Word Order and Case Marking in Language Acquisition and Processing
Maryia Fedzechkina, T. Florian Jaeger, Elissa L. Newport
Department of Brain & Cognitive Sciences, University of Rochester
{mashaf@bcs.rochester.edu; fjjaeger@bcs.rochester.edu; newport@bcs.rochester.edu}

Abstract
To understand a sentence, comprehenders must identify its actor and patient. In principle, these relationships can be signaled using a single cue, but most languages employ several redundant cues, including word order and case marking. In artificial language learning experiments we investigate word order and case as cues in processing and learning. Flexible word order languages are potentially ambiguous if no case-marking (or other cues) are employed to identify the doer of the action. We explore whether language learners have a bias towards ambiguity reduction and tend to fix word order in non-case-marking languages (cf. the historic change from Old English to contemporary English).

The Experiment

An artificial language learning study
- 19 monolingual English native-speakers
- over 4 consecutive days:

<table>
<thead>
<tr>
<th>Case marker</th>
<th>Word Order</th>
<th>Number of verbs associated with dominant word order</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Case</td>
<td>LG 1 (SOV)</td>
<td>100% 83% 58% 0%</td>
</tr>
<tr>
<td></td>
<td>LG 2 (SVO)</td>
<td>80% 3% 2% 0%</td>
</tr>
</tbody>
</table>

Lexicon

Grammar

Word Order Fixing
In the absence of case-marking there is a tendency to generalize the majority order in both comprehension and production (for comprehension, p<0.05 and for production, p=0.09). The tendency to generalize the dominant word order appears most strongly in the no-case condition when the majority word order is most variable in the input language. In contrast, when case marking is present, the word order variability of the input language is more accurately reproduced in learners’ usage.

Discussion
In the absence of case-marking, learners tend to generalize the dominant word order both in production, suggesting a bias to reduce ambiguity in the language. This effect is partially driven by a group of learners of the non-case-marking language who show extreme word order fixing both in comprehension and production.

When case is present, learners use it to resolve ambiguity. We do not see any learners of the case-marking language who show word order generalization in the production task. If anything, they overgeneralize (one interpretation being that learning word order is less important to them).

References
[1-3] also hold: Bias-consistent responses are significantly faster.

Fig. 1: Comprehension, final day of training
Fig. 2: Production, final day of training
Fig. 3: Total mistakes
Fig. 4: Mean Reaction Time
Fig. 5: Mean RTs, overgeneralizers removed