On the salience of linguistic alternatives in the inference task for scalar implicatures

Background. Variability in rates of Scalar Implicatures (SIs) has been observed across many studies: between contexts, individuals, participant groups, and scalar expressions. Here we focus on another kind of variability – the fact that inference tasks tend to result in higher rates of with-SI response than comparable verification tasks. [1] explicitly demonstrates this fact using the two tasks with the same sentences. It can also be observed in comparing outcomes for scalar expressions appearing both in inference tasks like [2] and verification tasks like [3]. A by now standard inference task stimulus is shown in Fig.1a (based on [2]). To account for the raised rates in inference tasks, [1] conjectures that, by asking the participant if the speaker excludes the alternative, the probe question strongly suggests that it is relevant. Another factor that may be at play is that the probe question references the linguistic alternative. According to some views, salience of the alternative expression itself can impact positively on SI availability [4, 5]. This view has recently been challenged in [6, 7]. [7] argues that mere salience of the scalar expression is relatively inert in promoting SI. We report on a study that tested these competing ideas on the efficacy of alternative salience by manipulating whether the alternative was explicitly mentioned, implicitly present, or entirely absent in the probe, while holding constant the meaning of the question asked and thereby the relevance of alternatives.

Experiment. Our test trials are illustrated in Fig.1. In each condition, the lexical content of the target statement was manipulated to test whether the presence of the alternative has any effect on SI rates above that of making the proposition expressed by the alternative contextually relevant. For these purposes, we introduced two novel ANTONYM probes in addition to the standard NOT-ALT probe. Unlike in the NOT-ALT probe, the query in the ANTONYM probes expressed the falsity of the alternative of interest by other linguistic means than referencing the stronger alternative and embedding it under negation. We further distinguished between ANTONYM and ANTONYM* probes in order to detect if implicit activation of the alternative promotes SI. This is possible in the former, as opposed to the latter, since the scalar expression is employed in the probe and this itself may trigger a SI, involving a representation of the alternative in its derivation. ANTONYM* probes were variants of the ANTONYM probes in which neither the weak scalar expression, nor its stronger scalemate appeared. These probes were created mainly by using a blank paraphrase of the weak scalar expression (as in Fig.1c), or else by replacing that expression with a lexical antonym (e.g., replacing *tried to* with *failed to*). We tested 12 lexical scales (see Table 1). Probe was a between-group factor. Participants ($n = 164$) were assigned to one of three lists containing 36 target items (3 instances of 12 scales) plus 10 control items. We hypothesized that if raising the salience of an alternative has a boosting effect on SI rates above that of suggesting its relevance, the proportions of Yes-responses in the test trials should be lower in either of the ANTONYM conditions than in the NOT-ALT conditions.

Main results. The distribution of by-participant mean rates was very similar in all three probe conditions, as shown in Fig.2. We fitted a Bayesian mixed effects logistic regression model to the data. The hypothesis that ANTONYM(*) should yield lower rates of acceptance than NOT-ALT was tested using the hypothesis function of `brms`. The posterior probability of ANTONYM yielding lower rates of acceptance than NOT-ALT was 49% with an evidence ratio of 0.96, and the difference was estimated to be 0.01 with 90% quantiles being [-0.54,0.55]. For ANTONYM*, the posterior probability was 40% with an evidence ratio of 0.67, and the difference was estimated to be 0.08 with 90% quantiles being [-0.47,0.64]. Fig.3 shows the mean rates by Scale and Probe type. For each scale, we fitted a GLMER model with a logit link function, predicting participants' responses from the fixed effect of Probe (treatment coded). The results of the model comparison tests showed that including Probe as a predictor led to a significantly improved fit over the null model for only two scales, ⟨permit, require⟩ and ⟨few, lot⟩. For both these scales, the estimated marginal means were significantly higher in the ANTONYM conditions than in the NOT-ALT conditions. We conclude that the by-scale rates of SIs were largely unaffected by the Probe manipulation, consistent with the results of the global analysis.

Discussion. Our results show that SI rates are much the same across all three probe conditions and they provide evidence against the hypothesis that making the alternative contextually salient has a boosting effect on SI rates above that of merely raising the relevance of that alternative. These findings, on the other hand, are in line with the idea that the probe question generally biases participants to think that the alternative is relevant, enhancing the likelihood that the SI reading be endorsed and accounting in turn for the inflated rates of SIs yielded by the inferential paradigm.

(a) NOT-ALT

George says:

Some of the students passed the exam.

Would you conclude from this that, according to George,
not all of the students passed the exam?

Yes No

(b) ANTONYM

George says:

Some of the students passed the exam.

Would you conclude from this that, according to George,
some of the students failed the exam?

Yes No

(c) ANTONYM*

George says:

Some of the students passed the exam.

Would you conclude from this that, according to George,
there were students who failed the exam?

Yes No

Figure 1: Example test trials in the (a) NOT-ALT, (b) ANTONYM and (c) ANTONYM* conditions, here for the scale ⟨some, all⟩. A Yes-response in these trials indicates that an SI is drawn.

Table 1: Scales tested in the experiment by category.

Category	Scales
Adjective	⟨possible, certain⟩, ⟨good, excellent⟩, ⟨difficult, impossible⟩
Adverb	⟨sometimes, always⟩
Connective	⟨or, and⟩
Determiner	⟨some, all⟩, ⟨a few, a lot⟩
Verb	⟨allow, require⟩, ⟨may, have to⟩, ⟨permit, require⟩, ⟨try, succeed⟩, ⟨participate, win⟩

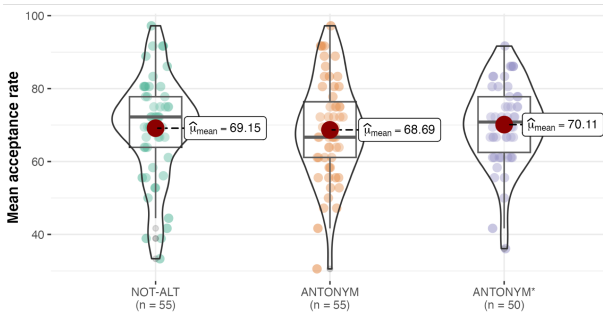

Figure 2: Percentage of Yes-responses to the test trials by Probe condition.

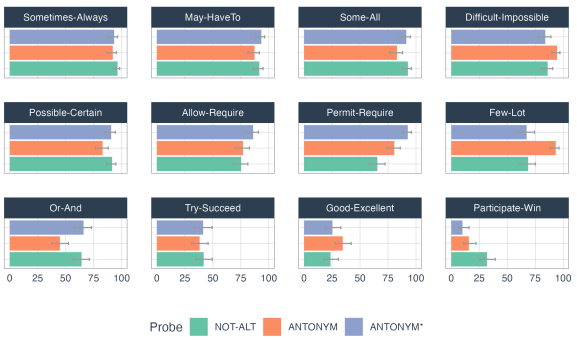

Figure 3: Percentage of Yes-responses to the test trials by Scale and Probe condition.

References [1] Geurts, B. & Pouscoulous, N. (2009). *Embedded implicatures?!?* [2] van Tiel, B., van Miltenburg, E. Zevakhina, N. & Geurts, B. (2016). *Scalar diversity* [3] van Tiel, B., Pankratz, E. & Sun, C. (2019). *Scales and scalarity: Processing scalar inferences.* [4] Barner, D., Brooks, N. & Bale, A. (2011). *Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference* [5] Rees, A. & Bott, L. (2018). *The role of alternative salience in the derivation of scalar implicatures.* [6] Skordos, D. & Papafragou, A. (2016). *Children's derivation of scalar implicatures: Alternatives and relevance.* [7] Marty, P., Romoli, J., Sudo, Y., & Breheny, R. (2024). *Implicature priming, salience, and context adaptation.*