Statistical inferences and linguistic knowledge in early phonological acquisition

Sharon Peperkamp

Emmanuel Dupoux, Rozenn Le Calvez, Jean-Pierre Nadal, Katrin Skoruppa

Workshop on Phonological Systems and Complex Adaptive Systems, July 4-6 2005, Lyon
Early language acquisition: evidence for statistical inferences

• **segments**: distribution of tokens within the acoustic space (Maye, Werker & Gerken 2002)
  – exposure: monomodal or bimodal [ta]-[da] continuum
  – testing: discrimination of [ta]-[da]

• **phonotactics**: distribution of segments in onsets vs. codas (Chambers, Onishi & Fischer 2002; Saffran & Thiessen 2003)
  – exposure: CVC syllables with different sets of onset and coda consonants
  – testing: listening time for new syllables in which the consonant phonotactics are respected or not

• **word segmentation**: transitional probabilities (Saffran, Aslin & Newport 1996)
  – exposure: continuous speech stream consisting of 4 trisyllabic non-words (tupirobidakupadotigolabubidaku...)
  – testing: listening time for words (bidaku) and part-words (kupado)
Early language acquisition: evidence for linguistic inferences

• **segments**: generalization within a natural class (Maye & Weiss 2003)
  – exposure: monomodal or bimodal [ta]-[da] continuum
  – testing: discrimination of [ka]-[ga]

• **phonotactics**: better learning in case of natural classes than unnatural classes (Saffran & Thiessen 2003)
  – natural classes: /p,t,k/ in onsets, /b,d,g/ in codas
  – unnatural classes: /p,d,k/ in onsets, /b,t,g/ in codas
This talk

• Examine the respective roles of statistical and linguistic interferences for the acquisition of underlying representations

• Two complementary approaches
  – *modeling*: simulation on phonetically-transcribed speech
  – *experiments*: artificial language-learning paradigm
Acquisition of underlying representations

- establish *phoneme* inventory:

<table>
<thead>
<tr>
<th>Spanish</th>
<th>bilabial</th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>postalveolar</th>
<th>palatal</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td></td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td>fricatives</td>
<td>f</td>
<td>v</td>
<td>θ</td>
<td>s</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>nasals</td>
<td>m</td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>trills</td>
<td></td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flaps</td>
<td></td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tf</td>
<td>dʒ</td>
</tr>
<tr>
<td>laterals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
<td>ŋ</td>
</tr>
</tbody>
</table>
Acquisition of underlying representations

- establish *phoneme* inventory:

- on the basis of a *segment* inventory:

<table>
<thead>
<tr>
<th>Spanish</th>
<th>bilabial</th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>postalveolar</th>
<th>palatal</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td></td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td>fricatives</td>
<td>β</td>
<td>f</td>
<td>v</td>
<td>θ</td>
<td>δ</td>
<td>s</td>
<td>z</td>
</tr>
<tr>
<td>nasals</td>
<td>m</td>
<td>m̃</td>
<td>n</td>
<td></td>
<td></td>
<td>j̃</td>
<td>η</td>
</tr>
<tr>
<td>trills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>r</td>
<td></td>
</tr>
<tr>
<td>flaps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l̃</td>
<td></td>
</tr>
<tr>
<td>affricates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tʃ</td>
<td>ɗʒ</td>
</tr>
<tr>
<td>laterals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ʎ</td>
</tr>
</tbody>
</table>
Acquisition of underlying representations

- establish *phoneme* inventory:
- on the basis of a *segment* inventory:

<table>
<thead>
<tr>
<th>Spanish</th>
<th>bilabial</th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>postalveolar</th>
<th>palatal</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops</td>
<td>p b</td>
<td>t d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricatives</td>
<td>ß f v δ s z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasals</td>
<td>m nj</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trills</td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flaps</td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tʃ dʒ</td>
</tr>
<tr>
<td>laterals</td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ʎ</td>
</tr>
</tbody>
</table>
Acquisition of segment inventory

- **age:** between 6-12 months (Polka & Werker 1994; Werker & Tees 1984)

- **method:** prototype formation (Kuhl 1991; Kuhl *et al.* 1997; Maye, Werker & Gerken 2002)
Acquisition of *phoneme* inventory

- **age:** unknown

- **method:**
  - semantics
    - [el̥disko] ‘the disk’
    - [mi̥disko] ‘my disk’
  - **distributional analysis**
    - [ɔ]: intervocally
    - [d]: elsewhere

- **objective:** test the feasibility of the distributional mechanism
  - **algorithm:** look for complementary distributions of segments
A statistical algorithm

• Problems with basic algorithm
  – not robust to noise (production and/or perception errors)
  – fails to detect optional rules

• Solution: look for near-complementary distributions
  – for each segment, list the contexts in which it appears
  – for each pair of segments, compare the distributions of their contexts, by means of the Kullback-Leibler dissimilarity measure:

\[
m_{KL}(s_1, s_2) = \sum_c \left( P(c|s_1) \log \left( \frac{P(c|s_1)}{P(c|s_2)} \right) + P(c|s_2) \log \left( \frac{P(c|s_2)}{P(c|s_1)} \right) \right)
\]

s: segment   c: context
A statistical algorithm

• For segment pairs with a KL number above some threshold, determine the default phone
  – the default segment is more frequent and appears in more contexts than the allophone
  – criterion of relative entropy:

\[ s_d = \min_s \left[ \sum_c P(c|s) \log \frac{P(c|s)}{P(c)} \right] \]

\( s_d \): default segment \hspace{1cm} c: context
Validation on artificial corpora

- 46 phonemes with equal relative frequencies
- 1 phoneme has an allophone in 8 contexts
- utterances composed of random strings
Validation on artificial corpora

- 46 phonemes with equal relative frequencies
- 1 phoneme has an allophone in 8 contexts
- utterances composed of random strings
A statistical + linguistic algorithm

• Problem with statistical algorithm:
  – false alarms due to phonotactics (e.g. French: [ʃ] only before vowels (pluie), [µ] only before consonants (peur))

• Solution: add a linguistically motivated filter
  – default phone and allophone are phonetically close
  – the context of a rule spreads a phonetic feature onto its targets
A statistical + linguistic algorithm

• Define each segment as a numerical vector encoding five articulatory properties (place, sonority, voicing, nasality, lip rounding)

• Criteria for detecting false alarms
  – there is a segment between default segment and the allophone:
    \[
    \exists s, \forall i \in \{1, \ldots, 5\}, v_i(s_a) \leq v_i(s) \leq v_i(s_d)
    \]
    \[
    \text{or } \forall i \in \{1, \ldots, 5\}, v_i(s_d) \leq v_i(s) \leq v_i(s_a)
    \]
  – the allophone is more distant from its contexts then the default segment:
    \[
    \exists i \in \{1, \ldots, 5\}, \sum_{s \in C[s_a]} (v_i(s_a) - v_i(s)) > \sum_{s \in C[s_d]} (v_i(s_d) - v_i(s))
    \]
Test on natural corpus

• CHILDES corpus
  – 42,000 short utterances of French parents to their children
  – transcribed phonemically

• Implementation of two allophonic rules:
  – palatalisation of /k/ and /g/ before /i,y,♭,♯,e,★,j,◊/
  – devoicing of /r,l,m,n,♭,♯,j/ before /p,t,k,f,s,▪/

• Corpus statistics:
  – Total number of segments: 35 default segments + (2+7) allophones = 44
  – Total number of segment pairs: 946
$0 \leq Z \leq 1$:
hits: 7
misses: 2

/●,●/, relative frequencies .02 and .01%

Peperkamp et al., submitted
number of FA’s at Z=1: 129 (=13.6%)
$0.5 \leq Z \leq 1$:
- max. number of hits
- no FA’s

Peperkamp et al., submitted
Summary

• Allophonic rules can be discovered in the absence of lexical knowledge on the basis of distributional information

• Linguistic knowledge concerning the nature of phonological rules is sufficient to discard false alarms

• Possible extensions
  – rule interaction
  – linguistic filter based on acoustic rather than articulatory distance (cf. Mielke 2005)

• Next step: test if infants are sensitive to complementary distributions and if it matters if the allophonic groupings are natural or not (work in progress with Jim Morgan)
Experiments

• Test if adults use linguistic knowledge when learning novel phonological rules

• Method: artificial language learning paradigm
  – natural versus unnatural allophonic rule
Experiment 1

- Two artificial languages:

<table>
<thead>
<tr>
<th></th>
<th>Language A</th>
<th>Language B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>stops</strong></td>
<td>/p t k b d g/</td>
<td>/p t k/</td>
</tr>
<tr>
<td><strong>fricatives</strong></td>
<td>/f s j/</td>
<td>/f s j v z z/</td>
</tr>
</tbody>
</table>
## Experiment 1

- Two artificial languages:

<table>
<thead>
<tr>
<th></th>
<th>Language A</th>
<th>Language B</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops</td>
<td>/ p t k b d g/</td>
<td>/ p t k/</td>
<td>/ p t k b d g/</td>
</tr>
<tr>
<td>fricatives</td>
<td>/ f s ŋ/</td>
<td>/ f s ŋ v zʒ/</td>
<td>/ f s ŋ v zʒ/</td>
</tr>
</tbody>
</table>
## Experiment 1

- **Two artificial languages:**

<table>
<thead>
<tr>
<th></th>
<th>Language A</th>
<th>Language B</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops</td>
<td>/p t k b d g/</td>
<td>/p t k/</td>
<td>/p t k b d g/</td>
</tr>
<tr>
<td>fricatives</td>
<td>/f s s/</td>
<td>/f s s v z 3/</td>
<td>/f s s v z 3/</td>
</tr>
</tbody>
</table>

- **Two natural allophonic rules:**
  
  Language A: intervocalic *fricative* voicing  
  Language B: intervocalic *stop* voicing

- **Determinant + noun phrases:**

  - *nel* ‘two’
  - *ra* ‘three’

  Nouns begin with a stop or fricative
Exposure: phrase-picture pairings

<table>
<thead>
<tr>
<th>Language A</th>
<th>Language B</th>
</tr>
</thead>
<tbody>
<tr>
<td>nel pemuš</td>
<td>nel pemuš</td>
</tr>
<tr>
<td>ra pemuš</td>
<td>ra bemuš</td>
</tr>
<tr>
<td>nel fo3am</td>
<td>nel fo3am</td>
</tr>
<tr>
<td>ra vo3am</td>
<td>ra fo3am</td>
</tr>
</tbody>
</table>
### Exposure: phrase-picture pairings

<table>
<thead>
<tr>
<th>Language A</th>
<th>Language B</th>
</tr>
</thead>
<tbody>
<tr>
<td>nel bovi</td>
<td>nel povi</td>
</tr>
<tr>
<td>ra bovi</td>
<td>ra bovi</td>
</tr>
<tr>
<td>nel fulek</td>
<td>nel vulek</td>
</tr>
<tr>
<td>ra vulek</td>
<td>ra vulek</td>
</tr>
</tbody>
</table>
Test I: phrase production, known items

Language A: ra pemus

Language B: ra bemus
Test II: phrase production, new items

Language A:  ra pura
Language B:  ra bura
Experimental details

• Exposure phase (15 minutes):
  – 8 lexical items
    • 4 stop-initial
    • 4 fricative-initial
  – each item appears in 2 phrases (one with nel one with ra), repeated 16 times each

• Test phase:
  – 8 old items
  – 32 novel items
Results

N = 16
Results

% response "no change"

Old items
Novel items

N = 16

Phonemic
Allophonic
Experiment 2

- Same segment inventories, different phoneme inventories:

<table>
<thead>
<tr>
<th></th>
<th>Language A</th>
<th>Language B</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops</td>
<td>[p t k b d g]</td>
<td>[p t k b d g]</td>
</tr>
<tr>
<td>fricatives</td>
<td>[f s (\tilde{v} z \tilde{3})]</td>
<td>[f s (\tilde{v} z \tilde{3})]</td>
</tr>
</tbody>
</table>

- Two unnatural rules:
  
  Language A: /z/ \(\rightarrow\) [t], /g/ \(\rightarrow\) [f], /p/ \(\rightarrow\) [3] / V_V
  
  Language B: /v/ \(\rightarrow\) [k], /\(\tilde{s}\)/ \(\rightarrow\) [b], /d/ \(\rightarrow\) [s] / V_V
Results

% response "no change"

Old items

N = 16
Results

% response "no change"

Old items

Novel items

N = 16

N = 16
Summary

• Adults can learn the distinction between phonemic and allophonic contrasts within 15 minutes of exposure to an artificial language, but only if the allophonic groupings are phonetically natural.

• Experiments 3 and 4:
  – as experiments 1 and 2, but with a perception rather than a production task.
Test I: phrase-picture matching, known items

Language A

Language B
Test II: phrase-picture matching, new items

Language A

Language B

nel pura

ra bura
Results exp. 3: natural rule

% pairings with novel object

Old items

- Phonemic
- Allophonic

N = 12

Peperkamp & Dupoux, to appear
Results exp. 3: natural rule

Peperkamp & Dupoux, to appear
Results exp. 4: unnatural rule

% pairings with novel object

Old items

N = 12

Peperkamp & Dupoux, to appear
Results exp. 4: unnatural rule

Peperkamp & Dupoux, to appear
Discussion

• Effect of phonetic naturalness with a production but not with a perception task

• Two possible explanations
  – task difference: free response in production, forced choice in perception
    → make the perception task harder
  – perception is not constrained by UG, but by a general algebraic learning system (Marcus et al, 1999)
    → validate the present results with pre-school children and, ultimately, with infants
Conclusion

• Both statistical and linguistic inferences seem necessary to acquire underlying representations

• Further research is necessary to
  – empirically demonstrate the presence of both types of inferences in infants
  – determine the precise nature of the linguistic inferences