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

Speech Perception

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

Cross-language Differences

![Sound examples]
Developmental Questions

• How does the native/non-native difference emerge?
• What is the nature of the ‘learning’ that is involved?
• When does the native/non-native difference emerge?

Possibility #1: Adding Features

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

Possibility #1: Predictions

• Poor discriminators at birth
• Better and better with age

Roman Jakobson, 1896-1982
English VOT Perception

To Test Children
Not so easy!
High Amplitude
Sucking

Universal Listeners

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

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

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!

Connecting Hearing & Speaking

McGurk Effect


Connecting Hearing & Speaking

Infants start to connect auditory and visual information early

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

• How does the native/non-native difference emerge?
• What is the nature of the ‘learning’ that is involved?
• When does the native/non-native difference emerge?

When does Change Occur?

• About 10 months

Janet Werker
U. of British Columbia

When does Change Occur?

• Hindi and Salish contrasts tested on English kids

Janet Werker
U. of British Columbia

Developmental Questions

• How does the native/non-native difference emerge?
• What is the nature of the ‘learning’ that is involved?
• When does the native/non-native difference emerge?
Possibility #2: Maintenance & Loss

- Infants maintain features they hear being used in their language
- They lose all others

Patricia Kuhl
University of Washington

Possibility #2: Predictions

- Loss of discriminability should be permanent and absolute

Possibility #2: Maintenance & Loss

- Another schematic:

Possibility #2: Reality Check

- Loss of discriminability IS NEITHER permanent NOR absolute
Training improves adult performance

Some non-native contrasts are easy for adults to discriminate

Adults perform better at non-native contrasts if they think the sounds are not language sounds!

We seem to have different “ears” for language sounds vs. other sounds
**Developmental Questions**

- How does the native/non-native difference emerge?
- What is the nature of the ‘learning’ that is involved?
- When does the native/non-native difference emerge?

**Possibility #3: Functional Reorganization**

- Changes in performance with development do not reflect changes in the hard-wiring of the brain
- They reflect the operation of postperceptual processes that allow the infant to attend to only phonetic information that distinguishes meaning

**What 1-year-olds 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’

![Diagrammatic representations of experiments 1-4](image)

Why 1-year-olds 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?

![Figure 1. Results showing the conditions under which infants show significant dressing on the switch task. Grains show mean looking times on the ‘bih’ and ‘lif’ trials with standard error bars.](image)

Word learning results

• Exp 2 vs 4

![Table 1. Correctly pronounced (CP) target words and their mispronounced (MP) versions](image)

<table>
<thead>
<tr>
<th>CP</th>
<th>MP-close</th>
<th>MP-distant</th>
</tr>
</thead>
<tbody>
<tr>
<td>apple</td>
<td>apple</td>
<td>opal</td>
</tr>
<tr>
<td>baby</td>
<td>vaby</td>
<td>raby</td>
</tr>
<tr>
<td>ball</td>
<td>gall</td>
<td>shaw</td>
</tr>
<tr>
<td>car</td>
<td>kier</td>
<td>kier</td>
</tr>
<tr>
<td>dog</td>
<td>tegg</td>
<td>mog</td>
</tr>
<tr>
<td>kitty</td>
<td>pity</td>
<td>yitty</td>
</tr>
</tbody>
</table>

Note. International Phonetic Alphabet transcriptions are provided in parentheses. Syllabic words were stressed on the first syllable.

Swingley & Aslin, 2002

• 14-month olds did recognize mispronunciations of familiar words
Werker et al. 2002

More…

- 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 early 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

One-Year Olds Again

- One-year olds know the *surface sound patterns* of the language
- One-year olds do not yet know which sounds are used contrastively in the language…
- …and which sounds simply reflect variation that does not affect meaning
- One-year olds need to learn *contrasts*

We need another level!
### Same or different?

<table>
<thead>
<tr>
<th>Word</th>
<th>Original</th>
<th>Altered</th>
</tr>
</thead>
<tbody>
<tr>
<td>light</td>
<td>lied</td>
<td>lajt</td>
</tr>
<tr>
<td>tight</td>
<td>tied</td>
<td>tajt</td>
</tr>
<tr>
<td>site</td>
<td>sighed</td>
<td>sajt</td>
</tr>
<tr>
<td>life</td>
<td>live</td>
<td>lajf</td>
</tr>
<tr>
<td>knife</td>
<td>knife(s)</td>
<td>najf</td>
</tr>
<tr>
<td>lice</td>
<td>lies</td>
<td>lajs</td>
</tr>
<tr>
<td>dice</td>
<td>dies</td>
<td>dajs</td>
</tr>
</tbody>
</table>

### Some people have this system:

<table>
<thead>
<tr>
<th>Word</th>
<th>Original</th>
<th>Altered</th>
</tr>
</thead>
<tbody>
<tr>
<td>light</td>
<td>lied</td>
<td>lajd</td>
</tr>
<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>knife(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>

### Some people have this one:

<table>
<thead>
<tr>
<th>Word</th>
<th>Original</th>
<th>Altered</th>
</tr>
</thead>
<tbody>
<tr>
<td>light</td>
<td>lajt</td>
<td>lied</td>
</tr>
<tr>
<td>tight</td>
<td>tajt</td>
<td>tied</td>
</tr>
<tr>
<td>site</td>
<td>sajt</td>
<td>sighed</td>
</tr>
<tr>
<td>life</td>
<td>lajf</td>
<td>live</td>
</tr>
<tr>
<td>knife</td>
<td>najf</td>
<td>knife(s)</td>
</tr>
<tr>
<td>lice</td>
<td>lajs</td>
<td>lies</td>
</tr>
<tr>
<td>dice</td>
<td>dajs</td>
<td>dies</td>
</tr>
</tbody>
</table>

### What’s the pattern?

<table>
<thead>
<tr>
<th>Word</th>
<th>Original</th>
<th>Altered</th>
</tr>
</thead>
<tbody>
<tr>
<td>light</td>
<td>lajt</td>
<td>lied</td>
</tr>
<tr>
<td>tight</td>
<td>tajt</td>
<td>tied</td>
</tr>
<tr>
<td>site</td>
<td>sajt</td>
<td>sighed</td>
</tr>
<tr>
<td>life</td>
<td>lajf</td>
<td>live</td>
</tr>
<tr>
<td>knife</td>
<td>najf</td>
<td>knife(s)</td>
</tr>
<tr>
<td>lice</td>
<td>lajs</td>
<td>lies</td>
</tr>
<tr>
<td>dice</td>
<td>dajs</td>
<td>dies</td>
</tr>
</tbody>
</table>
What’s the pattern?

<table>
<thead>
<tr>
<th>Voiceless</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alveolar Stop</td>
<td>Alveolar Stop</td>
</tr>
<tr>
<td>Labiodental Fricative</td>
<td>Labiodental Fricative</td>
</tr>
<tr>
<td>Alveopalatal Fricative</td>
<td>Alveopalatal Fricative</td>
</tr>
</tbody>
</table>

So these speakers have a rule ...

Before a voiceless consonant
aj → ʌj

Isn’t it just two sets of words?
Isn’t it just two sets of words?

Nope, it’s a rule ...

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 “store” in your head

/lajt/
/lajd/

The sounds you actually produce

Two “levels” of speech sounds

The sounds you actually produce

/lajt/
/lajd/

[lajt]
[lajd]

Two “levels” of speech sounds

Terminology

“phonemes”

/lajt/ /lajd/

The sounds you “store” in your head
The sounds you actually produce are called **phones**. We call the phones **allophones** of the phoneme.

So:

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

Or, more succinctly:

aj
Or, more succinctly:

This phoneme

aj

becomes

aj →

Or, more succinctly:

this sound (allophone)

aj → Δj

where?

aj → Δj /
Or, more succinctly:

\[ aj \rightarrow \Delta j / \_\_\_ \]

immediately before

Another rule:

\[ t \rightarrow d / V \_\_\_ V \]

The whole rule:

\[ aj \rightarrow \Delta j / \_\_\_ \text{ [-voice]} \]

a voiceless sound

\[ aj \rightarrow \Delta j / \_\_\_ \text{ [-voice]} \]
Another rule:

- “sit”  [sit]
- “sitter”  [sidər]
- “heat”  [hit]
- “heater”  [hidər]
- “at”  [æt]
- “attic”  [ædɪk]

What about these?

- “attack”
- “atone”
- “determine”
- “detect”

Is there a pattern?

| [sidər]  | [ətæk] |
| [hidər]  | [ətɔn] |
| [ædɪk]  | [dɪtəkt] |

Is there a pattern?

| [sidər]  | [ətæk] |
| [hidər]  | [ətɔn] |
| [ædɪk]  | [dɪtəkt] |
So we need a slight revision

t → d /ʊ/ V

1 phoneme

/sound stored in the brain/

What do you have in your head?

/sounds (phonemes)/

1 phoneme 2 allophones

/t/

[d] [t]

sound(s) actually produced
What do you have in your head?

/t, f, p.../

Rules

[278x395]What do you have in your head?

/t/

[d]

[t]

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.

A way to think about phonemes

- You have a “mental dictionary” where you store the words you have learned.
- What information does it have to contain for each word?
  - At least:
    - the meaning of the word
    - the sound of the word

A way to think about phonemes

- 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

Voiced stops /b, d, g/

Voiceless stops /p, t, k/

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] & [r]

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[irumi] “name”
[ratio] “radio”
[mul] “water”
[pal] “big”
[saul] “Seoul”
[ilkop] “seven”
[ipalsa] “barber”

But [r] doesn’t show up everywhere...

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[irumi] “name”
[ratio] “radio”
[mul] “water”
[pal] “big”
[saul] “Seoul”
[ilkop] “seven”
[ipalsa] “barber”
But [r] doesn’t show up everywhere...

[r] is always in front of a vowel

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[irumi] “name”
[ratio] “radio”
[mui] “water”
[pal] “big”
[saul] “Seoul”
[iukop] “seven”
[ipalsa] “barber”

And nor does [l]...

[l] is never in front of a vowel

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[irumi] “name”
[ratio] “radio”
[mui] “water”
[pal] “big”
[saul] “Seoul”
[iukop] “seven”
[ipalsa] “barber”

And nor does [l]...

[rupi] “ruby”
[kiri] “road”
[saram] “person”
[irumi] “name”
[ratio] “radio”
[mui] “water”
[pal] “big”
[saul] “Seoul”
[iukop] “seven”
[ipalsa] “barber”

So: Korean has only 1 liquid phoneme

(Koreans don’t have to remember if a word has [l] or [r])
... but that phoneme has 2 allophones

So in Korean, [l] and [r] are “the same”

So Korean works like this:

1 phoneme

(sound stored in the brain)
1 phoneme  2 allophones

While English works like this:

2 phonemes

(sound stored in the brain)
Even more schematically:

<table>
<thead>
<tr>
<th>Stored</th>
<th>Produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>/l/</td>
<td>[l]</td>
</tr>
<tr>
<td>/r/</td>
<td>[r]</td>
</tr>
</tbody>
</table>

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.
Korean has [l] & [r]

- [rupi] “ruby”
- [kiri] “road”
- [saram] “person”
- [irumi] “name”
- [ratio] “radio”
- [muli] “water”
- [pal] “big”
- [saül] “Seoul”
- [ilkop] “seven”
- [ipalsa] “barber”

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

- [rupi] “ruby”
- [kiri] “road”
- [saram] “person”
- [irumi] “name”
- [ratio] “radio”
- [muli] “water”
- [pal] “big”
- [saül] “Seoul”
- [ilkop] “seven”
- [ipalsa] “barber”

[r] is always in front of a vowel

And nor does [l] ...

- [rupi] “ruby”
- [kiri] “road”
- [saram] “person”
- [irumi] “name”
- [ratio] “radio”
- [muli] “water”
- [pal] “big”
- [saül] “Seoul”
- [ilkop] “seven”
- [ipalsa] “barber”

[l] is never in front of a vowel

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
One-Year Olds Again

- One-year olds know the surface sound patterns of the language
- One-year olds do not yet know which sounds are used contrastively in the language...
- ...and which sounds simply reflect pronunciation rules
- One-year olds need to learn contrasts

Maybe not so bad after all...

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

Roman Jakobson, 1896-1982

Word Learning

- Stager & Werker 1997
  - ‘bih’ vs. ‘dih’
  - ‘lif’ vs. ‘neem’

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
Where’s Whorf?

• We do “hear” language sounds differently depending on what language we speak. But when we fail to hear a contrast that a speaker of another language does hear, it isn’t because our physical ability to register that difference has disappeared. It’s because we have learned that that type of contrast is not a meaningful contrast for language.

• Our mental representations of the sounds of words are an abstraction from the physical signal. Allophones of /t/ are all the same sound in our minds.