

Incremental understanding of conjunctive generic sentences

Michael Henry Tessler, Karen Gu, and Roger Levy

Department of Brain & Cognitive Sciences, Massachusetts Institute of Technology

Generic statements convey generalizations about categories, but how generic predications should combine is unclear. “Elephants live in Africa and Asia” does not mean that individual elephants live on both continents; instead, the sentence should be understood as *elephants live in Africa, and elephants live in Asia*, but this is impossible if each individual generic sentence means that the majority hold the property (i.e., it is impossible for more than half of elephants to live in Africa and more than half of elephants to live in Asia; Nickel, 2008). The puzzle of understanding conjunctive generic sentences deepens when one considers that actual linguistic input is processed incrementally: Listeners ubiquitously form expectations about the intended meaning before the speaker finishes their sentence (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). For conjunctive generics about mutually exclusive (ME) properties, strongly incremental language understanding might produce non-monotonic interpretation updates: after the sentence prefix “Elephants live in Africa. . .”, a comprehender might infer a higher property *prevalence* than after hearing the sentence completion “. . . and Asia”.

To explore these ideas, we extend a model for interpreting generic language to incorporate an incremental processing mechanism that allows a listener to understand partial utterances. The original model of Tessler and Goodman (in press) can already adjust the meaning of a generic utterance by reasoning about an underspecified threshold θ . The model can then interpret multiple generics in succession, using the posterior distribution over prevalence as the prior for the next utterance. The model makes the prediction that for ME properties, conjunctive generics will non-monotonically update beliefs about the property prevalence in the target category (e.g., first believing that almost all elephants live in Africa and then believing that only half live in Africa).

If listeners parse and interpret utterances incrementally at the level of individual words, then incorporating probabilistic syntactic expectations (Levy, 2008) into this model predicts that when a comprehender encounters the conjunction “and”, they will infer the prevalence of elephants in Africa as a mixture of the inferences derived from different possible continuations (e.g., if the full sentence will be “live in Africa and eat bugs” vs. “live in Africa and Asia”). If, however, listeners instead wait for meaningful pieces of an utterance (e.g., content words like *Asia*) to compute interpretations, then we would not expect such an intermediate degree of interpretation: “Elephants live in Africa and. . .” should mean the same thing as “Elephants live in Africa. . .” (Fig. 1).

We design two experiments to test the mutual exclusivity (ME) and incremental predictions of the model. The first experiment tests the ME predictions that “Elephants live in Africa and

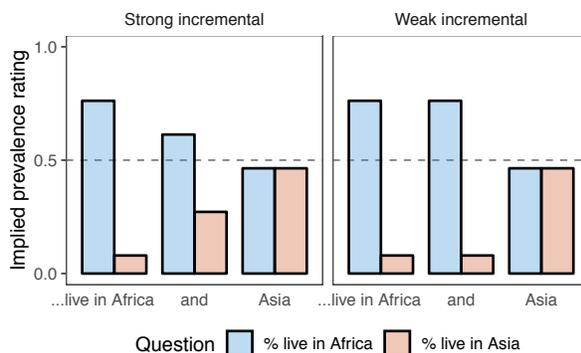


Figure 1: A model that incorporates syntactic expectations at the level of individual words predicts intermediate mutual-exclusivity inferences part-way through the conjunction (at “and”), whereas a model that waits for content-words (e.g., “Asia”) to begin to understand the utterances does not show a difference in expected prevalence at the word “and.”

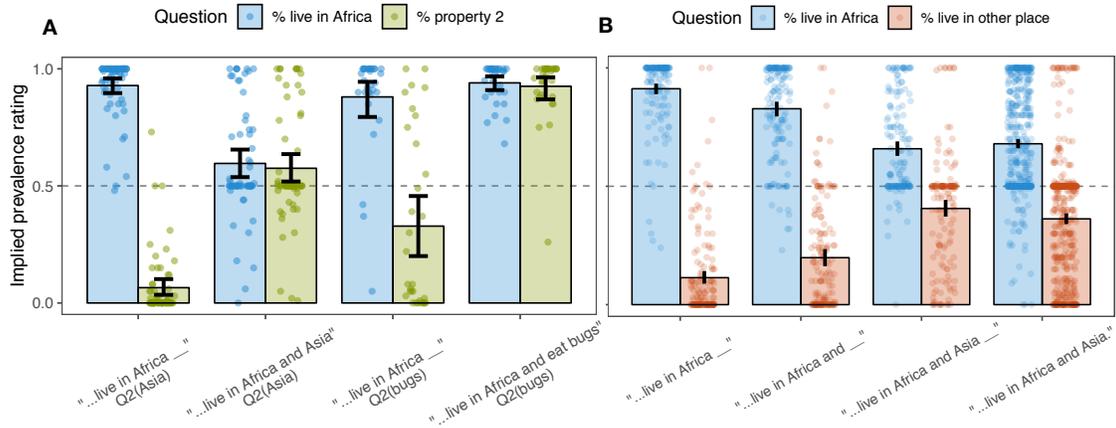


Figure 2: Experimental results. A: Participants rate prevalence for mentioned property (% live in Africa) and “Property 2”, either the mutually exclusive property (left two bars) or non-mutually exclusive property (right two bars). B: Participants are interrupted at various stages of the sentence (denoted by “_”) or uninterrupted (right-most bars) to be asked about the prevalence of *living in Africa* and *living in some other place*. Error-bars denote bootstrapped 95% confidence intervals.

Asia” means roughly that half live in Africa and half live in Asia. The second experiment is a pre-registered study that uses the gating paradigm to test the fine-grained incremental predictions of the model. The experiments and list of materials can be viewed at tinyurl.com/elephants-cogsci.

In both experiments, participants read a storybook about creatures on a faraway planet. Chapters of the book consist of a few sentences teaching about the category. Critical trials end with a conjunctive generic sentence (“Ks F and G”). Participants are asked to rate the % of Ks with F and the % of Ks with G using two slider bars ranging from 0% to 100%. Filler trials contain sentences using quantifiers *all*, *most*, and *none*.

Expt. 1 contains four conditions, randomized within-participant and within-item. Participants are either asked about two ME properties (e.g., live in Africa and Asia) or two non-ME (NME) properties (e.g., live in Africa and eat bugs). Participants are either asked at the end of the chapter or before the final page of the chapter, where the conjunctive generic is broken before the “and” (i.e., participants only read “Elephants live in Africa”). We find that when asked about mutually exclusive properties, participants exhibit non-monotonic belief updating, first believing that almost all elephants live in Africa and then believing that only half live in Africa; this phenomenon is not observed when asking about non-mutually exclusive properties predicated of a category (Fig. 2A).

Expt. 2 contained three conditions, corresponding to the point in the conjunctive generic sentence at which the page-break and prevalence questions occurred: “Elephants live in Africa... and... Asia...” (where ... denotes the point at which the page broke and the prevalence questions were presented). In a pre-registered study ($n=80$), we found that participants provided substantially lower estimates of prevalence when provided with the conjunct “and” in a partial sentence (posterior mean estimate and 95% credible interval: $\beta = -0.08(-0.13, -0.04)$) and that these ratings were also substantially higher than when participants read the full conjunctive predicate ($\beta = 0.17(0.12, 0.23)$; Fig. 2B). These results support a strong view of incrementality, suggesting that listeners begin to draw pragmatic interpretations of generics before the end of the sentence and even in the absence of additional content words.

References

- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177. doi: 10.1016/j.cognition.2007.05.006
- Nickel, B. (2008). Generics and the ways of normality. *Linguistics and Philosophy*, *31*(6), 629–648. doi: 10.1007/s10988-008-9049-7
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217).
- Tessler, M. H., & Goodman, N. D. (in press). The language of generalization. *Psychological Review*. Retrieved from <http://arxiv.org/abs/1608.02926>