Longitudinal Studies of Variation Sets in Child-directed Speech

Mats Wirén, Kristina Nilsson Björkenstam, Gintarė Grigonytė & Elisabet Eir Cortes
Department of Linguistics
Stockholm University
CogACLL, Berlin, 11 August 2016
Child-directed speech

• Key characteristic: repetitiousness
  – Also: shorter utterances, lower speech rate, fewer disfluencies, greater $F_0$ variation, etc.

• Variation sets: sequences of repetitious utterances with a constant intention
Variation sets: Example

You can put the animals there.
You can take the pig and the cat and put them there.
Can you put them there?
Good.
Can you put the pig there too?
Variation sets: Motivation

• Appears to facilitate language acquisition
  – Correlated with children’s language production (Waterfall 2006)
  – Benefits artificial language learning in adults (Onnis et al. 2008)
Variation sets: Previous work

• Lack of studies based on video
  – Studies so far made off transcriptions
• Lack of longitudinal studies
  – Not known if decrease studies involving infants under 14 months
• Very few languages studied
  – English and Turkish
The problem

• Generalised definition of variation sets motivated from videos

• Younger age group: 7–9 months
  – How is the decrease in variation sets as a function of age displayed here?

• Variation sets in additional languages
  – Manual annotation for Swedish
  – Automatic algorithm evaluated on Swedish, applied to Croatian and Russian (and English)
Data: MINGLE-3 corpus

• 18 parent–child sessions in Swedish (Nilsson Björkenstam & Wirén 2014)
  – Video, audio, transcriptions
  – 3 children, 7–33 months old, each participating in 6 sessions
  – Scenario: free play, but involving two target objects
Criteria for variation sets

• Sequences of similar utterances with a constant intention
  – Successive *pairs* of utterances
  – Up to two intervening utterances by parent
  – Intervening vocalisations/utterances by the child allowed
Criteria for variation sets

• **Extension 1**: “Similar” with respect to surface or semantics
  – Motivation: Semantic repetitiousness
    “Titta här då!”
    *But look here!*
    “Har du sett vilka tjusiga byxor?”
    *Have you seen the fancy pants?*
    “Kolla!”
    *Look!*
Criteria for variation sets

• **Extension 2:** “Similar” includes verbatim repetition
  – Motivation: Verbatim repetition seems to involve systematic *prosodic* (or even multimodal) variation
"Var är gummi-ankan?"

Where is the rubber duck?
Annotation of variation sets

• Annotation made from video
  – Previous work seems to have been made off transcriptions
  – ELAN video annotation tool
  – Using timelines to code extension of variation sets across utterances
Utterances in variation sets as function of age (Swedish)
Utterances in variation sets as function of age (Swedish)
Automatic approximation of variation sets

• Algorithm based solely on form
  – Compares successive \textit{pairs} of utterances
  – Returns sequences of utterances judged to be in variation sets
  – Uses Python string matcher for replicability: \texttt{difflib.SequenceMatcher}
  – Returns similarity measure between 0 and 1
  – Optimal threshold with respect to Swedish gold standard was 0.55
Automatic extraction: Evaluation

- **Strict matching**: Exact match of extracted utterances with utterances in the gold standard
- **Fuzzy matching**: Partial overlaps of these allowed
Evaluation: Example

- **Strict matching**: false positive
- **Fuzzy matching**: true positive
  - Average precision =
  - Recall = 0.5

<table>
<thead>
<tr>
<th></th>
<th>Member of gold set</th>
<th>Extracted by algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utt. 1</td>
<td>–</td>
<td>Yes</td>
</tr>
<tr>
<td>Utt. 2</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Utt. 3</td>
<td>Yes</td>
<td>–</td>
</tr>
</tbody>
</table>
Evaluation of automatic extraction on gold standard

<table>
<thead>
<tr>
<th>Age group</th>
<th>Strict F-score</th>
<th>Fuzzy F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0;7–0;9</td>
<td>0.56</td>
<td>0.82</td>
</tr>
<tr>
<td>1;0–1;2</td>
<td>0.41</td>
<td>0.69</td>
</tr>
<tr>
<td>1;4–1;7</td>
<td>0.39</td>
<td>0.60</td>
</tr>
<tr>
<td>2;3–2;9</td>
<td>0.26</td>
<td>0.43</td>
</tr>
</tbody>
</table>
Automatically extracted variation sets as function of age

<table>
<thead>
<tr>
<th>Language</th>
<th>0;7–0;9</th>
<th>1;0–1;2</th>
<th>1;4–1;7</th>
<th>2;3–2;9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Croatian</td>
<td>85%</td>
<td>54%</td>
<td>50%</td>
<td>–</td>
</tr>
<tr>
<td>English</td>
<td>–</td>
<td>–</td>
<td>54%</td>
<td>44%</td>
</tr>
<tr>
<td>Russian</td>
<td>–</td>
<td>–</td>
<td>35%</td>
<td>24%</td>
</tr>
</tbody>
</table>

**Age group**

- 0;7–0;9
- 1;0–1;2
- 1;4–1;7
- 2;3–2;9
The problem

• Generalised definition motivated by video
• Manually extracted variation sets from MINGLE-3 (gold standard)
  – Shows consistent decrease as function of age
• Surface algorithm for automatic extraction of variation sets
  – Evaluated on Swedish gold standard
  – Applied to Croatian, Russian and English
Conclusions (1)

• Proportion of utterances in variation sets decreases as a function of age
  – Age 7–9 months not previously investigated
  – Firm result for Swedish, indicative results for Croatian and Russian (and English)

• Verbatim repetitions display systematic prosodic and/or multimodal variation
  – Hence included in variation sets
Conclusions (2)

• Surface-based automatic extraction of variation sets can be useful approximation
  – At least for younger age groups
Thanks!

• The Swedish Research Council for funding the project *Modelling the emergence of linguistic structures in early childhood*, 2011-675-86010-31

• Lisa Tengstrand and Claudia Eklås Tejman for annotation work

• Annika Otsa for extracting and preparing the CHILDES data

• Three anonymous reviewers for valuable comments