Cognitive explanations in linguistic typology: Contemporary insights from language processing and language acquisition.

12th Conference of the Association for Linguistic Typology, Australian National University, Canberra, December 12-15, 2017

Deriving argument ordering biases from expectation-based processing

Thomas Hörberg
Department of Linguistics, Stockholm University

T. Florian Jaeger
Brain and Cognitive Sciences, University of Rochester
• Background and assumptions
• Corpus-based model of incremental argument interpretation
• Experimental test of model predictions
Background and assumptions

Performance-grammar correspondence hypothesis:

“Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments”

(Hawkins 2003)
Grammars prefer short dependency lengths

(Futrell et al. 2015)
The "subject-first" preference

SO massively preferred over OS
The "subject-first" preference

Initial NPs preferrably interpreted as subjects

- **Condition**: OLex-V-SPro, SLex-V-OPro
- **Example sentence**:
  - OLex-V-SPro: Bagarna (DEF.PL) visade (show-PRT) **ni** (2PL.SBJ) till (to) köket (DEF)
  - SLex-V-OPro: Bagarna (DEF.PL) visade (show-PRT) **oss** (1PL.OBJ) till (to) köket (DEF)

(Hörberg et al. 2013)
The "subject-first" preference

Also in ergative constructions in Hindi

- No subject function / \{S/A\} alignment
- Sentence-initial actors often omitted

(Bickel et al. 2015)
The "subject-first" preference (in comprehension)
- different views on its cognitive underpinnings

Language-independent, universal bias for interpreting initial NPs as subjects / {S/A}
(e.g., Bickel et al. 2015; Demiral et al. 2008; Wang et al. 2012; Bornkessel-Schlesewsky & Schlesewsky, 2008)

Expectation-based, dependent on statistical regularities in the input language
(e.g., Hörberg 2016; Gildea & Jaeger 2015; see also MacDonald 2013; Levy 2008; Trueswell et al. 1994)

- here, we test predictions of the latter view
Assumptions – argument interpretation

• Assignment of argument FUNCTIONS (Actor / Undergoer)

• Highly incremental process that is expectation-based, i.e. draws upon statistical regularities in the input

• Based on the distribution of morphosyntactic, NP prominence and verb semantic information which serve as Argument Interpretation Cues (AICs)

• Diachronically, distributional patterns shape grammars over time (e.g. loss of overt case marking on low-prominent direct objects)

(MacWhinney & Bates 1989)
Objectives

• Expectation-driven model of incremental argument interpretation on basis of distributional patterns of AICs in Swedish transitive sentences

• Experimentally test model predictions using SPR
Corpus-based model of incremental argument interpretation

- Use corpora to estimate distribution of AICs over subjects and objects in Swedish transitive sentences
- Calculate probabilities for a given word order / GF assignment based on AICs provided by constituents over time (i.e., NP1, verb, NP2)
- Estimate incremental change in the expectation for a given WO on basis of these probabilities (in terms of Bayesian surprise)
Target sentences

<table>
<thead>
<tr>
<th>Word order</th>
<th>Example</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVO</td>
<td>Barnen får inte äta upp all glass innan middan</td>
<td>“The kids can’t eat all ice cream before dinner”</td>
</tr>
<tr>
<td></td>
<td>Kid-PL can.PRS not eat up all ice.cream before dinner.DEF</td>
<td></td>
</tr>
<tr>
<td>VSO</td>
<td>Innan middan får barnen inte äta upp all glass</td>
<td>“Before dinner, the kids can’t all ice cream”</td>
</tr>
<tr>
<td></td>
<td>Before dinner.DEF can.PRS kid-PL not eat up all ice.cream</td>
<td></td>
</tr>
<tr>
<td>OVS</td>
<td>All glass får barnen inte äta upp innan middan</td>
<td>“All ice cream, the kids can’t eat it before dinner”</td>
</tr>
<tr>
<td></td>
<td>all ice.cream can.PRS kid-PL not eat up before dinner</td>
<td></td>
</tr>
</tbody>
</table>

- NP:s of any length
- Up to 4 verbs
- Adverbials + verb particles optional

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
## Corpus properties

**Svensk Trädbank: balanced written Swedish texts**

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Genre</th>
<th>N texts</th>
<th>N sentences</th>
<th>N hits</th>
<th>SVO</th>
<th>VSO</th>
<th>OVS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Press: reportage</td>
<td></td>
<td>44</td>
<td>7278</td>
<td>1495</td>
<td>1495</td>
<td>1149</td>
<td>68</td>
</tr>
<tr>
<td>Press: Editorial</td>
<td></td>
<td>17</td>
<td>2385</td>
<td>473</td>
<td>473</td>
<td>385</td>
<td>32</td>
</tr>
<tr>
<td>Press: Reviews</td>
<td></td>
<td>27</td>
<td>3961</td>
<td>712</td>
<td>712</td>
<td>536</td>
<td>52</td>
</tr>
<tr>
<td>Skills, Trades and Hobbies</td>
<td></td>
<td>58</td>
<td>8933</td>
<td>1840</td>
<td>1840</td>
<td>1343</td>
<td>118</td>
</tr>
<tr>
<td>Popular Lore</td>
<td></td>
<td>48</td>
<td>6525</td>
<td>1503</td>
<td>1503</td>
<td>1160</td>
<td>55</td>
</tr>
<tr>
<td>Belles Letters, Biography, Memoirs</td>
<td></td>
<td>26</td>
<td>3598</td>
<td>805</td>
<td>805</td>
<td>627</td>
<td>51</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td></td>
<td>70</td>
<td>10847</td>
<td>1540</td>
<td>1540</td>
<td>1239</td>
<td>50</td>
</tr>
<tr>
<td>Learned and Scientific Writing</td>
<td></td>
<td>83</td>
<td>9633</td>
<td>1809</td>
<td>1809</td>
<td>1398</td>
<td>64</td>
</tr>
<tr>
<td>General fiction</td>
<td></td>
<td>82</td>
<td>13028</td>
<td>3110</td>
<td>3110</td>
<td>2527</td>
<td>186</td>
</tr>
<tr>
<td>Mysteries and Science fiction</td>
<td></td>
<td>19</td>
<td>4070</td>
<td>826</td>
<td>826</td>
<td>665</td>
<td>58</td>
</tr>
<tr>
<td>Light reading</td>
<td></td>
<td>20</td>
<td>2908</td>
<td>749</td>
<td>749</td>
<td>611</td>
<td>53</td>
</tr>
<tr>
<td>Humor</td>
<td></td>
<td>6</td>
<td>1071</td>
<td>248</td>
<td>248</td>
<td>183</td>
<td>27</td>
</tr>
<tr>
<td>Brochure texts</td>
<td></td>
<td>25</td>
<td>1733</td>
<td>390</td>
<td>390</td>
<td>298</td>
<td>20</td>
</tr>
<tr>
<td>Newspaper texts</td>
<td></td>
<td>28</td>
<td>1669</td>
<td>361</td>
<td>361</td>
<td>277</td>
<td>16</td>
</tr>
<tr>
<td>Educational texts</td>
<td></td>
<td>14</td>
<td>1624</td>
<td>374</td>
<td>374</td>
<td>292</td>
<td>10</td>
</tr>
<tr>
<td>Debate articles</td>
<td></td>
<td>18</td>
<td>1134</td>
<td>316</td>
<td>316</td>
<td>259</td>
<td>12</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td>585</td>
<td>80397</td>
<td>16551</td>
<td>12949</td>
<td>2730</td>
<td>872</td>
</tr>
</tbody>
</table>
Estimation of AIC distributions over GFs

- Logistic mixed effects modeling
- Estimates individual AIC strengths in terms of predicting WO / GF assignment
- Predict probability for OS order at NP1, verb, and NP2 conditional on AICs

<table>
<thead>
<tr>
<th>NP properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Animacy</td>
</tr>
<tr>
<td>- Givenness</td>
</tr>
<tr>
<td>- Definiteness</td>
</tr>
<tr>
<td>- Number</td>
</tr>
<tr>
<td>- Egophoricity</td>
</tr>
<tr>
<td>(1st / 2nd vs. 3rd person)</td>
</tr>
<tr>
<td>- Pronominality</td>
</tr>
<tr>
<td>- Case</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb semantic properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Volitionality</td>
</tr>
<tr>
<td>- Causation</td>
</tr>
<tr>
<td>- Sentience</td>
</tr>
<tr>
<td>- Possession</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>- NP length</td>
</tr>
<tr>
<td>- Adverbial-initial</td>
</tr>
<tr>
<td>- Embedded</td>
</tr>
<tr>
<td>- Auxiliary verbs</td>
</tr>
</tbody>
</table>
Expectation-based modeling of incremental argument interpretation

- We model the incremental change in the expectation of OS word order over time.
- Expectations based upon AICs provided by constituents over time (i.e., NP1, verb(s) and NP2).
Expectation-based modeling of incremental argument interpretation

- expectation change for OS modelled in terms of Bayesian surprise / relative entropy / the Kullback–Leibler divergence:
  \[ D_{KL}(P \parallel Q) = \sum_i \log(P_i/Q_i)P_i \]

- Similar to word / parse surprisal \((-\log(p(W_i \mid w_{i-1}\ldots w_1)))\) but applies to constituents with respect to predicting WO / GF assignment
Expectation-based modeling of incremental argument interpretation

**Probabilities**

- **baseline model:**
  \[ p(\text{OS}) : \sim 0.05 \]

- **NP1 model:**
  \[ p(\text{OS} \mid \text{NP1}) \]

- **NP1 + verb model:**
  \[ p(\text{OS} \mid \text{NP1 + verb}) \]

- **full model:**
  \[ p(\text{OS} \mid \text{NP1 + verb + NP2}) \]

**Bayesian surprise**

- **Bayesian surprise NP1**
  \[ D_{KL}(p(\text{OS} \mid \text{NP1}) \mid \mid p(\text{OS})) \]

- **Bayesian surprise verb**
  \[ D_{KL}(p(\text{OS} \mid \text{NP1 + verb}) \mid \mid p(\text{OS} \mid \text{NP1})) \]

- **Bayesian surprise NP2**
  \[ D_{KL}(p(\text{OS} \mid \text{NP1 + verb + NP2}) \mid \mid p(\text{OS} \mid \text{NP1 + verb})) \]
Bayesian surprise in original data

- OVS sentences with initial lexical NP and final case marked NP

[De levande $D_{KL} = 0.02$] [fångade $D_{KL} = 0.00$] [jag $D_{KL} = 5.64$]

"The living, I caught them"

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
Bayesian surprise in original data
- OVS sentences with initial lexical NP and final case marked NP

[En liknande strävan $D_{KL} = 0.31$] [urskiljer $D_{KL} = 1.35$] [han $D_{KL} = 0.17$] nu...

A similar endavour discerns he now...

"A similar endavour he now discerns"

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
Bayesian surprise at NP1
Bayesian surprise at verb
Bayesian surprise at NP2

Volitional

Experiencer

Prominence feature

Bayesian surprise in bits

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
Experimental test of model predictions

- Testing the strongest model predictions
- Self-paced reading
- Reading times assumed to reflect processing
Self paced reading
Self paced reading

Bollen ####### ### #### ### # #######
Self paced reading

####### sparkar ####### # # # # #
Self paced reading

######## ######## jag ####### ### # ########
Self paced reading

####### ########### ### mitt ### # #######

2017-12-15 / Thomas Hörberg, Department of Linguistics
Self paced reading

########  ############  ###  ###  ###  upp  #  ########
Self paced reading

######## # ####### # i #######
Self paced reading

##### ####### ### #### # krysset
Self paced reading

**The ball, I kick it right up into the top corner**

- Dependent variable: time latency between button presses
- Analyses done on region RTs rather than word RTs
- Task: Comprehension question following each sentence

**Sparkar jag bollen mitt upp i krysset?**

“Do I kick the ball right up into the top corner?”

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
## Experimental sentences

<table>
<thead>
<tr>
<th>WO</th>
<th>Verb class</th>
<th>Obj. animacy</th>
<th>Example</th>
</tr>
</thead>
</table>
|    |            | Inanimate    | Bollen sparkar jag mitt upp i krysset  
|    |            |              | Boll-DEF kick-PRS 1SG middle up in top.corner  
|    |            |              | ”The ball, I kick it right up into the top corner”  
| Volitional |              | Animate      | Killen sparkar jag mitt på smallbenet  
|    |            |              | Guy-DEF kick-PRS 1SG middle up chin-DEF  
|    |            |              | ”The guy, I kick him in the middle of the shin”  

| OVS |            | Inanimate    | Bollen glömmer jag mitt på fotbollsplanen  
|     |            |              | Boll-DEF forget-PRS 1SG middle on football.field-DEF  
|     |            |              | ”The ball, I forget it in the middle of the football field”  
| Experiencer |              | Animate      | Killen glömmer jag sent på kvällen  
|    |            |              | Guy-DEF forget-PRS 1SG late at night-DEF  
|    |            |              | ”The guy, I forget him late at night”  

| SVO |            | Inanimate    | Jag sparkar bollen mitt upp i krysset  
|     |            |              | 1SG kick-PRS ball-DEF middle up in top.corner  
|     |            |              | ”I kick the ball right up into the top corner”  
| Volitional |            | Animate      | Jag sparkar killen mitt på smallbenet  
|    |            |              | 1SG kick-PRS guy-DEF middle up chin-DEF  
|    |            |              | ”I kick the guy in the middle of the shin”  

|      |            | Inanimate    | Jag glömmer bollen mitt på fotbollsplanen  
|      |            |              | 1SG forget-PRS ball-DEF middle on football.field-DEF  
|      |            |              | ”I forget the ball in the middle of the football field”  
| Experiencer |            | Animate      | Jag glömmer killen sent på kvällen  
|    |            |              | 1SG forget-PRS guy-DEF late at night-DEF  
|    |            |              | ”I forget the guy late at night”  

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
Results – RT differences vs predictions

The subject-first preference: faster RTs in SVO sentences vs. OVS sentences

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
Results – RT differences vs predictions

- Faster verb RTs for inanimate vs. animate objects in experiencer verb sentences
- Faster NP2 RTs for inanimate vs. animate objects in volitional verb sentences

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
Results – RT differences vs predictions

- No NP2 RT difference between inanimate vs. animate object sentences with experiencer verbs ....
- But NP2 RTs equally fast as in volitional verb sentences with inanimate objects

2017-12-15 / Thomas Hörberg, Department of Linguistics, Stockholm University
Conclusions

• The subject-first preference can be predicted on the basis of expectation-based modeling
• Expectation-based modeling also predicts influences of animacy - verb semantic interactions on RTs
• These findings provide evidence for the expectation-based perspective on language comprehension
• Cross-linguistic research on language processing should take into account context-conditioned expectations
Thanks for your attention!
References


