectrogram
Today: Acoustic-phonetics

Intended learning objectives
Students will be able to:
• Define acoustic phonetics, spectrogram, and phoneme
• Name types of speech sounds
• Identify vowels and consonants in spectrograms
• Visualize sounds as a spectrogram using Wavesurfer
Definitions

• **Phonetics**: branch of linguistics that studies *sounds* of human *speech*

• **Acoustic phonetics**: subfield of phonetics that focuses on *acoustic* characteristics of speech, including *frequency*, intensity, and duration
Outline

I. Course overview
II. Today: acoustic-phonetics
III. Spectrograms
IV. Speech sounds
V. Reading spectrograms
VI. In-class activity: make a spectrogram
What is a spectrogram?

- **Spectrogram**: a visual representation of the frequencies in a speech sound
- Time on horizontal axis, frequency on vertical, darkness indicates energy
- Made using FFT (Fast Fourier Transform)
Spectrogram
Spectrogram
Outline

I. Course overview
II. Today: acoustic-phonetics
III. Spectrograms
IV. Speech sounds
V. Reading spectrograms
VI. In-class activity: make a spectrogram
Speech Sounds

• ~40 speech sounds (phonemes) in American English
• Written using IPA (International Phonetic Alphabet) symbols
• Vowels and consonants
• Consonants can be broken down into several categories: fricatives, stops, nasals, semivowels, and affricates
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Sound</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>/iː/</td>
<td>iy</td>
<td>beat</td>
</tr>
<tr>
<td>/ɪ/</td>
<td>ih</td>
<td>bit</td>
</tr>
<tr>
<td>/eɪ/</td>
<td>ey</td>
<td>bait</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>eh</td>
<td>bet</td>
</tr>
<tr>
<td>/æ/</td>
<td>ae</td>
<td>bat</td>
</tr>
<tr>
<td>/ɑː/</td>
<td>aa</td>
<td>Bob</td>
</tr>
<tr>
<td>/ɔː/</td>
<td>ao</td>
<td>bought</td>
</tr>
<tr>
<td>/ʌ/</td>
<td>ah</td>
<td>but</td>
</tr>
<tr>
<td>/ɔʊ/</td>
<td>ow</td>
<td>boat</td>
</tr>
<tr>
<td>/u/</td>
<td>uh</td>
<td>book</td>
</tr>
<tr>
<td>/uː/</td>
<td>uw</td>
<td>boot</td>
</tr>
<tr>
<td>/ɜː/</td>
<td>er</td>
<td>Bert</td>
</tr>
<tr>
<td>/ɔɪ/</td>
<td>ay</td>
<td>bite</td>
</tr>
<tr>
<td>/ɔː/</td>
<td>ow</td>
<td>bout</td>
</tr>
<tr>
<td>/ʊ/</td>
<td>uh</td>
<td>book</td>
</tr>
<tr>
<td>[ə]</td>
<td>ax</td>
<td>about</td>
</tr>
<tr>
<td>[ɪ]</td>
<td>ix</td>
<td>roses</td>
</tr>
<tr>
<td>[ɑː]</td>
<td>axr</td>
<td>butter</td>
</tr>
</tbody>
</table>
### Consonants

- **8 fricatives**

<table>
<thead>
<tr>
<th>Unvoiced</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>/f/ f fee</td>
<td>/v/ v v</td>
</tr>
<tr>
<td>/θ/ th thief</td>
<td>/ð/ dh thee</td>
</tr>
<tr>
<td>/s/ s see</td>
<td>/z/ z z</td>
</tr>
<tr>
<td>/š/ sh she</td>
<td>/ʒ/ zh Gigi</td>
</tr>
</tbody>
</table>

- **6 stops**

<table>
<thead>
<tr>
<th>Unvoiced</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/ p pot</td>
<td>/b/ b bought</td>
</tr>
<tr>
<td>/t/ t tot</td>
<td>/d/ d dot</td>
</tr>
<tr>
<td>/k/ k cot</td>
<td>/g/ g got</td>
</tr>
</tbody>
</table>
Consonants

- 3 nasals
  
  | /m/ | m  | me |
  | /n/ | n  | knee |
  | /ŋ/ | ng | sing |

- 4 semivowels
  
  | /w/ | w  | wet |
  | /y/ | y  | yet |
  | /r/ | r  | red |
  | /l/ | l  | let |
Outline

I. Course overview
II. Today: acoustic-phonetics
III. Spectrograms
IV. Speech sounds
V. Reading spectrograms
VI. In-class activity: make a spectrogram
Victor Zue
How to Read a Spectrogram

- **Formants**: frequencies with high energy used to identify vowels

- **Vowels**: high energy, no constriction, formants at lower frequencies

- **Consonants**: less energy due to constriction, formants at higher frequencies
Happy Little Vowel Chart

F₂ Increases

F₁ Increases

HIGH MID LOW

FRONT BACK
boot
/but/
thief
/θiːf/
see
/siː/
sinner
/sinə/
Outline

I. Course overview
II. Today: acoustic-phonetics
III. Spectrograms
IV. Speech sounds
V. Reading spectrograms
VI. In-class activities
Make a Spectrogram!

1. Download Wavesurfer:  
   http://www.speech.kth.se/wavesurfer/

2. Record yourself, and inspect the spectrogram!

3. Right-click -> Create pane -> Spectrogram (or Waveform)
Sneak Preview: Language Modeling

• Which is more likely?
  – They’re interested in the class.
  – Their interested in the class.
Give us your feedback!