The role of relevance, competence and priors for scalar implicatures

Polina Tsvilodub (Osnabrück University, ptsvilodub@uos.de), Bob van Tiel (Radboud University), Michael Franke (University of Tübingen)

If someone says "Anna ate some cookies", the hearer might infer the upper-bounded reading that Anna ate *some, but not all* cookies. Similarly, given "Donald ate a donut or a pretzel.", one might infer that Donald ate either *the donut or the pretzel, but not both* (i.e., an exclusive interpretation). Both inferences are usually explained as a variety of *scalar implicature* (SI). SIs rely on lexical scales consisting of words ordered in terms of informativeness, like ⟨some, all⟩ and ⟨or, and⟩. If a speaker uses an informationally weaker term (e.g., "some"), they may imply that the corresponding stronger alternative (e.g., "all") is false [3,5]. Crucially, prior research suggests that the robustness of SIs is influenced by different contextual factors [1]. We investigate the effects of three factors: (1) the *competence* of the speaker about the truth of the stronger alternative, (2) its *relevance* to the listener, and (3) the *prior probability* that it is true [2,4,6]. We explore how these three factors *interactively* influence the robustness of SIs associated with the triggers "some" and "or".

In our web-based rating study, participants read background stories which were designed to vary in terms of the strength of the three factors (`high` or `low` with respect to: prior probability$\times$competence$\times$ relevance, for each trigger), manipulated within-subjects. On critical trials, participants were asked to rate three sentences on a scale ranging from "certainly true" to "certainly false" ($\hat{=}$0–100), one per factor. The story ended with one of the characters in the story making an utterance containing "some" or "or". Participants then had to indicate the probability of an SI-enriched paraphrase of that utterance. We thus obtained judgements on the contextual factors, and on the robustness of the SI (see `https://tinyurl.com/3ru9sdja` for experiment details). Based on the literature, we expected higher likelihood ratings for the SI-enriched paraphrase if the alternative was perceived as highly relevant, the speaker was judged as highly competent, and the stronger alternative was viewed as a priori unlikely [2,4,6]. Each participant saw four stories per trigger, sampled from 32 stories/triggers, randomly shuffled with eight structurally similar attention checks and comprehension questions.

We analysed data from 206 participants recruited on Prolific. Their ratings were z-scored within each factor by-participant. We regressed the implicature likelihood ratings against the fixed effects of trigger, all factor ratings within-story, all interactions and maximal random effects, using a Bayesian linear mixed effects model. Participants' factor ratings by-story agreed well with the designed classification of the stories (Fig. 1, red vs. blue color on x-axis). As predicted, participants were more likely to derive the SIs of "some" and "or" when judging speaker competence as high ($P = 0.999$ for "some", $P = 0.993$ for "or" for effect sizes being > 0.05, Fig. 2 for all results). They were also more likely to derive the SI of "some" when judging the prior probability of "all" as low ($P = 1$ for effect < 0.05), which was not the case for "or". Finally, we did not observe credible effects of relevance for either trigger. Given the unexpected absence of prior effects for "or", we computed exploratory pairwise correlations of all predictors. While no correlations were found for "some", we found a significant correlation between the explanatory factors prior and relevance for "or" ($R^2 = -0.106, p < 0.01$). An exploratory model comparison of two models, one containing the relevance effect over one containing the prior effect, each combined with competence, revealed mild evidence in favor of the prior as a better explanatory factor than relevance (Bayes Factor $= 3.70$). While supporting the view that SIs for both triggers rely on epistemic reasoning affected by speaker competence, our results indicate that prior and relevance might be closely connected for "or". Ultimately, our results suggest that the interdependence of the three factors is more complex than just the sum of the effects anticipated in the literature, and provide further insights into how they might enter into people's decision as to whether or not to derive an SI.

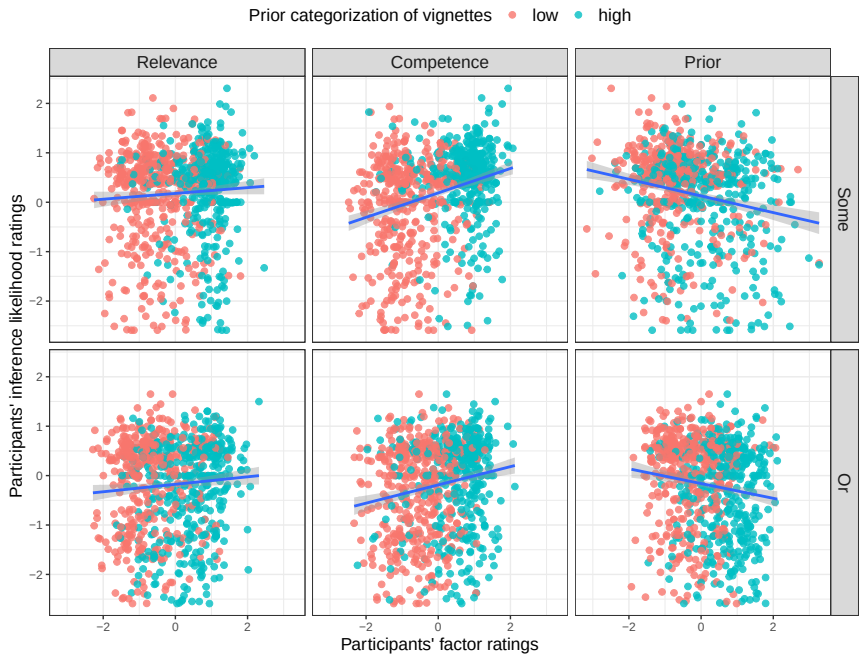

Figure 1: Relating ratings for relevance, competence and prior statements (x-axis) to ratings for the strength of pragmatic enrichments (y-axis). The top row shows ratings for "some" (enriched to "some, but not all"). The bottom row shows ratings for "or" (enriched to "A or B, but not both"). Ratings for stories categorized as low (red) w.r.t. a given factor are on average lower (x-axis) than for those categorized as high (blue). The apparent effect of prior for "or" is mitigated by collinearity with relevance.

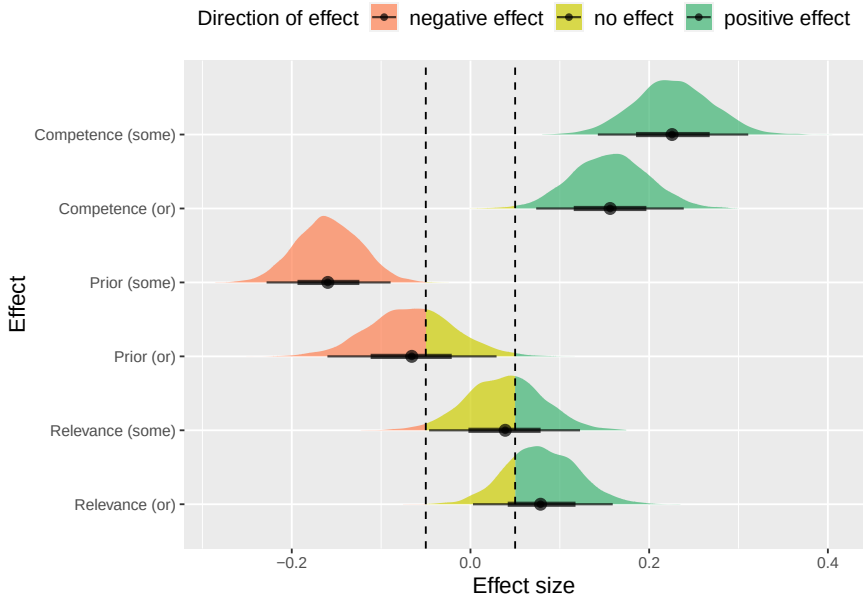

Figure 2: Distributions of posterior samples for the effects (x-axis) of each predictor for each trigger (y-axis). The colors indicate the fraction of the distribution density corresponding to a positive, no or a negative effect. The dashed lines indicate the Region Of Practical Equivalence that we defined as (-0.05, 0.05) in which posterior samples are considered equivalent to 0 (i.e., no credible effect).

References: [1] Degen, J., In *Semantics & Pragmatics*, 2015 [2] Degen, J., Tessler, M. H., & Goodman, N. D., In *Proceedings of CogSci*, 2015 [3] Geurts, B., *Quantity Implicatures*, 2010 [4] Goodman, N. D., & Stuhlmüller, A., In *Topics in Cognitive Science*, 5(1), 2013 [5] Horn, L. R., *On the semantic properties of logical operators in English*, 1972 [6] Sperber, D., & Wilson, D., *Relevance: communication and cognition*, 1995