The Early Lexical Bias Revisited: A Longitudinal Study of L1 Acquisition of Mandarin Chinese

LingLang Lunch Lite
Feb. 8th
Youtao Lu
Intro

• Lexical Bias:
  • for children at early stages of language acquisition, some lexical category, usually noun or verb, appears earlier and outnumbers other categories in children’s production.

• Universal Noun Bias
  • Earlier and more
  • A variety of languages

• Language Specific
  • Frequency, saliency and transparency
  • Mainly verb bias in Korean and Mandarin
Doubts

• Age is not well controlled both within and between studies
  • e.g. For Mandarin, Gentner’s 1;5 vs. Tardif’s 1;10
  • Bassano (1998): a transition at 1;8

• The bias of language and the bias of early production is tangled
The Study

• Mandarin
  • Conflicting Empirical Evidence
  • Issues and doubts remain unsolved

• To solve the problems
  • Age is not well controlled → Longitudinal approach
  • Language is biased → Set a baseline to eliminate it

• Research Questions
  • What bias exists in Mandarin children’s output (longitudinally)?
  • What is the relationship between adults’ input and children’s output?
  • Does the tentative removal of the bias of language change the answer to Q1?
• Beijing Chinese Early Acquisition (BJCELA) corpus

<table>
<thead>
<tr>
<th>Name (Initial letters)</th>
<th>CY</th>
<th>ZTX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>female</td>
<td>male</td>
</tr>
<tr>
<td>Age range</td>
<td>1;6-2;4</td>
<td>1;6-2;4</td>
</tr>
<tr>
<td>Selection of age points</td>
<td>every two months</td>
<td>every two months</td>
</tr>
<tr>
<td>Sessions at each age point</td>
<td>two</td>
<td>two</td>
</tr>
<tr>
<td>Total sessions analyzed</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

• Only count common nouns and main verbs (Tardif, 1996)
• Deleted: Repetition, utterances under request (“say ...”), quoted speech (nursery rhymes, songs, poems).
I. Cross-sectional Analysis
- Taken as independent observations
- Comparable to existing studies

Type: pairwise $t(11) = 2.316, p = 0.041$ (*). (Noun)
Token: pairwise $t(11) = -2.585, p = 0.025$ (*). (Verb)
Result 1

Nouns and Verbs in Adults' Input

Type: pairwise $t(11) = -1.259$, $p = .234$. (ns)
Token: pairwise $t(11) = -7.251$, $p = .000$ (***) (Verb)
II. Longitudinal: Noun/Verb Ratio

- Use the number of nouns to divide the number of verbs
- Ratio > 1: Noun bias

<table>
<thead>
<tr>
<th>Name: CY</th>
<th>Age</th>
<th>1;6</th>
<th>1;8</th>
<th>1;10</th>
<th>2;0</th>
<th>2;2</th>
<th>2;4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td>Type</td>
<td>1.56</td>
<td>1.24</td>
<td>0.94</td>
<td>1.35</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>Token</td>
<td>1.37</td>
<td>1.11</td>
<td>0.68</td>
<td>0.85</td>
<td>0.65</td>
<td>0.72</td>
</tr>
<tr>
<td>Adults</td>
<td>Type</td>
<td>1.02</td>
<td>0.97</td>
<td>1.02</td>
<td>1.08</td>
<td>1.04</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Token</td>
<td>0.49</td>
<td>0.51</td>
<td>0.44</td>
<td>0.44</td>
<td>0.45</td>
<td>0.38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name: ZTX</th>
<th>Age</th>
<th>1;6</th>
<th>1;8</th>
<th>1;10</th>
<th>2;0</th>
<th>2;2</th>
<th>2;4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td>Type</td>
<td>1.76</td>
<td>1.20</td>
<td>0.93</td>
<td>1.10</td>
<td>1.27</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Token</td>
<td>1.91</td>
<td>1.12</td>
<td>0.74</td>
<td>0.71</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td>Adults</td>
<td>Type</td>
<td>0.86</td>
<td>1.09</td>
<td>0.89</td>
<td>1.18</td>
<td>1.12</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Token</td>
<td>2.95</td>
<td>0.54</td>
<td>0.42</td>
<td>0.57</td>
<td>0.49</td>
<td>0.61</td>
</tr>
</tbody>
</table>
III. Developmental Trend of N/V Ratio

Types, CY

Types, ZTX

A Trend of Approximation
Result III

Approximation Mutual Influence?
IV. N/V Ratio Division

- eliminate the lexical bias in language itself

\[
\frac{N/V \text{ ratio in children's output}}{N/V \text{ ratio in adults' input}}
\]

- Interpretation:
  - What if... a balanced distribution in language itself?

<table>
<thead>
<tr>
<th>Age</th>
<th>1:6</th>
<th>1:8</th>
<th>1:10</th>
<th>2:0</th>
<th>2:2</th>
<th>2:4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>1.53</td>
<td>1.28</td>
<td>0.92</td>
<td>1.25</td>
<td>0.95</td>
<td>1.24</td>
</tr>
<tr>
<td>Token</td>
<td>2.81</td>
<td>2.18</td>
<td>1.53</td>
<td>1.91</td>
<td>1.44</td>
<td>1.90</td>
</tr>
<tr>
<td>ZTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>2.04</td>
<td>1.10</td>
<td>1.04</td>
<td>0.93</td>
<td>1.13</td>
<td>1.51</td>
</tr>
<tr>
<td>Token</td>
<td>0.65</td>
<td>2.06</td>
<td>1.75</td>
<td>1.24</td>
<td>1.42</td>
<td>1.20</td>
</tr>
</tbody>
</table>
• Answers to the Questions:
  • Verb bias while age is not controlled, Noun bias while it is and while early enough
  • Children seem take adult’s input as the paradigm of language using
  • To unravel the tangled bias: Mandarin is verb biased, but children’s production is noun biased

• Evidence for:
  • A Universal Noun Bias
  • Maybe nouns are easier
Discussion & Conclusion

- Longitudinal Finding: 1;8, is it a coincidence?
  - Also Bassano (1998)
Ambicategorical words (Conwell & Morgan, 2012)
Polysemy/Homophony
  Cross-linguistic differences?
Homophones and phoneme inventory
  larger → shorter (Nettle, 1995; 1998) → more (Piantadosi et al., 2012)
  larger → fewer (?)
Thank You!