Language & the Mind  
LING240  
Summer Session II 2005

Lecture 5  
Sounds

How you look to a phonetician

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

How you look to a phonetician

Nasal Cavity  
Oral Cavity

0 ‘thick’  
δ ‘the’  
ʃ ‘sh’  
ʒ ‘azure’  
ʧ ‘ch’  
ʤ ‘j’  
ŋ ‘thing’  
ʔ ‘uh-oh’

Forget Spelling!  
Sounds ≠ Spelling
One Sound - Many Characters

- he
- believe
- Caesar
- see
- people
- seas
- amoeba
- key
- machine
- seize

IPA: [i]

One Sound - Many Characters

- too
- to
- clue
- through
- threw
- lieu
- shoe

IPA: [u]

One Character - Many Sounds

- dame
- dad
- father
- call
- village
- many
- c
- ð
- ñ
- ı
- ñ
- ι
- ʇ
- ŕ

One Sound - Multiple Letters

- shoot
- either
- character
- deal
- Thomas
- physics
- rough

IPA: [ʃ]

One Letter - 0, 1, 2 Sounds

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

Differences across Languages

- English: judge, juvenile, Jesus [dʒ]
- Spanish: jugar, Jesus [h]
- German: Jugend, jubeln, Jesus [j]
- French: Jean, j’accuse, jambon [ʒ]

IPA: [kjuwt]
Major division: consonants vs vowels

- Consonantal sounds: narrow or complete closure somewhere in the vocal tract.
- Vowels: very little obstruction in the vocal tract. Can form the basis of syllables (also possible for some consonants).

Describing Speech Sounds

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

Where does the Air Flow?

Your vocal tract again

Block it at the velum

Where does the air go?
Block it at the velum

Where does the air go?

Now raise the velum to block the air...

Quickly drop your tongue again...

Tongue against velum again

Now raise the velum to block the air...

Quickly drop your tongue again...
Where does the air go this time?

So far we have:

Nasal stop:
[ŋ]

Non-nasal (oral) stops:
[g] [k]

Where can you stop the airstream?

(bi)labial
[b] [p] [m]
Where can you stop the airstream?

Labiodental
[v] [f]

Interdental
[θ] [ð]

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

Palatal
[z] [ʃ] [ʒ] [tʃ]

Velar
[g] [k] [ŋ]

Uvular
Where can you stop the airstream?

Manner - How the Air is Flowing

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

Fricatives & Affricates

- Palatal sounds [ʒ] [ʃ] [dʒ] [tʃ]
- Palatal Fricatives - [ʒ] [ʃ]
  [note: according to IPA chart these are strictly 'post-alveolar']
- Affricates - combination of stop + fricative - [dʒ] [tʃ], as in judge, church

Voiced & Voiceless Consonants

- Consonants either voiced or voiceless.
- English pairs:

  b p  v f  d t
  z s  ð θ

Describing Consonants

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

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
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 of the languages of the world

- **Different languages choose different feature combinations**
What can you do to alter the shape of your vocal tract?
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?

- i sound “sheep, sleep”
- I sound “ship, slip”
- e sound “laid, spade, trade”
- æ sound “led, sped, tread”
- ð sound “bat, lad”
So what vowels do you have?

“Luke, who’d, suit”

“look, hood, soot”

“coat, wrote, hoed”

“caught, wrought, hawed”

“bah, father, cot, Don”

“metallic, Texas”

“but, putt, rut”

So here they are!
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]).

Cross-language Differences

- Feature Combinations
  - English: back vowels are rounded, others are not
  - German/French has high, front, rounded vowel [y]
  - Russian has high back unrounded vowel [u̯]

  - Many languages don’t make the tense/lax distinction found in English (ex: Spanish [i])
  - Many languages distinguish short and long vowels (unlike English), ex: Japanese

Cross-language Differences

Diphthongs:

a

Diphthongs:

“side, my, kind”

aj

Diphthongs:

a
Diphthongs:

“loud, brow, hour”

aw

Diphthongs:

“boy, annoy, toil”

çj

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

Speech Perception

• Speech production processes must be undone by the ear
• Motions of articulators must be reconstructed from patterns of air vibration
• Requires extremely precise hearing, possibly a system specialized for hearing speech
• Substantially developed at birth

Acoustic Information

• Frequency
• Timing
• 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)
Synthesized Speech

• Allows for precise control of sounds
• Valuable tool for investigating perception

Voice Onset Time (VOT)

English VOT production

• Not uniform
• 2 categories
Perceiving VOT

‘Categorical Perception’

Discrimination

Why is this pair difficult?
(i) Acoustically similar?
(ii) Same Category?

Discrimination

A More Systematic Test

D 0ms 20ms D
D 20ms 40ms T
T 40ms 60ms T

Within-Category Discrimination is Hard

Cross-language Differences

English vs. Japanese R-L

Cross-Language Differences

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

Cross-Language Differences

?
Cross-language Differences

Participants: Thai – native English - second (>3 years in the US)

[d₁a] [d₂a] DIFFERENT

Imsri & Idsardi (2001)

What’s a Syllable?

• Another phonological unit of words
• Every vowel is at the center of a syllable
• Syllables have hierarchical structure

Phonotactic Constraints: Constraints on Syllable Structure

• Every language has restrictions on what sequences of phonemes may occur (*ktleem)
• These constraints are language specific
  English: *zleem Polish: zlev ‘sink’
• Illegal sequences are illegal within a single syllable
  English: *[kspl] [ɛk spl̩ sɪt]

Japanese Phonotactic Constraints

Japanese doesn’t allow consonant clusters within a syllable!

Japanese Syllable Structure

• Toyota
• Honda

Toyota, Honda...
Japanese Syllable Structure

Japanese Syllable Structure

Japanese Syllable Structure

Japanese Syllable Structure

Japanese Syllable Structure

Japanese Syllable Structure
Japanese Syllable Structure

Phonemic Level: /mækdonald/

Phonetic Level: [mækudonalo]d

Behavioral Results

• Japanese speakers have trouble hearing the difference

Speech Perception

It seems that a language speaker is a prisoner of his/her language phonemic alphabet

Additional Findings

<table>
<thead>
<tr>
<th>Language</th>
<th>RT Mean (ms)</th>
<th>SE (ms)</th>
<th>LT Mean (ms)</th>
<th>SE (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese</td>
<td>1002 65 3.9%</td>
<td>1102 75 5.2%</td>
<td>1002 54 5.8%</td>
<td></td>
</tr>
<tr>
<td>French</td>
<td>1173 73 21%</td>
<td>1087 75 39%</td>
<td>1002 54 5.8%</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Mean reaction time (ms), standard error, and error rate in AAN judgments on an non-speech contrast and a vowel length contrast in French and Japanese participants (Experiment 1).
A Puzzle...

- Korean speakers use the sounds [r] and [l] e.g. Korea Seoul
- Korean babies hear the difference between [ra] and [la] ... they don't know Korean yet
- Korean adults know Korean ... but they have difficulty hearing the [ra] vs. [la] contrast

Developmental Questions

- How does the native/non-native difference emerge?
- Does native-language discrimination improve as a result of native language input?

Possibility #1: Adding Features

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

Predictions of Possibility #1

- Poor discrimination at birth
- Better and better with age

Roman Jakobson, 1896-1982

What’s innate?

- Auditory abilities

Evidence from neonates?

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

To Test Adults
Simply ask:
same or different?
or
Is it a [b] or a [p]?

To Test Children
Not so easy!
High Amplitude Sucking

Reality Check for Possibility #1

- 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!

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!
What’s innate?

- Auditory abilities
- Articulatory rudiments
  - Not developed
  - Desire to coo and babble even in deaf infants

Connecting Hearing & Speaking

McGurk Effect


Evidence for connection

- Infants know connection between visual and auditory speech stimuli
- Mix and match [a] vs. [i]

Connecting Hearing & Speaking


What’s innate?

- Auditory abilities
- Articulatory rudiments
- Connection between them
  - Phonetic level
  - Universal Grammar (UG)

When does change occur?
When Does Change Occur?

• About 10 months

Janet Werker  
U. of British Columbia

What has Werker found?

• Is this the beginning of memory?

• Are the infants learning words?

• ...Or something else?

Possibility #2: Maintenance & Loss

• Infants maintain features being used in their language

• They lose all others

Patricia Kuhl  
University of Washington

Learning the surface

- Auditory
- Phonetic
- Articulatory
- Innate

Possibility #2: Schematic
Possibility #2: Predictions

• Loss of discriminability should be **permanent** and absolute

But...

• Training improves adult performance

Possibility #2: Reality Check

• Loss of discriminability is neither permanent nor absolute!

Possibility #3: Functional Reorganization

• Changes in performance with development do not reflect changes in the hard-wiring of the brain
What does Development Involve?

- Change - non-native categories *lost* (structure-changing)
- Growth - non-native categories *hidden* (structure-adding)

What does Development Involve?

- Evidence for *Growth*
  1. Some discrimination retained when sounds presented close together (e.g. Hindi d-D contrast)
  2. Discrimination abilities better when people hear sounds as non-speech
  3. Adults do better than 1-year olds on some sound contrasts
- All evidence comes from consonants

What does Development Involve?

- Evidence for *Change*
  1. No evidence of preserved non-native category boundaries in vowel perception
  2. Best evidence for change comes from vowels and vowel-like categories

What yearlings can’t do

- Recognize minimal pairs while relating them to real words
  - “bear” versus “pair”
  - “Piglet” versus “Biglet”
- More Werker experiments

Word Learning

- Stager & Werker 1997
  - “bih” vs. “dih” and “lif” vs. “neem”
- Procedure: familiarize with sound-object pairs, then test using same or different pairings

Word learning results

- Exp 2 vs 4
Key Findings

• 14-month olds can discriminate the minimally contrasting words (Expt. 4)
• But they fail to notice the minimal change in the sounds when they are paired with objects, i.e., when they are words (Expt. 2)
• They can perform the task, when the words are more distinct (Expt. 3)
• Therefore, 14-month olds use more detail to represent sounds than they do to represent words

Why Yearlings Fail on Minimal Pairs

• They fail specifically when the task requires word-learning
• They do know the sounds
• But they fail to use the detail needed for minimal pairs to store words in memory
• What is going on?
  – Is this true for all words?
  – When do they learn to do this?
  – What triggers the ability to do this?

Swingley & Aslin, 2002

• 14-month year olds did recognize mispronunciations of familiar words

Approximate Ages

• Surface 10 months
• Memory 18 months

Approximate Ages

Memory
Lexical
Surface

Innate

Articulatory
Phonetic

Swingley & Aslin, 2002

Possibility #1 Again...

• Children learn the feature contrasts of their language
• Children may learn gradually, adding features over the course of development
• Phonetic knowledge does not entail phonological knowledge

Roman Jakobson, 1896-1982
Word-learning & phonological detail

- Word-learning is very hard for younger children, so detail is initially missed when they first learn words
- Many exposures are needed to learn detailed word forms at earliest stages of word-learning
- Success on the Werker/Stager task seems to be related to the **vocabulary spurt**, rapid growth in vocabulary after ~50 words

Back to 1-year olds

- 1-year olds know the *surface sound patterns* of the language
- 1-year olds do not yet know which sounds are used *contrastively* in the language (which sound variations affect meaning and which don’t)
- 1-year olds still need to learn **contrasts**

Vowels Same or Different?

<table>
<thead>
<tr>
<th>light</th>
<th>lied</th>
</tr>
</thead>
<tbody>
<tr>
<td>tight</td>
<td>tied</td>
</tr>
<tr>
<td>site</td>
<td>sighed</td>
</tr>
<tr>
<td>life</td>
<td>live</td>
</tr>
<tr>
<td>knife</td>
<td>knive(s)</td>
</tr>
<tr>
<td>lice</td>
<td>lies</td>
</tr>
<tr>
<td>dice</td>
<td>dies</td>
</tr>
</tbody>
</table>

Some people have this system:

<table>
<thead>
<tr>
<th>light</th>
<th>lajt</th>
</tr>
</thead>
<tbody>
<tr>
<td>tight</td>
<td>tajt</td>
</tr>
<tr>
<td>site</td>
<td>sajt</td>
</tr>
<tr>
<td>life</td>
<td>lajf</td>
</tr>
<tr>
<td>knife</td>
<td>najf</td>
</tr>
<tr>
<td>lice</td>
<td>lajs</td>
</tr>
<tr>
<td>dice</td>
<td>dajs</td>
</tr>
</tbody>
</table>

Some people have this one:

<table>
<thead>
<tr>
<th>light</th>
<th>I A j t</th>
</tr>
</thead>
<tbody>
<tr>
<td>tight</td>
<td>t A j t</td>
</tr>
<tr>
<td>site</td>
<td>s A j t</td>
</tr>
<tr>
<td>life</td>
<td>I A j f</td>
</tr>
<tr>
<td>knife</td>
<td>n A j f</td>
</tr>
<tr>
<td>lice</td>
<td>I A j f</td>
</tr>
<tr>
<td>dice</td>
<td>d A j s</td>
</tr>
</tbody>
</table>
What’s the pattern?

<table>
<thead>
<tr>
<th>light</th>
<th>tied</th>
<th>lajd</th>
</tr>
</thead>
<tbody>
<tr>
<td>tight</td>
<td>tied</td>
<td>tajd</td>
</tr>
<tr>
<td>site</td>
<td>sighed</td>
<td>sajd</td>
</tr>
<tr>
<td>life</td>
<td>live</td>
<td>lajv</td>
</tr>
<tr>
<td>knife</td>
<td>knive(s)</td>
<td>najvz</td>
</tr>
<tr>
<td>lice</td>
<td>lies</td>
<td>lajz</td>
</tr>
<tr>
<td>dice</td>
<td>dies</td>
<td>dajz</td>
</tr>
</tbody>
</table>

What’s the pattern?

<table>
<thead>
<tr>
<th>light</th>
<th>lied</th>
<th>lajd</th>
</tr>
</thead>
<tbody>
<tr>
<td>tight</td>
<td>tied</td>
<td>tajd</td>
</tr>
<tr>
<td>site</td>
<td>sighed</td>
<td>sajd</td>
</tr>
<tr>
<td>life</td>
<td>live</td>
<td>lajv</td>
</tr>
<tr>
<td>knife</td>
<td>knive(s)</td>
<td>najvz</td>
</tr>
<tr>
<td>lice</td>
<td>lies</td>
<td>lajz</td>
</tr>
<tr>
<td>dice</td>
<td>dies</td>
<td>dajz</td>
</tr>
</tbody>
</table>

What’s the pattern?

| voiceless t | voiced d |
| alveolar t   | stop d  |
| stop t       | d       |
| voiceless f  | voiced v |
| labiodental f | alveolar v |
| fricative f  | fricative v |
| voiceless s  | voiced z |
| alveopalatal fricative s | alveopalatal z |

So these speakers have a rule ...

Before a voiceless consonant
a j --> ʌ j

Isn’t it just two sets of words?

| ʌ js lajz |
| d ʌ js dajz |
| l ʌ jf lajz |
| s ʌ jt sajd |
| l ʌ jt lajd |

Nope, it’s a rule ...

| stied |
| stight |
Nope, it’s a rule ...

Two “levels” of speech sounds

The sounds you “store” in your head

Two “levels” of speech sounds

The sounds you actually produce

Terminology

“phones”

1 phoneme; more than 1 phone

We call the phones allophones of the phoneme
So...

- In some dialects of English, the phoneme /aj/ has two allophones: [aj] and [ʌ j].
- The allophone [ʌ j] occurs whenever the phoneme precedes a voiceless sound.

The whole rule:

aj → ʌ j /_____ [-voice]

1 phoneme 2 phones

Another rule:

t → d /N_____V

sound(s) actually produced

Another rule:

What about these?

- "sit" [sɪt]
- "sitter" [sɪˈdɛr]
- "heat" [hɪt]
- "heater" [hɪˈdɛr]
- "at" [æt]
- "attic" [əˈdɪk]
- "attack"
- "atone"
- "determine"
- "detect"
Is there a pattern?

[sídər]  [ətæk]
[hídər]  [ətόn]
[ædɪk]  [dιτεkt]

Is there a pattern?

[sídər]  [ətæk]
[hídər]  [ətόn]
[ædɪk]  [dιτεkt]

So we need a slight revision

\[ t \rightarrow d /\tilde{\eta} _____ V \]

1 phoneme 2 allophones

sound(s) actually produced

What do you have in your head?
Points to note:

- Sequence becomes “easier to say”

  BUT

- This process is a specific rule of a particular dialect of English

In what sense a specific rule?

- doesn’t apply to all instances of “t” between vowels
- isn’t a part of the grammar of other dialects of English
- is only one way to make sequences of vowels and voiceless consonants easier to say

Moral:

The rules that we discover are often “natural” in that one can find an explanation for many of them in terms of ease of articulation, but they are not inevitable/innate: they are specific rules of particular dialects or languages, and had to be learned.

How much detail do you have to remember about the sound of each word?

If you can predict something by a rule, you don’t have to remember it

Just remember:

- the rule

- the things that can’t be predicted

Allophonic differences ignored by hearers

\[ /aj/ \]

\[ [\alpha j] \]
Varying Pronunciations

- Voiceless stops /p, t, k/
  - pit \(\approx\) tack \(\approx\)
  - spit \(\approx\) stack \(\approx\)
  - 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)

Languages can differ in what is predictable

Korean has [l] and [r]

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[rutmu] “name”
[ratio] “radio”
[mul] “water”
[pal] “big”
[soul] “Seoul”
[iikop] “seven”
[ipalsa] “barber”

But [r] doesn’t show up everywhere ...

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[rutmu] “name”
[ratio] “radio”
[mul] “water”
[pal] “big”
[soul] “Seoul”
[iikop] “seven”
[ipalsa] “barber”

[r] is always in front of a vowel

And nor does [l] ...

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[rutmu] “name”
[ratio] “radio”
[mul] “water”
[pal] “big”
[soul] “Seoul”
[iikop] “seven”
[ipalsa] “barber”

[l] is never in front of a vowel

So: Korean has only 1 liquid phoneme

(Koreans don’t have to remember if a word has [l] or [r])
So in Korean, [l] and [r] are “the same”

So Korean works like this:

1 phoneme 2 allophones

While English works like this:

2 phonemes 2 phones

Even more schematically:

Stored  |  Produced
English  | 
/l/    |  [l]  
/r/    |  [r]  
Korean  | 
/l/    |  [l]  
/r/    |  [r]  

sound(s) actually produced
Minimal Pairs

- In English, [r] and [l] can occur in the same position in a word
  
rake  lake
  ramp  lamp
  rim   limb
  ripper ripple

- In English, [r] and [l] can be used to mark a meaning contrast
- In English, /r/ and /l/ are two phonemes

Minimal Pairs

- Korean works differently

- [r] and [l] are two allophones of a single phoneme in Korean
- It’s impossible to create minimal pairs which contrast r/l in Korean
- [r] and [l] cannot be used contrastively in Korean

Puzzle Solved!

- Korean speakers use the sounds [r] and [l] e.g. Korea Seoul
- Korean babies hear the difference between [ra] and [la] ... they don't know Korean yet
- Korean adults know Korean ... but they have difficulty hearing the [ra] vs. [la] contrast

So they don't know that they are pronunciations of the same phoneme

Phonemic contrasts are easier to hear