The view from cognitive psychology: human infants and speech learning

Daniel Swingley
Department of Psychology
University of Pennsylvania

NIH R01 HD049681
The view from cognitive psychology: human infants and speech learning

the **What**: In their first year, infants learn about their native language:

- Its phonetic categories (roughly, its consonants and vowels)
- Some (50? 500?) frequent sequences / chunks
- Some meaningful aspects of some words
- Bits and pieces relevant to grammatical word order

the **How**: What they have to work with:

- Skills of intention reading, & so some comprehension
- A medium-sized corpus
- Capacity for grouping / chunking
- Ears
- Sounds, faces, and some innate links between them
The What:
methods; discrimination and categorization

“Is this stimulus new?”
Habituation (sucking, looking)
Oddball (EEG)

“Does this stimulus match a trained target?”
Conditioned Headturn
Anticipatory eye-movements (2AFC)

Duration of training in a typical training study: 3 minutes
The What: some results

Testing VOT perception in infants:
Eimas, Siqueland, Jusczyk, & Vigorito, 1971

6 syllables ba/pa: VOT -20, 0, +20, +40, +60, +80
(Adults perceive -20, 0, and 20 as [b], others [p])

1-month-olds and 4-month-olds recover from habituation for +20 / +40, but not the other contrasts
The What: some results

Testing VOT perception in infants: 
*Eimas, Siqueland, Jusczyk, & Vigorito, 1971*

1-month-olds and 4-month-olds recover from habituation for
+20 / +40, but not the other contrasts

<table>
<thead>
<tr>
<th>some contrasts infants discriminated in studies</th>
<th>some contrasts that seem more difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>b/p</td>
<td>s/z</td>
</tr>
<tr>
<td>b/d</td>
<td>s/θ</td>
</tr>
<tr>
<td>r/l</td>
<td>f/θ</td>
</tr>
<tr>
<td>b/w</td>
<td>d/g</td>
</tr>
<tr>
<td>b/m</td>
<td>a/i</td>
</tr>
<tr>
<td>w/j</td>
<td>u/y</td>
</tr>
</tbody>
</table>
The What: some results

Native-language learning: decreasing attention to non-native contrasts

*Werker & Tees 1984*, using conditioned headturn (CHT)
Hindi dental / retroflex [t]; Nthlkampmx velar / uvular [k] vs [q]

6-8 months: distinguish the sounds; 10-12 mos did not; 8-10 intermed.

*Polka & Werker 1994*
German [u] vs [y] and [U] vs [Y]

CHT: 6-8 mos, ~35% reach 7/8 criterion; 10-12 mos, ~10%
Vis. Habituation: 4 mos distinguish the sounds; 6 mos did not

*Bosch & Sebastián-Gallés 2003*
Catalan [e] vs [E], Catalan and Spanish – native infants, visual habituation

4 mos distinguish the sounds; 8 mos only the Catalans
The What: some results

Native-language learning: decreasing attention to non-native contrasts

*Kuhl et al., 2005*, Conditioned headturn

7 month olds’ [t] vs [d] discrimination is correlated (+) with vocab. at 18 mos.

/çi/, /tçʰi/ discrimination is correlated (−) with vocab at 18 mos.
Infants start by being able to distinguish clear instances of almost any contrastive speech sounds (from any language).
Over the first year they begin to conflate similar sounds that are not used contrastively in their native language.
The What: **words**

Infants prefer listening to words, rather than sequences that are not words

*Hallé & de Boysson-Bardies, 1994*

Words presented in lists, either potentially familiar, or rare:

<table>
<thead>
<tr>
<th>Familiar</th>
<th>Rare</th>
</tr>
</thead>
<tbody>
<tr>
<td>bonjour, gâteau, biberon,</td>
<td>busard, cobaye, berline,</td>
</tr>
<tr>
<td>lapin, poupée, ballon,</td>
<td>licence, diffus, caduc,</td>
</tr>
<tr>
<td>voiture, canard, chaussure,</td>
<td>soudard, tangage, enzyme,</td>
</tr>
<tr>
<td>encore, chapeau, oiseau</td>
<td>bigot, volute</td>
</tr>
</tbody>
</table>

Result: 11-month-olds (sometimes 8-mos) listen longer to familiar words

*Replications*: Vihman et al., 2004 (Engl.), Hallé & de. B-B, 1996 (Fr.), Swingley, 2005 (Dutch); 8 months: Jusczyk & Hohne, 1997

*Caution*: Ngon et al., (2011): freq. syllable pairs just as good as bisyll words
The What: words

Infants know what some words mean (at least roughly)

where’s the apple?

Bergelson & Swingley, 2012
The What: words
Infants know what some words mean (at least roughly)

Subject Means, 6-7 month olds
(Paired-Picture Trials)

Item-Pair Means, 6-7 month olds
(Paired-Picture Trials)

Bergelson & Swingley, 2012
The What: words
Infants know what some words mean (at least roughly)

n.b. non-mom talker? Same result (in progress)

Bergelson & Swingley, 2012
abstract words (i.e., not objects)

“hi”
“all gone”

“eat”
“hug”

Bergelson & Swingley, u.r.
abstract words (i.e., not objects)
abstract words (i.e., not objects)
knowledge relevant to grammar

• frequent syllables or sounds are “dissociable” from their contexts
  nouns are identified in article+noun pairs (but not for made-up articles). Likewise verb+inflection lists.
  e.g. Hallé, Durand, & de Boysson-Bardies, 2008; Marquis & Shi, 2012

• frequency and unit grouping:
  Given streams of alternating frequent and infrequent syllables, Japanese 8 month olds parse them as infreq-freq, Italian 8 month olds as freq-infreq
  Gervain et al., 2008
knowledge relevant to grammar

Familiarization Phase

This one is *blickish*. Do you like the *blickish* one? This one is a *blick*. Do you like the *blick*? Look here. Look at this. Do you like this?

Test Phase

“See what I have?”

**Result**: greater attention to novel-property item, only Adjective condition

**Interpretation**: hearing adjective guides attention to property of object (more than object kind).

*Waxman, 1999: 13 month olds (with 3+ words)*
The view from cognitive psychology: human infants and speech learning

the **What:** In their first year, infants learn about their native language:

- Its phonetic categories (roughly, its consonants and vowels)
- Some (50? 500?) frequent sequences / chunks
- Some meaningful aspects of some words
- Bits and pieces relevant to grammatical word order

the **How:** What they have to work with:

- Skills of intention reading, & so some comprehension
- A medium-sized corpus
- Capacity for grouping / chunking
- Ears
- Sounds, faces, and some innate links between them
The baby’s dataset

(LENA foundation, 2010)
Words heard by child, per day, by age in months

cumulative word tokens:
2 months, 0.75 mil wds; 4 mos, 1.5 mil wds; 6 mos, 2.25 mil wds

Google: 87 000 hours = 24 years of 10-hour days
how many very-frequent words do infants hear?

<table>
<thead>
<tr>
<th>freq</th>
<th>types with freq ≥ n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>about 1500 words</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Lots of short sentences

Words per utterance, Brent corpus
How much exposure needed to retain a consonant contrast?

12 sessions over 4 weeks, starting at 9 months; 4 talkers (men & women). Mandarin speaker talks to, plays with infant 15 min; book reading 10 min. About 2150 instances (total) of /ɕi/, /tɕʰi/.

Conditioned headturn test on /ɕi/, /tɕʰi/.

Kuhl, Tsao, & Liu 2003
Grouping / chunking

Language A
Torno a casa con le bici cariche di frutta in biliaco sulla sella.
La zia Carola si è esibita in una fuga colla bici verde.
Se porti il melo sulla bici forse cali un po’ di chili.
La bici ha subito un danno dentro la casa del capo di Lara.
La cavia Bida è in fuga da casa per aver giocato con le biliace blu.
La biscia in lenta fuga dal giardino capita in casa mia.
Il tuo melo arcana fuga l’afa che debilita la folla.
[ … ]

fuga, melo, bici, casa: occur 6 times each.
Lang. A: fu, ga, me, lo only occur in fuga, melo
bi, ca occur in other contexts; thus p(sa | ca) = .33
Lang. B: bi, ci, ca, sa only occur in bici, casa, […]

Result: preference for high trans. prob. words, p<.001
[b] and [w] differ, in part, in how speedy the transition is

  Fast transition: [b]
  Slow transition: [w].

But “fast” vs “slow” are not absolute; they depend on speaking rate (or syllable duration)

  80 ms duration syllable: transition 16 ms = b, 40 ms = w
  296 ms duration syllable: transition 40 ms = b, 64 ms = w
Having ears

Eimas & Miller 1980

[b] and [w] differ, in part, in how speedy the transition is

Fast transition: [b]
Slow transition: [w].

But “fast” vs “slow” are not absolute; they depend on speaking rate (or syllable duration)

5 groups:
- short syll, 16ms -> 40ms [b-w] .... +4.0 *
- short syll, 40ms -> 64 ms [w-w] -2.5
- long syll, 16 ms -> 40 ms [b-b] -3.8
- long syll, 40ms -> 64 ms [b-w] .... +4.3 *
- control: no change -4.6
sounds and faces

*Kuhl & Meltzoff, 1982*

One face says /i/ … /i/ …
Other face says /a/ … /a/ …
Infant hears either /i/ or /a/
19-week-olds (~4.4 months)
Looking to match: 73.6%, 24/32 Ss

Extension to 2 month olds: *Patterson & Werker 2003*


/a/ vs /u/, /i/ vs /u/, /i/ vs /wi/
sounds and faces

Chen, Striano, & Rakoczy, 2004, 1- to 7-day-olds

Model says: aaahhh ... aaahhh ... (4x / trial, 8 trials); then, model says: mmmm ... mmmm ... (4x, 8 trials) {or reverse order}

Q: does infant make like mouth movements?

Figure 1a  Mouth opening.

Figure 1b  Mouth clenching.
sounds and faces

Chen, Striano, & Rakoczy, 2004

13 infants who kept eyes closed:  

<table>
<thead>
<tr>
<th>Model</th>
<th>Opening</th>
<th>Clutching</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/</td>
<td>9.5</td>
<td>5.1</td>
</tr>
<tr>
<td>/m/</td>
<td>1.5</td>
<td>10.4</td>
</tr>
</tbody>
</table>

6 infants who kept eyes open:

<table>
<thead>
<tr>
<th>Model</th>
<th>Opening</th>
<th>Clutching</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/</td>
<td>12.3</td>
<td>1.5</td>
</tr>
<tr>
<td>/m/</td>
<td>7.5</td>
<td>7.2</td>
</tr>
</tbody>
</table>

All model /a/ vs /m/ comparisons p < .05
Language differentiation from visual information

Weikum, Vouloumanos, Navarra, Soto-Faraco, Sebastián-Gallés, & Werker, 2007

4, 6, 8 months old, English background

Videos of 3 bilingual French/English speakers reading sentences; No audio presented to infants

1. Show clips from each speaker, 1 language, till visual habituation.
2. Switch language (expt) or not (control).
Language differentiation from visual information

Weikum, Vouloumanos, Navarra, Soto-Faraco, Sebastián-Gallés, & Werker, 2007

4, 6, 8 months old, English background

4, 6 mos: increase to switched lang., decrease to same language
8 mos: no significant difference

(But: Fr/Eng bilingual 6, 8mos did dishabituate to change.)
So, what information have infants got, beyond the acoustic signal?

• **a lexicon**, or at least the start of one. Meanings, at least for some words, that may anchor some phonetic variation. And word-forms, that may provoke “acquired equivalence” effects for speech sounds.

• **your face**, as you talk, and the correlated acoustic data. [plus some complex “see you, I do” intermodal skills]