

Higher-probability content is more projective than lower-probability content

Author(s)

March 18, 2020

Abstract

abstract

1 Introduction

In making utterances, speakers present themselves as committed to the truth of some utterance content and as not committed to the truth of other utterance content. For example, if a speaker utters the polar question in (1), they may be committed to the truth of the content of the complement, that Sam has a new hat, but they are likely not committed to the main clause content, that Jane knows of Sam's new hat; rather, this is the content they are asking about.

- (1) Does Jane know that Sam has a new hat?

One cue that listeners rely on in identifying which content speakers present themselves as committed to are the expressions uttered. If, for example, a speaker utters the variant of (1) with *think*, given in (2), then the speaker is typically committed neither to the content of the complement nor to the main clause content.

- (2) Does Jane think that Sam has a new hat?

The inference to the truth of the content of the complement has long been taken to categorize clause-embedding predicates: the inference is typically taken to arise with factive predicates, like *know*, but not with non-factive predicates, like *think* (Kiparsky and Kiparsky 1970, Karttunen 1971, i.a.). However, Tonhauser, Beaver, and Degen 2018 recently showed that there is variability among factive predicates in the inference to the truth of the content of the complement. One source of variability was the factive predicate: for instance, the inference was less likely to arise with *reveal* than with *discover*; for both of these predicates it was less likely than with *know*. Another source of variability was the lexical content of the clausal complement, that is, whether content of the complement was 'Sam has a new hat', as in (1), or 'Sam has a new BMW'. The authors hypothesized that the inference to the truth of the content of the complement may depend on its prior subjective probability, such that the inference to the truth of the content of the complement is more likely to arise the higher its prior probability.¹ This paper provides support for this hypothesis in an experimental investigation of the contents of the complements of 20 English clause-embedding predicates for which the prior probability was systematically manipulated.

¹JT: discuss edinburgh MA thesis results

We argue that this finding provides support for constraint-based projection analyses that derive the projectivity of utterance content from the integration of multiple cues, including prior CC probability and the meanings of clause-embedding predicates.

2

The 20 predicates we investigated are listed in (3) with the categories they are typically taken to fall into: factive predicates in (3a), non-factive predicates in (3b) and optionally factive predicates in (3c). Specifically, the 5 factive predicates in (3a) include the cognitive predicate *know*, the change-of-states predicates *discover* and *reveal*, the sensory predicate *see*, and the emotive predicate *be annoyed*. The 6 non-factive predicates in (3b) include 4 ‘non-veridical non-factive’ predicates *pretend*, *suggest*, *say* and *think*, whose CC is typically taken to be neither presupposed nor entailed, as well as 2 ‘veridical non-factive’ predicates *be right* and *demonstrate*, whose CC is typically taken to be entailed but not presupposed.³ The 9 optionally factive predicates in (3c) include *acknowledge*, *admit* and *announce*, which Kiparsky and Kiparsky 1970 listed as optionally factive, as well as *confirm*, *inform*, *confess*, *establish*, *hear* and *prove*.⁴

(3) 20 clause-embedding predicates

- a. factive: *be annoyed*, *discover*, *know*, *reveal*, *see*
- b. non-factive: *pretend*, *suggest*, *say*, *think*, *be right*, *demonstrate*
- c. optionally factive: *acknowledge*, *admit*, *announce*, *confess*, *confirm*, *establish*, *hear*, *inform*, *prove*

2 Methods

Participants 300 participants with U.S. IP addresses and at least 99% of previous HITs approved were recruited on Amazon’s Mechanical Turk platform (ages: 21-72, median: 36; 145 female, 154 male, 1 undeclared). They were paid \$0.85 for participating.

Materials 20 clauses realized the complements of the 20 clause-embedding predicates, for a total of 400 predicate/clause combinations. These 400 predicate/clause combinations were realized as polar questions, as shown in the sample target stimuli in (4), in which Carol’s polar questions realize the clause-embedding predicate *know* and the complement clause *Julian dances salsa*. Each complement clause was paired with two facts, for a total of 800 predicate/clause/fact combinations: the content of the clausal complement has a higher prior probability with one fact and a lower prior probability with the other. For instance, in (4), the content that Julian dances salsa has a higher prior probability when Julian is Cuban, as in (4a), than when Julian is German, as in (4b). See Appendix A for the full set of target stimuli and Appendix B for the norming study that established the mean prior probabilities of the contents of the 20 clauses given their two facts.

²The experiment, data and R code for generating the figures and analyses of the experiment reported on in this paper are available at [redacted for review].

³Anand and Hacquard (2014) assumed that *demonstrate* is veridical, in contrast to Anand et al. ms. This latter work also takes *reveal* to be non-factive, in contrast to, for instance, Egré 2008, Wyse 2010 or Tonhauser et al. 2018.

⁴Different categories have been assumed for these 6 predicates. For *confirm* and *inform*, Anand and Hacquard (2014) took them to be optionally factive, but recall from above that Schlenker (2010) took *inform* to be factive. For *prove*, White and Rawlins (2018) suggested that it is optionally factive, but Anand and Hacquard (2014) took it to be non-veridical and non-factive. For *confess*, Swanson (2012) took it to be factive, Karttunen (2016) only took it to commit the speaker to the subject of the attitude being committed to the CC, and Wyse (2010) listed it under the non-factive predicates. For *establish*, Swanson (2012) took it to be non-factive, but Wyse (2010) listed it under the factive predicates. Finally, we also included *hear* in this class: even though it is often considered a factive sensory predicate (e.g., Beaver 2010, Anand and Hacquard 2014), it has been observed that *hear* also has a non-factive reportative evidential sense (see, e.g., Anderson 1986, Simons 2007), especially when it is combined with complements that describe events that cannot be auditorily perceived, as is the case in our experiments.

- (4) a. **Fact (which Carol knows):** Julian is Cuban.
Carol: Does Sandra know that Julian dances salsa?
 b. **Fact (which Carol knows):** Julian is German.
Carol: Does Sandra know that Julian dances salsa?

To assess whether participants were attending to the task, the experiment included 6 main clause control stimuli, whose content does not project: in the sample control stimulus in (7), for example, the speaker is not taken to be committed to the main clause content, that Zack is coming to the meeting tomorrow. The control stimuli were presented with a fact that we hypothesized did not influence the prior probability of the main clause content. For the full set of control stimuli see Appendix A.

- (5) **Fact (which Margaret knows):** Zack is a member of the golf club.
Margaret: Is Zack coming to the meeting tomorrow?

Each participant saw a random set of 26 stimuli: each set contained one target stimulus for each of the 20 clause-embedding predicates, each with a unique complement clause, and the 6 control stimuli. Trial order was randomized.

Procedure Participants were told to imagine that they are at a party and that, on walking into the kitchen, they overhear somebody ask somebody else a question. Participants were asked to rate whether the speaker was certain of the content of the complement, taking into consideration the fact that was presented. They gave their responses on a slider marked ‘no’ at one end (coded as 0) and ‘yes’ at the other (coded as 1), as shown in Figure 1.

Fact (which Dorothy knows): Jon lives 10 miles away from work.

Dorothy asks: *"Did Ruth say that Jon walks to work?"*

Is Dorothy certain that Jon walks to work?

noyes

Figure 1: A sample trial in the experiment

After completing the experiment, participants filled out a short, optional survey about their age, their gender, their native language(s) and, if English is their native language, whether they are a speaker of American English (as opposed to, e.g., Australian or Indian English). To encourage them to respond truthfully, participants were told that they would be paid no matter what answers they gave in the survey.

Data exclusion Prior to analysis, the data from 23 participants who did not self-identify as native speakers of American English were excluded. To assess whether the remaining 277 participants attended to the task, we inspected their responses to the 6 control stimuli, for which we expected low responses. There were 11

participants whose response means on the controls were more than 2 standard deviations above the group mean. We excluded the data from these participants, too, leaving data from 266 participants (ages 21-72; median: 36; 129 female, 136 male, 1 undeclared).

3 Results

Figure 2 shows the mean certainty ratings for the target stimuli by predicate and by the prior probability of the content of the complement; the mean certainty rating for the non-projective main clause controls are also included. As shown, mean certainty ratings were higher when the content had a higher prior probability than when it had a lower prior probability, as hypothesized by Tonhauser et al. (2018). This finding suggests that prior content probability influences projection. Furthermore, the effect of prior content probability on projectivity was present across all 20 clause-embedding predicates.

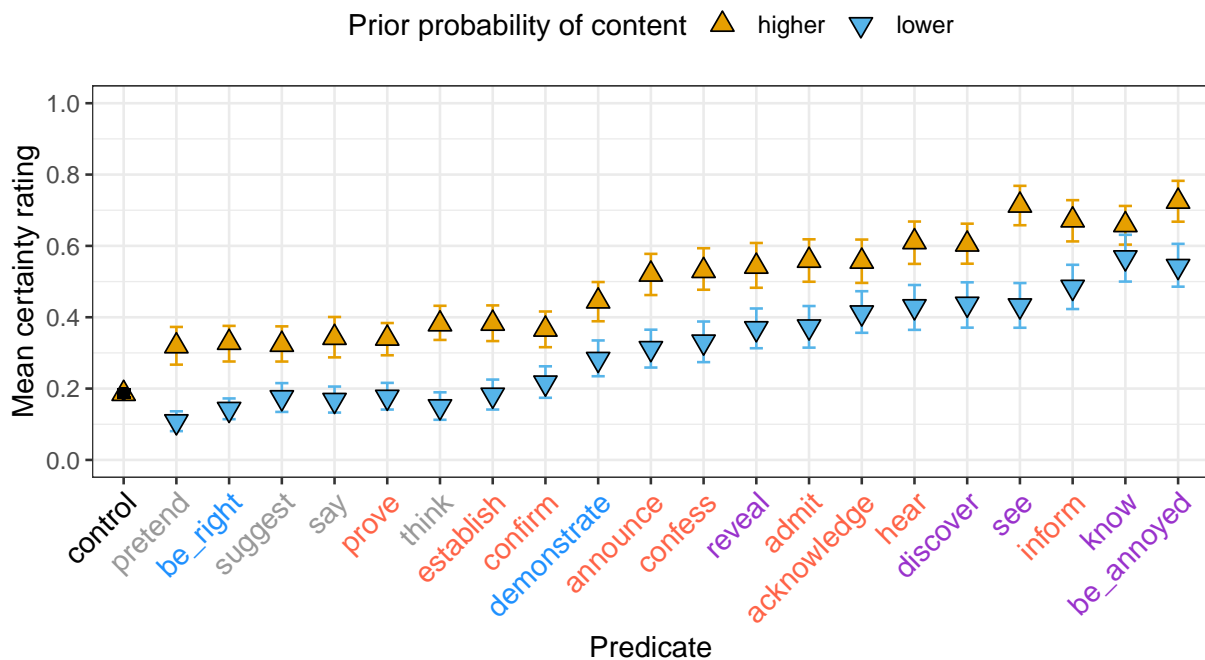


Figure 2: Mean certainty ratings by predicate and prior probability of the content of the complement. Error bars indicate 95% bootstrapped confidence intervals.

JT: need bayesian model here

4 Discussion

as well as for the main clause stimuli (abbreviated ‘MC’), in increasing order from left to right. The mean certainty ratings are largely consistent with impressionistic judgments reported in the literature. First, the ratings for main clause content are lowest overall, as expected for non-projective content. Second, the ratings for factive predicates are among the highest overall, suggesting comparatively high projectivity of the CC. Third, the mean certainty ratings of many optionally factive predicates are lower than those of many factive predicates and higher than those of main clauses as well as of non-veridical non-factives. However, Figure ?? also shows that the CCs of the 5 predicates assumed to be factive are not categorically more projective

than the CCs of the optionally factive predicates, contrary to what is expected under the first definition of factive predicates. Specifically, the CCs of the optionally factive predicates *acknowledge*, *hear* and *inform* are at least as projective as the CCs of *reveal* or *discover*. This finding suggests that projectivity alone does not categorically distinguish factive predicates from optionally factive and non-factive ones.

These results support Tonhauser et al.’s (2018) hypothesis that prior content probability influences projectivity. The finding that the CC of many non-factive predicates is at least weakly projective, even with low prior probability CCs, confirms intuitions reported in, e.g., Schlenker 2010, Anand and Hacquard 2014 and Spector and Egré 2015. These findings motivate the development of projection analyses that derive the influence of prior content probability and make predictions for the CCs of a broad range of both factive and non-factive predicates.

Current projection analyses, while limited to the CCs of factive predicates (e.g., Heim 1983, van der Sandt 1992, Abrusán 2011, Simons et al. 2010, Beaver et al. 2017), are compatible with the finding that prior content probability influences projectivity. Heim 1983, for instance, assumes default global accommodation when a presupposition is not entailed by the common ground (CG) when the trigger is uttered. This default is overridden when the presupposition is inconsistent with the CG. If we can assume that Julian dancing salsa is more likely to be consistent with the CG when Julian is from Cuba than when he is from Germany, Heim 1983 predicts that the presupposition that Julian dances salsa is more projective when it has a higher prior probability.

As shown in Fig. ??, the CCs of several non-factives, including *inform*, *hear*, *acknowledge* and *admit*, are at least as projective as that of factive *reveal*. This finding challenges the long-standing assumption that the CCs of factives are more projective than those of non-factives. We suggest that this motivates constraint-based analyses that derive the projectivity of utterance content from the integration of multiple cues, including prior CC probability, the meanings of clause-embedding predicates, at-issueness (Tonhauser et al. 2018), and information structure (Tonhauser 2016).

discussion: measure prior and projection at participant-level (as we do in next work)

5 Conclusions

A Target and control stimuli

The target stimuli of the experiment consisted of 800 combinations of a predicate, a complement clause and a fact relative to which the content of the complement clause had a higher or lower probability. The following list gives the 20 complement clauses together with the two facts that the clause was paired with: first the lower probability fact, then the higher probability one. The target stimuli of the norming study consisted of 40 combinations of a complement clause and a fact.

1. Mary is pregnant (Mary is a middle school student / Mary is taking a prenatal yoga class)
2. Josie went on vacation to France (Josie doesn’t have a passport / Josie loves France)
3. Emma studied on Saturday morning (Emma is in first grade / Emma is in law school)
4. Olivia sleeps until noon (Olivia has two small children / Olivia works the third shift)
5. Sophia got a tattoo (Sophia is a high end fashion model / Sophia is a hipster)
6. Mia drank 2 cocktails last night (Mia is a nun / Mia is a college student)
7. Isabella ate a steak on Sunday (Isabella is a vegetarian / Isabella is from Argentina)
8. Emily bought a car yesterday (Emily never has any money / Emily has been saving for a year)
9. Grace visited her sister (Grace hates her sister / Grace loves her sister)
10. Zoe calculated the tip (Zoe is 5 years old / Zoe is a math major)

11. Danny ate the last cupcake (Danny is a diabetic / Danny loves cake)
12. Frank got a cat (Frank is allergic to cats / Frank has always wanted a pet)
13. Jackson ran 10 miles (Jackson is obese / Jackson is training for a marathon)
14. Jayden rented a car (Jayden doesn't have a driver's license / Jayden's car is in the shop)
15. Tony had a drink last night (Tony has been sober for 20 years / Tony really likes to party with his friends)
16. Josh learned to ride a bike yesterday (Josh is a 75-year old man / Josh is a 5-year old boy)
17. Owen shoveled snow last winter (Owen lives in New Orleans / Owen lives in Chicago)
18. Julian dances salsa (Julian is German / Julian is Cuban)
19. Jon walks to work (Jon lives 10 miles away from work / Jon lives 2 blocks away from work)
20. Charley speaks Spanish (Charley lives in Korea / Charley lives in Mexico)

In the target stimuli of the experiment, eventive predicates, like *discover* and *hear*, were realized in the past tense and stative predicates, like *know* and *be annoyed*, were realized in the present tense. The direct object of *inform* was realized by the proper name *Sam*. The subject of the clause-embedding predicate and the speaker of the target stimuli were realized by a proper name.

Six control stimuli

1. **Fact:** Zack is a member of the golf club. Is Zack coming to the meeting tomorrow?
2. **Fact:** Mary visited her aunt on Sunday. Is Mary's aunt sick?
3. **Fact:** Todd goes to the gym 3 times a week. Did Todd play football in high school?
4. **Fact:** Vanessa won a prize at school. Is Vanessa good at math?
5. **Fact:** Trish sent Madison a card. Did Madison have a baby?
6. **Fact:** Hendrick just bought a car. Was Hendrick's car expensive?

B Norming study

Prior probability was measured for the content of the 20 clauses given two facts.

Participants 95 participants with U.S. IP addresses and at least 99% of previous HITs approved were recruited on Amazon's Mechanical Turk platform (ages: 21-75, median: 33; 45 female, 50 male). They were paid \$0.55 for participating.

Materials The 20 clauses were realized as polar interrogative sentences and paired with one of the two facts, for a total of 40 target stimuli. See Appendix A for the full set of target stimuli. As shown in the sample target stimuli in (6), the facts were identified by '**Fact:**'. We hypothesized that the prior probability of the content of the clause would be relatively high (though not at ceiling) with one fact and relatively low (though not at floor) with the other: for instance, in (6a), we hypothesized that the prior probability of Julian dancing salsa is higher if Julian is Cuban than if Julian is German.

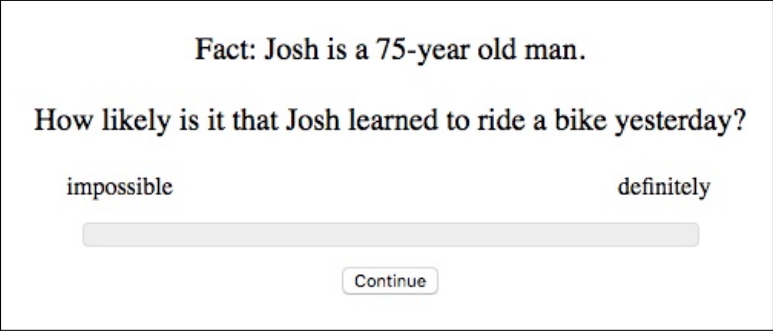
- (6) a. **Fact:** Julian is Cuban.
How likely is it that Julian dances salsa?
- b. **Fact:** Julian is German.
How likely is it that Julian dances salsa?

The experiment included the two control stimuli in (7), which were used to assess whether participants were attending to the task. We expected participants to assess the prior probability of the content in (7a) as high and that of the content in (7b) as low.

- (7) a. **Fact:** Barry lives in Germany.
How likely is it that Barry lives in Europe?
- b. **Fact:** Tammy is a rabbit.
How likely is it that Tammy speaks Italian and Greek?

Each participant saw a random set of 22 stimuli: each set contained one target stimulus for each of the 20 clauses and the 2 control stimuli. Trial order was randomized.

Procedure Participants were told to read facts and to assess the likelihood of events, given those facts. They gave their responses on a slider marked ‘impossible’ at one end (coded as 0) and ‘definitely’ at the other (coded as 1), as shown in the sample trial in Figure A1



Fact: Josh is a 75-year old man.

How likely is it that Josh learned to ride a bike yesterday?

impossible definitely

Continue

Figure A1: A sample trial in the norming study

After completing the experiment, participants filled out a short, optional survey about their age, their gender, their native language(s) and, if English is their native language, whether they are a speaker of American English (as opposed to, e.g., Australian or Indian English). To encourage them to respond truthfully, participants were told that they would be paid no matter what answers they gave in the survey.

Data exclusion Prior to analysis, the data from 8 participants who did not self-identify as native speakers of American English were excluded. To assess whether the remaining 87 participants attended to the task, we inspected their responses to the 2 control stimuli. We excluded the data from 12 participants whose response means was more than 2 standard deviations below the group mean on the high prior probability control in (7a) or more than 2 standard deviations above the group mean on the low prior probability control in (7b). We excluded the data from these participants, too, leaving data from 75 participants (ages 21-75; median: 35; 34 female, 41 male).

Results Figure A2 plots the mean likelihood ratings for the 20 contents by fact: for each content, the mean likelihood rating is higher with the higher probability fact than with the lower probability fact. Thus, the prior probability of each content is successfully manipulated by the two facts.

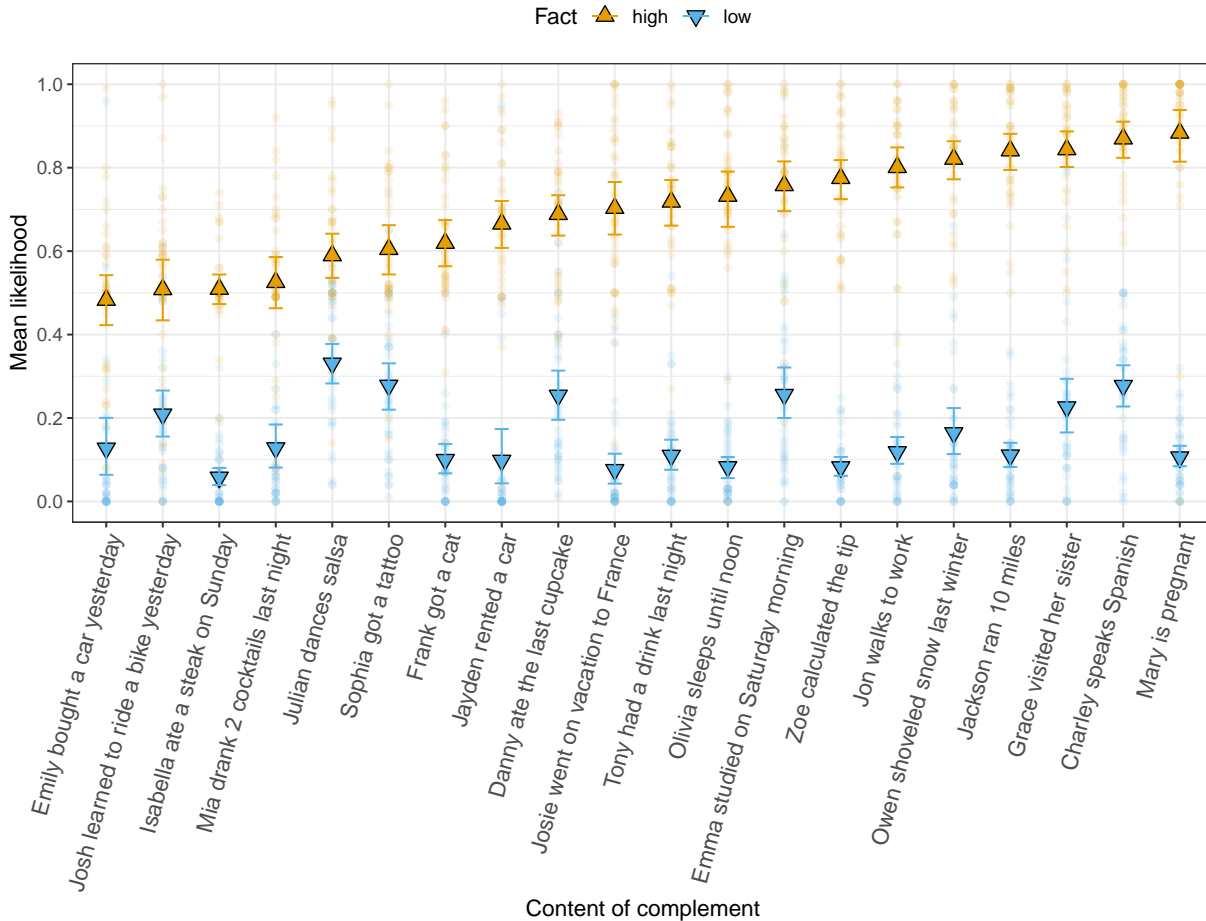


Figure A2: Mean likelihood ratings by content and fact in norming study. Error bars indicate 95% bootstrapped confidence intervals. Light dots indicate participants' responses.

References

- Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics & Philosophy* 34:491–535.
- Anand, Pranav, Jane Grimshaw, and Valentine Hacquard. ms. Sentence embedding predicates, factivity and subjects. In C. Condoravdi, ed., *A Festschrift for Lauri Karttunen*. Stanford: CSLI Publications.
- Anand, Pranav and Valentine Hacquard. 2014. Factivity, belief and discourse. In *The Art and Craft of Semantics: A Festschrift for Irene Heim*, pages 69–90. MIT Working Papers in Linguistics.
- Anderson, Lloyd B. 1986. Evidentials, paths of change, and mental maps: Typologically regular asymmetries. In W. Chafe and J. Nichols, eds., *Evidentiality: The Linguistic Coding of Epistemology*, pages 273–312. New Jersey: Ablex Publishing Corporation.
- Beaver, David. 2010. Have you noticed that your belly button lint colour is related to the colour of your clothing? In R. Bäuerle, U. Reyle, and E. Zimmermann, eds., *Presuppositions and Discourse: Essays Offered to Hans Kamp*, pages 65–99. Oxford: Elsevier.
- Beaver, David, Craig Roberts, Mandy Simons, and Judith Tonhauser. 2017. Questions Under Discussion: Where information structure meets projective content. *Annual Review of Linguistics* 3:265–284.
- Egré, Paul. 2008. Question-embedding and factivity. In L. Lihoreau, ed., *Grazer Philosophische Studien*, pages 85–125. Amsterdam: Rodopi.

- Heim, Irene. 1983. On the projection problem for presuppositions. In M. Barlow, D. Flickinger, and M. Westcoat, eds., *West Coast Conference on Formal Linguistics (WCCFL) 2*, pages 114–125.
- Karttunen, Lauri. 1971. Some observations on factivity. *Papers in Linguistics* 4:55–69.
- Karttunen, Lauri. 2016. Presupposition: What went wrong? In *Proceedings of Semantics and Linguistic Theory (SALT) XXVI*, pages 705–731. Ithaca, NY: CLC Publications.
- Kiparsky, Paul and Carol Kiparsky. 1970. Fact. In M. Bierwisch and K. Heidolph, eds., *Progress in Linguistics*, pages 143–173. The Hague: Mouton.
- Schlenker, Philippe. 2010. Local contexts and local meanings. *Philosophical Studies* 151:115–142.
- Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 117:1034–1056.
- Simons, Mandy, Judith Tonhauser, David Beaver, and Craige Roberts. 2010. What projects and why. In *Semantics and Linguistic Theory (SALT) XXI*, pages 309–327. Ithaca, NY: CLC Publications.
- Spector, Benjamin and Paul Egré. 2015. A uniform semantics for embedding interrogatives: *An* answer, not necessarily *the* answer. *Synthese* 192:1729–1784.
- Swanson, Eric. 2012. Propositional attitudes. In K. von Heusinger, C. Maienborn, and P. Portner, eds., *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, pages 1538–1561. Berlin: Mouton de Gruyter.
- Tonhauser, Judith. 2016. Prosodic cues to speaker commitment. In *Semantics and Linguistic Theory (SALT) XXVI*, pages 934–960. Ithaca, NY: CLC Publications.
- Tonhauser, Judith, David Beaver, and Judith Degen. 2018. How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics* 35:495–542.
- van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.
- White, Aaron S. and Kyle Rawlins. 2018. The role of veridicality and factivity in clause selection. In *48th Meeting of the North East Linguistic Society*, page ??
- Wyse, Brendan. 2010. *Factive/non-factive predicate recognition within Question Generation systems*. Master's thesis, Open University.