Learning discourse patterns through exposure: Mixed input helps identify informative categories

While much of language is learned during childhood, adults continue to adapt to the most frequent patterns in local contexts. Speakers tend to imitate the structure of syntactic primes (e.g., Bock, 1986) and comprehenders are biased toward recently-heard syntactic structures (e.g., Thothathiri & Snedeker, 2008). Adaptation also occurs at the discourse level, where the interpretation of ambiguous pronouns is biased toward recently encountered patterns (e.g., Kaiser, 2009; Contemori, 2019). Johnson & Arnold (2023, exp. 2) tested the interpretation of ambiguous pronouns like "{Ana sent a text to Liz / Ana got a text from Liz} and then she took a screenshot." Here people favor the subject (Ana) as the referent of "she", following the well-known subject-bias (e.g., Stevenson et al., 1994), but they also favor the goal (Liz for "send", Ana for "got"), so the preference for Ana is stronger for "got" than "send" (Langlois & Arnold, 2020). Johnson & Arnold showed that these biases are malleable. If people have recently read numerous examples of unambiguous pronouns referring to the nonsubject (half goal, half source), they are somewhat more likely to pick nonsubject antecedents (Exp. 2a). But if people have recently read many pronouns referring to the source (half subject, half nonsubject), they shift their interpretation in favor of the source (Exp. 2b). This shows that people adapt to the property of antecedents that is most informative. In Exp. 2a, the subject/nonsubject distinction was informative, and goal/source was not; the reverse pattern held for Exp. 2b.

This raises questions about how people respond to linguistic input that could be informative about multiple patterns. Given "Matt got a book to Ana and he…", do people learn that pronouns refer to goals? Or to subjects? We hypothesize that over a lifetime of input, people may abstract across exemplars to learn biases related to both syntactically-driven categories (e.g., subject antecedents) and semantically-driven categories (e.g., goal antecedents). When either one varies in the local context, people shift their biases to adapt.

In two experiments we tested how people respond to input that is either uninformative about the relevant category to learn (Exp. 1, 116 subjects), or informative (Exp. 2, 80 subjects). Experiment 1 used Johnson and Arnold's methods and stimuli, but all the exposure stories used goal-source verbs ("sent" type; see Table 1). In the subject-exposure condition, all 32 exposure stories used pronouns referring to the subject/goal; in the nonsubject-exposure condition, all exposure pronouns referred to the nonsubject/source. Interspersed were 12 stories with ambiguous pronouns using either goal-source or source-goal verbs, and we probed interpretation with questions (Table 2). The key question was whether exposure to goal-source stories would influence pronoun interpretation for both verb types or not.

Results showed it did not (Figure 1). For stories with matching verbs, there was a strong exposure effect: more subject/goal interpretations for subject/goal-exposure than for nonsubject/source-exposure. There was no exposure effect for the mismatching verbs. We know that exposure effects are not specific to thematic role, because Ye & Arnold (2023) found that exposure generalizes across verbtype. Thus, participants may have learned both syntactically- and semantically-conditioned patterns that canceled out for the source-goal verbs, or failed to learn either, or a mix.

Experiment 2 tested whether mixed input can direct participants' attention to the syntactic dimension of pronoun antecedents. Using the same goal-source exposure stimuli as Exp. 3, we replaced 8 exposure trials with "joint action" verbs (see Table 1), where the pronoun refers to either a subject/agent or a nonsubject/comitative role. This thematic role variability may signal that the informative dimension is syntactic role, and not thematic role. If so, exposure should generalize to test trials with the source-goal verb.

Results showed that indeed exposure generalized (Figure 1). Both experiments contrasted exposure stories with 100% subject vs. 100% objPP antecedents. But in Exp. 1, with only one verbtype, people didn't learn anything special about syntactic position per. In Exp. 2 we

varied the thematic roles (subject antecedents were 75% goal/25% agent while nonsubject antecedents were 75% source/25% comitative), and exposure modulated the subject bias for both verbtypes.

This study shows that discourse patterns are inferred from the input by abstracting over multiple exemplars, and not just through immediate priming from the previous trial. It also shows that people can extract generalizations like "pronouns tend to refer to subjects" from exposure to complex inputs.

==

Table 1. Example exposure stimuli:
<u>Goal-source verb; Subject pronoun:</u> Ana and Matt were taking an English lit class. Ana borrowed the book from Matt and then she looked up a reference.
<u>Goal-source verb; Nonsubject pronoun:</u> … and then he looked up a reference.
<u>(Exp. 2 only) Joint-action verb; Subject pronoun:</u> Liz and Will were spending the weekend together. Liz set up a picnic in the park with Will and then she ate some sandwiches.
<u>(Exp. 2 only) Joint-action verb; Nonubject pronoun:</u> … and then he ate some sandwiches.

Table 2. Example critical (ambiguous) stimuli:
<u>Goal-source verb:</u> Will and Matt were taking an exam in class. Will borrowed a pencil from Matt and then he began his exam. Did Matt begin his exam? (no = subject interpretation)
<u>Source-goal verb:</u> Will and Matt were taking an exam in class. Will loaned a pencil to Matt and then he began his exam. Did Matt begin his exam? (no = subject interpretation).

Figure 1. Results from Exp. 1 and Exp. 2
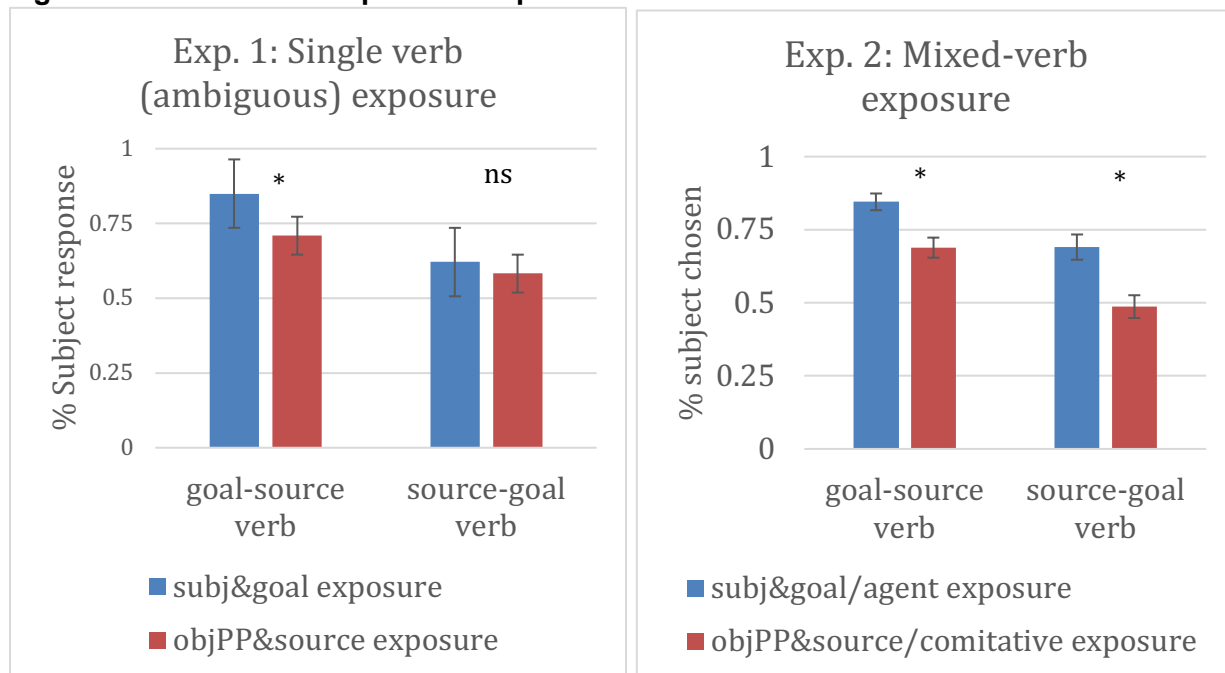
References:
Contemori, C. (2019). Changing comprehenders' pronoun interpretations… Second Language Research. ✠ Johnson, E., & Arnold, J. E. (2023). The Frequency of referential patterns…. JEP:LMC. ✠ Kaiser, E. (2009). Effects of anaphoric dependencies … Discourse anaphora…(pp. 121–129). Springer. ✠ Langlois, V. J., & Arnold, J. E. (2020). Print exposure .... Cognition. ✠ Stevenson, R. J., et al.. (1994). Thematic roles… LCP. ✠ Thothathiri, M., & Snedeker, J. (2008). Give and take… Cognition. ✠ Ye, Y. & Arnold, J. E. (2023). Learning the statistics … *Cognition.*