Learnability and constraints on the semantics of clause-embedding predicates

Mora Maldonado, Jennifer Culbertson and Wataru Uegaki

Summary. Responsive predicates (RPs) are clause-embedding predicates like English *know* and *guess* that can take both declarative and interrogative clausal complements. The meanings of RPs when they take a declarative complement and when they take an interrogative complement are hypothesized to be constrained in systematic ways. Here we investigate whether one such constraint—C(lausal)-distributivity—is reflected in learning. To preview, we find that adults learning a novel clause-embedding predicate in the lab infer the constraint without explicit evidence.

Constraints on RP meanings. Since Karttunen (1977), a major question for the semantics of question-embedding is the relationship between the interpretation of a given RP when it embeds a **declarative** complement (e.g., *Jo knows **that** it is raining*) and when it embeds an **interrogative** complement (e.g., *Jo knows **whether** it is raining*). A number of proposals have been made in the form of constraints on the meanings of RPs. Two examples of such constraints are given below.

(1) **Veridicality constraint:** An RP is veridical w.r.t. declarative complements iff it is veridical w.r.t. interrogative complements (Spector & Égré 2015, i.a.), where *V* is veridical w.r.t. interrogative complements iff ⌜*x Vs Q*⌝ together with ⌜*p*⌝ entails ⌜*x Vs that p*⌝.

(2) **C-distributivity:** ⌜*x Vs Q*⌝ ⇔ there is an answer *p* to *Q* s.t. ⌜*x Vs p*⌝ (Theiler et al. '18).

Compared to the rich theoretical literature on these constraints (e.g., Spector & Égré; Theiler et al. 2018), relatively few attempts have been made to assess the validity of these constraints from empirical grounds. Notably, Sterinert-Threlkeld (2019) tested (1) in learnability experiments using neural nets, and Roelofsen & Uegaki (2021) surveyed the cross-linguistic validity of several constraints including (1) and (2). Nevertheless, it remains unclear whether human learners are sensitive to these kinds of constraints. In this study, we tested the hypothesis that RPs satisfy (2). From this hypothesis, we derive a novel learning-based prediction: when learning a new RP, learners will infer that it is C-distributive. We tested this hypothesis for two different RPs: 'falsely believe' (FALSEBEL) and 'has a correct belief that *p* is false' (KNOWFALSE). The former would be C-distributive if *Jo* FALSEBEL *whether it's raining* is true only in situations where Jo believes a false answer to the question of whether it's raining. The latter, if *Jo* KNOWFALSE *whether it's raining* is true only in situations where Jo believes a true answer.

Experimental design. Participants are randomly assigned to one of two possible conditions. In both conditions, they learn a new verb *lem*, which can be combined with declarative and interrogative complements. Conditions differ on whether *lem* means KNOWFALSE or FALSEBEL. Participants are first trained on how to use the predicate *lem* with declarative complements, in sentences of the form *Jo lems that p*, where *p* is one of [*it's raining outside*, *it's sunny outside*, *it's snowing outside*]. The training consists of: (a) Exposure phase: participants are shown the situations where they can use a sentence of the form *Jo lems that p* (positive evidence only; Fig.1A); and (b) Acceptability phase: Participants are shown different situations and asked to decide whether a sentence of the form *Jo lems that p* could be used to describe them (Fig.1B). The situations illustrate where *lem* can be used and where it cannot be used. Participants are given feedback on their answers, so they get both positive and negative evidence. For example, in the FALSEBEL condition, participants are shown that they cannot use *lem* in a situation where Jo has a true belief about the weather. Participants are then tested on their interpretation of sentences of the form *Jo lems Q*, where *Q* is *what the weather is like* (Fig.1C). Participants are asked whether the sentence *Jo lems Q* can be used in the following three situations: (i) When Jo believes a true answer to *Q* (True answer); (ii) When Jo believes a false answer to *Q* (False answer); (iii) When Jo has no belief

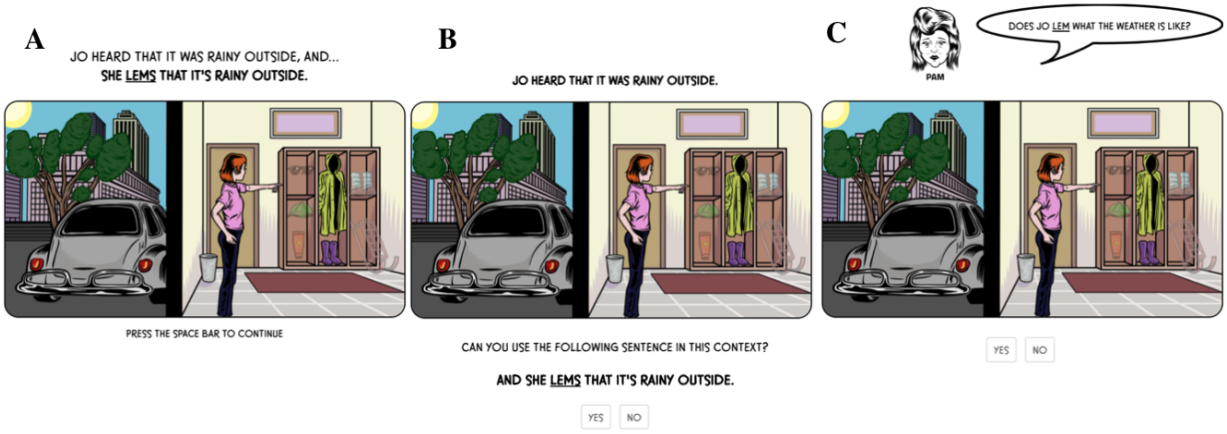

Figure 1: Example trials for Exposure (**A**), Acceptability (**B**) and Testing (**C**) in the FALSEBEL condition.

(No answer). No feedback was given in this part. Learners who infer that *lem* is C-distributive in the FALSEBEL/KNOWFALSE condition are expected to accept the sentence *Jo lems Q* in False/True answer situations, and reject it otherwise (in No answer and True/False answer situations).[1]

Results. 61 English-speaking participants were recruited on Prolific and successfully trained on the use of *lem* with declarative complements (FALSEBEL=40; KNOWFALSE=21). Fig.2 shows the proportion of responses compatible with C-distributivity during testing for each condition, across situations (True, False and No answers). A logit mixed-effects model, including random intercepts per subject (nested by condition) and situation, revealed that the proportion of trials in which *lem* is treated as satisfying C-distributivity is significantly above chance (β =3.73; p =.0024).

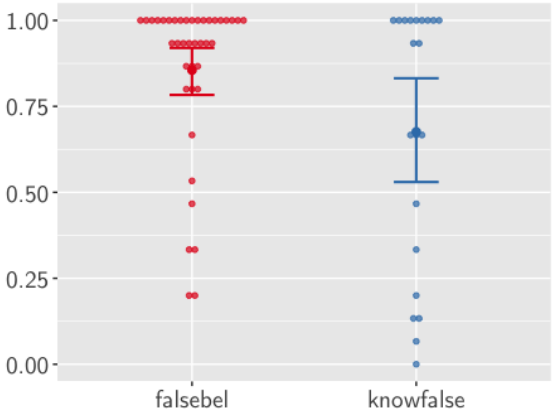

Figure 2: Responses compatible with C-distributivity.

Discussion. Our results suggest that the learning-based prediction derived from the hypothesis that RPs must satisfy (2) is borne out for the novel RPs in our experiment. Note, however, that this finding is mainly driven by the FALSEBEL condition, as the sample in the KNOWFALSE group is too small to confidently infer a pattern.[2] Notably, our results cannot be explained by (1) because (1) doesn't make any prediction about the participants' choices in Testing (both predicates are non-veridical w.r.t. interrogative complements regardless of participants' choices). While our results concerning KNOWFALSE are still tentative, they align Roelofsen & Uegaki ('21) who observe that RPs tend to obey (a version of) (2) cross-linguistically. Importantly, the results also suggest that this constraint might drive inferences during natural language acquisition, thus providing a mechanism for explaining this cross-linguistic tendency. // **References.** Karttunen. 77. Syntax and semantics of questions • Roelofsen & Uegaki. 21. Searching for a universal constraint on ... • Spector & Égré. 15. A uniform semantics for embedded interrogatives • Steinert-Threlkeld. 19. An Explanation of the Veridical Uniformity Universal • Theiler, Roelofsen, & Aloni. 18. A uniform semantics for declarative and interrogative complements.

[1] This experiment, including predictions, design, and analysis was preregistered here.

[2] KNOWFALSE turned out to be very difficult to learn w.r.t. declarative complements to begin with, and for this reason we have not been able to collect our target sample size.