Acquiring Recursive Structures through Distributional Learning

Daoxin Li, Kathryn Schuler (University of Pennsylvania)

This study investigates the learning mechanism that enables the acquisition of recursive structures. Languages differ regarding the depth, structure, and syntactic domains of recursive structures (Pérez-Leroux et al. 2018). For example, English allows infinite free embedding of genitive *-s*, (1), whereas German restricts this structure to only one level and to a limited set of items, (2), (Weiss 2008). Thus, while the ability for recursion may be innate and universal (e.g. Hauser et al. 2002), speakers need to learn in which syntactic domains this ability can be applied.

It has been proposed that explicit evidence for deep embedding in the input is necessary for the acquisition of recursive structures (e.g. Roeper 2011), but experiments have reported early acquisition of recursive structures even though evidence for deep embedding is rarely attested in young children's input (e.g. Giblin et al. 2019). This present study tests an alternative proposal that does not require evidence for multiple-level embeddings. Instead, learners acquire recursion through distributional learning (e.g. Braine 1987; Maratsos & Chalkley 1980). Specifically, the proposal (Grohe et al. 2020; Li et al. 2020) suggests that productivity, defined as structural interchangeability, is the prerequisite for recursion; so the recursion of a structure (e.g. **X's-Y**) is licensed if a sufficiently large proportion of nouns attested in the **X** position in the input are also attested in the **Y** position in the input.

We used two artificial language learning experiments to test the proposal. In each experiment, 25 adults were exposed to 88 **X-ka-Y** phrases, where 12 different words were attested in the **X** position. In Experiment 1, only some of the words were also attested in Y position (6 out of the 12); in Experiment 2, nearly all were (10 out of the 12 words). The frequency of 12 words followed a Zipfian distribution, and the total frequency of each word was the same across two experiments. At test, we asked the participants to rate on a scale of 1 to 5 the acceptability of one-level (**X-ka-Y**) and two-level (**X-ka-Y-ka-Z**) attested phrases (i.e. phrases or combinations of two phrases heard during exposure), unattested phrases (i.e. phrases or combinations of two phrases whose post-**ka** position (**Y** or **Z**) was occupied by a word that never appeared in **Y** position the input), and ungrammatical phrases with wrong word order (e.g. **ka-X-Y, ka-X-Y-Z-ka**). The distributional learning proposal predicts participants from Exp2 would learn the **X-ka-Y** structure was productive and thus recursive, so they would rate unattested phrases higher than participants from Exp1 at both one- and two-level, even though two-level phrases were never attested in the input.

Results are shown in Figure 1 (one-level) and Figure 2 (two-level). We analyzed the results using ordinal regression. There was a main effect of sentence type (attested, unattested, or ungrammatical) for both one- ($\chi^2(2)=253.00$, $p<0.001$) and two-levels ($\chi^2(2)=323.82$, $p<0.001$); in particular, as predicted, unattested recursive phrases were rated significantly higher than ungrammatical phrases in Exp2 ($p<0.001$) but not in Exp1 ($p=0.47$). There was also a significant interaction between sentence type and experiment (Exp1, Exp2) for both one-level ($\chi^2(2)=8.67$, $p=0.01$) and two-level ($\chi^2(2)=52.74$, $p<0.001$). Comparison between experiments showed that unattested phrases were rated marginally lower in Exp1 than in Exp2 at one-level ($p=0.08$) and significantly lower at two-level ($p<0.01$). Overall, our results suggest that speakers can use distributional information at one level to learn whether a structure can be recursive.

(1) a. the man*'s* neighbor*'s* book

(2) a. Vater*s* Buch ('father's book') vs. *Mann*s* Buch ('man's book')

b. *das Mann*s* Nachbar*s* Buch ('the man's neighbor's book')

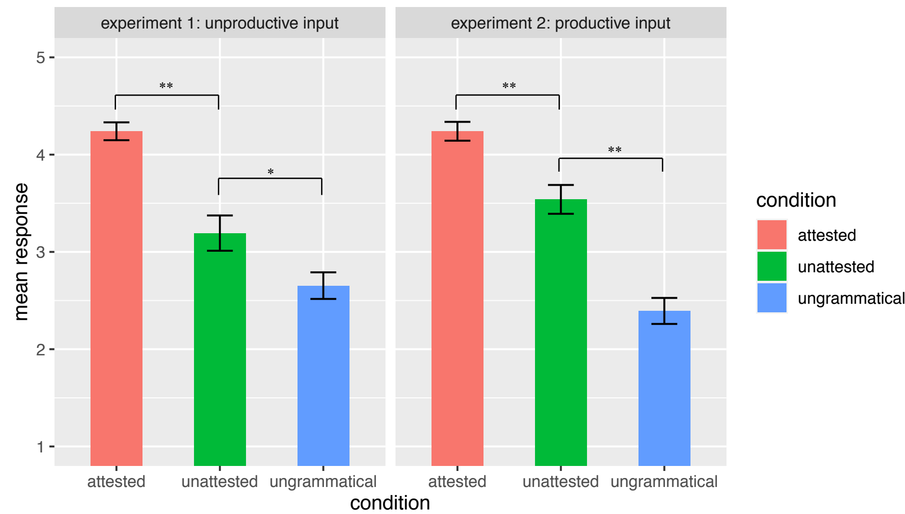

Figure 1. Mean rating response to one-level test phrases in each condition by participants in experiment 1 and experiment 2. Error bars indicate standard errors of the mean.

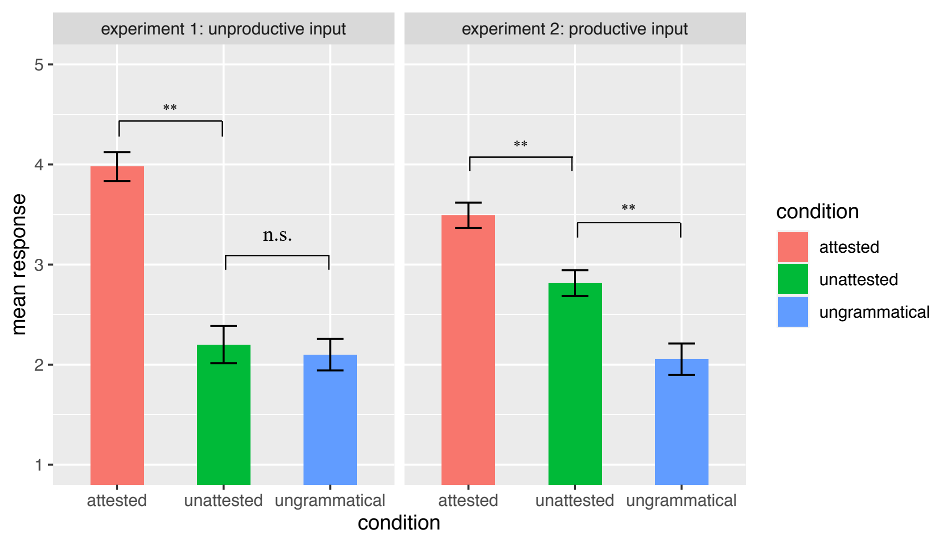

Figure 2. Mean rating response to two-level (recursive) test phrases in each condition by participants in experiment 1 and experiment 2. Error bars indicate standard errors of the mean.

Selected References

Daoxin Li, Lydia Grohe, Petra Schulz, & Charles Yang**.** (2020). *The distributional learning of recursive structures*. Paper presented at BUCLD-45.