I don't know if projection is binary.
Did Mandelkern et al. 2020 discover that it is?

Judith Tonhauser° and Judith Degen•

°University of Stuttgart
•Stanford University

November 4, 2023

Abstract

Presuppositions are taken to typically project out of entailment-canceling environments like the scope of negation (e.g., Frege 1892, Chierchia and McConnell-Ginet 1990). While projection is often characterized as binary (that is, content either typically projects or not), there is mounting empirical evidence that projection is gradient, with content being more or less projective (e.g., Karttunen 1971, Xue and Onea 2011, de Marneffe et al. 2019, Tonhauser et al. 2018, Degen and Tonhauser 2022). Mandelkern et al. 2020 critically evaluated the inference rating measures on which this empirical evidence is based and proposed that naturalness ratings in explicit ignorance contexts are more suitable to measure binary projection and, thereby, to distinguish presupposition triggers from nontriggers. This paper presents the results of an experiment designed to investigate whether the measure proposed by Mandelkern et al. 2020 can distinguish factive predicates (presumed presupposition triggers) from nonfactive ones (nontriggers). While the results do not support Mandelkern et al.'s 2020 hypothesis and are rather in line with results based on inference rating measures (e.g., Degen and Tonhauser 2022), the results suggest that naturalness ratings in explicit ignorance contexts are sensitive to felicity conditions of anaphoric projective content.

Keywords: Presuppositions, binary vs. gradient projection, inference ratings, naturalness ratings in explicit ignorance contexts.

1 Introduction

Presuppositions are defined as content that typically projects out of entailment-canceling environments, like the scope of negation or a polar interrogative (e.g., Frege 1892, Chierchia and McConnell-Ginet 1990).[1] Thus, a content like the content of the clausal complement (CC) of *know* in (1a) is taken to be a presupposition because a speaker who utters (1a) is typically taken to believe the CC, that Julian dances salsa, even though the clausal complement occurs in a polar interrogative. By contrast, the CC of *think* in (1b) is not assumed to be a presupposition because it does not typically project. Accordingly, *know* is analyzed as a presupposition trigger, but *think* is not.

(1) a. Does Cole know that Julian dances salsa?
     b. Does Cole think that Julian dances salsa?

To investigate whether a particular content is a presupposition, empirical research has long relied on inference rating measures. Such measures involve presenting participants with an utterance (like that in (1a)) and contents (like the content that Julian dances salsa), and asking for ratings that are taken to reveal the existence or strength of the inference to the content. Many theoretical works on projection assume that projection is a binary property of utterance content, that is, content either typically projects (in which case it is a presupposition) or it does not (in which case it is not a presupposition). There is, however, mounting empirical evidence (based on investigations that used inference rating measures) that projection is not a binary property of content but rather a gradient one, with content being more or less projective. Karttunen 1971, for instance, already suggested that the contents of the clausal complements (CCs) of *regret* and *discover*, which he took to be presuppositions, exhibit projection variability in examples like (2). Specifically, he suggested that a speaker who utters (2a) believes that the addressee has not told the truth, whereas a speaker who utters (2b) "is not sure about the truth of the complement" (p.63).

(2) a. Did you regret that you had not told the truth?
     b. Did you discover that you had not told the truth?   (Karttunen 1971:63)

More recently, experimental investigations have also observed projection variability between presuppositions (e.g., Xue and Onea 2011, de Marneffe et al. 2019, Tonhauser et al. 2018, Degen and Tonhauser 2022). Tonhauser et al. 2018, for instance, observed in their Exps. 1 that the CCs of *know, discover,* and *reveal* exhibit projection variability: The CC of *know* is more projective than that of *discover*, and the CC of *discover* is more projective than that of *reveal*. These results led Tonhauser et al. 2018:498 to propose that "projectivity [. . . is] a gradient property of content, rather than a binary, categorical one". Further support for this proposal comes from investigations of both factive and nonfactive predicates in Degen and Tonhauser 2022. In a series of experiments and analyses of existing datasets, this work not only found projection variability between the CCs of factive predicates, but also no empirical support for a binary distinction in the projection of the CCs of factive and nonfactive predicates. This is because the CCs of both factive and nonfactive predicates were projective to varying degrees compared to nonprojective main clause content. A sample illustration of these results is given in Fig. 1 from their Exp. 1a, which shows the mean projection ratings of 5 factive predicates (in orange) and 15 nonfactive predicates (in green).

Recently, Mandelkern et al. 2020:§6.2 challenged the hypothesis that projection is gradient by arguing that inference rating measures are not suitable to detect binary projection. Such measures are not suitable, they argued, because they do not distinguish between, on the one hand, contents that project because they are presuppositions and, on the other hand, contents that project because they are "a natural conclusion to draw for any of a variety of pragmatic reasons short of entailment or presupposition" (p.497). Inference

---

[1]Conventional implicatures also project out of these environments. They differ from presuppositions in that they contribute novel information. For discussion see, e.g., Chierchia and McConnell-Ginet 1990, Potts 2005.
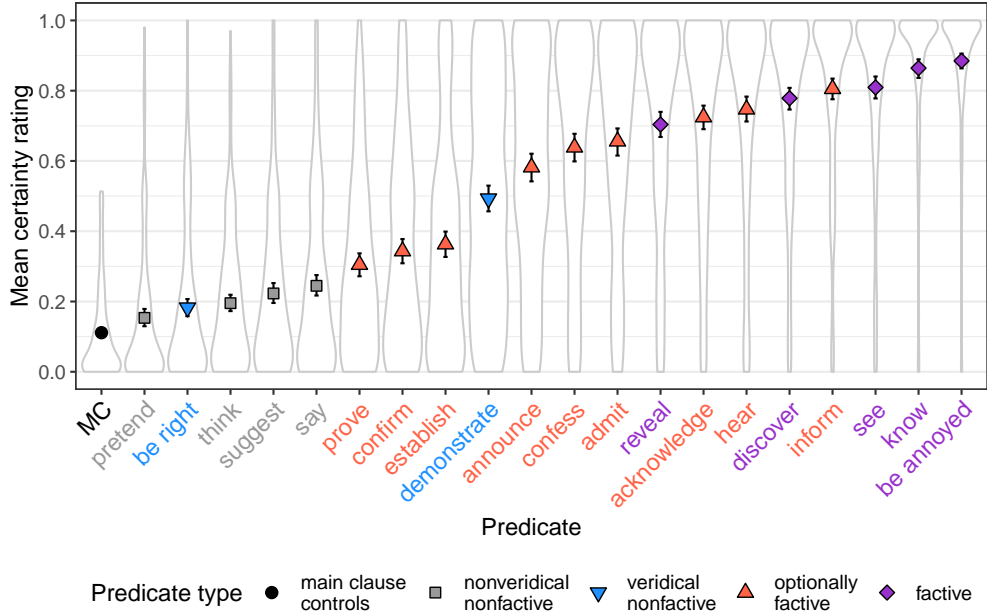
Figure 1: Mean certainty rating (measuring projection) of main clause content and the CCs of the 20 predicates investigated in Exp. 1a of Degen and Tonhauser 2022 (adapted by collapsing the nonfactive predicates into a single category). Error bars indicate 95% bootstrapped confidence intervals. Violin plots indicate the kernel probability density of the individual participants' ratings.

rating measures, they suggested, might invite participants to draw projection inferences even when not lexically required because "the content in question is right there before the [participants'] eyes" (p.498). They concluded that "we should think twice before embracing a notion of presupposition projection that is gradient based on results from inference tasks alone" (p.497).

Mandelkern et al. 2020:§6.2 suggested that naturalness ratings of utterances in explicit ignorance contexts are better suited to distinguish presuppositions from nonpresuppositional projection inferences.[2] On this measure, participants are presented with utterances in which a speaker states that they are ignorant about a content, as in (3a), and then follows up with an utterance in which the content occurs embedded in an entailment-canceling environment, such as the antecedent of a conditional in (3c). Mandelkern et al. 2020 assumed that if the content is a presupposition (such as the pre-state content of *stop* in (3c)), the utterance is rated as unnatural in the explicit ignorance context because presuppositions are contents that the speaker believes to be true. That is, in such constellations, participants "will not have an alternative to interpreting the presupposition at the utterance level, and in turn seeing the utterance as infelicitous and the speaker as incoherent" (p.497). If, on the other hand, the content is not a presupposition (such as the pre-state content of *now frown on* in (3c)), the utterance is assumed to be natural: Even if participants might, in a different context, draw a projection inference "for any of a variety of pragmatic reasons short of entailment or presupposition, [...] they will tend to relinquish that inclination when there is pragmatic pressure to do so" as in the explicit ignorance contexts (p.497).[3] Both presuppositions and nonpresuppositions were expected to be judged as natural in the support contexts, which entail the relevant content, illustrated in (3b).

(3)   Mandelkern et al. 2020:490f.

---

[2]Mandelkern et al. 2020 refer to the measure as an acceptability rating measure, but since participants were asked to rate the naturalness of utterances, we refer to it as a naturalness rating measure.

[3]Presuppositions won't be relinquished, modulo local accommodation

a. Explicit ignorance context:
   Mary always was involved in a lot of sports, but I don't know whether she ever did any yoga.

b. Support context:
   Mary always was involved in a lot of sports, and she used to do yoga, too.

c. Sentence with presupposition / nonpresupposition:
   If Mary {has stopped / now frowns on} doing yoga, then Matthew will interview her for his story.

Mandelkern et al.'s 2020 Exp. 3, which used the proposed measure, only provided limited empirical evidence for the hypothesis that projection is a binary property of content. This is primarily because the goal of this work (including Exp. 3) was not to investigate this hypothesis, but rather whether presupposition projection in conjunctions is symmetric. The evidence provided in this work is also limited, however, because only four presupposition triggers and nontriggers each were investigated: The presupposition triggers were the change-of-state verbs *stop* and *continue*, and the two clause-embedding predicates *be aware* and *be happy*; the nontriggers were *now frown on, enjoy, hope* and *be sure*.[4] In the control items of Exp. 3, these eight expressions were realized in (single-clause) antecedents of conditionals, as shown in (3c). Naturalness ratings were elicited in explicit ignorance and support contexts.

Mandelkern et al. 2020 found (as shown in their Fig. 3) that the naturalness ratings for the nontriggers in the control items were relatively high in both contexts (with means just below 5, where 1 meant "completely unnatural" and 7 meant "completely natural" on their 7-point Likert scale). The presupposition triggers, on the other hand, had lower mean naturalness ratings in the explicit ignorance context (mean around 3) than in the support context (mean just above 5). Mandelkern and his colleagues took this to "establish[] the effectiveness of the methodology" (p.492). They proposed (p.497):

> [. . .] comparing contexts which support the inference to contexts in which it has been made clear that the speaker is ignorant about the inference provides a way to distinguish a broad class of natural and invited pragmatic inferences from those that are really encoded as presuppositions, and thus have no choice but to project. [...] In methodological terms, we strongly recommend at least a two-pronged approach, with careful attention paid to results stemming from the evaluation of acceptabilty [sic] in different contexts.

However, an inspection of the naturalness ratings for the eight (non)triggers provided in Fig. 8 in their Appendix 3 raises doubt about whether the measure indeed provides evidence for a binary distinction between presuppositions and nonpresuppositions. First, considering the mean naturalness ratings in the explicit ignorance contexts, we find ratings for the presuppositions range from just above 1 (for the pre-state of *continue*) to about 3.5 (for the CC of *be aware*). As the ratings for the nonpresuppositions range from just below 3 (for the CC of *hope*) to 4.5 (for the CC of *be aware*), it is not clear whether naturalness ratings in explicit ignorance contexts provide evidence for a binary distinction between the four presuppositions and the four nonpresuppositions investigated in Mandelkern et al. 2020. Second, the effect of context (explicit ignorance vs. support) also does not appear to be consistent for presuppositions or nonpresuppositions. Whereas the difference between the mean naturalness ratings in the explicit ignorance and support contexts was quite large for the (presuppositional) pre-state of *continue*, the difference was much less pronounced for the (presuppositional) CC of *be aware*, where it was roughly comparable to that of the (nonpresuppositional)

---

[4]While their Exps. 1 and 2 investigated *frown on*, Exp. 3 investigated *now frown on*, as shown in (i). It is possible that the pre-state content contributed by this expression is projective: (i), for instance, appears to give rise to the projective content that Mary did not frown on doing yoga in the past.

(i) If Mary now frowns on doing yoga, then Matthew will interview her for his story.   (Mandelkern et al. 2020:491)

pre-state of *frown on*. In sum, it is not clear that comparisons of ratings in support and explicit ignorance contexts provides empirical support for a binary projection either.

There also appears to be disagreement in the literature about the naturalness of presuppositions in explicit ignorance contexts. Whereas Mandelkern and his colleagues assumed that all presuppositions are unnatural in explicit ignorance contexts, Simons 2001 and Abusch 2010 assumed that only hard (nondefeasible) presuppositions are, but that 'soft' (defeasible) ones are natural. For instance, *stop* and *win* in (4) were assumed to be natural in explicit ignorance contexts, and *again*, the *it*-cleft and *too* in (5) unnatural.[5] This disagreement in the literature necessitates further investigation.

(4) a. I have no idea whether Jane ever smoked, but she hasn't stopped smoking. (Simons 2001:443)

    b. I have no idea whether John ended up participating in the Road Race yesterday. But if he won it, then he has more victories than anyone else in history. (Abusch 2010:39)

(5) a. #I don't know if Jane ever rented "Manhattan" before, but perhaps she's renting it again. (Simons 2001:443)

    b. #I have no idea whether anyone read that letter. But if it is John who read it, let's ask him to be discreet about the content. (Abusch 2010:40)

    c. #I have no idea whether John read that proposal. But if Bill read it too, let's ask them to confer and simply give us a yes-no response. (Abusch 2010:40)

This paper presents the results of an experiment designed to investigate whether the measure proposed by Mandelkern et al. 2020 provides evidence that projection is a binary property of content rather than a gradient one.[6] To allow for comparison with the results of Degen and Tonhauser 2022, the CCs of the same 20 (non)factive clause-embedding predicates (see Fig. 1) were investigated. The experiment also included control stimuli with *stop, continue, again, too, also*, and an *it*-cleft, to allow for comparison with the results of Mandelkern et al. 2020 and the assumptions made in Simons 2001 and Abusch 2010.

One change we made to the design of Mandelkern et al.'s 2020 Exp. 3 was that our experiment implemented a three-level context condition instead of a two-level one. Specifically, we replaced Mandelkern et al.'s 2020 support context (which entailed the relevant context, see (3b)) with two contexts that were compatible with the relevant content but did not entail it. The two contexts differed in the prior probability of the relevant content: In the 'lower prior probability' context, illustrated in (6a), the relevant content (in (6c), that Julian dances salsa) had a comparatively lower prior probability, whereas it had a comparatively higher prior probability in the 'higher prior probability' context, illustrated in (6b). These two contexts were normed in Degen and Tonhauser 2021.

(6) a. Julian is German. [lower prior probability]

    b. Julian is Cuban. [higher prior probability]

    c. Did Cole discover that Julian dances salsa?

This change was implemented to allow for comparison to the result of Degen and Tonhauser 2021, namely that the projection of a content is sensitive to the content's prior probability. Specifically, this work found that the higher the prior probability of a content, the stronger the projection inference. Under standard analyses of presuppositions, like Heim 1983 or van der Sandt 1992, presuppositions are by default globally accommodated in such contexts. Accordingly, such analyses predict that presuppositions are rated as natural in both lower and higher probability contexts, just as in Mandelkern et al.'s 2020 support contexts.

---

[5]Abusch 2010 marked (5b) and (**??**c) with '??' instead of '#'.

[6]Link to experiment, materials, and analysis scripts: xxx

2 Experiment

To investigate whether the measure proposed in Mandelkern et al. 2020 provides support for a binary distinction between presuppositions and nonpresuppositions, participants read two-sentence discourses consisting of a declarative (which provided the context) and an interrogative (which realized a (non)factive clause-embedding predicate), and rated the naturalness of the interrogative in the context of the declarative.

2.1 Methods

2.1.1 Participants

We recruited 425 participants on Prolific. Due to a server error, the data of only 398 participants was recorded (ages: 19-73, mean: 40.8; 187 women, 201 men, 8 non-binary, 2 preferred to not disclose). The recruited participants were required to live in the USA, to speak English as their first language, to have completed at least 100 tasks, and to have an approval rating of at least 99%. The median time spent on the task was 6:24 minutes. Participants were paid $1.78 (corresponding to an hourly pay of $16.6).

2.1.2 Materials

Participants read two-sentence discourses consisting of a declarative followed by an interrogative, as shown in (7). In the target stimuli, the interrogatives combined the 20 (non)factive predicates of Degen and Tonhauser 2022 (see Fig. 1) with the 20 complement clauses of Degen and Tonhauser 2022, for a total of 400 interrogatives. The preceding declaratives implemented a three-level context condition: In the 'explicit ignorance' context (7a), the declarative sentence conveyed the speaker's ignorance about the CC. In the 'lower prior probability' context (7b), the CC had a comparatively lower prior probability, whereas it had a comparatively higher prior probability in the 'higher prior probability' context (7c). See Supplement A for the full set of 20 complement clauses and the two contexts for each complement clause.

(7) a. Explicit ignorance context:
         I have no idea if Julian dances salsa. Does Cole discover that Julian dances salsa?

    b. Lower prior probability context:
         Julian is German. Did Cole discover that Julian dances salsa?

    c. Higher prior probability context:
         Julian is Cuban. Did Cole discover that Julian dances salsa?'

In addition to the 1,200 target stimuli (400 interrogatives × 3 contexts), the materials also included six control stimuli, shown in (8). The interrogatives of the control stimuli featured six expressions typically analyzed as presupposition triggers, namely *stop, continue, again, too, also*, and an *it*-cleft. The preceding declaratives conveyed the speaker's explicit ignorance about the presuppositions associated with these expressions.[7] Three of these expressions, namely *stop, continue,* and *again*, were target expressions in the relevant experiments in Mandelkern et al. 2020 and Kalomoiros and Schwarz 2021, who assumed these expressions to be unnatural in explicit ignorance contexts. This is in contrast to Simons 2001, who took *stop* to be a soft trigger that is acceptable in explicit ignorance contexts. The other three expressions, namely *too, also,* and the *it*-cleft, were assumed to be hard triggers in Simons 2001 and Abusch 2010, that is, unnatural in explicit ignorance contexts. The control stimuli were included for comparison to the target stimuli, to the claims of Simons 2001 and Abusch 2010, and to the results of Mandelkern et al. 2020.

---

[7]As the stimuli were presented in writing, prosody was not controlled for. Since the projective contents of *too, also,* and the *it*-cleft cannot be accommodated, the filler stimuli with these expressions should receive low naturalness in explicit ignorance context under any focus assignment.
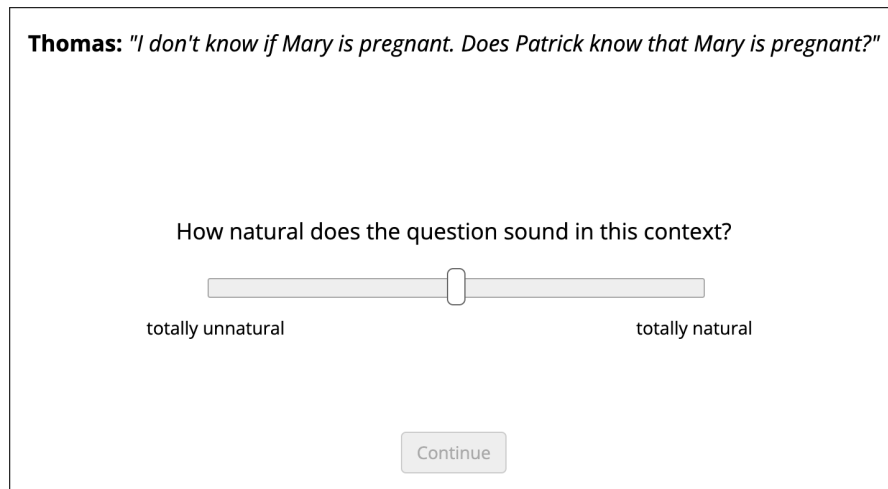
(8)   a. I don't know if Stephen was ever in the habit of vaping. Has Stephen recently stopped vaping?

        b. I don't know if John was ever reading "Dune". Has John recently continued reading "Dune"?

        c. I don't know if William was ever interested in history. Is William interested in history again?"

        d. I don't know if Ann plays any instrument. Does Ann play the flute, too?

        e. I don't know if Svenja plays any sport. Does Svenja also play soccer?

        f. I don't know if anyone was playing outside with the kids. Was it Jack who was playing outside with the kids?

The experiment also included 4 filler stimuli that were used to exclude participants not attending to the task (see Supplement B). The filler stimuli were expected to receive high naturalness ratings.

A random set of 30 stimuli was created for each participant. Each set contained 20 target stimuli in which each of the 20 complement clauses was paired with a unique clause-embedding predicate. Twelve of the target stimuli were presented in the explicit ignorance context, and the other eight in a low or a higher prior probability context (four each). Each participants' set of 30 stimuli contained the same six control stimuli and the same four filler stimuli. Each of the 30 stimuli were presented as utterances by a unique named speaker. Trial order was randomized for each participant.

2.1.3   Procedure

Participants were instructed to rate "how natural the question sounds in the context of the statement". As shown in Figure 2, they were asked give their rating on a slider from 'totally unnatural' (coded as 0) to 'totally natural' (coded as 1). The experiment began with four practice trials to familiarize participants with the task (see Supplement B for details). After responding to the 30 trials, participants filled out a short optional demographic survey. To encourage truthful responses, participants were told that they would be paid no matter what answers they gave in the survey.

Thomas: *"I don't know if Mary is pregnant. Does Patrick know that Mary is pregnant?"*

How natural does the question sound in this context?

totally unnatural   totally natural

Continue

Figure 2: Sample experiment trial with explicit ignorance context.

2.1.4   Data exclusion

We excluded the data of five participants who did not self-identify as native speakers of American English as well as the data of 23 participants whose mean response to the attention checks was more than 2 sd below

the group mean.[8] The data of 370 participants entered into the analysis (ages: 19-80, mean: 40.7; 175 women, 185 male, 8 non-binary, 2 preferred to not disclose). Each of the 20 predicates received at least 200 ratings in the explicit ignorance context (mean: 222 ratings), at least 59 ratings in the lower prior probability context (mean: 74), and at least 61 ratings in the higher prior probability context (mean: 74). The six control stimuli received 370 ratings each.

2.2 Results

We first address the question of whether naturalness ratings in explicit ignorance contexts provide empirical evidence for a binary difference between presuppositions and nonpresuppositions (§2.2.1). We then consider whether a comparison of naturalness ratings in the contexts provides such evidence (§2.2.2).

2.2.1 Naturalness ratings in explicit ignorance contexts

Recall that Simons 2001 and Abusch 2010 assumed that hard presuppositions are unnatural in explicit ignorance contexts, whereas Mandelkern et al. 2020 assumed that all presuppositions are unnatural in such contexts. Figure 3 shows the mean naturalness ratings in the explicit ignorance context by expression (in distinct colors: factive predicates, nonfactive predicates, controls). As shown, purported presuppositions, that is, the controls and the factive predicates, do not exhibit a unified pattern: The mean naturalness ratings of four controls (*continue, too, also, again*) is at floor, those of *stop* and *be annoyed* are a bit higher, those of the *it*-cleft and *know* are higher again,[9] and those of *see, discover*, and *reveal* are just as high as the means of some nonfactive predicates. These results do not provide empirical support for a binary distinction between presuppositions and nonpresuppositions, contrary to what Mandelkern et al. 2020 hypothesized. Furthermore, these results do not provide empirical support for for a class of factive predicates that is categorically distinct from nonfactive predicates, in line with the result of Degen and Tonhauser 2022 based on inference rating measures.

These observations were confirmed by a posthoc pairwise comparison of the naturalness ratings of the contents associated with the six controls and 20 target expressions using the 'emmeans' package (Lenth 2023) in R (R Core Team 2016). The input to the pairwise comparison came from a Bayesian mixed-effects beta regression model with weakly informative priors using the 'brms' package (Bürkner 2017). The model predicted naturalness ratings from a fixed effect of expression (with treatment coding and 'continue' as reference level) and included random by-participant and by-item intercepts (where an item is a complement clause). We thus obtain, for each predicate, a 95% credible interval for the mean rating for that predicate. This analysis provided us with with posterior distributions of estimated marginal means of pairwise differences. We assume that two expressions differ in their mean naturalness rating when the posterior distribution of their difference does not include 0. See Supplement C for the full model output.

As shown in Table 1, the analysis suggests difference between the naturalness ratings of the factive predicates in the explicit ignorance context: the ratings of *be annoyed* are lower than those of *know*, which in turn are lower than those of *see* and *discover*, which in turn are lower than those of *reveal*. Furthermore, the naturalness ratings for the CCs of the five factive predicates are indistinguishable from different sets of expressions: The naturalness ratings for *be annoyed* do not differ from those of *stop* (but are higher than those the prejacents of the controls *continue, too*, and *also*), the naturalness ratings for *know* do not differ from those of the *it*-cleft, the naturalness ratings for *see* and *discover* do not differ from those of the nonfactives *confess, acknowledge*, and *admit*, and those of *reveal* do not differ from those of the nonfactives *prove* and *announce*. These results confirm that naturalness ratings in explicit ignorance contexts do not

---

[8]Contrary to what was preregistered, one of the four controls was not used to exclude participants because it received a mean rating of only .5. See Supplement B for details.

[9]The distribution of the ratings for the CC of *know* is bimodal. We discuss this in §3.
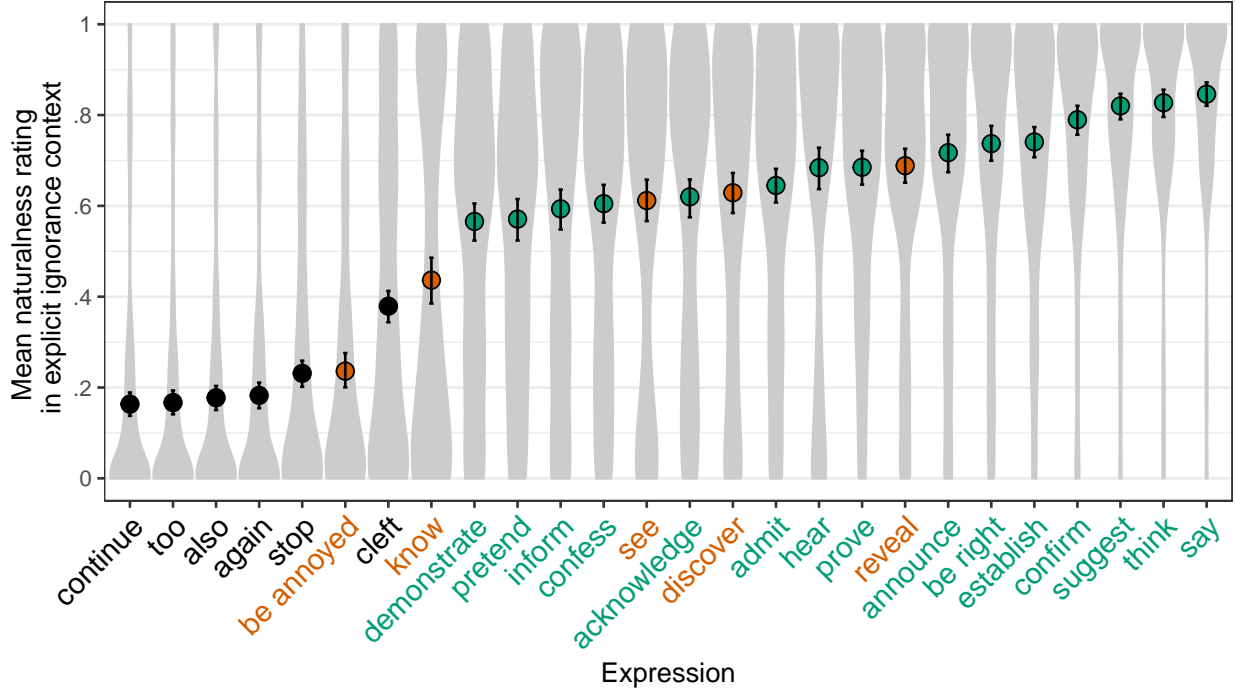
Figure 3: Mean naturalness rating in explicit ignorance context by expression (factive, nonfactive, filler). Error bars indicate 95% bootstrapped confidence intervals. Overlaid violin plots indicate the kernel probability density of the participants' ratings.

provide empirical support for a binary distinction between factive and nonfactive predicates, or between presuppositions and nonpresuppositions.

Contrary to what is assumed in Mandelkern et al. 2020 and Kalomoiros and Schwarz 2021, presuppositions are not invariably unnatural in explicit ignorance contexts. Rather, in line with what was assumed in Simons 2001 and Abusch 2010, some presuppositions seem to be rated as more unnatural in such contexts than others. On the assumption that naturalness ratings in explicit ignorance contexts differentiate between undefeasible/hard and defeasible/soft triggers, one might take the results of our experiment to suggest that the pre-state contents of *continue* and *again*, and the existence requirements of *too* and *also* are not defeasible. The CCs of *see, discover*, and *reveal*, on the other hand, might be taken to be defeasible. Contrary to what Abusch 2010 assumed, the existence requirement of the *it*-cleft is not as clearly undefeasible as that of other projective contents (Smith and Hall 2011 also observed that the *it*-cleft did not pattern like a hard trigger in their experiment.) The results of our experiment align with our observation about possible by-content variation in Mandelkern et al.'s 2020 Exp. 3. Recall from above that in their Exp. 3, the pre-state of *continue* was judged as (numerically) less natural in explicit ignorance contexts than the CC of the cognitive predicate *be aware*. This result is in line with the result of our experiment that the pre-state of *continue* is judged as less natural than the CC of the cognitive predicate *know*.

2.2.2 Comparison of naturalness ratings across contexts

We next consider whether presuppositions and nonpresuppositions differ with respect to the ratings in contexts in which the speaker is explicitly ignorant of the content and in contexts that are compatible with the content. Recall that Mandelkern et al. 2020 expected presuppositions to be sensitive to their context manipulation (explicit ignorance vs. support) but not nonpresuppositions. Fig. 4 shows mean naturalness ratings

9

	continue	too	also	again	stop	be annoyed	cleft	know	demonstrate	pretend	inform	confess	see	acknowledge	discover	admit	hear	prove	reveal	announce	be right	establish	confirm	suggest	think
too																									
also																									
again																									
stop																									
be annoyed																									
cleft																									
know																									
demonstrate																									
pretend																									
inform																									
confess																									
see																									
acknowledge																									
discover																									
admit																									
hear																									
prove																									
reveal																									
announce																									
be right																									
establish																									
confirm																									
suggest																									
say																									

Table 1: Posterior distributions of estimated marginal means of pairwise differences '***' indicates significance at .001, '**' at .01, '*' at .05, '.' marginal significance at .1, and *n.s* indicates no significant difference in means.

for the CCs of the 20 factive and nonfactive predicates in the three contexts featured in our experiment, with the predicates ordered by their mean naturalness ratings in the explicit ignorance context, as in Fig. 3.

As shown, the effect of context is not uniform for the factive predicates: For *be annoyed* and *know*, the mean naturalness rating is lower in the explicit ignorance context than in the higher prior probability context, the two ratings are equally high for *see, discover*, and *reveal*. This latter pattern is also observed for many of the nonfactive predicates, for instance *confess, hear, announce*, and *be right*. Thus, naturalness ratings in explicit ignorance vs. higher prior probability contexts do not provide support for a binary distinction between the CCs of factive and nonfactive predicates. A comparison between the explicit ignorance and lower prior probability contexts also suggests variation between the factive predicates: Whereas for *be annoyed* the mean naturalness ratings are lower in the explicit ignorance context than in the lower prior probability context, they are equally high for *know*, but lower in the lower prior probability context than the explicit ignorance context for *see, discover*, and *reveal*. This latter pattern is observed for all of the

nonfactive predicates, except possibly *pretend*. Thus, naturalness ratings in explicit ignorance vs. lower prior probability contexts also do not provide support for a binary distinction between the CCs of factive and nonfactive predicates.

Finally, we observe that the prior probability of content modulates its naturalness ratings for all predicates, except *demonstrate* and *pretend*): Content is judged as more natural in a context in which it has a higher prior probability than in a context in which it has a lower prior probability. This result is in line with the result of Degen and Tonhauser 2021 based on an inference rating task.
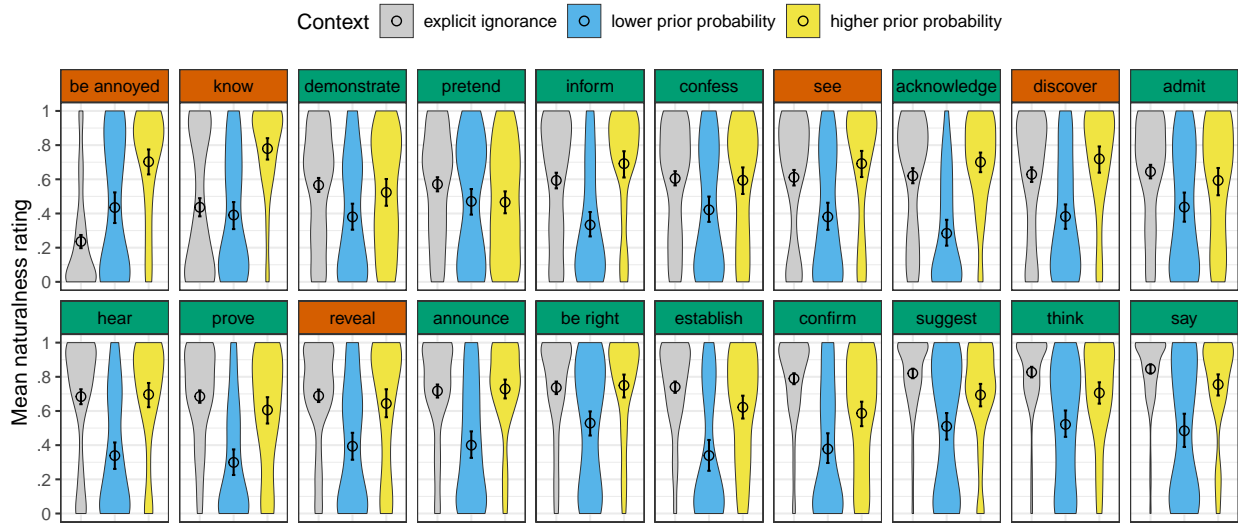

Figure 4: Mean naturalness rating by context and predicate (factive , nonfactive); predicates ordered as in Fig. 3. Error bars indicate 95% bootstrapped CIs. Violin plots indicate kernel probability density of participants' ratings.

These results were confirmed by a model...**BETA MODELS**

bla

3   Discussion

The experiment was designed to investigate Mandelkern et al.'s 2020 claim that the measure used in their Exp. 3 provides empirical support for the hypothesis that projection is a binary property of content and distinguishes presuppositions from nonpresuppositions. Neither the by-content comparison of naturalness ratings in explicit ignorance contexts nor the by-context comparison of naturalness ratings provided empirical support for a binary distinction between the CCs of factive and nonfactive predicates, or, more generally, between presuppositions and nonpresuppositions. In sum, the experiment did not find support for Mandelkern et al.'s 2020 claim nor for the hypothesis that projection is a binary property of content. Instead, the experiment provided further support for the hypothesis of Tonhauser et al. 2018 that projection is a gradient property of content and the result of Degen and Tonhauser 2022 that there is no well-defined class of factive predicates. In this section, we discuss theoretical and methodological implications of these results.

3.1   What are presuppositions?

Contemporary literature still frequently treats presuppositions as if they were a homogenous class, for instance by characterizing them as contents that typically project or as felicity requirements. Research on projective content has long recognized that the set of contents that have been called presuppositions is heterogeneous. Two dimensions of variation that were already mentioned in §1, namely defeasibility and strength of projection, fly in the face of this homogeneity assumption. A third dimension of variation is whether the presupposition is associated with what Tonhauser et al. 2013 referred to as a 'strong contextual felicity' constraint: Some presumed presuppositions are not associated with such a constraint, which means that they are judged to be acceptable in a context that is neutral with respect to the content (like the CC of *know*, which is informative in such a context), while others are associated with such a constraint, that is, judged to be unacceptable in a context that does not entail or satisfy the content (like the existence requirement of pronouns or *too*). A fourth dimension of variation is at-issue, that is, whether the content can address the Question Under Discussion: As shown in Tonhauser et al. 2018, the CCs of factive predicates vary in this regard, as the CC of *discover* is more at-issue than those of *be annoyed* and *know*.

Given that the set of contents called presuppositions is heterogeneous, it is not entirely surprising that these contents do not pattern alike in explicit ignorance contexts. For instance, a content that is associated with a strong contextual felicity constraint, like the existence requirement of *too*, might be expected to be judged as unnatural in a context in which the speaker is explicitly ignorant about the content. On the other hand, a content that is not associated with a strong contextual felicity constraint and that can be at-issue, like the CC of *discover*, might be expected to be judged as somewhat natural in a context in which the speaker is explicitly ignorant about the content, as the content can be interpreted as what the speaker is inquiring about. Our experiment suggests that both of these expectations were borne out by the data.

(9)   a.   If Julian dances salsa, then Cole knows that he does.
      b.   If Julian dances salsa, then

Against this background, it does not seem fruitful to us to

- When Mandelkern et al. 2020 talk about presuppositions, it's not even clear what they mean: +SCF or -SCF? They only investigated -SCF/+OLE expression but assumed an analysis for them that was developed for +SCF triggers.

need to discuss in the results what naturalness in EIC appears to be sensitive to, and also bimodal for know and low responses for be annoyed

3.2 Is projection a binary or a gradient property of content?

- We need to identify which level we're talking about

  - Speaker belief:
    Option 1: Speaker belief is binary, gradience is introduced by interpretation.
    Option 2: Speaker belief is gradient.
  - Lexical representation:
    Option 1: Representation codes projection
    Option 2: Representation does not code projection but projection is possible in particular utterances

- Typically, researchers ask this because they wonder about the linguistic representation we want to assume in our models. Simple picture:

  triggers introduce context requirement and hence content typically projects

  nontriggers do not introduce this and hence exhibit projection to varying degrees

- More complex picture, already represented by contemporary analyses:

  +SCF: introduce anaphoric requirement/felicity requirement

  soft triggers: lexical meaning typically gives rise to projection inference

  nontriggers: lexical meaning may give rise to projection inference

- To address this question, we have to distinguish whether we're talking about content associated with a particular expression (aggregated projection over multiple utterances) versus utterance content

- The characterization of presuppositions is at the level of content, but has to involve 'typically' to capture that a presupposition in some utterances does not project.

- We also have to distinguish the speaker and the hearer: The speaker may be binary or gradient, and even if binary, interpretation may vary because of different assumptions about speaker, different integration of other factors.

- Projection inferences, sources of gradience

	expression	utterance	speaker	interpreter
class 1	binary	binary	binary	binary
class 2	binary			
class 3				

- What does gradient projection mean?

  [H1] Projection is binary at the level of the expression: An expression either triggers a presupposition or not. The observed gradience at the utterance level is due to utterance-level factors that modulate projection.

  [H2] Projection is not binary at the level of the expression: The lexical representations of (classes of) expressions give rise to projection inferences of varying strength. The observed gradience at the utterance level is due to the lexical variation in projection strength as well as utterance-level factors that modulate projection.

14

- Contemporary presupposition analyses (e.g., Heim 1983, van der Sandt 1992, Abrusán 2011, Karttunen 2016, Simons et al. 2017) predict neither the heterogeneity of the naturalness ratings of factive predicates nor their sensitivity to the prior probability of the CC (though see Schlenker 2021 for the latter). These results, which mirror those based on inference ratings, suggest that projection analyses must incorporate more fine-grained distinctions between attitude predicates and the systematic effect of prior probabilities on projection.

3.3 Which measures are suitable to investigate projection?

- And while naturalness ratings in EICs may provide insight into the meanings of attitude predicates, the interpretation of the results depends on the presumed linking function (see, e.g., Sprouse 2018). Instead of the ratings directly reflecting whether the CC is presupposed, the very low mean rating for *be annoyed* in EICs could be due to its CC being not-at-issue while the EIC suggests that it is at-issue; likewise, reading *know* with a contrastive pitch accent in EICs may have resulted in about half of the participants giving a high rating.

- This result suggests that interrogatives are particularly acceptable when the speaker is either ignorant about the CC or the CC has a higher prior probability: to inquire about the CC, or to inquire about the matrix clause.

- We also observe that the mean naturalness rating is lower in the lower prior probability context than in the higher prior probability context, for all of the predicates except for *pretend*. This suggests that the prior probability of the CC modulates the naturalness of the interrogative regardless of whether the embedding predicate is factive, where the observed context-sensitivity is unexpected under standard analyses (e.g., Heim 1983, van der Sandt 1992) or nonfactive (where it is expected). This result mirrors that of Degen and Tonhauser 2021, who found in an inference rating task the strength of projection inferences to the CCs of factive and nonfactive predicates is modulated by the prior probability of the CC. One could, of course, propose to make the global accommodation of presuppositions to not be default but to be sensitive to the prior probability of content. But the fact of the matter is that the context-sensitivity is observed not just for factive but also for nonfactive predicates, thereby motivating a common projection mechanism.

- The talk also discusses that the results do not align with assumptions about hard and soft triggers; Simons 2001, Abusch 2010.

3.4 What does an empirically adequate analysis of projection need to capture?

4 Conclusions

References

Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics & Philosophy* 34:491–535.

Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27:37–80.

Bürkner, Paul-Christian. 2017. brms: An R package for bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1):1–28.

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

de Marneffe, Marie, Mandy Simons, and Judith Tonhauser. 2019. The CommitmentBank: Investigating projection in naturally occurring discourse. *Sinn und Bedeutung* 23:107–124.

Degen, Judith and Judith Tonhauser. 2021. Prior beliefs modulate projection. *Open Mind* 5:5970.

Degen, Judith and Judith Tonhauser. 2022. Are there factive predicates? An empirical investigation. *Language* 98:552–591.

Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* pages 25–50.

Heim, Irene. 1983. On the projection problem for presuppositions. *West Coast Conference on Formal Linguistics* 2:114–125.

Kalomoiros, Alexandros and Florian Schwarz. 2021. Presupposition projection from disjunction is symmetric. *Linguistic Society of America* 6:556–571.

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.

Lenth, Russell V. 2023. *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.8.8.

Mandelkern, Matthew, Jérémy Zehr, Jacopo Romoli, and Florian Schwarz. 2020. We've discovered that projection across conjunction is asymmetric (and it is!). *Linguistics & Philosophy* 43:473–514.

Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.

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

Schlenker, Philippe. 2021. Triggering presuppositions. *Glossa* 6:1–28.

Simons, Mandy. 2001. On the conversational basis of some presuppositions. *Semantics and Linguistics Theory* 11:431–448.

Simons, Mandy, David Beaver, Craige Roberts, and Judith Tonhauser. 2017. The best question: Explaining the projection behavior of factives. *Discourse processes* 54(3):187–206. Publisher: Taylor & Francis.

Smith, E. Allyn and Kathleen Currie Hall. 2011. Projection diversity: Experimental evidence. *2011 ESSLLI Workshop on Projective Content.* pages 156–170.

Sprouse, Jon. 2018. Acceptability judgments and grammaticality, prospects and challenges. In N. Hornstein, H. Lasnik, P. Patel-Grosz, and C. Yang, eds., *Syntactic Structures after 60 Years*, pages 195–223. De Gruyter Mouton.

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.

Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. Toward a taxonomy of projective content. *Language* 89:66–109.

van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.

Xue, Jingyang and Edgar Onea. 2011. Correlation between projective meaning and at-issueness: An empirical study. *2011 ESSLLI Workshop on Projective Content.* pages 171–184.

Supplemental materials

A   Experiment stimuli

The twenty predicates used in the experiment are the same as in Degen and Tonhauser 2021, 2022:

(10) a. Factive: be annoyed, discover, know, reveal, see
      b. Non-factive: acknowledge, admit, announce, be right, confess, confirm, establish, hear, inform, pretend, prove, say, suggest, think

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*. Each clause-embedding predicate was paired with a unique subject proper name. The speaker of the target stimuli was realized by a randomly sampled unique proper name.

The following list shows the 20 clauses that realized the complements of the predicates in the target and filler stimuli, together with their lower and higher probability facts, respectively, that realized the preceding declarative sentences in the filler stimuli.

1. Mary is pregnant. Facts: Mary is a middle school student / Mary is taking a prenatal yoga class
2. Josie went on vacation to France. Facts: Josie doesn't have a passport / Josie loves France
3. Emma studied on Saturday morning. Facts: Emma is in first grade / Emma is in law school
4. Olivia sleeps until noon. Facts: Olivia has two small children / Olivia works the third shift
5. Sophia got a tattoo. Facts: Sophia is a high end fashion model / Sophia is a hipster
6. Mia drank 2 cocktails last night. Facts: Mia is a nun / Mia is a college student
7. Isabella ate a steak on Sunday. Facts: Isabella is a vegetarian / Isabella is from Argentina
8. Emily bought a car yesterday. Facts: Emily never has any money / Emily has been saving for a year
9. Grace visited her sister. Facts: Grace hates her sister / Grace loves her sister
10. Zoe calculated the tip. Facts: Zoe is 5 years old / Zoe is a math major
11. Danny ate the last cupcake. Facts: Danny is a diabetic / Danny loves cake
12. Frank got a cat. Facts: Frank is allergic to cats / Frank has always wanted a pet
13. Jackson ran 10 miles. Facts: Jackson is obese / Jackson is training for a marathon
14. Jayden rented a car. Facts: Jayden doesn't have a driver's license / Jayden's car is in the shop
15. Tony had a drink last night. Facts: Tony has been sober for 20 years / Tony really likes to party with his friends
16. Josh learned to ride a bike yesterday. Facts: Josh is a 75-year old man / Josh is a 5-year old boy
17. Owen shoveled snow last winter. Facts: Owen lives in New Orleans / Owen lives in Chicago
18. Julian dances salsa. Facts: Julian is German / Julian is Cuban
19. Jon walks to work. Facts: Jon lives 10 miles away from work / Jon lives 2 blocks away from work
20. Charley speaks Spanish. Facts: Charley lives in Korea / Charley lives in Mexico

B   Control and practice stimuli

The following list shows the four control stimuli where the interrogative was expected to receive high naturalness ratings in the context of the preceding declarative sentence. The values in parentheses indicate the mean naturalness rating for each of the four controls. As shown, the third control did not receive the

expected high naturalness mean ratings (probably because participants were unwilling to accommodate that Hendrick has a car in a context in which Hendrick was looking to buy a car). This control was therefore not used to exclude participants' data.

1. I don't know if Samantha has a new hat. Does Samantha have a new hat? (.89)
2. I don't know if this pizza has mushrooms on it. Does this pizza have mushrooms on it? (.87)
3. Hendrick was looking to buy a car. Was Hendrick's car expensive? (.5)
4. Mary visited her aunt yesterday. Is Mary's aunt sick? (.91)

The following list shows the four practice stimuli in the order in which they were presented to the participants. Participants were able to advance to the experiment only if they gave a naturalness rating higher than .6 for the first and third stimulus, and a naturalness rating lower than .4 for the second and fourth stimulus.

1. I have no idea where Natalie is from. Is Natalie from the USA?
2. I don't have any sisters. Have you met my sister yet?
3. I am going on vacation to Ireland. Does Fritz realize that Joe is going with me?
4. I have no idea if Anna has any dogs. Is Samuel glad that Anna fed her dogs?

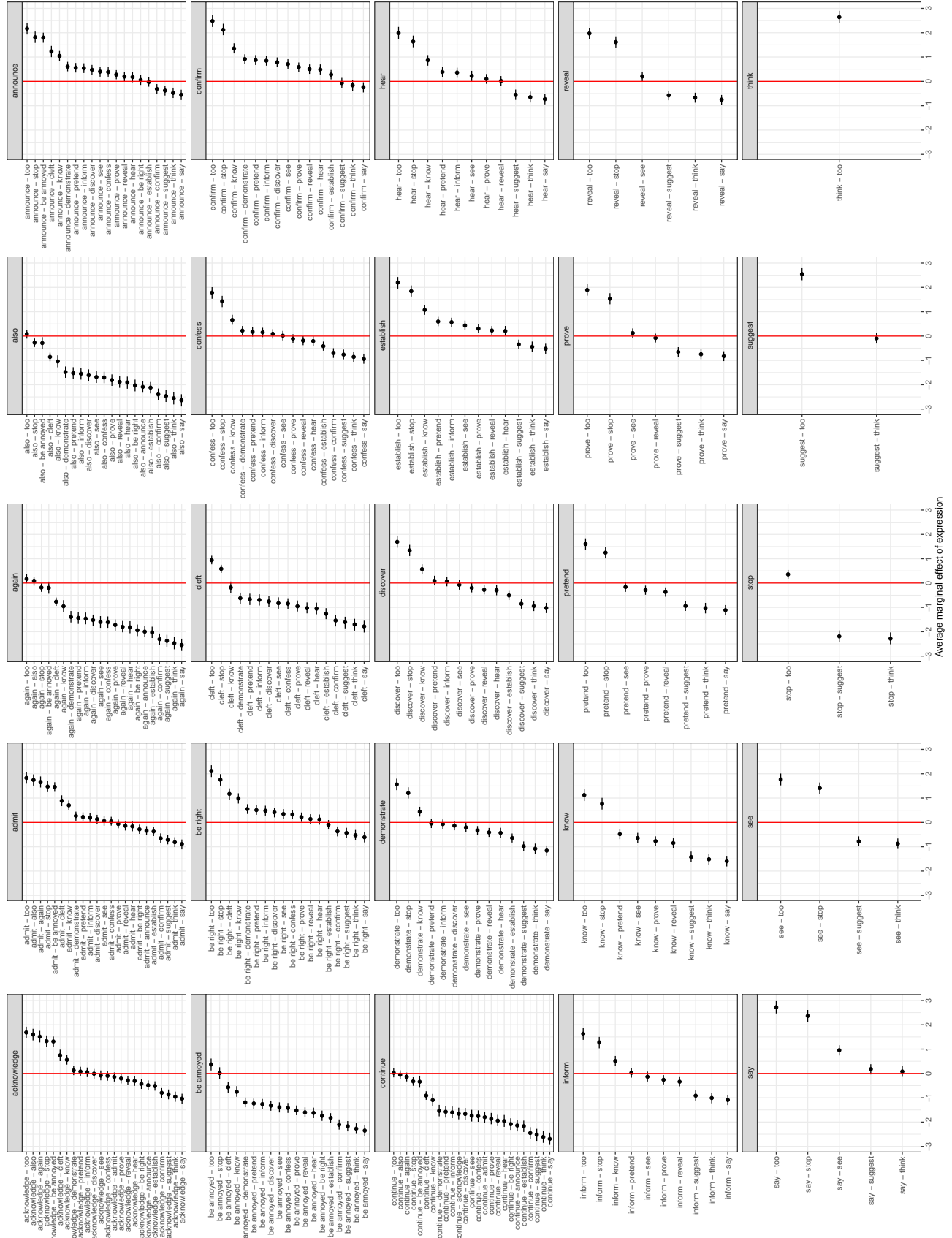# C   Model output: Pairwise comparison of expressions

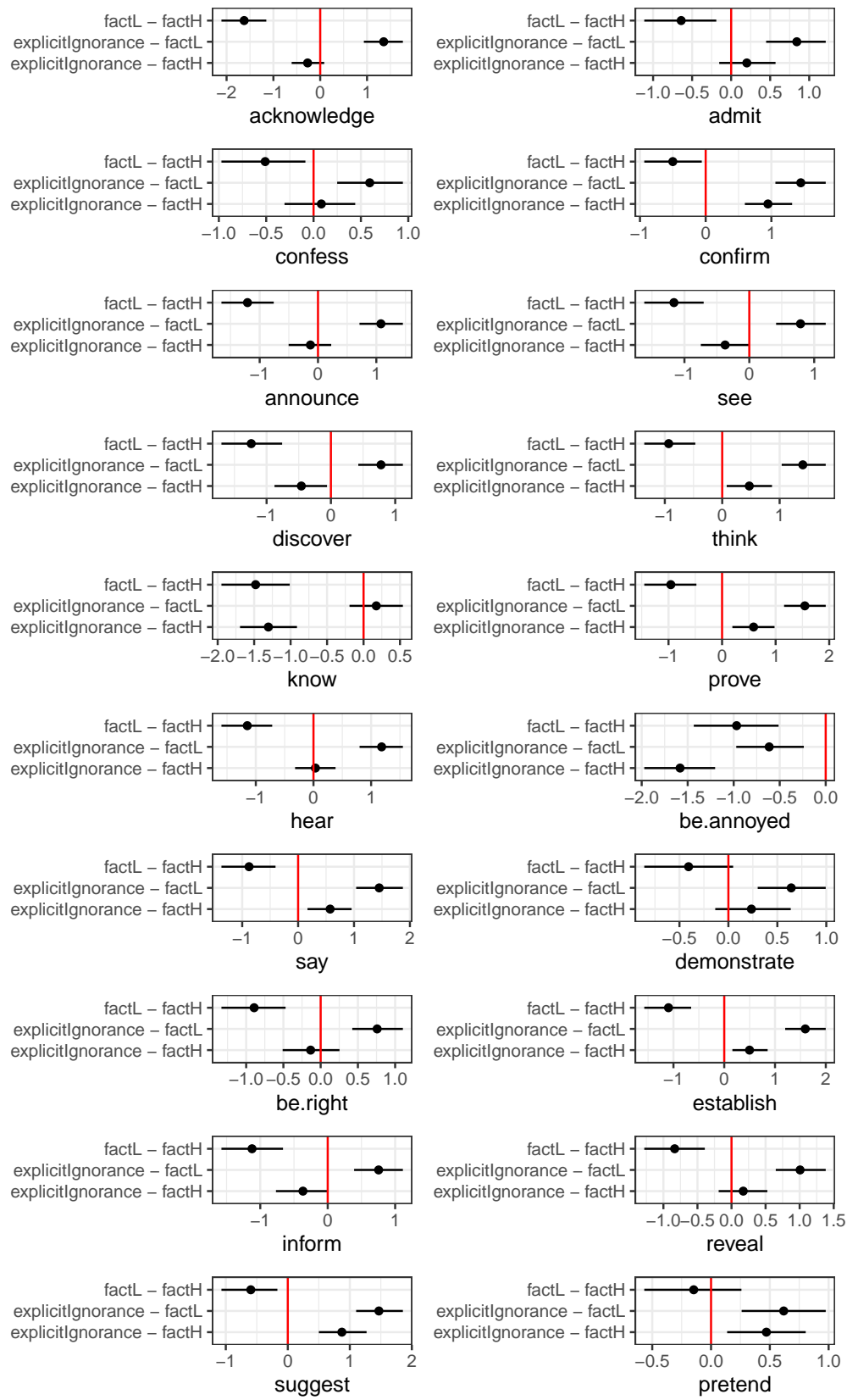# D   Model output: Pairwise comparison of contexts

Figure A1: .

Figure A2: .