Language & Mind
HONR 218L

Class #5
Sound Categories

Sound Production

How you look to a phonetician

Tongue
Palate
Velum
Tongue (vocal folds)
Lips, teeth etc.

Nasal Cavity
Oral Cavity

Frequency - Tones

Frequency - Tones
Frequency - Vowels

- Vowels combine acoustic energy at a number of different frequencies
- Different vowels ([a], [i], [u] etc.) contain acoustic energy at different frequencies
- Listeners must perform a ‘frequency analysis’ of vowels in order to identify them (Fourier Analysis)
Frequency - Female Vowels

Forget Spelling!

Sounds ≠ Spelling

One Sound - Many Characters

<table>
<thead>
<tr>
<th>he</th>
<th>see</th>
<th>caesar</th>
<th>believe</th>
<th>sea</th>
<th>people</th>
<th>infinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>eas</td>
<td>aegba</td>
<td>oe</td>
<td>iea</td>
<td>machine</td>
<td>seize</td>
</tr>
</tbody>
</table>

IPA: [i]

One Sound - Many Characters

<table>
<thead>
<tr>
<th>too</th>
<th>oo</th>
<th>threw</th>
<th>ew</th>
</tr>
</thead>
<tbody>
<tr>
<td>to</td>
<td>o</td>
<td>lieu</td>
<td>ieu</td>
</tr>
<tr>
<td>clue</td>
<td>ue</td>
<td>shoe</td>
<td>oe</td>
</tr>
<tr>
<td>through</td>
<td>ough</td>
<td>two</td>
<td>wo</td>
</tr>
</tbody>
</table>

IPA: [u]

One Character - Many Sounds

<table>
<thead>
<tr>
<th>dame</th>
<th>dad</th>
<th>father</th>
<th>call</th>
<th>village</th>
<th>many</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>æ</td>
<td>a</td>
<td>o</td>
<td>æ, i</td>
<td>ø</td>
</tr>
</tbody>
</table>

One Sound - Multiple Letters

<table>
<thead>
<tr>
<th>shoot</th>
<th>either</th>
<th>character</th>
<th>deal</th>
<th>Thomas</th>
<th>physics</th>
<th>rough</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>ð</td>
<td>k</td>
<td>i</td>
<td>t</td>
<td>f</td>
<td>f</td>
</tr>
</tbody>
</table>
One Letter - 0, 1, 2 Sounds

- mnemonic
- psychology
- resign
- ghost
- island
- whole
- debt

Differences across Languages

- English: judge, juvenile, Jesus
- Spanish: jugar, Jesus
- German: Jugend, jubeln, Jesus
- French: Jean, j’accuse, jambon

Timing - Voicing

Voice Onset Time (VOT)

60 msec
Speech is continuous

• No boundaries between sounds/words
• Continuous movement of articulators between sounds
• Continuous variation in acoustic signal

• ARTICULATION AND ACOUSTICS OF SPEECH DO NOT CONTAIN NEATLY ORGANIZED CHUNKS

Speech Perception is Discontinuous

Voice Onset Time (VOT)

Perceiving VOT

‘Categorical Perception’

Discrimination

Categories differ across languages

A More Systematic Test

Within-Category Discrimination is Hard
Cross-language Differences

English vs. Hindi
alveolar [d] vs. [D]
retroflex

Describing Speech Sounds

• Is the air-flow blocked?
  vowel vs. consonant
• What are the vocal folds doing?
  voiced vs. voiceless
• Where/how is the air flowing?
  nasal/oral, stop, fricative, liquid etc.
• Where is the air-flow blocked?
  labial, alveolar, palatal, velar etc.

Voiced & voiceless consonants

• Consonants either voiced or voiceless.
• English pairs:
  – b/p
  – v/f
  – d/t
  – z/s
  – ð/θ
Your vocal tract again.

Where does the air go?

Block it at the velum.

Now raise the velum to block the air....

Quickly drop your tongue again....

So so far we have:

Nasal stops:
[m] [n] [ŋ]

Non-nasal (oral) stops:
[b] [p] [d] [t] [g] [k]
Where can you stop the airstream?

(bi)labial
[b] [p] [m]

labiodental
[v] [f]

interdental
[s] [θ]

alveolar
[d] [t] [n] [s] [z] [l] [r]

palatal
[ʃ] [ʃ] [ʃ] [θ]

velar
[ɡ] [k] [ŋ]
Where can you stop the airstream?

Features

• Ways of describing sounds
e.g., [t] = voiceless, alveolar, stop
• Stronger claim: features are the smallest building blocks of language, used to store sounds in the mind
• Atoms of Speech

Roman Jakobson, 1896-1982

Manner

• Stops
[p] [k] [t] [d] [k] [g]...
• Fricatives
[f] [v] [θ] [ð] [s] [z]
• Approximants/Glides
[w] [j]
• Liquids
[r] [l]

Voiced & voiceless consonants

• Consonants either voiced or voiceless.
• English pairs (voiced-left; voiceless-right):
  – b/p
  – v/f
  – d/t
  – z/s
  – ð/θ
  – etc.

Fricatives & Affricates

• Palatal sounds [ʒ] [ʃ] [ʃ] [ʒ]
• Palatal Fricatives - [ʒ] [ʃ]
  [note: according to IPA chart these are strictly “post-alveolar”]
• Affricates - combination of stop + fricative - [ʃ] [ʃ], as in judge, church
Describing Consonants

- What are the vocal folds doing? *voiced* vs. *voiceless*
- Where/how is the air flowing? *nasal/oral, stop, fricative, liquid* etc.
- Where is the air-flow blocked? *labial, alveolar, palatal, velar* etc.

Features

- Prediction: by combining a small number of *atomic* features, it should be possible to create a larger number of speech sounds
- Goal: a set of universal features should make it possible to describe the speech sounds of all the languages of the world
- Different languages choose different feature combinations

- Ways of describing sounds e.g., [t] = voiceless, alveolar, stop
- Stronger claim: features are the *smallest building blocks of language*, used to store sounds in the mind
- *Atoms of Speech*

Roman Jakobson, 1896-1982

<table>
<thead>
<tr>
<th>Features</th>
<th>Bilingual</th>
<th>Labial</th>
<th>Inter-dental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral stop</td>
<td>p b</td>
<td>t d</td>
<td>k g</td>
<td>(h)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal stop</td>
<td>m n</td>
<td>n n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f v j č</td>
<td>f v j č</td>
<td></td>
<td>(h)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>g h k j</td>
<td>g h k j</td>
<td></td>
<td>(h)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>l r m l</td>
<td>l r m l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Cross-language Variation

- Another type of cross-language variation
  - Some places of articulation not used in all languages (e.g., dental, retroflex, palatal, pharyngeal stops not in English)
  - Some manners of articulation used only rarely (e.g., clicks in southern African and in Australian languages; aspirated vs. unaspirated stops in English vs. Spanish/Russian)

What can you do to alter the shape of your vocal tract?
So. You can:

- Raise or lower your tongue
- Advance or retract your tongue
- Round or spread your lips
- Tense or not tense your mouth

So what vowels do you have?

- \( \text{æ} \) “sheep, sleep”
- \( \text{i} \) “ship, slip”
So what vowels do you have?

- i
- e
  - “laid, spade, trade”

So what vowels do you have?

- i
- e
  - “led, sped, tread”

So what vowels do you have?

- i
- e
  - “bat, lad”

So what vowels do you have?

- i
- e
  - “Luke, who’d, suit”

So what vowels do you have?

- i
- e
  - “look, hood, soot”

So what vowels do you have?

- i
- e
  - “coat, wrote, hoed”
So what vowels do you have?

• coat, wrote, hoed
• caught, wrought, hawed

So what vowels do you have?

• bah, father, cot, Don

So what vowels do you have?

• but, putt, rut
• metallic, Texas

So what vowels do you have?

• side, my, kind

Some dialectal differences

• caught/cot [Mid back lax vowel and mid back tense vowel]: many American speakers do not have both of these.
• pot/father: some British and (fewer) American dialects have different vowels in these words (“pot” has a low back rounded vowel [ɒ]).

Diphthongs:

“side, my, kind”
More dialect differences

- [aj] in Standard US/UK English
  - \( \rightarrow \) long [a] in south eastern US
  - \( \rightarrow \) [i:] in central UK and Australia

- [aw] in Standard US/UK English
  - \( \rightarrow \) [aw] in Canada

Cross-language Differences

- Feature combinations
  - English: back vowels are rounded, others are not
  - German/French have high front rounded vowel [y]
  - Russian has high back unrounded vowel
  - Many languages don’t make the tense/lax distinction found in English (e.g., Italian [i])
  - Many languages distinguish short and long vowels, unlike English (e.g., Japanese)

Cross-language Differences

Speech Production - Summary

- Airflow set in vibration by vocal folds
  - Airflow modified by vocal tract
- Vowels: shaping of oral cavity
- Consonants: narrowing or blocking of oral/nasal cavity
- Different languages choose different selections of articulatory gestures
The Next Step...

- All languages use a similar ‘toolkit’ to produce speech sounds, but vary in how they carve up the space of possible speech sounds.

- By adulthood, speech perception is highly sensitive to the native language of the speaker.

- …how do children get there?

- …and how deep is the bias to classify speech sounds according to your own language?

Speech Perception

A speaker of a language is a prisoner of its ‘alphabet’ of sounds.

Voice Onset Time (VOT)

Perceiving VOT

'Categorical Perception'

 Discrimination

A More Systematic Test

R  L

R

Within-Category Discrimination is Hard

Cross-language Differences
Cross-Language Differences

English vs. Japanese R-L

• American English pronunciation
  \[ \text{mæk dan ald} \]
• Japanese pronunciation
  \[ \text{mækudonardu} \]

Speech Perception

A speaker of a language is a prisoner of its ‘alphabet’ of sounds

Developmental Questions

• How does the native/non-native difference emerge?
• Does native-language discrimination improve with practice?

One Answer...

• Children learn the feature contrasts of their language
• Children learn gradually, adding features over the course of development

Evidence from neonates?

• How do we know babies can hear differences in speech?
• What can babies do?
• High-amplitude sucking

Roman Jakobson, 1896-1982
English VOT Perception

To Test Children
Not so easy!
High Amplitude
Stacking

General Infant Abilities

- Infants’ show *Categorical Perception* of speech sounds - at 2 months and earlier
- Discriminate a wide range of speech contrasts
- Discriminate Non-Native speech contrasts
e.g., Japanese babies discriminate r-l
e.g., Canadian babies discriminate d-D

Universal Listeners

- Infants may be able to discriminate all speech contrasts from the languages of the world!

Varying Pronunciations

- Voiceless stops /p, t, k/

<table>
<thead>
<tr>
<th>pit</th>
<th>tack</th>
</tr>
</thead>
<tbody>
<tr>
<td>spit</td>
<td>stack</td>
</tr>
<tr>
<td>spit</td>
<td>stack</td>
</tr>
<tr>
<td>bit</td>
<td>dack</td>
</tr>
</tbody>
</table>

- Aspirated at start of syllable; unaspirated after [s]
- 6 month olds easily distinguish bottom 2 rows; 1 year olds do not (adults aren’t great either)

How can they do this?

- Innate speech-processing capacity?
- General properties of auditory system?

What About Non-Humans?

- Chinchillas show categorical perception of voicing contrasts!