Every ambiguity isn't syntactic in nature: Testing the Rational Speech Act model of Scope Ambiguity

Sherry Yong Chen & Bob van Tiel

XPRAG.de

Question & Goal

Ambiguity resolution:
Every horse didn't jump over the fence.
a. None of the horses jumped over the fence.
b. Not all of the horses jumped over the fence; some may have.

Syntactic explanation:
Different syntactic parses are postulated to yield a wide scope and a narrow scope reading of *every*
a. ∀ >> ¬ b. ¬ >> ∀

Pragmatic explanation:
Pragmatic cues such as the Question Under Discussion, salient alternatives, and world knowledge modulate the accessibility of these readings

An unaddressed question:
What is the division of labor between syntax and pragmatics in shaping people's interpretation of *every-not* utterances?

We experimentally test the predictions made by the Rational Speech Act framework, in particular the ambiguity resolution model by Scontras and Pearl (2020) (also see Savinelli et al., 2017, 2018). Our results suggest that variability in the interpretation of 'every-not' utterances can be explained almost entirely in terms of pragmatics, suggesting only a marginal role for syntax.

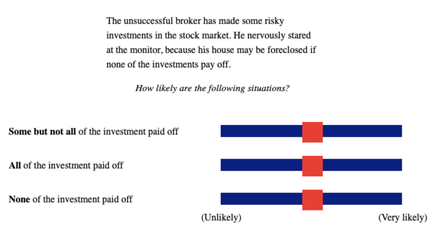
Experimental Setup

	All-prior	**Some**-prior	**None**-prior
Context	Rachel is an industrious teaching assistant who has to grade student essays. After a week, her boss asks…	Barney has been submitting recipes to a famous cooking website for some time now. His partner asked…	Jack is an incompetent homicide detective trying to solve three recent murders. His boss tells him that…
All?-QUD	if she is ready to enter the grades into the school system.	whether all of his submissions had been posted on the website.	he will have to look for a new job if he doesn't solve these murders.
HowMany? QUD	how much progress she's made so far.	how successful his submissions had been.	he will get a bonus depending on his performance.
Any? QUD	if she has already started grading the essays.	whether any of his submissions had even been posted on the website.	the newspapers will be all over him if he fails to solve any of these murders.
'Every-not' utterance	Rachel says: Every essay hasn't been graded.	Barney says: Every submission hasn't been published.	Jack says: Every case hasn't been solved.

- 72 short stories separately manipulating the prior probability and biases toward certain types of QUDs.
  - See examples 👈
- 3 experiments, each with 45 participants recruited on Prolific.
- For each story, determine (i) prior probabilities (Exp. 1), (ii) the most likely QUD (Exp. 2), and (iii) the interpretation of the "every-not" utterance (Exp. 3), by moving continuous sliders.

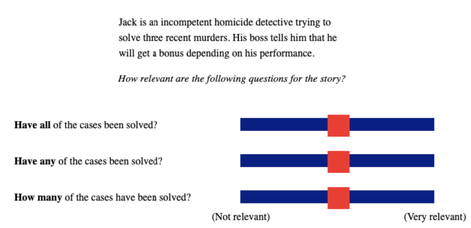
Exp 1: Prior Testing

- Purpose: test the prior probabilities of the various situations.



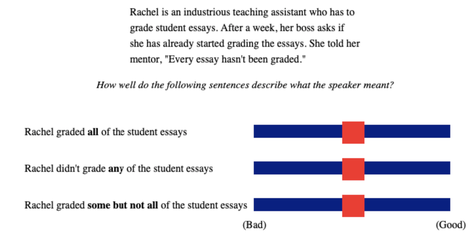
Exp 2: QUD Testing

- Purpose: test the most likely QUD in each story.



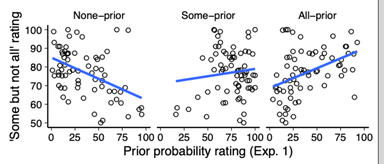
Exp 3: Ambiguity Resolution

- Purpose: test participants' interpretation of "every-not" utterances.

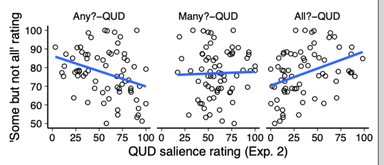


Testing the Scontras and Pearl (2020) Model

- **The Scontras & Pearl Model of Ambiguity Resolution:**
  - In a nutshell, the interpretation of "*every-not*" utterances is determined by (1) **prior probabilities**, (2) the **QUD**, and (3) the **syntactic scope assignment**.
  - Each utterance is indexed with a **scope** parameter *i*
    - "every-not" utterance: *surface* vs. *inverse*
    - *null* utterance is always true
  - Messages are produced and interpreted relative to a **QUD**
    - whether *all* individuals satisfy the predicate ($q_{all?}$)
    - whether *any* of the individuals satisfy the predicate ($q_{any?}$)
    - *how many* individuals satisfy the predicate ($q_{how\text{-}many?}$)
  - Recursive reasoning with a *literal listener* L_0, a *speaker* S_1, and a *pragmatic listener* L_1.



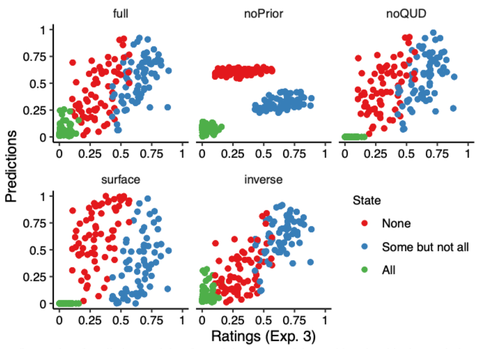
(a) Correlation between the rating for the 'some but not all' situation (Exp. 3) and the ratings for prior probabilities of the three situations (Exp. 1).



(b) Correlation between the rating for the 'some but not all' situation (Exp. 3) and the ratings for QUD salience of the three possible QUDs (Exp. 2).

- Crucially, in Scontras and Pearl's model, the syntactic parse ***constrains but does not determine*** listeners' interpretation
- **Parametrizing the model:**
  - Feed context priors and QUD priors obtained from Exp 1 & Exp 2 into the model(s) to make prediction.
  - Take results of Exp 3 to be the ***PragmaticListener*** inferring true states of the world. Compare with predictions.
- **Model Evaluation:** considering 4 models, each nullifying the effect of one of the components
  - $L_{noPrior}$ did not take into account prior probabilities over possible states (Exp. 1).
  - L_{noQUD} did not take into account the QUD (Exp. 2).
  - $L_{surface}$ always assigned surface scope.
  - $L_{inverse}$ always assigned inverse scope.

Finding 1: Participants' interpretation was mostly shaped by prior expectations, and only marginally by QUD.
Finding 2: We also observed an interaction between prior expectations and QUD salience. S&P's model assumes that they are completely independent; our results call that assumption into question.
Finding 3: The data was best described by a model that always assigned an inverse scope parse. But this may indicate low accessibility, not necessarily a entire lack of ambiguity



Scatterplot of predictions and data for each item, state of the world, and ambiguity resolution model.

	L_{full}	$L_{noPrior}$	L_{noQUD}	$L_{surface}$	$L_{inverse}$
per item	.82	.42	.81	.54	**.88**
per state	**.39**	.22	.34	.31	.36
overall	.78	.39	.79	.48	**.85**
average	.67	.34	.65	.45	**.70**

Correlations between data from Exp. 3 and predictions for each ambiguity resolution model.

Contact

Email: sychen@mit.edu
Website: sherrychen.org
Twitter: @linguistsherry

Email: bobvantiel@gmail.com
Website: https://sites.google.com/site/bobvantiel/

Selected References

[1] Buring, D. (1997). The great scope inversion conspiracy. [2] Conroy, A., Fults, S., Musolino, J., & Lidz, J. (2008). Surface scope as a default: The effect of time in resolving quantifier scope ambiguity. [3] Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. [4] Gualmini, A., Hulsey, S., Hacquard, V., & Fox, D. (2008). The question–answer requirement for scope assignment. [5] Kadmon, N., & Roberts, C. (1986). Prosody and scope: the role of discourse structure. [6] Musolino, J. (1998). Universal grammar and the acquisition of semantic knowledge. [7] Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. [8] Savinelli, K., Scontras, G., & Pearl, L. (2017). Modeling scope ambiguity resolution as pragmatic inference: Formalizing differences in child and adult behavior. [9] Savinelli, K., Scontras, G., & Pearl, L. (2018). Exactly two things to learn from modeling scope ambiguity resolution: Developmental continuity and numeral semantics. [10] Scontras, G. & Pearl, L. (2020). When pragmatics matters more for truth-value judgments: An investigation of quantifier scope ambiguity. [11] Syrett, K., Simon, G., & Nisula, K. (2014). Prosodic disambiguation of scopally ambiguous quantificational sentences in a discourse context.

MUQTASP