The influence of lexical meaning and prior belief on projectivity[*]

Judith Tonhauser[•] and Judith Degen[▻]

[•]The Ohio State University
[▻]Stanford University

January 30, 2018

Abstract

Projective content, like presuppositions or conventional implicatures, is utterance content that a speaker may be taken to be committed to even when the expression associated with the content occurs embedded under an entailment-canceling operator (e.g., Chierchia and McConnell-Ginet 1990). It has long been observed that non-entailed content may also be taken to be a commitment of the speaker when the expression associated with the content is embedded under such operators (e.g., Simons 2001; Schlenker 2010; Simons et al. 2017). Some authors, like Anand and Hacquard 2014, assume that speaker commitment in such cases is a different phenomenon/should be given a different analysis ("illusion of factivity" and "illusion of projection"), whereas others, like Schlenker 2010 talks of "part-time triggers". Former authors assume that entailed content = presupposition, and non-entailed content = something else. Entailment is a binary property but is it really the way in which language users treat it?

The goal of this paper was to investigate whether projectivity is influenced by entailment and by the prior probability of the event described by the projective content.

We show that...and that....

1 Introduction

- This paper looks at the influence of lexical meaning on the projectivity of the content of the clausal complement of clause-embedding predicates.

  Predicates: lexical meaning is entailment

  Clausal complement: lexical meaning is prior probability of the event described by the clausal complement

- **Motivation to look at beliefs about event probability**

  - Tonhauser, Beaver, and Degen accepted found that lexical content of projective content matters for projection; hypothesized to be about prior probability of the events described by clauses that lexicalize the projective content.

  - Non-conventionalist analyses of projection would presumably generally consider prior event probability in considering whether speaker is taken to be committed to content.

- Question-based analysis of projection: Utterance content that is less likely is more likely to be at-issue than utterance content that is highly likely. Or: utterance content that is highly likely, given facts about the world, are more likely to be backgrounded than utterance content that is less likely, given facts about the world. Assuming that less likely events are more likely to be at-issue, and that at-issue content does not project (Simons et al. 2010; Tonhauser et al. accepted), prior event probability is predicted to influence projectivity.

- Accommodation: more difficult with less likely events than with highly likely events, i.e., highly likely events are more likely to be taken as commitments of the speaker than less likely events (my car / my fire engine didn't start)

- **THUS:** The idea that prior probabilities of events may play a role in projection has been around but not empirically investigated yet.

- **Motivation to look at entailment**

  - Some theoretical analyses of presuppositions assume that presuppositions are entailments of unembedded sentences. GAZDAR

  - Not all theoretical analyses of projective content do so: POTTS CONVENTIONAL IMPLICA-TURES, NON-CONVENTIONAL ANALYSES OF PROJECTION

  - What's the empirical purview of analyses of projective content? Entailed content that is projective, or any content that is projective?

  - However, also, can we distinguish entailing from non-entailing predicates?
    Schlenker 2010 takes *inform* to be entailing, in contrast to *announce*, but Anand and Hacquard 2014 take both to be apparently-entailing

  - Schlenker 2010 has this funny thing about *announce* only sometimes entailing the truth of the content of its complement. Entailment is context-sensitive??

    (1) a. Context: Mary is a responsible 30-year old.
             Mary has announced to her parents that she is pregnant.
             it follows: Mary is pregnant   (Schlenker 2010:139)

         b. Context: Mary is playful 7-year old.
             Mary has announced to her parents that she is pregnant.
             it does not follow: Mary is pregnant   (Schlenker 2010:140)

         c. John has announced that he has met Elvis
             it does not follow: John has met Elvis.   (Schlenker 2010:140)

  - 7 of these predicates are typically taken to entail the complement (*be annoyed, know, discover, reveal, see, establish, be right*), 5 are typically taken to not entail the complement (*pretend, think, suggest, say, hear*), and the remaining 8 are typically taken to not entail the content of the complement even though they may sometimes appear to (*prove, demonstrate, confess, inform, announce, acknowledge, admit, confirm*); see e.g., Schlenker 2010; Swanson 2012; Anand and Hacquard 2014.
    which of these are found with *falsely*?

    * "falsely annoyed that": no hits
    * "falsely annoyed NP" HITS!

- CCMG on presuppositions:

  p.349f.: "the main empirical characteristics of presuppositions can be taken to be the following two: being backgrounded and being taken for granted." THIS DOES NOT INCLUDE BEING ENTAILED!!

  p.350: "What the P family test essentially tests for is backgroundedness of implications: it marks out implications that are attached to S, not only when it is asserted but also when it is denied, questioned, or offered as a hypothetical assumption."

  p.352: "The hallmark of a presupposition is that it is taken for granted in the sense that its assumed truth is a precondition for felicitous utterance of the sentence and places a kind of constraint on discourse contexts that admit the sentence for interpretation. [...] The P family test does not definitively identify presuppositions, because background status of an implied proposition is compatible with its being presented as not already assumed. If there is no suggestion of infelicity in using S in a discourse where p is clearly not taken to be part of the common ground, then S does not presuppose p even if p is backgrounded by S (as in the case of nonrestrictive relative clauses)."

  SPENADER: many presuppositions are new information in discourse

- This paper presents experiments that..[1]

---

[1]The data and R code for generating the figures and analyses of the norming studies and the experiment reported on in this paper are available at `https://github.com/judith-tonhauser/projective-probability`.

2 Entailment

Chierchia and McConnell-Ginet 1990:19f.: "Asked to define entailment, you might come up with any of the following:

(2) A entails B = $_{df}$
- whenever A is true, B is true
- the information that B conveys is contained in the information that A conveys
- a situation describable by A must also be a situation describable by B
- *A and not B* is contradictory (can't be true in any situation)

"Judgments about entailment relations can be defended and supported by evidence. As in the case of any linguistic phenomenon, there may be areas of real diversity within the community of language users, dialectal and even idiolectal differences. This complication must not, however, obscure the important fact that judgments about semantic phenomena are interconnected, and thus that there is relevant evidence to be offered in support of such judgments." (p.22)

2.1 Experiment 1a: Is *A and not B* contradictory?

This experiment explored whether the content of the clausal complement of 20 clause-embedding English predicates is entailed. Specifically, this experiment implemented the diagnostic for entailment from Chierchia and McConnell-Ginet 1990 according to which *A* entails *B* if and only if *A and not B* is contradictory. Gradient contradictoriness ratings were collected.

2.1.1 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: 18-72, median: 35; 137 female, 162 male, 1 other). They were paid 75 cents for participating in the experiment.

Materials Whether the content of the clausal complement was entailed was tested for the following 20 clause-embedding predicates: *be annoyed, know, discover, reveal, see, establish, be right, pretend, think, suggest, say, hear, prove, demonstrate, confess, inform, announce, acknowledge, admit* and *confirm*. As discussed in the introduction, 7 of these predicates are typically taken to entail the complement (*be annoyed, know, discover, reveal, see, establish, be right*), 5 are typically taken to not entail the complement (*pretend, think, suggest, say, hear*), and the remaining 8 are typically taken to not entail the content of the complement even though they may sometimes appear to (*prove, demonstrate, confess, inform, announce, acknowledge, admit, confirm*).

Each predicate was paired with each of the 20 event-describing sentences from section **??** and a proper name subject, for a total of 400 sentences.[2] The 400 target stimuli consisted of one of these 400 sentences followed by an elliptical clause that denied the truth of the complement clause, as shown in the sample stimuli in (3). The speaker of the target stimuli was realized by a (bold-faced) proper name. The proper names that realized the speakers, the subjects of the 20 predicates or that occurred in the complement clauses were all unique. The gender of the proper name subject of the predicate was distinct from the gender of the proper name in the complement clause, to ensure that the pronoun in elliptical clause unambiguously referred to the individual referred to in the complement clause of the predicate.

---

[2]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 indirect object of *inform* was realized by the proper name *Sam*.

(3)   a. **Christopher:** Melissa knows that Danny ate the last cupcake, but he didn't.

       b. **Susan:** Jerry pretended that Emma studied on Saturday morning, but she didn't.

The experiment also included eight control stimuli, which were used to assess whether participants were attending to the task and to allow participants to use the full response scale. The four control stimuli in (4) were hypothesized to be non-contradictory, and the four control stimuli in (5) were hypothesized to be contradictory.

(4)   a. Zack believes that I'm married, but I'm actually single.

       b. Tara wants me to cook for her and I'm a terrific cook.

       c. Frederick is both smarter and taller than I am.

       d. Vanessa is really good at math, but I'm not.

(5)   a. Dana has never smoked in her life and she stopped smoking recently.

       b. Hendrick's car is completely red and his car is not red.

       c. Madison laughed loudly and she didn't laugh.

       d. Sebastian lives in the USA and has never been to the USA.

Each participant saw a random set of 28 stimuli: each set contained one stimulus for each of the 20 predicates (each with a unique complement clause) and the same 8 control stimuli. Trial order was randomized.

Procedure.   Participants were told that they would read utterances made by a speaker and were asked to assess whether the speaker's utterance is contradictory. On each trial, participants read the speaker's utterance and then gave their response on a slider marked 'definitely no' at one end (coded as 0) and 'definitely yes' at the other (coded as 1), as shown in Figure 1.

Figure 1: A sample trial in Experiment 1

To familiarize participants with the task, they first completed the two familiarization items in (6): participants who rated (6a) in the lower half of the scale or (6b) in the upper half of the scale were given an explanation for why their answer was wrong. Participants could only advance to the 28 stimuli if they gave the correct rating to the two training stimuli.

(6)   a. **Bill:** Drew is aware that Patty lives in Canada, but she doesn't.

       b. **Bob**: Drew thinks that Patty lives in Canada, but she doesn't.

After responding to the 28 stimuli, 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 19 participants who did not self-identify as native speakers of American English were excluded. For the remaining 281 participants, we inspected their responses to the 8 control stimuli: we expected low responses to the non-contradictory stimuli in (4) and high responses to the contradictory stimuli in (5). The response means of 12 participants were more than 3 standard deviations above the group mean for the non-contradictory control stimuli or below the group mean for the contradictory control stimuli (the group means were .08 and .94, respectively).[3] Closer inspection revealed that these participants' responses to the control stimuli were systematically higher or lower, suggesting that these participants did not attend to the task or interpreted the task differently. The data from these 12 participants were also excluded, leaving data from 269 participants (ages 18-72; median: 36; 128 female, 140 male, 1 other).

2.1.2   Results

The boxplot in Figure 2 shows the contradictoriness ratings by predicate, collapsing over the 20 complement clauses that each predicate was paired with. Predicates whose clausal complement is typically taken to not be an entailment are given in brown, predicates whose clausal complement is typically taken to be an entailment are given in blue, and apparently-entailing predicates are given in black. As expected, stimuli with the 5 non-entailing predicates, namely *pretend, suggest, think, hear* and *say*, received the lowest contradictoriness ratings. Likewise, the median rating and the mean rating of stimuli with *be right* were at and close to ceiling, respectively, in line with the assumption that the content of the clausal complement of this predicate is entailed. However, the median and mean ratings of stimuli with the other entailing predicates were not consistently high and, crucially, many were lower the median and mean ratings of stimuli with apparently-entailing predicates.

To determine which predicates differed from one another in the contradictoriness ratings, we conducted post hoc pairwise comparisons using Tukey's method (allowing for by-participant variability), using the `lsmeans` package (Hothorn et al., 2008) in R (R Core Team, 2016). P-values for each pair of predicates are displayed in Table 1. This analysis confirmed that the 5 non-entailing predicates received significantly lower contradictoriness ratings than the 7 entailing predicates, which is in line with the assumption that non-entailing and entailing predicates differ in the status of the content of their clausal complement. The 5 non-entailing predicates also received significantly lower contradictoriness ratings than the 8 apparently-entailing predicates. This finding reflects the observation, made in the literature, that the content of the clausal complement is more likely be taken to follow from utterances of unembedded sentences with apparently-entailing predicates than from utterances of unembedded sentences with non-entailing predicates.

We also observe, however, that the contradictoriness ratings do not neatly separate entailing from apparently-entailing predicates. For instance, the contradictoriness ratings for stimuli with the entailing predicate *know* were indistinguishable from those with the apparently-entailing predicate *prove*, as were those with the entailing predicate *be annoyed* and the apparently-entailing predicate *confess*. Crucially, contradictoriness ratings for stimuli with entailing predicates were not, as a whole, higher than for stimuli with apparently-entailing predicates. For instance, stimuli with the entailing predicate *reveal* received signifi-

---

[3]The response mean of one of the non-contradictory control stimuli, (4a) *Zack believes that I'm married, but I'm actually single*, was higher, at .17, than the response means of the remaining three non-contradictory control stimuli (which were .05 or .06). We hypothesize that some participants gave higher responses to this stimulus because the speaker can be taken to contradict Zack's belief.
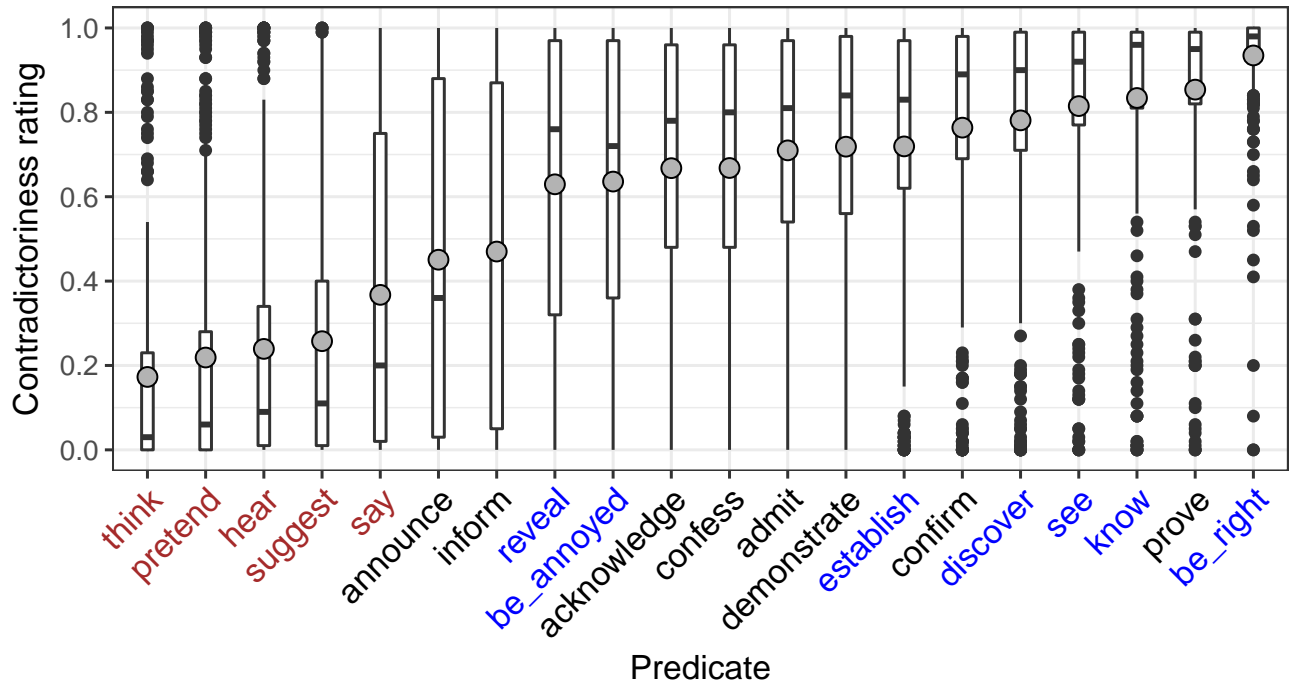
Figure 2: Boxplot of contradictoriness ratings by predicate, collapsing across complement clauses. Grey dots indicate means and notches indicate medians. Non-entailing predicates are given in brown, entailing predicates in blue, and the remaining predicates are given in black.

cantly lower contradictoriness ratings than those with the apparently-entailing predicates *demonstrate* and *confirm*, and stimuli with the entailing predicates *establish* received significantly lower ratings than those with the apparently-entailing predicate *prove*. What's more, the contradictoriness diagnostic is suggestive of significant differences between entailing predicates: the contradictoriness ratings for stimuli with *reveal* and *be annoyed* were significantly lower than those with *establish*, which in turn were significantly lower than those with *discover, see* and *know*, which in turn were significantly lower than those with *be right*. Finally, note that stimuli with *inform* are statistically indistinguishable from those with the apparently-entailing predicate *announce*. As a consequence, the experiment results are inconclusive about whether *inform* is an entailing predicate (as maintained in Schlenker 2010) or an apparently-entailing predicate (as maintained in Anand and Hacquard 2014). However, at the very least, the contradictoriness diagnostic does not provide empirical motivation for treating the two predicates differently based on whether they entail the content of the clausal complement.

Overall, then, the contradictoriness diagnostic for entailment does not provide empirical evidence for a categorical distinction between non-entailing and apparently-entailing predicates, on the one hand, and entailing predicates, on the other hand. Rather, the findings of Exp. 1a suggest that the extent to which the content of the clausal complement follows from an utterance of an unembedded sentence is a gradient property.

	think	pretend	hear	suggest	say	announce	inform	reveal	be annoyed	confess	acknowledge	admit	establish	demonstrate	confirm	discover	see	know	prove
pretend	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
hear	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
suggest	*	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
say	***	***	***	**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
announce	***	***	***	***	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
inform	***	***	***	***	**	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-
reveal	***	***	***	***	***	***	***	-	-	-	-	-	-	-	-	-	-	-	-
be annoyed	***	***	***	***	***	***	***	n.s.	-	-	-	-	-	-	-	-	-	-	-
confess	***	***	***	***	***	***	***	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-
acknowledge	***	***	***	***	***	***	***	n.s.	n.s.	n.s.	-	-	-	-	-	-	-	-	-
admit	***	***	***	***	***	***	***	*	.	n.s.	n.s.	-	-	-	-	-	-	-	-
establish	***	***	***	***	***	***	***	**	*	n.s.	n.s.	n.s.	-	-	-	-	-	-	-
demonstrate	***	***	***	***	***	***	***	**	*	n.s.	n.s.	n.s.	n.s.	-	-	-	-	-	-
confirm	***	***	***	***	***	***	***	***	***	**	**	n.s.	n.s.	n.s.	-	-	-	-	-
discover	***	***	***	***	***	***	***	***	***	**	**	n.s.	n.s.	n.s.	n.s.	-	-	-	-
see	***	***	***	***	***	***	***	***	***	***	***	**	**	**	n.s.	n.s.	-	-	-
know	***	***	***	***	***	***	***	***	***	***	***	***	***	***	n.s.	n.s.	n.s.	-	-
prove	***	***	***	***	***	***	***	***	***	***	***	***	***	***	**	n.s.	n.s.	n.s.	-
be right	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	**	*

Table 1: P-values associated with pairwise comparison of contradictoriness ratings of predicates using Tukey's method. '***' indicates significance at .0001, '**' at .01, '*' at .05, '.' marginal significance at .1, and 'n.s' indicates no significant difference in means. Non-entailing predicates are given in brown, entailing predicates are given in blue, and apparently-entailing predicates are given in black.

In Figure 2, the large box sizes and long whisker lengths for many of the predicates was already suggestive of by-participant variation in the extent to which stimuli with such predicates were assessed to be contradictory. By-participant variation in contradictoriness ratings is also revealed in the top panel of Figure 3, which gives participants' mean contradictoriness ratings over all 20 predicates. The bottom panel of Figure 3 shows that there is also by-participant variability in participants' mean contradictoriness ratings over the 7 entailing predicates.
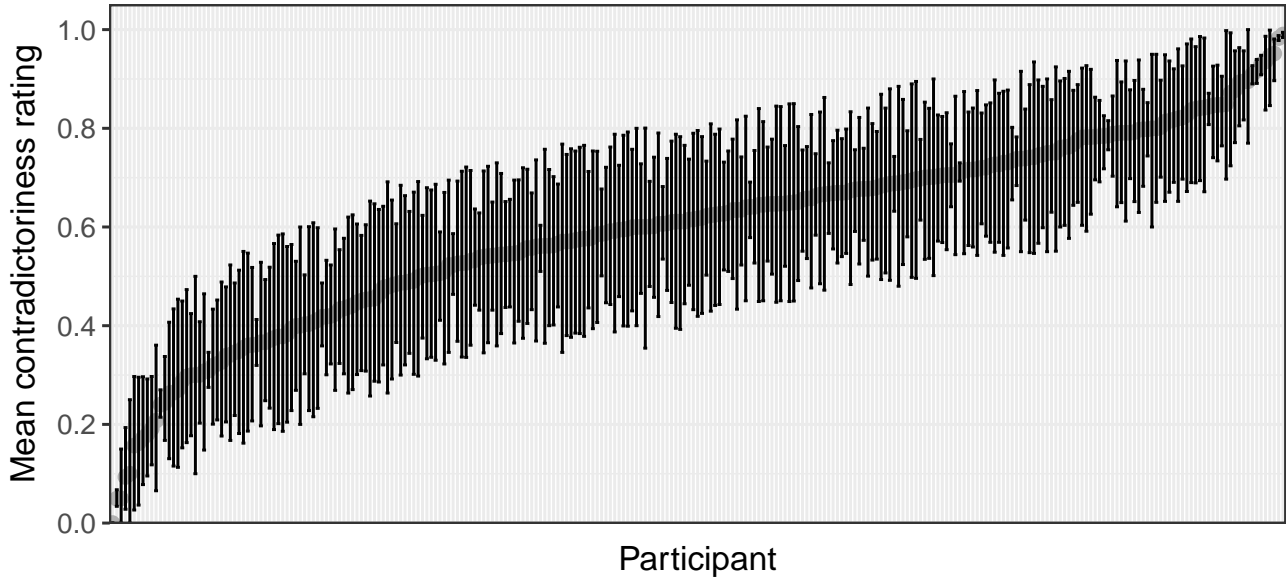
(a) Mean contradictoriness ratings over all predicates.
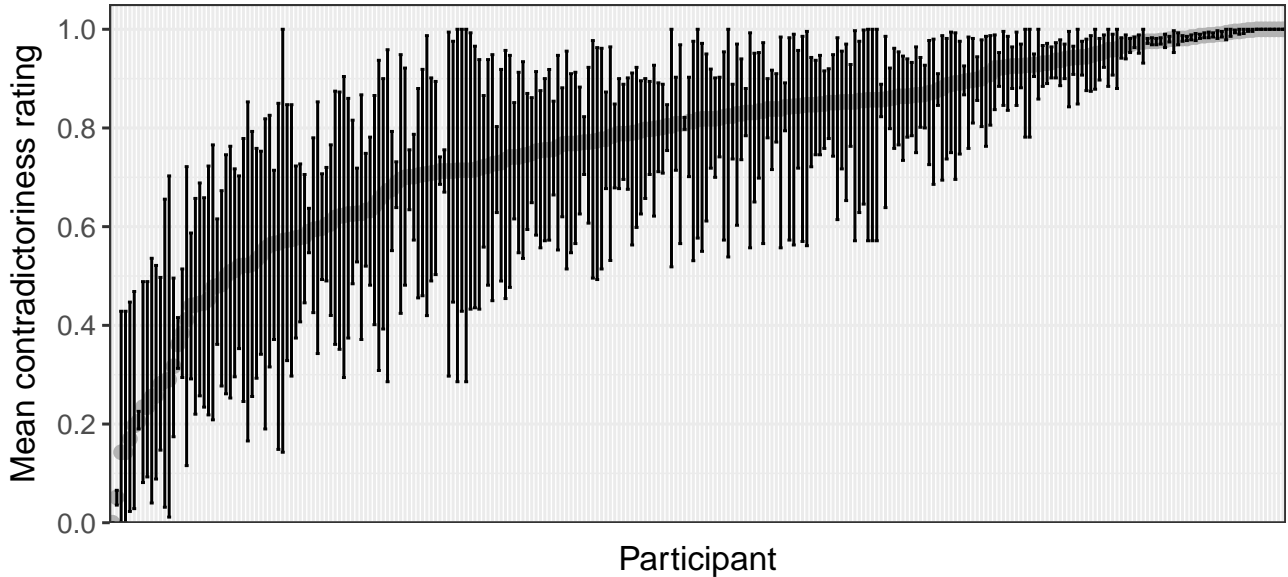
(b) Mean contradictoriness ratings over entailing predicates.

Figure 3: Mean contradictoriness ratings by participant, collapsing across predicates and complement clauses. Grey dots indicate means and error bars indicate bootstrapped 95% confidence intervals.

2.1.3 Discussion

- Contradictoriness diagnostic for entailment distinguishes non-entailing predicates from entailing predicates but does not distinguish apparently-entailing from entailing predicates

- Possible that participants interpreted "contradictory" differently especially for the verbs of saying

2.2 Experiment 1b: How likely is B given A?

This experiment explored whether the content of the clausal complement of 20 clause-embedding English predicates is entailed. Specifically, this experiment implemented the diagnostic for entailment from Chierchia and McConnell-Ginet 1990 according to which A entails B if and only if B follows from A. Gradient inference ratings were collected.

2.2.1 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: XX-XX, median: XX; XX female, XX male, X other). They were paid 75 cents for participating in the experiment.

Materials   Whether the content of the clausal complement was entailed was tested for the same 20 clause-embedding predicates as in Exp. 1. Each predicate was paired with each of the 20 event-describing sentences from section **??** and a proper name subject, for a total of 400 sentences.[4] The 400 target stimuli consisted of one of these 400 sentences, as shown in the sample stimuli in (7). The speaker of the target stimuli was realized by a (bold-faced) proper name. The proper names that realized the speakers, the subjects of the 20 predicates or that occurred in the complement clauses were all unique.

(7) a. Fact: Melissa knows that Danny ate the last cupcake.
     How likely is it that Danny ate the last cupcake?
   b. Fact: Jerry pretended that Emma studied on Saturday morning.
     How likely is it that Emma studied on Saturday morning?

The experiment also included eight control stimuli, which were used to assess whether participants were attending to the task and to allow participants to use the full response scale. The four control stimuli in (4) were hypothesized to be non-contradictory, and the four control stimuli in (5) were hypothesized to be contradictory.

(8) a. Zack believes that I'm married, but I'm actually single.
   b. Tara wants me to cook for her and I'm a terrific cook.
   c. Frederick is both smarter and taller than I am.
   d. Vanessa is really good at math, but I'm not.

(9) a. Dana has never smoked in her life and she stopped smoking recently.
   b. Hendrick's car is completely red and his car is not red.
   c. Madison laughed loudly and she didn't laugh.
   d. Sebastian lives in the USA and has never been to the USA.

---

[4]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 indirect object of *inform* was realized by the proper name *Sam*.

Each participant saw a random set of 28 stimuli: each set contained one stimulus for each of the 20 predicates (each with a unique complement clause) and the same 8 control stimuli. Trial order was randomized.

Procedure   Participants were told that they would read utterances made by a speaker and were asked to assess whether the speaker's utterance is contradictory. On each trial, participants read the speaker's utterance and then gave their response on a slider marked 'definitely no' at one end (coded as 0) and 'definitely yes' at the other (coded as 1), as shown in Figure 7.

Figure 4: A sample trial in Experiment 2

To familiarize participants with the task, they first completed the two familiarization items in (10): participants who rated (10a) in the lower half of the scale or (10b) in the upper half of the scale were given an explanation for why their answer was wrong. Participants could only advance to the 28 stimuli if they gave the correct rating to the two training stimuli.

(10)   a. **Bill:** Drew is aware that Patty lives in Canada, but she doesn't.
       b. **Bob**: Drew thinks that Patty lives in Canada, but she doesn't.

After responding to the 28 stimuli, 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 XX participants who did not self-identify as native speakers of American English were excluded. For the remaining XXX participants, we inspected their responses to the XX control stimuli: we expected low responses to the non-contradictory stimuli in (8) and high responses to the contradictory stimuli in (9). The response means of XX participants were more than 3 standard deviations above the group mean for the non-contradictory control stimuli or below the group mean for the contradictory control stimuli (the group means were .08 and .94, respectively). Closer inspection revealed that these participants' responses to the control stimuli were systematically higher or lower, suggesting that these participants did not attend to the task or interpreted the task differently. The data from these XX participants were also excluded, leaving data from XX participants (ages XX-XX; median: XX; XX female, XX male, XX other).

2.2.2 Results

The boxplot in Figure 5 shows the contradictoriness ratings by predicate, collapsing over the 20 complement clauses that each predicate was paired with. Predicates whose clausal complement is typically taken to not be an entailment are given in brown, predicates whose clausal complement is typically taken to be an entailment are given in blue, and apparently-entailing predicates are given in black.
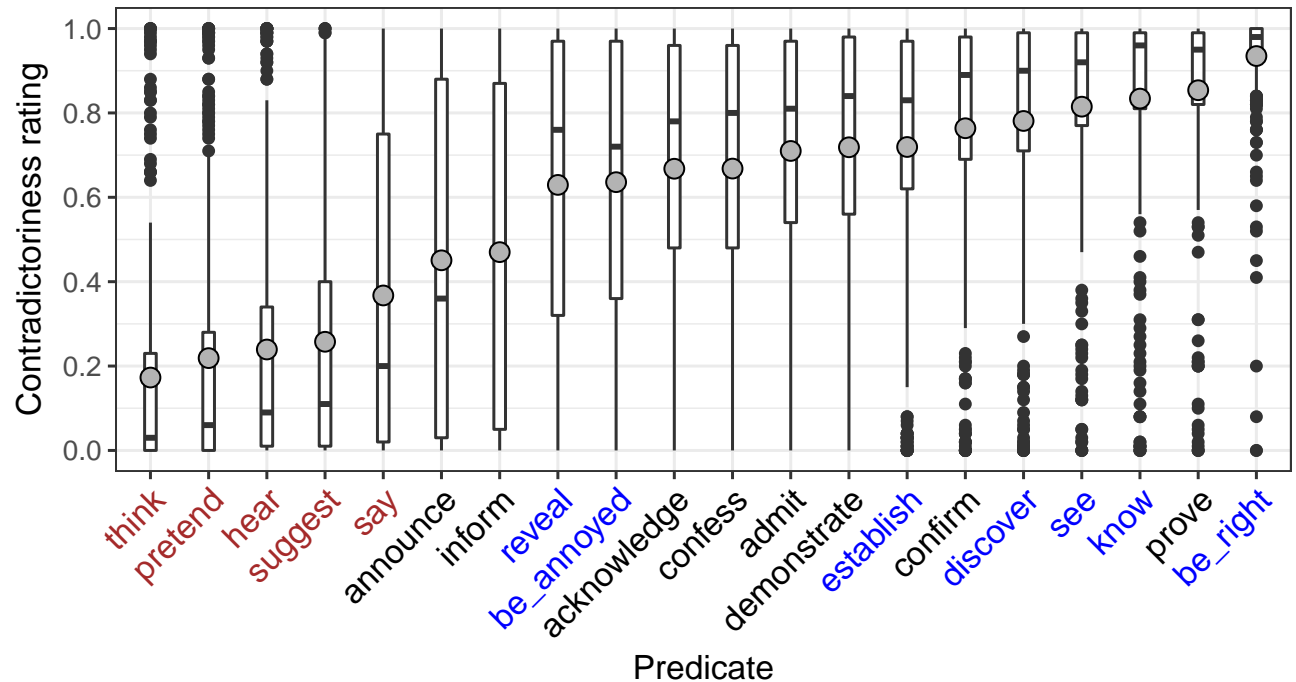

Figure 5: Boxplot of XXX ratings by predicate, collapsing across complement clauses. Grey dots indicate means and notches indicate medians. Non-entailing predicates are given in brown, entailing predicates in blue, and the remaining predicates are given in black.

	think	pretend	hear	suggest	say	announce	inform	reveal	be annoyed	confess	acknowledge	admit	establish	demonstrate	confirm	discover	see	know	prove
pretend	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
hear	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
suggest	*	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
say	***	***	***	**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
announce	***	***	***	***	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
inform	***	***	***	***	**	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-
reveal	***	***	***	***	***	***	***	-	-	-	-	-	-	-	-	-	-	-	-
be annoyed	***	***	***	***	***	***	***	n.s.	-	-	-	-	-	-	-	-	-	-	-
confess	***	***	***	***	***	***	***	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-
acknowledge	***	***	***	***	***	***	***	n.s.	n.s.	n.s.	-	-	-	-	-	-	-	-	-
admit	***	***	***	***	***	***	***	*	.	n.s.	n.s.	-	-	-	-	-	-	-	-
establish	***	***	***	***	***	***	***	**	*	n.s.	n.s.	n.s.	-	-	-	-	-	-	-
demonstrate	***	***	***	***	***	***	***	**	*	n.s.	n.s.	n.s.	n.s.	-	-	-	-	-	-
confirm	***	***	***	***	***	***	***	***	***	**	**	n.s.	n.s.	n.s.	-	-	-	-	-
discover	***	***	***	***	***	***	***	***	***	**	**	n.s.	n.s.	n.s.	n.s.	-	-	-	-
see	***	***	***	***	***	***	***	***	***	***	***	**	**	**	n.s.	n.s.	-	-	-
know	***	***	***	***	***	***	***	***	***	***	***	***	***	***	n.s.	n.s.	n.s.	-	-
prove	***	***	***	***	***	***	***	***	***	***	***	***	***	***	**	n.s.	n.s.	n.s.	-
be right	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	**	*

Table 2: P-values associated with pairwise comparison of XXX ratings of predicates using Tukey's method. '***' indicates significance at .0001, '**' at .01, '*' at .05, '.' marginal significance at .1, and 'n.s' indicates no significant difference in means. Non-entailing predicates are given in brown, entailing predicates are given in blue, and apparently-entailing predicates are given in black.

In Figure 5, the large box sizes and long whisker lengths for many of the predicates was already suggestive of by-participant variation in the extent to which stimuli with such predicates were assessed to be contradictory. By-participant variation in contradictoriness ratings is also revealed in the top panel of Figure 6, which gives participants' mean contradictoriness ratings over all 20 predicates. The bottom panel of Figure 6 shows that there is also by-participant variability in participants' mean contradictoriness ratings over the 7 entailing predicates.
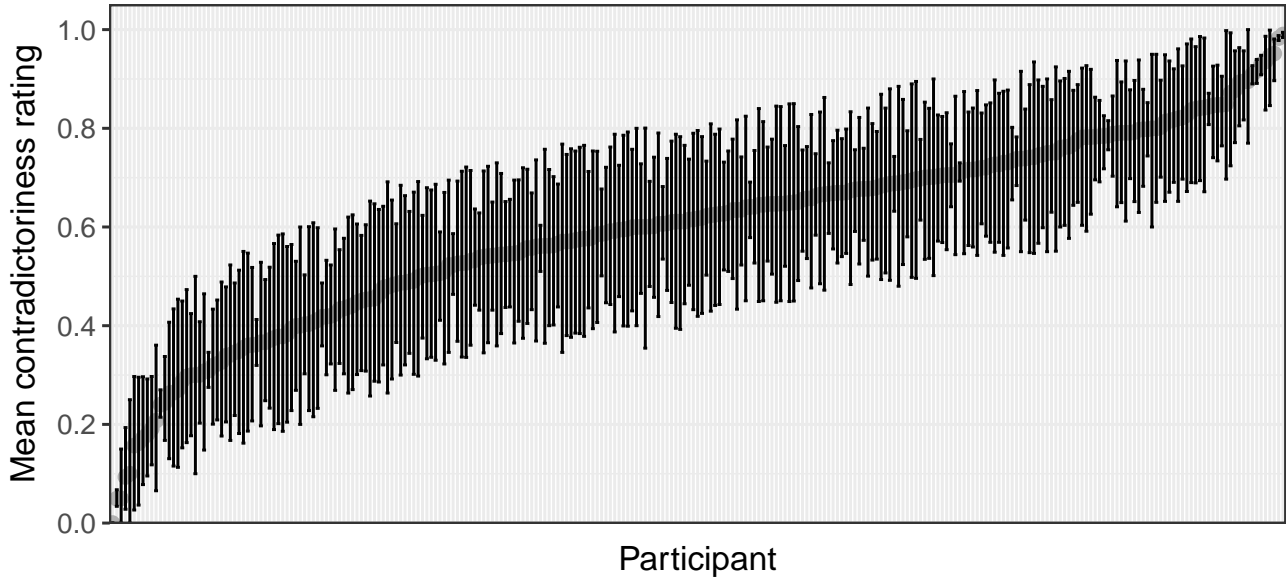
(a) Mean contradictoriness ratings over all predicates.
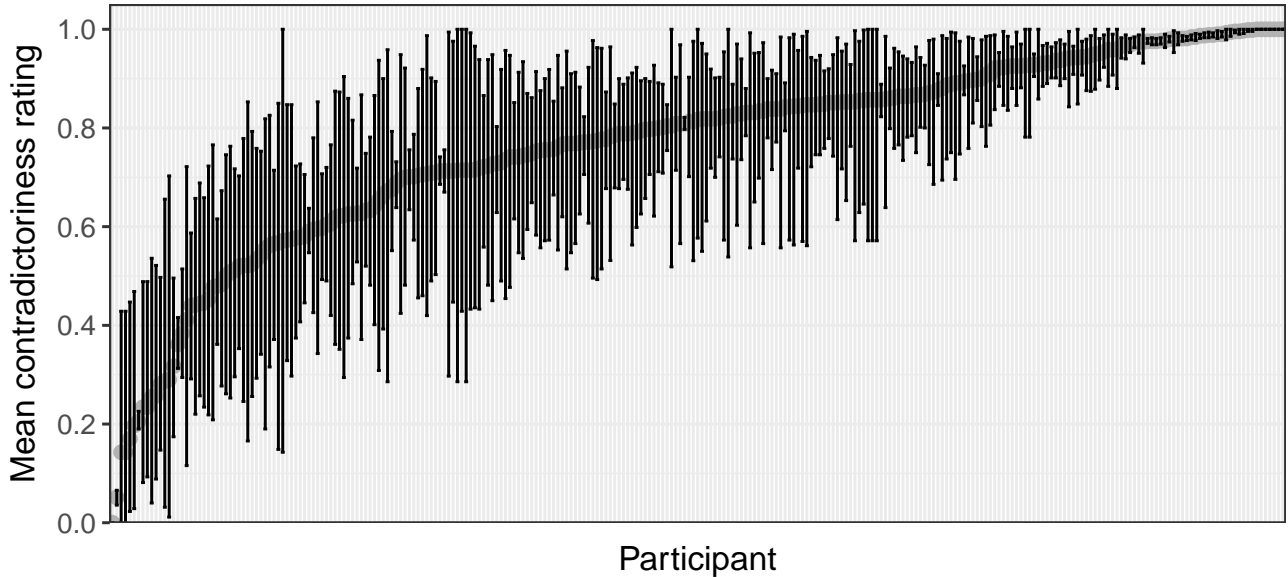
(b) Mean contradictoriness ratings over entailing predicates.

Figure 6: Mean contradictoriness ratings by participant, collapsing across predicates and complement clauses. Grey dots indicate means and error bars indicate bootstrapped 95% confidence intervals.

14

2.2.3 Discussion

-

2.3 Interim summary

3 Experiment 2: Projectivity

This experiment explored the influence of the clause-embedding predicate and of the prior probability of the event described by the embedded clause on the projectivity of the contents of clausal complements of the 20 predicates under investigation. Gradient projectivity ratings were collected.

3.1 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). They were paid 85 cents for participating in the experiment.

Materials   400 polar questions were formed by combining the 20 predicates with a proper name subject and the 20 event-describing clauses from the experiments in section 2. As discussed in the introduction, 6 of the 20 predicates are typically taken to be factive (*be annoyed, know, discover, reveal, see, hear*), 7 are typically taken to be non-factive (*be right, pretend, think, suggest, say, prove, demonstrate*), and the remaining 7 predicates have been suggested to be "part-time triggers" or to give rise to the "illusion of factivity" (*establish, confess, announce, acknowledge, admit, confirm, inform*).[5] For each of the 20 events described by the clausal complements of the predicates, we identified a fact that made the event more likely and a fact that made the event less likely. (The norming study is described in Appendix A.) The 800 target stimuli consisted of a polar question uttered by a speaker who was identified by a proper name and either the fact that made the event described by the complement clause more likely or the fact that made the event less likely. In the two sample stimuli in (11), the predicate *discover* is paired with a clausal complement that describes the event of Julian dancing salsa. This event has a higher prior probability given the fact in (11a) than given the fact in (11b). As shown, the speaker of the target stimuli was realized by a (bold-faced) proper name. The proper names that realized the speakers, the subjects of the 20 predicates or that occurred in the complement clauses were all unique.

(11) a. **Fact (which Carol knows):** Julian is German.
         Carol: Did Sandra discover that Julian dances salsa?

      b. **Fact (which Paul knows):** Julian is Cuban.
         Paul: Did Sandra discover that Julian dances salsa?

The experiment also included six control stimuli, which consisted of a speaker's utterance of a polar question and a fact about the world that was hypothesized to not affect the prior probability of the event described by the polar question. A sample control stimulus is shown in (12); the remaining stimuli can be found in Appendix B. The six control stimuli were included to assess whether participants were attending to the task and to allow participants to use the full response scale.

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

---

[5]As in the experiments in section 2, eventive predicates were realized in the past tense and stative predicates were realized in the present tense. The indirect object of *inform* was realized by the proper name *Sam*.

Each participant saw a random set of 28 stimuli: each set contained one stimulus for each of the 20 predicates (each with a unique complement clause) and the same 8 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. They were also told to consider the fact that was presented with the speaker's question. Participants were asked to identify whether the speaker was certain of the content of the clausal complement. They gave the response on a slider marked 'no' at one end (coded as 0) and 'yes' at the other (coded as 1), as shown in Figure 7.

> **Fact (which Betty knows):** Grace loves her sister.
>
> **Betty asks:** *"Did Andrew say that Grace visited her sister?"*
>
> Is Betty certain that Grace visited her sister?
>
> **no** ⸻⸻⸻⸻⸻⸻⸻⸻⸻ **yes**
>
> Next

Figure 7: A sample trial in Experiment 2

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. For the remaining 277 participants, we inspected their responses to the 6 control stimuli, for which we expected low responses because the main clause content is not expected to project from these questions. The response means of 25 participants were more than 1.5 standard deviations above the group mean of .21 for the control polar question stimuli. Closer inspection revealed that these participants' responses to the control stimuli were systematically higher or lower, suggesting that these participants did not attend to the task or interpreted the task differently. The data from these 25 participants were also excluded, leaving data from 252 participants (ages 21-72; median: 37; 124 female, 128 male).

3.2   Results

Mean certainty ratings by predicate and by fact type are shown in Fig. 8. Overall, mean certainty ratings are higher for for events with a higher prior probability than for events with a lower prior probability. We also observe by-predicate variability in mean certainty ratings: for instance, the content of the clausal complement of *pretend* had the lowest mean certainty rating (low prior: .09; high prior: .3) and that of *be*

annoyed had the highest mean certainty rating (low prior: .55; high prior: .73). These findings suggest that participants considered both the lexical meaning of the predicate and the the prior probability of the event described by the complement clause in identifying whether the speaker was committed to the content of the clausal complement.
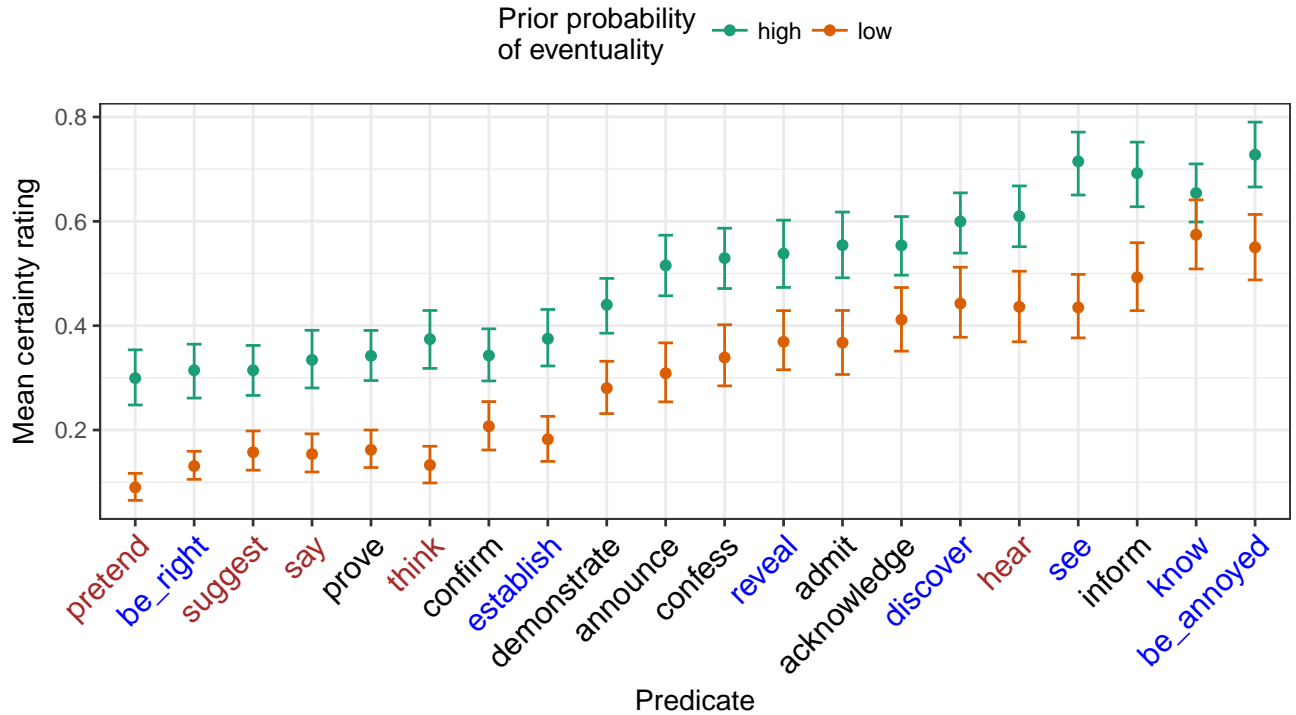

Figure 8: Mean certainty ratings by predicate and fact type, collapsing across complements. Error bars indicate 95% confidence intervals. Non-entailing predicates are given in brown, entailing predicates are given in blue, and apparently-entailing predicates are given in black.

The qualitative observation about the influence of the prior probability of events on projectivity was borne out statistically. We conducted a mixed-effects linear regression predicting certainty rating from fixed effects of the mean prior probability of events (norming study), the mean contradictoriness rating of the predicates (Exp. 1a), and their interaction. The model included the maximal random effects structure justified by the data and the theoretical assumptions: random by-participant intercepts (capturing differences in projectivity between participants), random by-content intercepts (capturing differences in projectivity between contents of clausal complements), random by-predicate intercepts (capturing differences in projectivity between predicates), random by-fact intercepts (capturing differences in projectivity between the facts) and random slopes for prior event probability by participant, content and predicate (capturing that the effect of prior event probability may vary across participants, contents, and predicates). **WHAT SENSE DOES PRIOR EVENT PROB SLOPE MAKE FOR CONTENT?** The analysis was conducted on target trials only (5,040 data points) because we were specifically interested in the projectivity of the content of the clausal complement. Analyses were conducted using the `lme4` package (Bates et al., 2015); *p*-values were obtained using the `lme4Test` package.

The contents of clausal complements that describe events with a higher probability were more likely to be taken as a commitment of the speaker than the contents of clausal complements that describe events with a lower probability ($\beta = .33$, $SE = .03$, $t = 10.95$, $p < .0001$). Neither the contradictoriness rating of the predicate nor its interaction with the prior event probability reached significance ($p > .26$).

3.3 Discussion

Replicating Tonhauser et al. accepted, we observe variability in projectivity.

Although projectivity is influenced by the lexical meaning of clause-embedding predicates (e.g., Karttunen 1971; Tonhauser et al. accepted), our findings do not support the assumption that entailment determines the empirical purview of analyses of projection. For instance, even though the veridicality of *reveal, be annoyed* and *confess* is statistically indistinguishable, the content of the complement of *be annoyed* is significantly more projective than that of *reveal* or *confess*. Our findings are compatible with analyses of projection that are not restricted to entailed content and according to which listeners integrate multiple sources of information, including prior event probabilities, in determining what speakers are committed to. Which lexical meaning properties of predicates influence projectivity and how these properties can be reliably diagnosed is a pressing question for future research.

4 Discussion

5 Conclusions

A Norming study: Prior probabilities of events

This norming study explored the prior probability of 20 events described by English sentences given a fact about the world that made the event more likely and given a fact about the world that made the event less likely. Gradient likeliness ratings were collected for each of the 40 fact/event pairings.

A.1 Methods

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 55 cents for participating in the experiment.

Materials   The 20 events were described by English sentences, as were the 2 facts about the world that each event was presented with. As illustrated with the two sample fact/event pairs in (13), the fact in each fact/event pair was preceded by the label 'Fact:'. For each event, we hypothesized that one of the two facts that the event was paired with makes the event more likely than the other fact. For instance, for the event of Julian dancing salsa in (13), we hypothesized that this event would be more likely given the fact that Julian is from Cuba than given the fact that Julian is from Germany. A second consideration in choosing the two facts was that the prior probability of an event given a fact would be neither at ceiling nor at floor. If the prior probability of an event given a fact was at ceiling, one could argue that the proposition denoted by the sentence that describes the event is entailed by the fact, i.e., a speaker might be taken to be committed to the proposition just because they are committed to the fact about the world. Likewise, if the prior probability of an event given a fact was at floor, one could argue that the negation of the proposition denoted by the sentence that describes the event is entailed by the fact, i.e., a speaker might be taken to be committed to the negation of the proposition just because they are committed to the fact about the world. To be able to explore the influence of the prior probability of events on the commitment of the speaker to the content of the clausal complement of an attitude predicate, we avoided contents that were entailed by or whose negation was entailed by facts about the world. See Appendix C for the remaining 38 fact/event pairs.[6]

---

[6]Need to say that we use 'event' for events and states, rather than 'eventuality'.

(13)   a. Fact: Julian is from Cuba.
           Julian dances salsa.
        b. Fact: Julian is from Germany.
           Julian dances salsa.

The experiment also included two control stimuli, which were used to assess whether participants were attending to the task. Like the target stimuli, both control stimuli consisted of an English sentence describing a fact about the world and an English sentence describing an event, as shown in (14). The two control stimuli differ in the probability of the event given the fact about the world: the prior probability of the control event in (14a) is 1 (the proposition denoted by the sentence is entailed by the fact) and the prior probability of the control event in (14b) is 0 (the negation of the proposition denoted by the sentence is entailed by the fact).

(14)   a. Fact: Barry lives in Germany.
           Barry lives in Europe.
        b. Fact: Tammy is a rabbit.
           Tammy speaks Italian and Greek.

The 40 target stimuli were distributed across two lists of 20 target stimuli each so that each list included the 20 event descriptions as well as 10 fact/event pairs for which the event was hypothesized to be likely and 10 fact/event pairs for which the event was hypothesized to be less likely. The two control stimuli were added to both lists, for a total of 22 stimuli per list.

Procedure   Participants were randomly assigned to a list. They were told that they would read a fact about the world and were asked to assess how likely a particular event was, given the fact. The 22 stimuli were presented in random order to each participant. On each trial, participants read the fact and the corresponding response question, which was formed from the question *How likely is it that. . . ?* with the sentence describing the event realizing the embedded clause of the question. Participants gave their response on a slider marked 'impossible' at one end (coded as 0) and 'definitely' at the other (coded as 1), as shown in Fig. A1.

Fact: Frank is allergic to cats.

How likely is it that Frank got a cat?

impossible   definitely

Continue

Figure A1: A sample trial in the prior probability 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. For the remaining 87 participants, we inspected their responses to the two control stimuli; the group means were .86 for (14a) and .03 for (14b). 19 participants gave responses

19

lower than .8 to (14a) or responses higher than .2 to (14b), suggesting that these participants did not attend to the task or interpreted the task differently. The data from these 19 participants were also excluded, leaving data from 68 participants (ages 21-75; median: 36; 31 female, 37 male).

A.2 Results and discussion

As expected, the likeliness ratings of events were influenced by facts about the world: the mean prior probability of the events was .7 (sd = .21) when presented with facts that made the events more likely and .16 (sd = .17) when presented with facts that make the events less likely. Figure A2 shows the likeliness ratings for the 20 events given facts that make the events more likely (brown dots) and given the facts that make the events less likely (blue dots). Mean likeliness ratings for each of the 40 fact/event pairs are shown as black dots. Error bars indicate 95% confidence intervals.

Figure A2 shows that, for each event, the likeliness of the event is higher given the fact that makes the event more likely than given the fact that makes the event less likely. The most likely event is that of Charley speaking Spanish given the fact that he lives in Mexico (mean: .89, sd: .14). The least likely event is that of Isabella eating a steak on Sunday given the fact that she is a vegetarian (mean: .05, sd: .07). Given that, for the 68 remaining participants, the mean prior probability of the control event in (14a) was .99 (sd: .01) and the mean prior probability of the control event in (14b) was .01 (sd: .01), we conclude that the likeliness of none of the 40 event/fact pairs is at ceiling or at floor.

B   Control stimuli in Exp. 2

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

        b. **Fact (which NAME knows): Mary visited her aunt on Sunday**
            NAME: Is Mary's aunt sick?

        c. **Fact (which NAME knows): Todd goes to the gym 3 times a week**
            NAME: Did Todd play football in high school?

        d. **Fact (which NAME knows): Vanessa won a prize at school**
            NAME: Is Vanessa good at math?

        e. **Fact (which NAME knows): Trish sent Madison a card**
            NAME: Did Madison have a baby?

        f. **Fact (which NAME knows): Hendrick just bought a car**
            NAME: Was Hendrick's car expensive?

C   Materials used in the event probability norming study

The list below provides the 20 English sentences that describe the events whose prior probability was explored in the norming study described in section A, together with the fact about the world that makes the event more likely and the fact about the world that make the event less likely.

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)
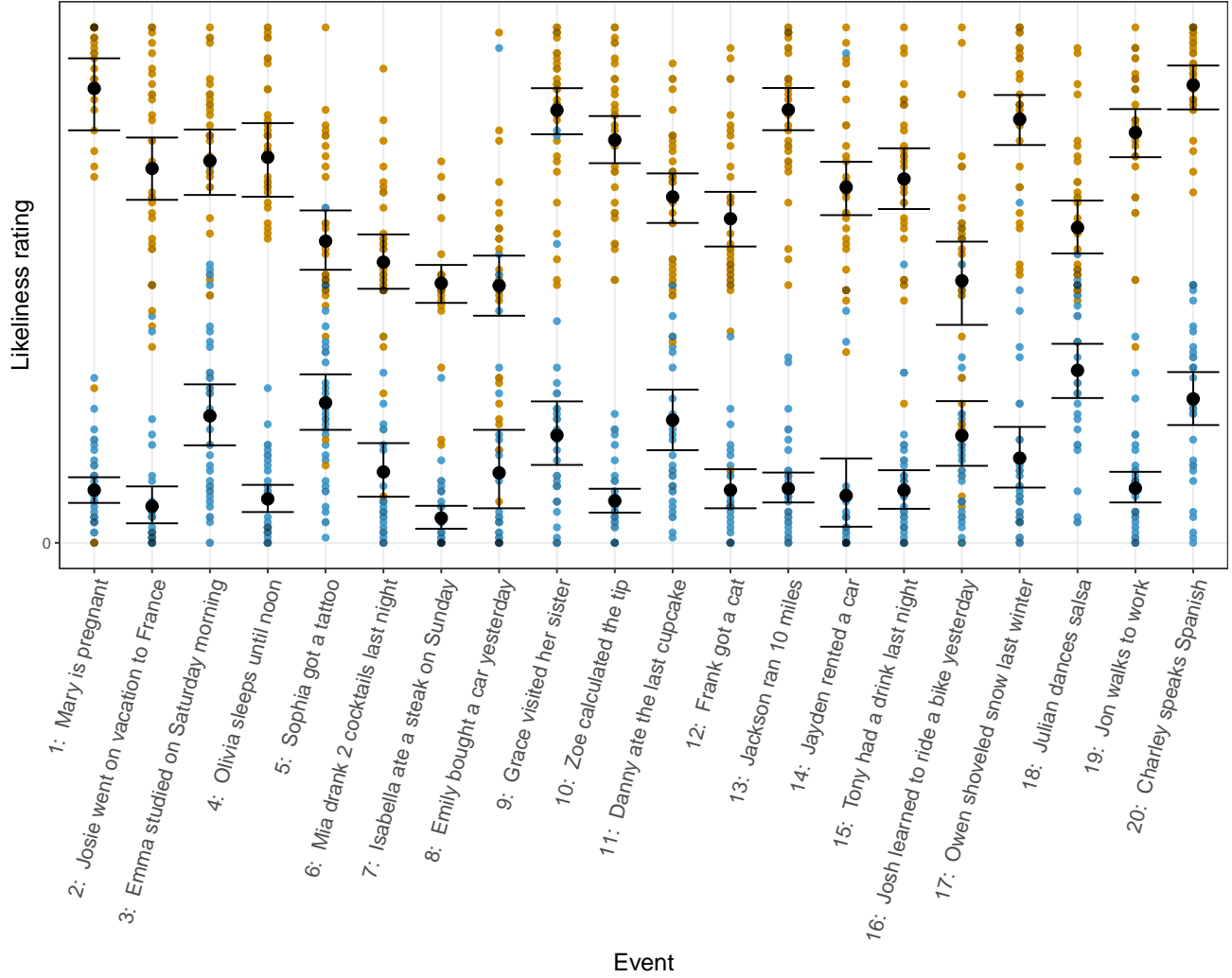5. Sophia got a tattoo (Sophia is a high end fashion model / Sophia is a hipster)

Figure A2: Likeliness ratings for 20 events given facts about the world that make the events more likely (brown dots) and facts about the world that make the events less likely (blue dots); darker dots indicate overlapping ratings. Mean likeliness ratings for the 40 fact/event pairs are given by black dots. Error bars indicate 95% confidence intervals.

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)

D Materials used in Exp 2

E Materials used in Exp 3

References

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.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Chierchia, Gennaro and Sally McConnell-Ginet. 1990. *Meaning and Grammar*. Cambridge, MA: MIT Press.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50(3):346–363.

Karttunen, Lauri. 1971. Some observations on factivity. *Papers in Linguistics* 4:55–69.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schlenker, Philippe. 2010. Local contexts and local meanings. *Philosophical Studies* 151:115–142.

Simons, Mandy. 2001. On the conversational basis of some presuppositions. In *Semantics and Linguistics Theory (SALT) XI*, pages 431–448. Ithaca, NY: CLC Publications.

Simons, Mandy, David Beaver, Craige Roberts, and Judith Tonhauser. 2017. The Best Question: Explaining the projection behavior of factive verbs. *Discourse Processes* 3:187–206.

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.

Swanson, Eric. 2012. Propositional attitudes. In *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, pages 1538–1561. Berlin: Mouton de Gruyter.

Tonhauser, Judith, David Beaver, and Judith Degen. accepted. How projective is projective content? Gradience in projectivity and at-issueness. Accepted for publication in *Journal of Semantics*.