Processing quantity implicatures under QUDs

Eszter Ronai & Ming Xiang

The University of Chicago

24th AMLaP
6-8 September 2018

THE UNIVERSITY OF CHICAGO

Objectives

Investigate the effect of Question Under Discussion on implicature calculation and processing, looking at *some-all* scalar inferences and *it*-cleft exhaustivity:

- Different **rates of calculation** based on QUD?
- Different **time cost of processing** based on QUD?

Background

Two types of quantity implicatures: **scalar inferences** and ***it*-cleft exhaustivity**.

Some of the shapes are blue. (SI)

- Literal meaning: Some and possibly all of the shapes are blue.
- Inference-enriched meaning: Some but not all of the shapes are blue.

It is the square that is blue. (EXH, see i.a. Byram Washburn, et al., 2014)

- Literal meaning: The square and possibly other things are blue.
- Inference-enriched meaning: It it only the square that is blue.

Implicatures: **integrating semantic and pragmatic information**.

- Semantic information is privileged → implicature calculation is costly, as seen in reaction times (i.a. Bott Noveck, 2004), as well as eye-tracking, ERPs.
- Implicature calculation is effortless and default (Grodner, et al., 2010).
- Constraint-based approaches (Degen & Tanenhaus, 2015): implicature calculation and processing results from the **interaction of multiple cues and constraints** → no categorical distinction between "costly" vs. "cost free".

Identifying and quantifying cues e.g. **Question Under Discussion** (Roberts, 1996):

- Zondervan, et al. (2008): manipulation via explicit questions (*some* vs. *all*).
- Degen (2013): manipulation via background story (*none* vs. *all*).
- Cummins & Rohde (2015): manipulation via focus prosody (e.g. *pretty* vs. *gorgeous*, scalar diversity).

→ More implicatures calculated under QUDs with the stronger alternative (e.g. *all*).

However, **questions remain about time cost** and **how QUDs can be empirically elicited** and verified in a more systematic way.

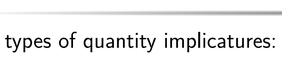
Design

Story: Anne is asking questions from Bob, about pictures that only Bob can see.

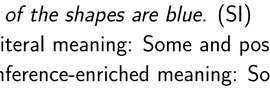
- Control: Bob's answers unambiguously good/bad descriptions of the picture.
- Target: descriptions either good (on literal reading) or bad (inference-enriched).

SI: *Some of the shapes are blue.*

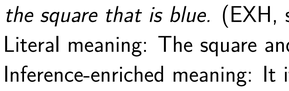
Control: Good



Control: Bad



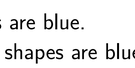
Target: Underinformative



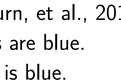
EXH: *It is the square that is blue.*

Control: Good

Control: Bad



Target: Underinformative



Design and materials modified from van Tiel & Schaeken (2017), who found that (when there is no QUD manipulation) SI calculation has a time cost, but EXH does not.

References

Bott, Lewis, & Noveck, Ira A. 2004. Some utterances are underinformative: The onset and time course of scalar inferences. Journal of Memory and Language, 51(3), 437-457. | Byram Washburn, Mary, Kaiser, Elsi, & Zubizarreta, Maria Luisa. 2014. The English It-Cleft: No Need to Get Exhausted. In: Proceedings of the 'Questions in Discourse Conference'. | Cummins, Chris, & Rohde, Hannah. 2015. Evoking Context with Contrastive Stress: Effects on Pragmatic Enrichment. Frontiers in Psychology, 6, 1779. | Degen, Judith. 2013. Alternatives in Pragmatic Reasoning. Ph.D. thesis, University of Rochester. | Degen, Judith, & Tanenhaus, Michael K. 2015. Processing Scalar Implicature: A Constraint-Based Approach. Cognitive Science, 39(4), 667-710. | Grodner, Daniel J., Klein, Natalie M., Carbary, Kathleen M., & Tanenhaus, Michael K. 2010. Some, and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. Cognition, 116(1), 42-55. | Roberts, Craige. 1996/2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. Semantics and Pragmatics, 5(6), 1-69. | van Tiel, Bob, & Schaeken, Walter. 2017. Processing conversational implicatures: alternatives and counterfactual reasoning. 41, 1119-1154. | Zondervan, Arjen, Meroni, Luisa, & Gualmini, Andrea. 2008. Experiments on the Role of the Question Under Discussion for Ambiguity Resolution and Implicature Computation in Adults. Pages 765-777 of: Friedman, Tova, & Ito, Satoshi (eds), Proceedings of SALT 18.

Experiment 1: QUD elicitation

Participants and procedure:

40 native monolingual speakers of American English.

- Participants saw SI and EXH target sentences paired with pictures, and were told that the sentences were Bob's answers to Anne's questions.
  - Anne: _____?
  - Bob: *Some of the..., It is the...*
  - Picture (Good Control or Target, between-participants)
- Task: guess what Anne's questions were.

Results:

Dominant SI questions:

- what: *What color are the shapes?*
- any: *Are any (of the) shapes black? Are there (any) red shapes?*
- all: *Are all of the shapes yellow?*
- some: *Are some of the shapes yellow?*

Dominant EXH questions:

- which: *Which/what shape is black? Which one (of them) is blue?*
- any: *Are any of the shapes yellow? Are there any black shapes?*
- what: *What color are the shapes? What color is the square?*

	SI				EXH		
	what	any	all	some	which	any	what
Target	42%	25%	6%	12%	54%	9%	8%
Good Control	32%	33%	20%	2%	67%	14%	6%

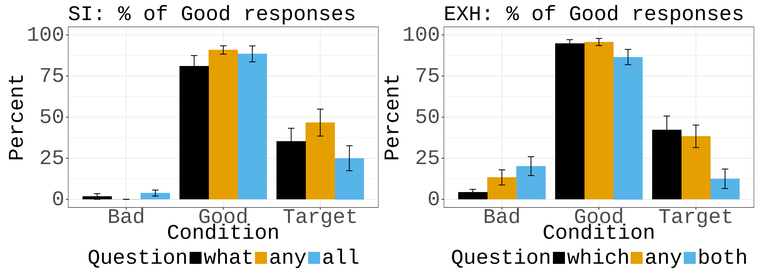
Experiment 2: QUD manipulation

Participants and procedure:

85 native monolingual speakers of American English (different from Exp. 1).

- **Sentence-picture verification task**: participants saw a **dialogue** between Anne and Bob, together with a picture.
- Task: make a **binary judgment** about whether Bob gave a good answer to Anne's question, given the picture he saw.
- We are interested in their response (Good/Not Good) and reaction time.
- 3 × 3 design: Picture (within-participants) × QUD (between-participants)
  - Anne's questions: most frequent questions elicited from Exp. 1.
    SI: What color are the shapes?, Are any shapes blue?, Are all shapes blue?
    EXH: Which shape is blue?, Are there any blue shapes?, Are both shapes blue?

Results: calculation rates



"Good" responses to Target: higher % indicate a lower rate of implicature calculation.

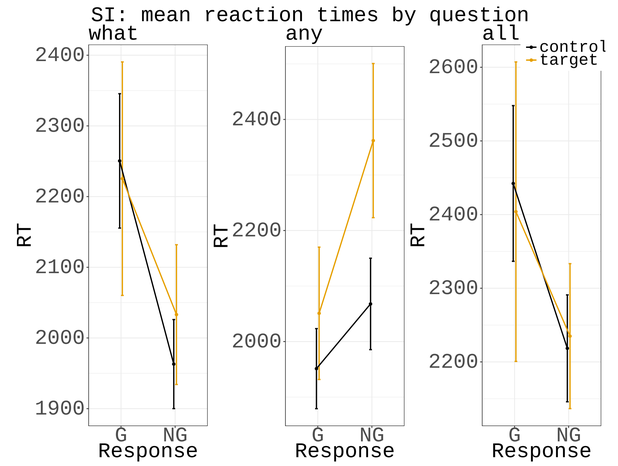
Fitting a generalized linear model; levels within the QUD variable treatment-coded:

- SI: any QUDs resulted in fewer implicatures than all ($p<0.001$) or what ($p<0.05$).
- EXH: both resulted in more implicatures than any ($p<0.001$) or which ($p<0.001$).

→ Interim finding: any and which are **Literal-biasing**, while what, all and both are **Implicature-biasing QUDs**.

Results: reaction times

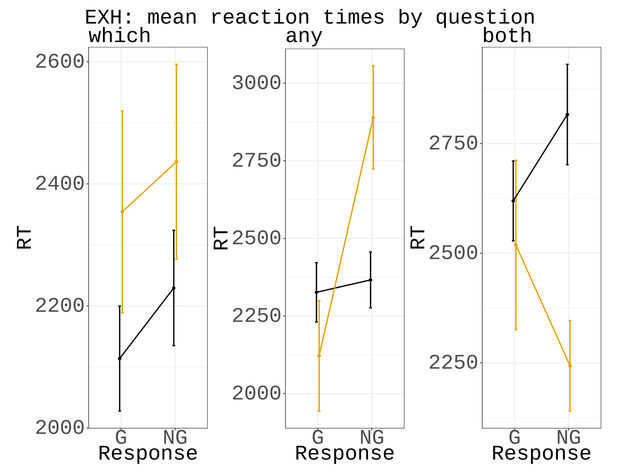
Cost of implicature calculation: longer reaction time when responding Not Good (NG) to Target, relative to the reaction time when responding NG to Control.



Mixed effects models were fit, predicting reaction time by Response, Picture and QUD, comparing any vs. all and any vs. what. We find a significant interaction of QUD and Response ($p<0.01$), such that:

- any: SI-enriched judgement takes longer than relevant (Bad) control, i.e. cost.
- what and all: no difference between NG to Target vs. NG to Control, i.e. no cost.

→ **SI computation is only costly** when preceded by **non-supportive QUDs**.



Similar, but more nuanced results:

- any: we also see a cost for inference computation, on NG to Target vs. Control (predicting reaction time by Response and Picture: significant interaction $p<0.05$)
- which (also Literal-biasing): similar, but not exactly the same pattern.
- both: unexpected cost for responding NG to Control.
  - Open question whether this is a side-effect of the picture stimuli.

Discussion and Conclusion

QUDs modulate implicature calculation rates and processing cost.

- QUDs that bias against implicature derivation make that derivation incur a reaction time cost.
- Under QUDs that bias towards implicature derivation, there is no cost.

→ No uniform cost or lack of cost for implicature derivation

Support for **constraint-based account** of implicature, and language processing more generally, where **QUD is one of many cues**.