Question availability as a predictor of scalar diversity

Eszter Ronai & Ming Xiang

The University of Chicago

AMLaP 26
3-5 September 2020

Objectives

Investigate the role of Questions Under Discussion (QUDs) in explaining scalar diversity.

Background

Listeners reason about what is not said: the stronger alternative—*all* in (1) and *brilliant* in (2).

(1) Mary ate some of the cookies. → SI: Mary ate some, but not all, of the cookies.

(2) The student is intelligent. → SI: The student is intelligent, but not brilliant.

→ Scalar inference (SI) calculation.

Scalar diversity

Considerable **variation across different scales in SI calculation rates**; e.g. the SI in (1) arises much more robustly than the one in (2) (Van Tiel et al. 2014; see also Doran et al. 2012; Beltrama & Xiang 2013).

What properties of scales can explain this variation?

- Distinctness of the stronger scalar term (Van Tiel et al. 2014).
- Local enrichability (Sun et al. 2018).
- But: still a lot of variance unaccounted for in the empirical results.

The role of context

QUDs have an effect on the rate of SI calculation: *Did Mary eat all of the cookies?* will lead to a higher SI rate in (1) than *Did Mary eat any/some of the cookies?* (i.a. Degen & Tanenhaus 2014; Ronai & Xiang 2019; Yang et al. 2018; Zondervan et al. 2008).

In previous work on scalar diversity, stimulus sentences were presented in the absence of any context.

Open question: is there variation across scalar terms in what kind of QUD they most naturally bring to mind?

Hypothesis

Scalar diversity, in the absence of an explicit QUD, arises (in part) **due to the differential availability of a polar question containing the stronger scalar term** from the scale.

Intuition: the more likely a question such as *Is the student brilliant?* is, the higher the rate of SI calculation from the corresponding statement *She is intelligent*.

Experiments

Experiment 1: replication of van Tiel et al. (2014)

- 37 native speakers of American English; MTurk; IbexFarm.
- Inference task to investigate the likelihood of deriving an SI from 43 different scales.
- Participants were presented with "Mary: *The student is intelligent.*" + asked the question "Would you conclude from this that, according to Mary, the student is not brilliant?".
  - "Yes" response = SI was calculated.
  - "No" response = SI was not calculated.

Experiment 2: inference task with Question manipulation

- 40 native speakers of American English; MTurk; IbexFarm.
- Basic inference task identical to Experiment 1.
- Two-condition Question manipulation: Mary's statement embedded in a dialogue context.
  - Question containing stronger scalar term: "Sue: *Is the student brilliant?*"; "Mary: *She is intelligent*".
  - Question containing weaker scalar term: "Sue: *Is the student intelligent?*"; "Mary: *She is intelligent*".

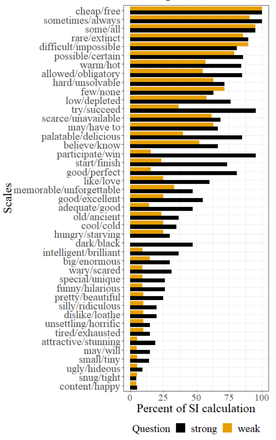
Experiment 3: question availability

- 35 native speakers of American English; MTurk; IbexFarm.
- Forced choice task: participants had to choose which of two polar questions (containing the stronger vs. the weaker scalar term) they would be more likely to ask.
  - "Compare the following two questions about a student. Which one are you more likely to ask?":
  - → choice between: *Is the student brilliant?* vs. *Is the student intelligent?*.
- **Prediction**: forced choice results (henceforth **Question Choice**) **should predict scalar diversity**.
  - The more preferred the stronger question in Exp. 3, the higher the SI rate for that scale in Exp. 1.

Experiment 2 results

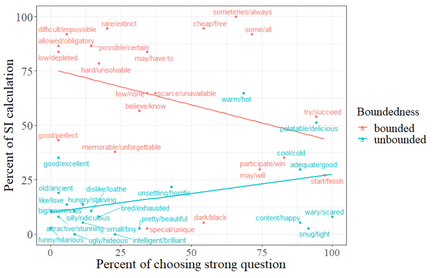
Across the board, more SIs derived when the preceding question contains the stronger scalar term than when it contains the weaker one (significant effect of Question, $p < 0.001$).

Explicit QUD influences SI calculation rates for a large number of scales.



Experiment 1 and 3 results

- Exp. 1 (y axis): replicated the scalar diversity effect.
- Exp. 3 (x axis): Question Choice not an overall predictor of SI calculation rates, contra our predictions.



Boundedness (Van Tiel et al. 2014)

Bounded scale: the stronger scalar denotes an endpoint, e.g. *all* → <*some, all*> is bounded.

- Bounded scales led to higher SI rates than unbounded ones ($p < 0.001$; replicates Van Tiel et al. 2014).

Interaction of Question Choice with Boundedness ($p < 0.05$):

- Unbounded scales: Question Choice showed a strong trend ($p < 0.08$) in predicting SI calculation.
  - The more likely participants were to choose the strong question (*Is the student brilliant?*), the higher the rate of calculating the relevant SI (*intelligent→not brilliant*).
- Bounded scales: no effect of Question Choice ($p = 0.14$).

Discussion

Unbounded scales: both scalar terms are vague; they denote intervals whose values vary according to context. → opportunity for contextual support.

- The more available a QUD based on the stronger term is, the more likely hearers will be to reason about that term as the stronger alternative.
- → More likely to derive the SI.

Bounded scales: the stronger scalar is not vague, but instead denotes a fixed point.

- Stronger scalar term is more salient as a stronger alternative to the vague, weaker term.
- → High rates of SI calculation; the QUD makes no difference.

Future work

Other types of QUDs for Exp. 2, setting up biasing contexts without explicitly mentioning the scalar terms.

- Right now, effects may be due to the strong question raising the contextual salience of the alternative.

Better empirical measure of question availability than Experiment 3?

- Which question a speaker is more likely to choose may itself be context-dependent.

Conclusions

- **QUDs robustly affect SI calculation rates for a large number of scales**: questions based on the stronger of two scalar terms lead to higher SI rates.
- **Likelihood of a question based on the stronger scalar contributes to scalar diversity**, but only **for unbounded scales**.

References

Beltrama & Xiang (2013). Is 'good' better than 'excellent'? An experimental investigation on scalar implicatures and gradable adjectives. Proceedings of SuB 17. | Degen & Tanenhaus (2014). Processing scalar implicature: A constraint-based approach. Cognitive Science. | Doran et al. (2012). A novel experimental paradigm for distinguishing between what is said and what is implicated. Language. | Ronai & Xiang (2019). Calculating scalar inference under QUDs. Proceedings of NELS 49. | Sun et al. (2018). A link between local enrichment and scalar diversity. Frontiers in Psychology | Van Tiel et al. (2014). Scalar diversity. Journal of Semantics. | Zondervan et al. (2008). Experiments on the role of the question under discussion for ambiguity resolution and implicature computation in adults. Proceedings of SALT 18. | Yang et al. (2018). Context-sensitivity and individual differences in the derivation of scalar implicature. Frontiers in Psychology.