**Psych 215L:**
Language Acquisition

Lecture 4
Speech Perception I

**Sounds of Language (Speech Perception)**

Learner’s job: parse continuous stream of speech into sentences, clauses, words, syllables, and phonemes

Phonemes are language-specific - /r/ is a phonemic contrast (changes word’s meaning) in English but not in Japanese.

Kids of the world require knowledge of phonemes before they can figure out what different words are - and when different meanings are signaled by different words.

**About Speech Perception**

Important: Not all languages use the same sounds. Languages draw from a common set of sounds.

Child’s task: Figure out what sounds their native language uses.

**Speech Perception: Computational Problem**

Divide sounds into contrastive categories

**Acoustic-Level Information**

Includes: timing and frequency

Tones: frequency (close-up)
Acoustic-Level Information

Language sounds

Vowels combine acoustic energy at a number of different frequencies

Different vowels (a) “ah”, (i) “ee”, (u) “oo” etc.) contain acoustic energy at different frequencies

Listeners must perform a ‘frequency analysis’ of vowels in order to identify them (Fourier Analysis)

Acoustic-Level Information

Language sounds

Male Vowels (close up)

Acoustic-Level Information

Language sounds

Female Vowels (close up)

Synthesized Speech

Allows for precise control of sounds

Valuable tool for investigating perception

Acoustic-Level Information

Language sounds

Timing: Voice Onset Time (VOT)

English VOT production

Not uniform - there are 2 categories

Figure 6-b: VOT productions of a single normal adult speaker of American English for vowels beginning with (i) and (u). (shape adapted with permission from Buracas, Cooper, Diveny, Kelso, & Liberman, 1993, Resolution Methods in Speech: A Voice Onset Time Analysis, plan and language, 9, 155-171). (copyright 1993 by Academic Press.)
Perceiving VOT

'Categorical Perception': 

Decision between \( d/t \) 

Time to make decision

Discrimination

<table>
<thead>
<tr>
<th>Same/Different</th>
<th>0ms</th>
<th>60ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>0ms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10ms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40ms</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Why is this pair difficult?

(i) Acoustically similar?
(ii) Same Category?

Discrimination

A More Systematic Test

<table>
<thead>
<tr>
<th>Same/Different</th>
<th>0ms</th>
<th>20ms</th>
<th>40ms</th>
<th>60ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>0ms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10ms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40ms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Within-Category Discrimination is Hard

Cross-language Differences

English vs. Japanese R-L

Cross-Language Differences

English vs. Hindi

alveolar [d]

retroflex [D]
Infant Speech Perception

How do we tell what infants know, or use, or are sensitive to?

Researchers use indirect measurement techniques.

Some information from the High Amplitude Sucking (HAS) paradigm

Infants have sophisticated discrimination abilities, but they don't abstract sounds into categories the way that adults do.

Infant perception

"It" "It" "It" "It" "It"

Adult perception

"It" "It"

Infant Speech Perception

How do we tell what infants know, or use, or are sensitive to?

Researchers use indirect measurement techniques.

Some information from the High Amplitude Sucking (HAS) paradigm

Infants have sophisticated discrimination abilities, but they don't abstract sounds into categories the way that adults do.

Infant speech perception

Kids...

This ability to distinguish sound contrasts extends to phonemic contrasts that are non-native. (Japanese infants can discriminate contrasts used in English but not in Japanese, like it.) This goes for both vowels and consonants.

...vs. adults

Adults can't, especially without training - even if the different is quite acoustically salient.

So when is this ability lost?

And what changes from childhood to adulthood?

Perceiving sound contrasts

Speech Perception of Non-Native Sounds

Comparing perceptual ability

Werker et al. 1981: English-learning 6-8 month olds compared against English & Hindi adults on English & Hindi contrasts

Conditioned Head Turn Procedure

Speech Perception of Non-Native Sounds


But when after 6-8 months is the ability to lost? Werker & Tees (1984)

Key into "critical period" hypothesis for language (Lenneberg 1967) - when language can be learned natively

"To test for this critical period, children of 12 and 8 years were tested, with the expectation that the 6-year-olds but not the 12-year-olds would be able to discriminate non-native contrasts. English-speaking children of both ages, however, performed like English-speaking adults... study was extended to 4-year-old children, who actually performed most poorly of all on nonnative contrasts... findings revealed that experience must begin to influence speech perception long before 4, certainly well before the critical period suggested by Lenneberg."
Speech Perception of Non-Native Sounds

But when after 6-8 months is the ability to lost?

Saiish & Hind contrasts

Change happens somewhere around 8-10 months, depending on the sound contrast.

How change happens

Maintenance & Loss Theory
Infants maintain contrasts being used in their language and lose all the others.

Patricia Kuhl

Language data

Contrasts remaining

“Perceptual Magnet”

How change happens

Maintenance & Loss Theory
Predictions for performance on non-native contrasts over time

Loss of discrimination ability is permanent and absolute

…”appears that the role of experience is to “maintain” those perceptual sensitivities that are already evident in the young infant. Without such exposure, initial abilities will be lost.”

How change happens

Another theory: functional reorganization
Changes attested experimentally reflect operation of postperceptual processes that kick in for language

Janet Werker

Non-linguistic level

Linguistic level

unconscious filter

conscious decision

http://hctv.humnet.ucla.edu/departments/linguistics/Vowel sounds/Consonants/course/chapter6/zulu/zulu.html

How change happens

Problems with the Maintenance & Loss Theory
If it doesn’t sound like speech, adults can tell the difference. - Werker & Tees (1984) showed this with truncated portions of syllables of non-native contrasts. They told subjects the sounds were water dropping into a bucket, and to tell them when the bucket changed.

Pisoni et al. (1982), Werker & Logan (1985): adults can be trained if given enough trials or tested in sensitive procedures with low memory demands

Decline and then recovery (after 4 years old) should never happen if this theory is correct... But there’s improvement for older speakers

And another problem
Some non-native contrasts are easy for older infants and adults to discriminate. (Click languages (Zulu) - click sounds like “tsk tsk” nonspeech)
How change happens

Another theory: functional reorganization
Changes attested experimentally reflect operation of postperceptual processes that kick in for language

Janet Werker

Explanatory power: the whole story
Very young infants respond to any detectable variation - so they can pick up any salient ones in surrounding language. Adults have bias for phonemic information since those are the ones relevant to language. If in non-language setting, adults can tell the nonphonemic differences.

Open question: but why can't 12-month-olds (up to 4 year olds) do the same?

Word Learning & Back to the Critical Period

The connection with word-learning
“Starting at around 1 year of age, infants are poised to begin to learn words, a task they will devote considerable energy to...”

Adults already have their vocabularies fairly stable
“Adults... have the cognitive “distance” and strategic skills to listen for whatever information is required in a particular task. Thus, if the task requires listening to nonphonetic distinctions, the adults will - with varying amounts of practice or training - be able to demonstrate such an ability.”

Linking to the critical period?
“Similarly, young children moving to a new linguistic environment would have the auditory sensitivity to listen to the relevant phonetic detail to acquire words in their new language.”

More on Critical Periods...

But a slight problem, with respect to the critical period... there is one functional reorganization would imply continued flexibility throughout life. Maybe the problem is that there’s a difference between perceptual accent (ability to perceive non-native differences) and productive accent (ability to produce non-native differences). Could be a separate critical period for each.

Also a problem with word-learning motivation - kids don’t seem to show phonetic distinction when word-learning.
12-18 month olds treat “dog” and “bog” as the same.

Distributional learning

Infants exposed to either unimodal or bimodal distribution

Possible Mechanism: Statistical Learning

9-month-old infants are sensitive to the frequency and distribution of perceptual input in speech.

Highly frequent distinctions are learned earlier.

Life’s easier when the distribution is bimodal, though
Possible Mechanism: Statistical Learning

Maye et al. 2002

Infants exposed to either unimodal or bimodal distribution
Alternating test: stimuli 1 and 8
Non-alternating baseline: stimuli 3 or stimuli 6

Bimodal children are sensitive to the presence of two categories

Dietrich, Swingley, & Werker (2007)

Dutch and English vowel categories differ
In English, the length of the vowel is not contrastive
“cat” = “caat”
In Dutch, the length of the vowel is contrastive
“tam” is a different word from “taam” (Japanese also has this distinction)

Tests with 18-month-old children

“Switch” Procedure: measures looking time
...this is a tam...look at the tam

Habitation

Same: look at the tam!

Test

Switch: look at the taam!
Distributional learning from real language data
Dietrich, Swingley, & Werker (2007)
Tests with 18-month-old children

What drives children to learn the distinction?

"One frequently raised hypothesis is that it is driven by contrast in the vocabulary. Dutch children might learn that [u:] and [a:] are different because the words [stut]... and [stut]... mean different things... however, children that young do not seem to know many word pairs that could clearly indicate a distinction between [u:] and [a:].”

What drives children to learn the distinction?

“The other current hypothesis is that children begin to induce phonological categories, “bottom-up”, based on their discovery of clusters of speech sounds in phonetic space... undoubtedly implicated in infants’ early phonetic category learning, which begins before infants know enough words for vocabulary-based hypotheses to be feasible...”

“An necessary condition for such learning to be the driving force behind Dutch children’s phonological interpretation in the present studies is that long and short vowels be more clearly separable in Dutch than in English... preliminary examination of this problem using corpora of Dutch child-directed speech indicated that the set of long and short instances formed largely overlapping distributions.”

Implication: Dutch children need other cues to help them out