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

Language processing is adaptive; even adults adjust their biases to the most frequent structures in the local context (e.g., Thothariri & Snedeker, 2008; Fine et al., 2013). Such adaptation even occurs at the discourse level, where the interpretation of ambiguous pronouns is biased toward recently encountered patterns (Kaiser, 2009; Contemori, 2019). But an open question is how the system figures out which of many possible patterns to adapt to.

For example, Johnson & Arnold (2023) examined the interpretation of ambiguous pronouns following transfer verbs, e.g. "{Ana sent a text to Liz / Ana got a text from Liz} and then she took a screenshot." Pronoun interpretation is biased toward both subjects and goals: people tend to assign pronouns to the subject (Ana), but more when the subject is the goal (for "got"). But these biases shift to follow the locally frequent pattern. When participants were exposed to unambiguous pronouns that frequently referred to the nonsubject (half goal, half source), nonsubject antecedents were chosen more often for ambiguous pronouns (Exp. 2a). But when prior pronouns frequently referred to the source (half subject, half nonsubject), people shifted their interpretation for ambiguous pronouns in favor of the source (Exp. 2b). This shows that people follow the most frequent pattern when frequency is conditioned on either the syntactic or semantic role of the antecedents, and the other category is uninformative.

But in real life antecedent properties may be correlated. How do people figure out which pattern to learn? In "Matt got a book to Ana and he…", do people learn that pronouns refer to goals? Or to subjects? When the input exhibits a tendency for pronouns to refer to either subjects or prepositional objects, people generalize this pattern from one verb type to another (Ye & Arnold, 2023), suggesting that input from different verb types is pooled. That is, people don't learn highly specific syntactic/thematic combinations. We hypothesize that people generalize across exemplars to identify the most informative properties of the input. Depending on what patterns are informative, people may follow frequency patterns that are based on both syntactically-driven categories (e.g., subject antecedents) and semantically-driven categories (e.g., goal antecedents). Over a lifetime of input, people may abstract across exemplars to learn biases related to both categories. In two experiments we tested the hypothesis that adaptation requires evidence from multiple exemplars about which dimension is informative.

Experiment 1 (N=116) used Johnson & Arnold's methods and stimuli, but all 32 exposure stories used goal-source verbs (Table 1), meaning the relevant dimension of adaptation was ambiguous. In the subject-exposure condition, 100% exposure pronouns referred to the subject/goal; in the nonsubject-exposure condition, 100% 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 source-goal stories too. **Results** showed it did not (Figure 1). Exposure modulated pronoun interpretation only for the matching verb type. Thus, participants may have learned both syntactically- and semantically-conditioned patterns that canceled out for the source-goal verbs.

Experiment 2 (N=80) tested whether mixed input can direct participants' attention to the syntactic dimension of antecedents. 32 exposure stories included 75% with goal-source verbs and 25% with "joint action" predicates (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, and not semantic. If so, exposure should generalize to test trials with the source-goal verb. **Results** showed that indeed it did, revealing exposure effects for both goal-source and source-goal test items (Figure 1).

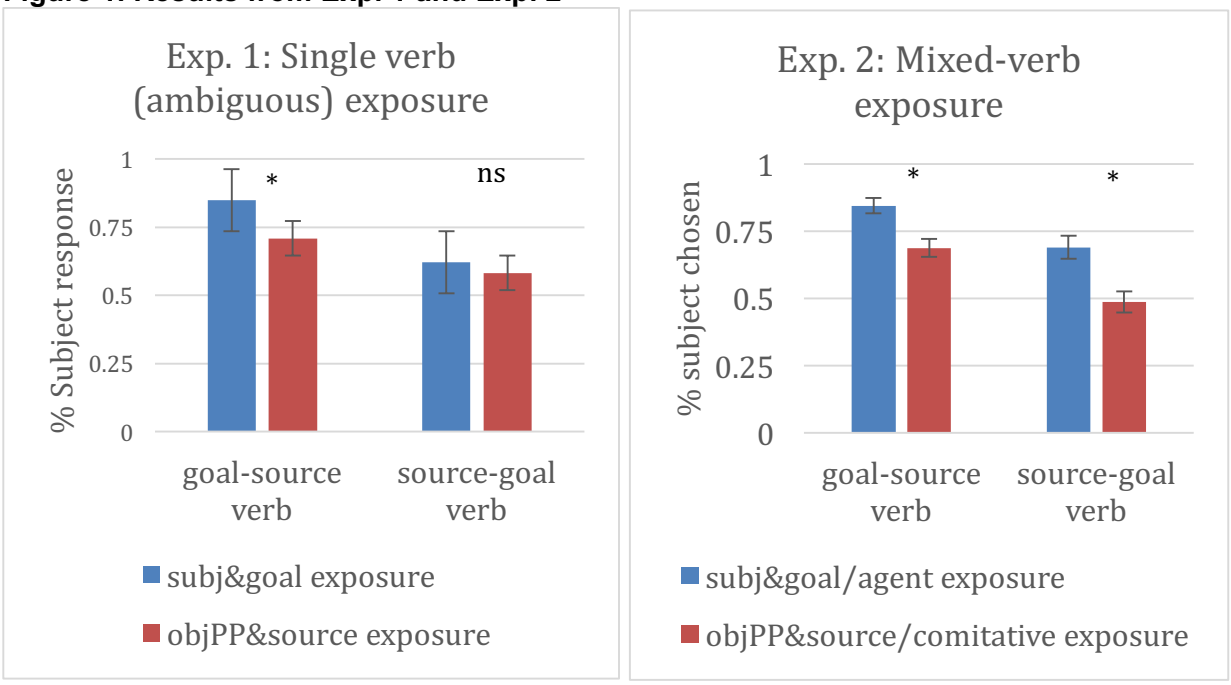
In sum, discourse patterns are inferred from the input by abstracting over multiple exemplars. Varied inputs are necessary to isolate the biases associated with one dimension such as the preference to assign pronouns to subject antecedents.

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

Table 2. Example critical (ambiguous) stimuli:
Goal-source verb: 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)
Source-goal verb: 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).

* other items used instead who questions (e.g., Who began his exam?, where Will = subject interpretation) or yes questions (e.g., Did Will begin his exam?, where yes = subject interpretation). Responses were categorized in terms of % subject interpretation.

Figure 1. Results from Exp. 1 and Exp. 2



References:
Contemori, C. (2019). Changing comprehenders' pronoun interpretations… Second Language Research. ✠ Fine, A. B., et al. (2013). Rapid expectation adaptation … PLoS ONE, 8. ✠ 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.*