A computational model for projection inferences in clause-embedding predicates

Dingyi (Penny) Pan[1], Judith Degen[2]
[1]UC San Diego, [2]Stanford University

Background

Projection inferences: inferences about speaker's commitment to the embedded content that projects through an entailment-canceling environment such as polar interrogatives.

Scott says:
"John knows/thinks that Julian dances salsa."
"Does John know/think that Julian dances salsa?"

know: the listener infers that the speaker (*Scott*) believe the content of the embedded clause (*Julian dances salsa*).
think: the listener does not make such inference.

Factors that modulate projection inferences

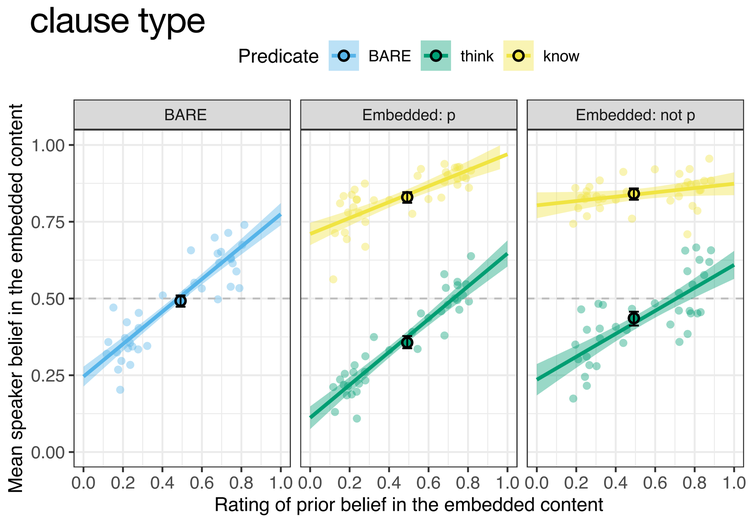
- **Predicates**: Predicates show different projection patterns [1,2]. The degree of projection is gradient and probabilistic [2,3].
- **At-issueness** (Gradient Projection Principle): The content projects to the extent that it is not at-issue with respect to the Question Under Discussion. [3-5]
- **Prior beliefs:** If the proposition *p* is more probable, then it is more likely to project. [6-8]

How do these factors interact to generate the observed probabilistic projection patterns?

Behavioral experiment

Participants
345 native English speakers recruited on Prolific.
Stimuli (10 critical items, 8 control items, 6 fillers)
2 facts x 2 predicates ("know" and "think") x 2 embedded clause types ("p" and "not p") + 2 facts x unembedded polar interrogatives ("BARE p")
Tasks: belief rating
Fact (that everyone knows): Julian is German.
Scott asks: "*Does John think that Julian dances salsa?*"
Does Scott believe that Julian daces salsa?
Results

- Main effects of prior and predicates
  - Anti-veridical effect of "think"
- Interaction between predicate and prior, prior and embedded clause type, predicate and embedded clause type



Computational model: Mixture RSA

- Partially couched in the Rational Speech Act (RSA) framework [9,10].
- meaning space: inferences modeled as how likely the speaker is taken to believe in the embedded content, $b_{SP} \in [0,1]$
- utterance set: $u \in$ U = {"know p", "know not p", "think p", "think not p", "BARE"}

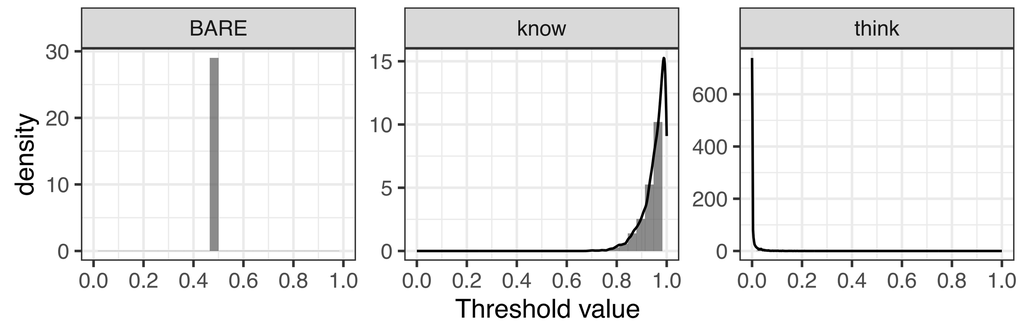
Model setup

- **Literal listener** reasons about the meaning of the utterance represented by a threshold semantics [11-14].
  - "know" and "think": the utterance is felicitous if the belief exceeds the threshold, each sampled from a beta distribution. $\theta_{know} \sim Beta(20,1)$, $\theta_{think} \sim Beta(0.1,20)$
  - "BARE": more likely to be used if it is close to $\theta_{BARE} = 0.5$

$$P_{L_0}(b_{SP}|u) \propto \begin{cases} 1 & \text{if } b_{SP} > \theta_u \text{ for } p \\ & \text{or } b_{SP} < 1-\theta_u \text{ for } \neg p \\ 0 & \text{otherwise} \end{cases}$$

$$P_{L_0}(b_{SP}|u) \propto P(1-|b_{SP}-\theta_u|^x)$$

$$\text{where } x = \begin{cases} 0.1 & \text{if } |b_{SP}-\theta_u| > 0.1 \\ 2 & \text{otherwise} \end{cases}$$



- **Pragmatic speaker** soft-maximizes the utility of the utterance, balancing the informativeness and the costs.
- **Pragmatic listener** samples from either the prior belief distribution or the speaker production distribution, weighing by how likely the embedded content is at-issue, given the predicate $P(q_{CC}|u)$.

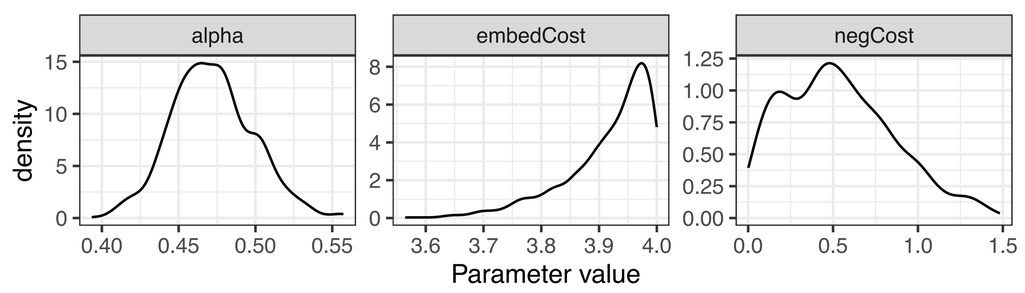
$$P_{S_1}(u|b_{SP}) \propto \exp(\alpha \cdot \underbrace{(\ln P_{L_0}(b_{SP}|u) - (C_{\text{Neg}}(u) + C_{\text{Embed}}(u)))}_{\text{utility of the utterance}})$$

$$P_{L_1}(b_{SP}|u) \propto \underbrace{P_{S_1}(u|b_{SP})}_{\text{speaker model}} \cdot P(q_{MC}|u) + \underbrace{P(b_{SP})}_{\text{prior belief}} \cdot P(q_{CC}|u)$$

$$\text{where } P(q_{MC}|u) + P(q_{CC}|u) = 1$$

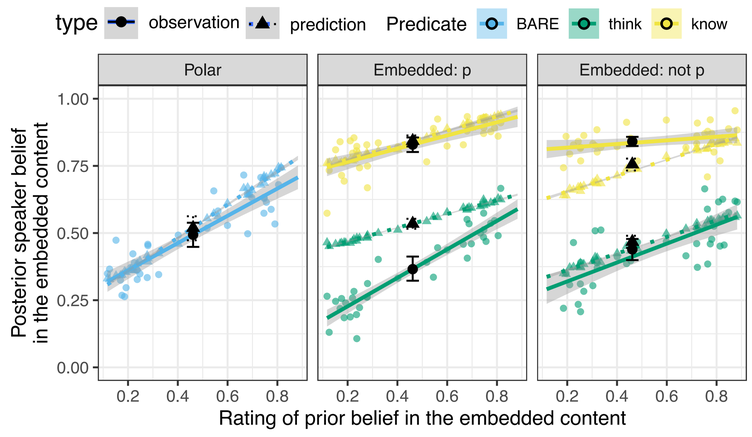
Paremeter estimation

- Bayesian data analysis (BDA) was conducted to estimate the values of the optimality parameter $\alpha \sim U(0,10)$, and the two cost terms $C_{Neg} \sim U(0,4)$ and $C_{Embed} \sim U(0,4)$
- The belief ratings collected in the behavioral experiment and the prior norms from previous experiments [2,7] were used to inform the model.



Model evaluation

- qualitatively captured: the main effects of predicate and prior, the interaction between predicate and prior for "know p" and "think not p."
- failed to capture: the anti-veridical of "think", the effect of prior for "know not p" and "think p."



Discussion

- This is a step towards a systematic analysis of projection inferences in polar interrogatives with clause-embedding predicates using probabilistic pragmatic models.
- The proposed mixture RSA model combines the speaker production distribution and the prior belief, and can account for some empirical patterns.
- Alternative ways to model at-issueness and prior beliefs are still to be explored.
- Speaker might consider the attitude holder's belief (i.e., how likely that John believes…) when choosing the utterance, which can be included in the model.

Links

Experiments, data, analysis script, and models: https://github.com/pennydy/Projectivity_RSA.
Preregistration: https://osf.io/gtdw5

[1] Kiparsky & Kiparsky (1970). [2] Degen & Tonhauser (2022). [3] Tonhauser et al. (2018). [4] Beaver et al. [2017]. [5] Roberts (2012). [6] Degen & Tonhauser (2021). [7] Tonhauser & Degen (under review). [8] Mahler (2020). [9] Frank & Goodman (2012). [10] Goodman & Frank (2016). [11] Lorson et al. (2021). [12] Schuster & Degen (2020). [13] Lassiter (2017). [14] Yalcin (2010).