What can we learn from bound learners?

Barend Beekhuizen

Leiden University Center for Linguistics
Leiden University

Institute of Logic, Language and Computation
University of Amsterdam

11 March 2013
Marr’s (1982) three levels:
- computational (CL)
- algorithmic (AL)
- implementational
Marr’s (1982) three levels:
- computational (CL)
- algorithmic (AL)
- implementational

Typical question: Can a function be calculated at all? (CL)
Does not (have to) consider AL properties of the system
Introduction

- Marr’s (1982) three levels:
  - computational (CL)
  - algorithmic (AL)
  - implementational

- Typical question: Can a function be calculated at all? (CL)
- Does not (have to) consider AL properties of the system
- But: can the function be calculated given AL properties?
- Moreover, are there situations where AL properties explain certain behavior?
- Let’s call a learner bound if it is constrained on the AL, i.e. in its representations and processing algorithms.
Constraints on the AL

- Are there situations where AL properties explain certain behavior (in language acquisition)?
- Evidence from different domains:

<table>
<thead>
<tr>
<th>domain</th>
<th>reference</th>
<th>constraints on</th>
</tr>
</thead>
<tbody>
<tr>
<td>decision making</td>
<td>Gigerenzer &amp; Selten (2001)</td>
<td>search, decision</td>
</tr>
<tr>
<td>garden-path sentences</td>
<td>Ferreira &amp; Patson (2007)</td>
<td>ability to track multiple analyses</td>
</tr>
<tr>
<td>word-meaning acquisition</td>
<td>Medina et al. (2012)</td>
<td>ability to track multiple lexical semantic analyses</td>
</tr>
<tr>
<td>production of root infinitives</td>
<td>Freudenthal et al. (2007)</td>
<td>working memory</td>
</tr>
<tr>
<td>grammar learning</td>
<td>Elman (1993)</td>
<td>working memory</td>
</tr>
</tbody>
</table>
What does this mean for modeling the acquisition of grammar?

- Learning is often taken to involve optimizing some function.
- Is this the correct way of looking at it?
Bound learners: theoretical perspective

Implications

What does this mean for modeling the acquisition of grammar?

- Learning is often taken to involve optimizing some function
- Is this the correct way of looking at it?
- Optimization is often algorithmically very heavy
- Herbert Simon’s (1955) idea of *satisficing*: making a decision that meets an *aspiration level* rather than optimizes the decision
- Bounded Rationality: domain-specific fast and frugal heuristics
Bound learners: theoretical perspective

Implications

What does this mean for modeling the acquisition of grammar?

- Learning is often taken to involve optimizing some function
- Is this the correct way of looking at it?
- Optimization is often algorithmically very heavy
- Herbert Simon’s (1955) idea of satisficing: making a decision that meets an aspiration level rather than optimizes the decision
- Bounded Rationality: domain-specific fast and frugal heuristics
- For language acquisition: using dumb, heuristic strategies to learn, that stop when an “aspiration level” is met
- Social approach: Learner tries to get by (communicatively)
Assumption
Assume an incremental learner trying to build up some sort of grammar.

- Types of constraints
  - Working memory: only process one new word per utterance and words to the right of that
    → Natural “starting small” heuristic (Elman 1993; Spitkovsky et al. 2009)
  - Search: only add one new syntactic rule per utterance
    → No full hypothesis space (U-DOP)
    → Danger of getting stuck in bad part of hypothesis space?
    → Starting small alleviates?
  - Parse: minimal attachment, late closure (with revision?)
  - Abstraction: lazy (only when analogy can be made)
### Types of phenomena

**Phenomena**

What kind of phenomena are likely candidates for an explanation in terms of a bound learner?

- Recent Minimalist explanations of certain production phenomena (Yang & Roeper 2012)
- E.g. argument drop within *wh*-questions; order asymmetries; argument realization
Types of phenomena

Phenomena

What kind of phenomena are likely candidates for an explanation in terms of a bound learner?

- Recent Minimalist explanations of certain production phenomena (Yang & Roeper 2012)
- E.g. argument drop within *wh*-questions; order asymmetries; argument realization
- Generally: deviations in production from the adult grammar and input
- And their development/convergence
- E.g. over- and undergeneralization, chunks that are compositional
- Explain as *interactions* between bound learning algorithms and input.

Barend Beekhuizen

What can we learn from bound learners?
Why take this approach?

- **Desirability** of this type of explanation: acknowledges limitations of processing system and explains behavior.
- Seems promising for developmental patterns
- Seems promising for deviations from input data
- Provides a general learning answer to parameter-setting explanations like Yang & Roeper’s
Thank you