Infant Speech Perception

(Part of the Prelinguistic Period)

A. Aslin & Pisoni (1980) describe four theoretical approaches
   1. Perceptual Learning Theory (behaviorist)
   2. Attunement Theory (constructionist)
   3. Universal Theory (innatist)
   4. Maturational Theory (restructuring)

B. Predictions

<table>
<thead>
<tr>
<th>Theory</th>
<th>Ability at birth</th>
<th>Non-native sounds</th>
<th>Role of experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptual Learning Theory</td>
<td>none</td>
<td>never</td>
<td>all</td>
</tr>
<tr>
<td>Attunement Theory</td>
<td>basic</td>
<td>basic</td>
<td>non-basic</td>
</tr>
<tr>
<td>Universal Theory</td>
<td>all</td>
<td>all</td>
<td>none</td>
</tr>
<tr>
<td>Maturational Theory</td>
<td>some?</td>
<td>as they mature</td>
<td>none</td>
</tr>
</tbody>
</table>

C. Methods

1. **High Amplitude Sucking (HAS)**
   a. measures infant sucking rate during exposure to auditory stimuli in three phases
      i. acquisition phase–infants increase their sucking rate during initial exposure
      ii. habituation phase–point where experimenter might change the auditory stimulus
      iii. dishabituation phase–period where infants react to the change/continued stimulus
   b. used with infants from birth to 6 months
2. Heart Rate (HR)
   a. measures infants heart rate during exposure to auditory stimuli
   b. used with infants from birth to 8 months
3. Visually Reinforced Infant Speech Discrimination (VRISD)
   a. measures head turn to anticipated visual reinforcer synchronized with auditory stimuli
   b. used with infants between 6 and 18 month

D. Stimuli

1. Voice Onset Time (VOT)–time between consonant release and voicing
   a. 0 msec for voiceless unaspirated consonants
   b. - for pre- or fully voiced consonants
   c. + for aspirated consonants
2. critical VOT times vary between **languages**
   a. English makes voiced/voiceless distinction ~ +25 msec (Lisker & Abramson 1967)
   b. Spanish makes its voiced/voiceless distinction ~ +10 msec
3. speakers make a “categorical” distinction between VOT stimuli

E. Results

1. Eimas, Siqueland, Jusczyk & Vigorito (1971) demonstrate categorical perception in infants
   a. Used HAS technique with 1 and 4-month-olds exposed to English
2. Eilers, Gavin & Wilson (1978) demonstrate differences between English and Spanish infants
   a. Used VRISD technique with 6-8-month-olds to allow for “experience”
   b. Results
      i. English infants correct on 92% of English stimuli and 46% of Spanish stimuli
      ii. Spanish infants correct on 86% of English stimuli and 80% of Spanish stimuli
   c. Conclude English contrast is basic; Spanish contrast is learned
3. Kuhl & Miller (1975) demonstrate chinchillas also make the English VOT discrimination
F. Interpretation
1. Mammalian auditory system naturally discriminates between certain stimuli
2. Human infants lose the ability to make some discriminations around 10 months (Werker & Tees 1984)

<table>
<thead>
<tr>
<th>Age</th>
<th>Hindi</th>
<th>Salish</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-8 months</td>
<td>95%</td>
<td>80%</td>
</tr>
<tr>
<td>8-10 months</td>
<td>68%</td>
<td>52%</td>
</tr>
<tr>
<td>10-12 months</td>
<td>20%</td>
<td>10%</td>
</tr>
</tbody>
</table>

3. Still lack good measures of infant discriminations of non-English sounds
4. Phonetic discrimination does NOT entail phonemic perception
5. Infant is remarkably adapted for speech perception

Infant Speech Production (Also part of the Prelinguistic Period)

A. Infant vocal tracts develop from birth to 8 months
   1. They have shorter vocal tracts than adults
   2. They have a shorter pharynx
   3. Their oral cavity is relatively wider and flatter (they lack teeth)
   4. They breathe through the nose; oral breathing begins around 6 months
B. Infant vocal tracts have distinct acoustic properties until 6 months
C. At 6 months infants enter an “Expansion Stage” of vocalization (Oller 1980), including:
   1. Fully Resonant Nuclei (FRN)—vowellike vocalizations
   2. Marginal Babbling (MB)—lacks reduplication, not regularly timed

Phonological Development

A. Jakobson’s Theory (Child Language, Aphasia and Phonological Universals 1941–in German)
   1. most famous, but now considered disproven
   2. based on an innate set of universal features

   The child possesses in the beginning only those sounds which are common to all the languages of the world, while those phonemes which distinguish the mother tongue from the other languages of the world appear only later.

3. predicted a discontinuity between babbling and a child’s first words
4. recognized an interaction between the “particularist spirit” and the “unifying force”
Accordingly, we recognize in the child’s acquisition of language the same two mutually opposed but simultaneous driving forces that control every linguistic event, which the great Genevan scholar (de Saussure) characterizes as the “particularist spirit”, on the one hand, and the “unifying force” on the other. The effects of the separatist spirit and the unifying force can vary in different proportions, but the two factors are always present. (Jakobson 1941/68: 16)

5. predicted an invariant developmental sequence (contradicting #4!)
6. based on production data, mostly Slavic languages (Czech, Bulgarian, Russian, Polish, Serbo-Croatian)
7. predicts the child’s sounds are constrained by her underlying linguistic system, not motor articulation
   e.g., Ament (1899) daughter initially varied between [k] and [t], later [k] → [t]
8. linguistic laws regulate the acquisition of phonemic contrasts (Table 6.23, p. 192)
   1. CV opposition syllables, e.g. pa, ma
   2. nasal/oral contrast m/b
   3. labial/dental contrast m/n
   4. narrow/wide contrast a/i
   5. front/back contrast i/u

9. Jakobson derived his acquisition predictions from a study of the world’s languages
   The laws of irreversible solidarity (implicational, Table 6.24, p. 194)

Consonants
   1. The existence of fricatives implies the existence of stops
   2. Back consonants (palatals and velars) imply front consonants (labials and dentals)
   3. If a language has one fricative, it will be /s/
   4. An affricate/stop contrast implies a fricative within the same series

Vowels
   5. A vowel contrast with the same aperture implies a contrast with a narrower aperture,
      e.g. /æ/ vs. /a/ implies /a/ vs. /e/.
   6. A rounded vowel contrast implies the same contrast between unrounded vowels
      e.g. /u/ vs. /o/ implies /i/ vs. /e/.

10. Jakobson’s predictions are incomplete; when is the first liquid acquired?
11. Jakobson, himself, confuses the acquisition of sounds with the acquisition of contrasts
12. The form of children’s oppositions are influenced by the structure of the adult phonology
13. Jakobson recognized an abstract level of representation
   a. Child’s [t]-[?] phonetic distinction represents an underlying dental/velar contrast
   b. Child’s [papa]-[dada] distinction represents an underlying /papa/-/dada/ contrast

14. Jakobson proposed the Principle of Maximal Contrast to explain phonemic differentiation
    ‘This sequence obeys the principle of maximal contrast and proceeds from the simple and undifferentiated to the stratified and differentiated.’ (p. 68; see Table 6.25, p. 196)
15. Jakobson finds independent evidence for his principle in
    a. data from language acquisition
    b. data from language disorders

B. Data (Shvachkin 1948/73) - Perceptual
   1. Method ‘... it was necessary to work out a method which would correspond to the actual course of development of phonemic perception in the child. This problem proved to be quite
difficult and required a great deal more time and effort than the actual study of the facts themselves.'

a. used nonsense pairs (‘bak’, ‘mak’) to avoid linguistic effects
b. used the novel words as names for geometric shapes (wooden pyramids, cones)
c. used a ‘clinical method’ to observe children’s responses (Table 6.20, p. 181)
   i. Day 1 teach a novel word, e.g., ‘bak’
   ii. Day 2 introduce a novel word, e.g., ‘zub’
   iii. test for non-minimal opposition (whole syllable), e.g., ‘bak’ vs. ‘zub’
   iv. teach a new novel word, e.g., ‘mak’
   v. test for new non-minimal opposition, e.g., ‘mak’ vs. ‘zub’
   vi. test minimal opposition, e.g., ‘bak’ vs. ‘mak’
d. used six tests of children’s ability (Table 6.21, p. 182); criterion was 3/6
   i. pointing to the object
   ii. giving the object
   iii. placing the object
   iv. finding the object
   v. relating one object to another
   vi. substitution of objects
e. Subjects–14 girls, 5 boys aged 1;3-1;9 (roughly the one-word stage)

2. Results (Table 6.22, p. 183)
a. vowel contrasts
   a vs. other vowels
   i-u, e-o, i-o, e-u
   i-e, u-o

b. presence/absence of consonant
   bok-ok, vek-ek

c. sonorant/obstruent
   m-b, r-d, n-g, j-v

d. palatalized/non-palatalized consonants
   n-n’, m-m’, b-b’, v-v’, z’ l-l’, r-r’
e. sonorant distinctions
   nasals vs. liquids and /j-/; nasals; liquids
f. sonorant/non-labial fricatives
   m-z, l-x, n-ž

g. labials/non-labials
   b-d, b-g, v-z, f-x

h. stops/spirants
   b-v, d-ž, k-x,

i. velars/non-velars
   d-g, s-x, š-x

j. voiced/voiceless
   p-b, t-d, k-g, f-v, s-z, š-ž

k. children showed rapid phonemic perceptual development between 1;0 and 2;0

C. Braine (1974a) - Production
1. studied Jonathan’s first words
   ‘that, there’ [da ~ dʌ ~ dæ ~ dɛ]
   ‘see’ [di]
   ‘no’ [do]
   ‘juice’ [du]
   ‘hi’ [ʔai]
2. hypothesis—the d/ʔ opposition is non-contrastive
3. taught two new words: ‘cat’ or ‘food’ [i] and a toy [dai] should result in contrast between
di/i and dai/ʔai
4. J changed new words to [di] and [da ~ dʌ] respectively

D. Methods of Phonological Analysis
1. Phone classes and phone trees (Ferguson & Farwell 1975)
a. phone class—words that begin with the same sounds, e.g. (6.9) Phone classes for T:
   [pʰ]  pat, please, pretty, purse

   ‘The notion of ‘phone class’ here is similar to the notion of ‘phoneme’ of American
structuralism, in that it refers to a class of phonetically similar speech sounds believed to contrast with other classes, as shown by lexical identification.’ (Ferguson & Farwell 1975: 425)

b. phone tree—development of a phone class over time (Figure 6.2, p. 202)
‘If successive phone classes did not contain the same word but were related to phone classes which did, dotted lines were drawn connecting them. For example in T’s /m/ class:

\[
\begin{align*}
/m/ & \quad \text{mama} \\
/m/ & \quad \text{milk} \\
/m/ & \quad \text{milk, mama} \quad (\text{Ferguson & Farwell 1975: 424})
\end{align*}
\]

c. problems
i. the analysis is hard to do
ii. the method is extremely sensitive to surface variability of lexical items
iii. sensitive to the level of phonetic transcription
iv. the method leads to a measurement sequence—lexically specific development

2. Phonetic inventories and phonological contrasts (Ingram 1981a, 1988)
a. establish the child’s **phonetic inventory**—the sounds used in the child’s words
   i. use a broad phonetic transcription to minimize transcriber variability
   ii. select a typical phonetic type for each lexical type
      a. select the phonetic type that occurs in the majority of the phonetic tokens
         e.g. T at VI (Ferguson & Farwell 1973: 34)
         \[
         \begin{align*}
         \text{pat} & \quad \text{p}^{h} \text{æt} (3 \text{ tokens}) \\
         \text{p}^{h} \text{æ} & \\
         \end{align*}
         \]
      b. select the phonetic type that shares the most segments with the other phonetic types
         \[
         \begin{align*}
         \text{bounce} & \quad \text{b} \text{æ} \\
         \text{b} \text{æ} & \\
         \text{bwæ} & \\
         \end{align*}
         \]
c. for two phonetic types, select the one that is not correctly pronounced

_book_  $\approx$  $\alpha g$

_bʌʃ_

d. if the other steps do not work, select the first phonetic type listed

_paper_  $\approx$  $\phi e t \theta$

_bæd u_

iii. analyze word-initial and word-final consonants separately

iv. determine the criterion frequency for the sample (Table 6.28, p. 205)

<table>
<thead>
<tr>
<th>Vocabulary Size</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of lexical types</td>
<td>marginal</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------</td>
</tr>
<tr>
<td>1-37</td>
<td>1</td>
</tr>
<tr>
<td>38-67</td>
<td>1</td>
</tr>
<tr>
<td>68-87</td>
<td>2</td>
</tr>
<tr>
<td>88-112</td>
<td>2,3</td>
</tr>
</tbody>
</table>

v. divide the child’s sounds into (6.12, p. 206)

a. **marginal**: if the sound does not meet the frequency criterion, (d-)

b. **used**: if the sound meets the frequency criterion, n-

c. **frequent**: if the sound is twice the frequency criterion, *b-

e.g. (6.12) T’s phonetic inventory at session VI

<table>
<thead>
<tr>
<th>Initial</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>n-</td>
<td>-g</td>
</tr>
<tr>
<td>*b-</td>
<td>(d-)</td>
</tr>
<tr>
<td>*p-</td>
<td>t-</td>
</tr>
<tr>
<td>(Φ-)</td>
<td>s-</td>
</tr>
<tr>
<td>(w-)</td>
<td>-ç</td>
</tr>
</tbody>
</table>

b. determine the child’s **patterns of substitution** (6.13, p. 206)

i. child matches adult target if consonants in over 50% of the child’s lexical types match

ii. child has a marginal match if there is only one lexical type with the correct consonant

e.g. (6.13)

<table>
<thead>
<tr>
<th>Lexical types</th>
<th>Proportion correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>C C 0 C ph</td>
<td>3/5</td>
</tr>
<tr>
<td>b- baby, ball, book, bounce, bye-bye</td>
<td>2/3</td>
</tr>
<tr>
<td>p- paper, pat, purse</td>
<td>1/1</td>
</tr>
<tr>
<td>d- dog</td>
<td>1/1</td>
</tr>
<tr>
<td>t- tea</td>
<td>1/1</td>
</tr>
<tr>
<td>s- cereal</td>
<td>0/1</td>
</tr>
<tr>
<td>ç cheese</td>
<td>0/1</td>
</tr>
<tr>
<td>h- hi</td>
<td>1/1</td>
</tr>
<tr>
<td>w- rock</td>
<td>0/1</td>
</tr>
</tbody>
</table>
c. determine the child’s **phonological contrasts** (6.14, p. 207)
   A sound is considered part of the child’s phonological system when
   i. it is frequent, or
   ii. it is used, and it appears as a match or substitute (207)

   e.g. (6.14) T’s phonology for initial consonants
   n-
   b- (d-)
   p- t-
   s- ç- ñ- (h-)
   (w-)

   d. Ingram’s method can also be applied longitudinally (Table 6.29, p. 208)

E. Characteristics of early phonological development (Ferguson & Farwell 1975) ‘Lexical Parameter’
   1. early phonological development is heavily influenced by the properties of individual words
   extended lexical oppositions, e.g., T only used [m-] in ‘mama’ and [n-] in ‘no’
   2. find gradual spread of contrasts to other words
   gradual lexical spread, e.g., T’s [t-]
   3. sudden emergence of some sounds
   sudden emergence, e.g., T’s [p-]
   4. the contrast between **stable** and **variable** word forms, although see Ingram’s assessment

5. **phonological idioms**–pronunciations which are superior to later pronunciations
   e.g., Hildegard Leopold’s ‘pretty’: [plti] (whispered) at 1;9 and [bldi] at 1;10
   6. children focus on words that contain sounds within their phonological system (salience), and
   avoid words with sounds outside of their system
   7. variation?

F. Cross-linguistic comparison (Pye, Ingram & List 1987)
   (6.17) basic phonetic inventories of K’iche’ and English
<table>
<thead>
<tr>
<th>K’iche’ (5 children)</th>
<th>English (15 children)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m) n (b’) p t tf k</td>
<td>(m) n b d (g) p t k</td>
</tr>
<tr>
<td>w</td>
<td>x</td>
</tr>
</tbody>
</table>

   Functional load–frequency of lexical types with specified sounds, e.g. Table 6.32 (p. 210)
   The rank-order frequencies for initial consonants common to K’iche’ and English

<table>
<thead>
<tr>
<th>Sounds</th>
<th>Language</th>
<th>K’iche’</th>
<th>English</th>
<th>K’iche’</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/tf</td>
<td>w</td>
<td>k</td>
<td>p</td>
<td>t</td>
</tr>
<tr>
<td>K’iche’</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>English</td>
<td>9</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

   Conclusion: articulatory and frequency effects are less important than functional load
Pye, Mateo, Pfeiler, Lopez & Gutierrez (2006) used data from five Mayan languages to test universalist and particularist accounts of phonological development. In the universalist camp, theorists have pointed to markedness constraints (Gnanadesikan 2004) and motorically based rules for syllable formation (MacNeilage, Davis & Matyear 1997) to account for similarities across children’s phonologies. The particularists counter that “the wide range of individual differences in this period [first words, CP] continues to resist coherent formulation in terms of universal phonetic or phonological principles, even across children acquiring a single language (Vihman 2006).”

We recorded three children learning the Mayan languages Yucatec, Chol, Q’anjob’al, Mam and K’iche’. The children were between the ages of 1;8 and 3;11 and had between 29 and 296 words that they produced in the recordings. We focused exclusively on the initial consonants in the children’s words. Using Ingram’s method of phonetic inventories and phonological contrasts, we found the children produced the initial consonants shown in the following table.

Table 7. Mayan Child Phonologies

<table>
<thead>
<tr>
<th>Language</th>
<th>ARM</th>
<th>SAN</th>
<th>DAV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yucatec</td>
<td>m n p t* (ch) k b d</td>
<td>m* n p* t* ch* k* b (d) ch’</td>
<td>m* n p* t* ch k b* (d) (t’) (ts’)</td>
</tr>
<tr>
<td>Ch’ol</td>
<td>m* n p</td>
<td>m* n p ty (ts) ch* (k) b</td>
<td>m* n p ty ch* k* b</td>
</tr>
<tr>
<td>Q’anjob’al</td>
<td>m* n* p t* (ts) (ch) (k)</td>
<td>m* n* p* t* ch* k*</td>
<td>m* n* p* t* ch* k*</td>
</tr>
<tr>
<td>Mam</td>
<td>m* n* p t* ch* k* (d) (ch’)</td>
<td>m* n* p* t* ch* k*</td>
<td>m* n* p* t* ch* k*</td>
</tr>
<tr>
<td>K’iche’</td>
<td>n* p t* ts ch* k* b</td>
<td>m* n* p* t* ch k* (q) b (k’)</td>
<td>m* n* p* t* ch* k*</td>
</tr>
</tbody>
</table>

Within each language, there are some children who produce initial consonants that are missing from the consonant sets produced by the other children. We used the chi-square method to quantitatively assess the degree of within language variation. We grouped the children’s responses by manner of articulation to adjust for the paucity of their responses. With this adjustment we found significant results for only Q’anjob’al and Mam, suggesting the Yucatec, Ch’ol and K’iche’ children produced initial consonants with similar frequencies:
Yucatec: $\chi^2 = 1.52$, ns  
Ch’ol: $\chi^2 = 5.84$, ns  
Q’anjob’al: $\chi^2 = 20.52$, $p < .05$  
Mam: $\chi^2 = 23.47$, $p < .01$  
K’iche’: $\chi^2 = 9.53$, ns

A mixed result supports the particularist contention that individual children can adopt their own strategy of consonant production, although our data raises the possibility that children acquiring some languages have initial phonologies that are more variable than children acquiring other languages.

To further understand the variation in the children’s consonant frequencies we employed a linear discriminant analysis to assess the consonant data. Discriminant analysis tries to create a set of linear functions that classify the children as speakers of Yucatec, Ch’ol, Q’anjob’al, Mam, or K’iche’. The analysis revealed four functions, but one was very dominant in capturing 94% of the variation between groups. This function was successful in grouping the subjects by language.

![Figure 2. Discriminant Analysis Function](image)

We tested the influence of the type frequency of consonants on the children’s phonological development by extracting samples of child-directed adult speech from each language. Once again we combined the children’s consonant type frequencies for each language, and then used the chi-square method to compare the children’s combined consonant frequencies with those of the adults. Our results reveal significant differences between the children’s consonant productions and the adult frequency for Yucatec, Q’anjob’al and Mam. The result for K’iche’ confirms a result that Pye et al. (1987) reported in a previous study of phonological development in K’iche’, but shows that the effect of ambient frequency on children’s phonology is limited to Ch’ol and K’iche’.
