The prosody of presupposition projection in naturally-occurring utterances[1]

Taylor MAHLER — *The Ohio State University*
Marie-Catherine de MARNEFFE — *The Ohio State University*
Catherine LAI — *The University of Edinburgh*

Abstract. In experimental studies, prosodically-marked pragmatic focus has been found to influence the projection of factive presuppositions of utterances like *these parents didn't know the kid was gone* (Cummins and Rohde, 2015; Tonhauser, 2016; Djärv and Bacovcin, 2017), supporting question-based analyses of projection (i.a., Abrusán, 2011; Abrusán, 2016; Simons et al., 2017; Beaver et al., 2017). However, no prior work has explored whether this effect extends to naturally-occurring utterances. In a large set of naturally-occurring utterances, we find that prosodically-marked focus influences projection in utterances with factive embedding predicates, but not those with non-factive predicates. We argue that our findings support an account where lexical semantics of the predicate contributes to projection to the extent that they admit QUD alternatives that can be assumed to entail the content of the complement.

Keywords: projective content, attitude predicates, (non-)factive predicates, prosody.

1. Introduction

Understanding what speakers mean requires listeners to determine which propositions conveyed by the speakers' utterance are intended as speaker commitments. In some cases, this task is relatively simple. For example, when a speaker asserts that a proposition is true, they convey their commitment to its truth. A speaker who utters (1) asserts the proposition "*Sinn und Bedeutung* is an excellent conference" and as a result, they commit themselves to its truth.

(1)   *Sinn und Bedeutung* is an excellent conference.

But assertion is not the only means by which speakers convey their propositional commitments. In (2), the speaker has not asserted that the content of the complement (henceforth CC) of *know* is true, and yet they seem to be committed to its truth.

(2)   John knows that *Sinn und Bedeutung* is an excellent conference.

The observation that the speaker is committed the truth of the CC of *know* is rather unsurprising, given that the CC of *know* is entailed by the (asserted) main clause proposition, and speakers are typically taken to be committed to the entailments of propositions that they assert. But the example in (3) demonstrates that speakers can appear committed to non-entailed propositions:

(3)   a.   John doesn't know that *Sinn und Bedeutung (SuB)* is an excellent conference.
        b.   Does John know that *SuB* is an excellent conference?
        c.   Perhaps John knows that *SuB* is an excellent conference.
        d.   If John knows that *SuB* is an excellent conference, he will submit an abstract.

In (3), the complement of *know* is embedded under four different entailment-cancelling operators. While the CC is not entailed by the main clause proposition, the speaker can never-

theless be taken to be committed to the content expressed by the complement. Content that displays such behavior is said to *project* (e.g., Karttunen, 1971; Heim, 1983; Chierchia and McConnell-Ginet, 1990). The observation that the speaker can be taken to be committed to the CC requires an explanation that does not hinge on entailment. The classical explanation is that the speaker's commitment to the CC arises from another property of the CC: that it is *presup-posed*. According to standard accounts, the presupposition of the CC arises due to the speaker's use of *know*, which is assumed to lexically encode its complement as a presupposition (e.g., Heim, 1983; van der Sandt, 1992). Such predicates, called *factives*, have been distinguished from *non-factive* clause-embedding predicates like *think* and *believe* which are not assumed to lexically-encode their complements as presuppositions (Kiparsky and Kiparsky, 1970).

However, the CCs of factive predicates do not always project. As Beaver (2010) pointed out based on examples like (4), projection (in English) is sensitive to prosodically-marked focus. When a constituent within the complement receives narrow focus, as *plagiarized* in (4a), the CC (that the work is plagiarized) seems not to project. However, in the absence of narrow focus within the complement as in (4b), the CC seems to project (capitals indicate narrow focus):

(4) a. If the T.A. discovers that your work is PLAGIARIZED, I will be forced to notify the dean.
     b. If the T.A. DISCOVERS that your work is plagiarized, I will be forced to notify the dean.

These observations are predicted by accounts that assume that presupposition projection arises from the set of alternatives evoked by the focal structure of the utterance (e.g., Abrusán, 2011; Abrusán, 2016; Simons et al., 2017; Beaver et al., 2017). Following Rooth (1992), these accounts assume that the interpretation of focus contributes a set of alternative propositions. For example, the interpretation of focus in (4a) introduces the alternatives in (5a), whereas the interpretation of focus in (4b) introduces the alternatives in (5b):

(5) a. $\{q:$ Your work is $P \mid P$ is a property$\}$
     b. $\{q:$ The T.A. R that your work is plagiarized $\mid R$ is a cognitive attitude$\}$

On most theories of questions, the semantic value of a question is the set of propositions that are possible answers to it (e.g., Karttunen, 1971; Rooth, 1992). Hence, questions and focal alternatives are semantically equivalent. An important related observation is that the focus-marking of an utterance must be congruent with the question it addressess (e.g., Roberts, 2012). For instance, an utterance with the focus marking in (4a) would be a felicitous response to a question about the status of the student's work, e.g., *What property does your work have?* This question is equivalent to the set of alternatives in (5a). But such an utterance would be an infelicitous response to a question about the T.A.'s attitude with respect to the complement proposition, e.g., *What cognitive attitude does the T.A. have to the proposition that your work is plagiarized?* This question is equivalent to the set of alternatives in (5b). On the other hand, an utterance with the focus marking in (4b) would be a felicitous response to the question represented in (5b), but not to the question represented in (5a).

Beaver et al. (2017) formalize the relation between focus, questions, and CC projection in sentences with factive predicates in terms of Roberts' (2012) notion of the Question Under Discussion (QUD). On their account, the focal structure of the utterance provides a set of alter-

native propositions, as outlined above, and a contextually-restricted subset of these alternatives corresponds to the QUD. In other words, focus-marking does not fully determine the set of propositional alternatives: in the context of the utterance, some alternatives will be ruled out. For example, the alternative set in (5a) introduced by the focus-marking in (4a) presumably does not include the alternative *The T.A. discovers that your work is publishable*, as there would be no reason that the professor would be "forced to notify the dean" were this alternative to be true. The QUD for (4a) is not simply (5a), but rather the set of propositions that are actually under consideration, that is propositions in which P is instantiated by properties such as e.g., *plagiarized, fabricated, unsatisfactory*, and other properties that would warrant a report to the dean. Beaver et al. (2017) hypothesize that CC projection is predicted by the CC's relation to the QUD. Specifically, they hypothesize that the CC will project when the QUD alternatives entail the CC, i.e., when each alternative in the QUD entails that the CC is true. Such content is characterized as "not-at-issue", while content that is not entailed by the QUD is "at-issue". This hypothesis is formulated in slightly different terms across publications. We use the version of Tonhauser et al. (2018) who recast the hypothesis as the Gradient Projection Principle (GPP), characterizing projection as a gradient rather than a binary property of content:

(6)   **Gradient Projection Principle (GPP):** If content C is expressed by a constituent embedded under an entailment-cancelling operator, then C projects to the extent that it is not at-issue.

In utterances with clausal complements, the GPP predicts that the CC projects less when a constituent within the complement is prosodically focused. For utterances in which complements are embedded under factive predicates, this prediction has been borne out in several experimental studies using laboratory speech (e.g., Cummins and Rohde, 2015; Tonhauser, 2016; Djärv and Bacovcin, 2017). Though work on QUD-based approaches to CC projection primarily investigates projection in utterances with factive predicates, Djärv and Bacovcin (2017) also explored prosodically-marked focus and projection in utterances with non-factive predicates. Surprisingly, they found that the CCs of non-factives actually project *more* when a complement constituent is prosodically focused. In addition, they found that the influence of prosodically-marked focus on the projection of both factive and non-factive CCs was small relative to the influence of factivity. Based on these findings, they argue that projection is constrained by the lexical semantics of clause-embedding predicates, with factives contributing inferences that can conflict with inferences based on the prosodically-marked QUD.

Given the minimal research on focus and non-factive CC projection as well as Djärv and Bacovcin's (2017) suggestive findings that focus interacts differently with factives and non-factives, it remains an open question how the lexical semantics of clause-embedding predicates interact with prosodically-evoked QUDs to influence projection. We shed light on this question by investigating projection in naturally-occurring utterances with both factive and non-factive predicates. Consistent with prior experimental work on laboratory speech and the predictions of the GPP, we find that factive CCs project more when the prosodically prominent constituent is within the matrix clause compared to the complement. However, for utterances with non-factive predicates, prosodically-marked focus did not influence projection.

We outline a preliminary analysis to account for our findings. We start from the assumption that there are two relevant strategies for backgrounding information: expressing the content as

the complement of a clause-embedding predicate, and using prosodically-marked information structure. However, content that is backgrounded by either of these strategies does not automatically become not-at-issue. We suggest that whether the CC is not-at-issue and potentially projective depends on the interaction between the lexical semantics of the embedding predicate used to background the CC together with information-structural backgrounding. The CCs of factives predicates are (by default) associated with a QUD in which each alternative proposition, including the CC itself, is true. However, the existing conversational background and the information structure of the utterance (via accenting of the complement) can introduce potentially false propositions into the QUD, thereby making the CC at-issue. Otherwise, the QUD alternatives under consideration are all true, and so the CC is not-at-issue. On the other hand, when the embedding predicate is non-factive, the associated QUD obligatory contains at least one alternative in which the CC is false. Even when the CC is backgrounded by virue of deaccenting the complement, it remains at-issue. The lexical semantics of the embedding predicate contributes to the perception of projection to the extent that they admit QUD alternatives that can be assumed to entail the CC. We sketch the details of this account in section 5.

2. Experimental investigations of prosody and projection with laboratory speech

To the best of our knowledge, prior comprehension experiments investigating prosodically-marked focus and CC projection have exclusively used laboratory speech. The stimuli were recorded by the experimenters, and their prosodic properties were manipulated either by the speaker of the utterance or by manipulating the recordings. Across these experiments, participants listened to the manipulated utterances and indicated the extent to which the CC projects using rating tasks that differed slightly. The stimuli in these experiments varied slightly too, but all used a third person subject for the matrix verb. Cummins and Rohde (2015) explored projective content associated with a variety of standard presupposition triggers, restricting their investigation of CC projection to the CCs of factive predicates within the scope of negation. Tonhauser's (2016) stimuli were utterances with factive predicates embedded within the scope of the epistemic modal *perhaps*. Djärv and Bacovcin (2017) used stimuli with both factive and non-factive predicates under the epistemic modal *might*. In this section, we review these three studies with the goal of highlighting the extent to which their findings are predicted by the GPP.

2.1. Cummins & Rohde 2015

Cummins and Rohde (2015) investigated contents associated with standard presupposition triggers, including factive predicates. Negation was used as the entailment-cancelling operator for all target sentences, as illustrated in the target sentence in (7).

(7)   Bill doesn't regret arguing with his boss.

Participants heard utterances of sentences like (7) in either a "focus condition" in which the final word of the sentence carried a pitch accent (*Bill doesn't regret arguing with his BOSS*), or a "neutral condition" (no details are provided about the prosodic properties of the neutral condition). The focus condition evokes a QUD roughly equivalent to *With whom does Bill regret arguing?* As this QUD does not entail the CC of *regret*, i.e., *Bill argued with his boss*, the CC is considered at-issue and the GPP therefore predicts that it should not project. In contrast, the QUD associated with the neutral condition (though not specifically identified in their paper) concerns the truth of the main clause proposition: *Does Bill regret arguing with*

his boss? As this QUD does not entail the CC, the CC is considered to be not-at-issue and predicted by the GPP to project.

After listening to a target sentence, participants responded to a question about the likelihood of the CC, on a 7-point Likert scale from 1/"unlikely" to 7/"likely." For (7), the question was *How likely is it that Bill argued with his boss?* Higher responses on the scale were taken to indicate that the CC was more likely to project. Participants provided higher ratings in the neutral condition than in the focus condition, suggesting that whether content addresses the prosodically-evoked QUD influences projection. Thus, these findings are consistent with the GPP. However, it is worth noting that the decrease in projectivity in the focus condition varied greatly depending on which presupposition trigger was used. In fact, in some cases, the rating in the focus condition was higher than that of the neutral condition (i.e., *forget, manage, stop, return*). Moreover, as the authors note, this study did not control for specific prosodic properties of the utterances: the speaker recording the target sentences attempted to communicate a particular meaning, not produce a particular prosodic contour. In particular, no details are provided about the prosodic properties of the neutral condition. Although we can assume this condition lacks narrow focus in the complement, it is not clear whether other prosodic effects could be playing a role here, e.g., the level of prominence on the matrix verb. This leaves open several questions about the role of prosody in these findings.

2.2. Tonhauser 2016

Building on Cummins and Rohde's (2015) findings, Tonhauser (2016) conducted several experiments in which the specific prosodic properties of the stimuli were carefully controlled. A native English speaker trained in the Tones and Break Indices (ToBI) annotation system (Beckman and Ayers, 1997) recorded the utterances. The target sentences were produced in three different prosodic conditions, one with a H* pitch accent on the predicate, one with a L+H* pitch accent on the final content word of the complement, and one with a L+H* pitch accent on the subject pronoun of the complement. The investigation was constrained to the CC of cognitive factive predicates under the epistemic modal *perhaps* as in (8).

(8) a. Perhaps he DISCOVERED that she's a widow.   (H*-on-predicate)
     b. Perhaps he discovered that she's a WIDOW.   (L+H*-on-content)
     c. Perhaps he discovered that SHE'S a widow.   (L+H*-on-pronoun)

Participants listened to an utterance and answered a question about the speaker's certainty with respect to the CC on a 7-point Likert scale from 1/"No, not certain" to 7/"Yes, certain". For example, for (8), participants were told that the speaker, Dana, was speaking about two individuals named Scott and Valeria, and was then asked *Is Dana certain that Valeria is a widow*?

Consistent with Cummins and Rohde's (2015) findings, participants provided lower ratings to complement accented compared to matrix accented utterances. Tonhauser (2016) interprets this finding as support for QUD-based accounts of projection. In the L+H*-on-content and L+H*-on-pronoun conditions, the CC addresses the QUD prosodically-evoked by the utterance and hence is at-issue. For example, in the L+H*-on-pronoun condition (8c), the CC addresses the induced QUD *Who did Scott discover is a widow?* In contrast, the QUD evoked by the H*-on-predicate condition (8c) is (assumed to be) about the matrix subject's attitude with respect to the CC, e.g., *What's the Scott's cognitive attitude with respect to the proposition that Valeria*

is a widow? The proposition *Valeria is a widow* does not address this QUD, and is hence not-at-issue. In sum, these results are consistent with the predictions of the GPP: the CC is more likely to project when it is not-at-issue (the H*-on-predicate condition) compared to when it is at-issue (the L+H*-on-content and L+H*-on-pronoun conditions).

A further experiment reported in Tonhauser (2016) investigated whether different pitch accents influence the projection of the CC. Pitch accents were placed on both the predicate and the final content word in each condition. More specifically, this experiment looked at the two conditions shown in (9), where the (L+)H* pitch accent is considered more prominent than H*. The experimental procedure was identical to the single pitch accent experiments.

(9)   a.   Perhaps he DISCOVERED that she's a WIDOW.
                        H*   (L+)H*
     b.   Perhaps he DISCOVERED that she's a WIDOW.
                    (L+)H*   !H

Tonhauser (2016) reports that participants rated the speaker as less certain about the CC when the pitch accent on the last content word was more prominent (9a), compared to the opposite case in (9b). In light of the complexity of the prosodic contours of these stimuli, Tonhauser (2016) points out that they might be consistent with more than one information-structural analysis. She analyzes the more prosodically prominent constituent as corresponding to the focus of the utterance. Thus, the focus of (9a) is the complement verb phrase *is a widow*, while the predicate *discover* is the focus of (9b). Hence, the QUD associated with (9b) is analyzed in the same way as for (8a): the QUD evokes alternatives about the subject's attitude toward the complement proposition. Since the CC is entailed by this QUD, it is not-at-issue. In contrast, the QUD associated with (9a) is about properties of Valeria, e.g., *What property did Scott discover about Valeria?* The CC addresses this QUD, and is therefore at-issue.

Again, as predicted by the GPP, these results indicate that the CC is more likely to project when it is not-at-issue as signalled by focus ((L+)H*) on the matrix predicate (9b) compared to when it is at-issue as signalled by focus ((L+)H*) in the complement (9a). Moreover, these findings suggest that listeners attend to information structure as signalled by relative prominence. Interestingly, Tonhauser (2016) notes that the secondary prominences in (9) partially fit the description of a contrastive topic (Büring, 2003), but argues that this is an unlikely interpretation for this data as both pitch accents are contained in a single intonational phrase. However, it is not clear that this is enough to rule out a contrastive interpretation (cf. Calhoun, 2010). In any case, we would still expect contrastive topic/focus configurations to be possible on these types of sentences. Assuming that contrastive topics generate alternatives in a similar way to foci, this could greatly impact the QUD and hence what projects for utterances collected in the wild.

2.3. Djärv & Bacovcin 2017

Whereas Cummins and Rohde's (2015) and Tonhauser's (2016) investigation of CC projection was limited to the CCs of factive predicates, Djärv and Bacovcin (2017) also investigated the influence of prosody on CC projection in utterances with non-factive predicates. As illustrated in (10), the epistemic modal *might* was used as the entailment-cancelling operator.

(10)   John might've discovered/believed that Anna left town.

For target sentences with embedding predicates, the subject of the complement was focused in one condition (11a), and the predicate was focused in the other condition (11b). For the unembedded target sentences, the matrix subject was focused in one condition (11c). It is not clear from the description of the unembedded stimuli whether and which constituent was focused for the other condition, but presumably none were narrowly focused (11d).

(11)   a.   John might've discovered that ANNA left town.
        b.   John might've DISCOVERED that Anna left town.
        c.   ANNA left town.
        d.   Anna left town.

The rating task was identical to the one in Tonhauser's (2016) study. For utterances with factive predicates, participants' ratings followed the pattern of results from Cummins and Rohde's (2015) and Tonhauser's (2016) experiments: participants provided higher certainty ratings when the predicate was focused than when a constituent within the complement was focused. However, for a subset of utterances with (verbal) non-factive predicates, the effect of focus was actually reversed: participants provided higher certainty ratings when a constituent within the *complement* was focused compared to when the predicate was focused. Djärv and Bacovcin (2017) also found that participants' ratings were influenced by the factivity of the predicate: the CCs of factive predicates were associated with higher ratings than the CCs of non-factive predicates. Crucial to their analysis is their observation that the effect of factivity was substantially greater than the effect of focus. They interpret this result as evidence against "strong versions of pragmatic accounts of projection" and specifically "the claim that such causal pragmatic effects are driven by prosodic signals of the QUD" (p.128).

This leads them to argue for an account of projective meaning in terms of interacting lexical constraints and prosodically-signalled pragmatic information about the QUD. They assume that factive predicates encode their complements as presuppositions, giving rise to the inference that the CC is in the common ground. For example, the use of *discover* in (12) is associated with the "factive" inference in (12a) that *Anna left town* is in the common ground. Additional inferences are derived from the QUD indicated by the prosody of the utterance. When the complement subject is focused as in (12a), the QUD is about the identity of the person who left town, i.e., (12c). This QUD generates the two additional inferences (i) that it is common ground that someone left town, and (ii) that the identity of this individual is not common ground.

(12)   John might've discovered/believed that ANNA left town.
        a.   inference from factive *discover*: That Anna left town is common ground.
        b.   inference from non-factive *believe*: nothing.
        c.   inferences from QUD 'Who left town?':
              (i)   That someone left town is common ground.
              (ii)   That the identity of the person who left town is not common ground.

The latter of these two inferences, i.e., (12c-ii), conflicts with the factive inference (12a) from *discover*. But no conflict with QUD inferences arises with non-factive predicates like *believe*, since such predicates do not give rise to an inference about whether the CC is in the common ground. The QUD-based inferences associated with focus on the predicate also do not conflict with the factive inferences. With focus on the predicate as in (13), the QUD is about the matrix subject's attitude toward the CC, i.e., (13c). This QUD is irrelevant to whether the CC is

common ground, so its associated inference (13c-i) does not conflict with the factive inference.

(13)   John might've DISCOVERED/BELIEVED that Anna left town.
   a.   inference from factive *discover*: That Anna left town is common ground.
   b.   inference from non-factive *believe*: nothing.
   c.   from QUD 'What cognitive attitude does John have (or might have) to the proposition that Anna left town?':
      (i)   John has some cognitive attitude to the proposition that Anna left town.

Djärv and Bacovcin (2017) suggest that the results of their experiment reflect how the participants' synthesized these various inferences. The increased ratings for utterances with factives compared to non-factives, regardless of prosodic realization, suggest that the factive inference carries more weight than inferences based on prosodically-evoked QUDs. When the QUD is irrelevant to the CC, as when the predicate is focused, participants only attend to the factivity of the predicate in evaluating the speaker's commitment to the CC. When one of the QUD-based inferences in (ii) conflicts with the factive inference, participants reduce their rating to reflect diminished confidence that the factive inference holds.

For the utterances with non-factives, Djärv and Bacovcin (2017) argue that the increased ratings associated with complement subject focus are due to the QUD-based inference in (12c-i). They suggest that the inference that someone left town makes the CC more probable. Once it is known that someone left town, it becomes more probable that Anna left town, and this higher probability is reflected in increased certainty ratings. In contrast, the inferences associated with matrix focus are completely independent of whether the CC is true and so they do not affect participants' certainty ratings. However, Djärv and Bacovcin (2017) do not specifically test whether this existential inference (12c-i) associated with the complement subject is actually available to listeners, and whether this can affect the probability of a particular alternative being true separate from other factors. It could be that the a priori probability that Anna left town is extremely low because, for example, it is known that she is cursed to die if she does, or simply that the listener believes the subject (*John* in (12)) to be unreliable.

In sum, the results of these experimental studies provide evidence for QUD-based analyses of CC projection, at least when the complement is embedded by a factive predicate. However, outside of impressionistic observations (e.g., Beaver, 2010; Simons et al., 2017), the relation between prosody and factive presupposition projection has not been investigated in naturally-occurring discourse. In the following, we extend the evidence base for this relation by collecting annotations of naturally-occurring utterances with clausal complements that target speaker intuitions about projection and focus.
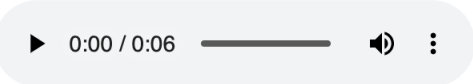
3. Data

The discourses selected for annotation come from the CommitmentBank (de Marneffe et al., 2019), which contains 1,200 discourses from 3 corpora: the British National Corpus, the Wall Street Journal, and the Switchboard Corpus of spontaneous telephone dialogues. As illustrated in (14), each discourse contains a target sentence featuring a clause-embedding predicate and a clausal complement, embedded under an entailment-cancelling operator (negation, modal, antecedent of conditional, or question), preceded by up to two context sentences. Of the 464 Switchboard discourses in CommitmentBank, we extracted the audio files corresponding to

392 discourses. Each discourse was annotated for projection, and the target sentences from 350 discourses were annotated for prominence in a separate annotation task. We restrict our analysis to the 350 discourses that were annotated for both projection and prominence.[2]

(14)   A: I work in the airline, marketing group.
         B: Oh.
         A: So, we do a lot of,
         B: I didn't even know they had anything like that.   (SWBD-270)

3.1. Projection annotations

To determine whether interpreters judge the speaker of the utterance to be committed to the CC, we asked participants to listen to the discourses and rate how certain the speaker is that the CC is true. For each discourse segment, judgments were elicited from at least eight self-reported native English speakers, using a questionnaire on Amazon's Mechanical Turk Platform.[3] The 392 discourse segments were distributed in groups of 8 across 49 questionnaires. As illustrated in the sample trial in Figure 1, participants were instructed to listen to the discourse segment and answer two questions. For the first question, the completion question, they were presented with an elided version of the target sentence that they had just listened to. They were then asked to indicate which of two minimally different utterances the speaker had actually produced: one utterance corresponded to the utterance on the recording, and one was subtly different from what the speaker had said. For example, the speaker in Figure 1 had actually said "I didn't even know they had anything like that", so the second of the two utterances was the correct option. The incorrect utterances were constructed in such a way that they were compatible with the elided version of the target sentence, i.e., participants could not tell which was the correct or incorrect utterance from looking at the elided sentence. Incorrect responses to this question were taken to indicate that the participant had not listened to the entire audio clip or not paid attention.

Listen to the whole snippet: ▶ 0:00 / 0:06 ———————— 🔊 ⋮

Victoria just said: "I didn't... had anything like that."

What did Victoria actually say? (You are allowed to listen again.)

○ "I didn't know the airline had anything like that."
○ "I didn't even know they had anything like that."

Now given what Victoria said, tell us how certain Victoria is that the airline had anything like that.

○   ○   ○   ○   ○   ○   ○
Victoria is certain   Victoria is not certain   Victoria is certain
that it is true   whether it is true or false   that it is false

Figure 1: Sample trial in the projection annotation task.

The second question probed participants' intuition with respect to the projection of the CC: participants were asked to indicate how certain the speaker is with respect to the CC, given what had been said. Participants responded on a 7-point Likert scale labeled at three points 3/"[The speaker] is certain that it is true, 0/"[The speaker] is not certain whether it is true or false", -3/"[The speaker] is certain that it is false.

In addition to the 8 discourse segments from the CommitmentBank, each questionnaire included 2 constructed "filler" discourse segments to ensure that participants were paying attention. For half of the fillers, participants were expected to rate the speaker as highly certain about the truth of the content in question, as in (15a); for the other half, participants were expected to rate the speaker as highly certain that the content was false, as in (15b).

(15)   a.   Megan: "so I love, you know, baseball teams. In fact, Nolan Ryan was on the news tonight."
Now given what Megan said, tell us how certain Megan is that Nolan Ryan was on the news tonight.

        b.   Judy: "But even after spending so much time studying, Olivia didn't manage to pass the test."
Now given what Judy said, tell us how certain Judy is that Olivia passed the test.

The annotation task was completed by 288 unique participants (some participants completed multiple questionnaires). We removed data from participants who responded incorrectly to one or both fillers or ≥ 2 completion questions (N=132) and participants who did not report American English as their native language (N=2). The final dataset included annotations from 154 unique participants.

3.2. Prominence annotations

Prominence annotations were collected using the Rapid Prosody Transcription methodology (RPT; e.g., Cole et al. 2017), presented via the Language Markup and Experiment Design software (LMEDS; Mahrt 2018). RPT was performed for the target sentences of 350 discourses that were annotated in the projection annotation task.[4] Participants listened to the utterance of the target sentence, saw its corresponding transcript, and were instructed to select the word that they heard as most prominent. For each utterance, we elicited such judgments from at least six self-reported native English speakers using a questionnaire linked from Amazon's Mechanical Turk Platform.[5] The 350 utterances were distributed in groups of 14 across 25 questionnaires. As illustrated in the sample trial in Figure 2, participants were instructed to listen to the audio recording and click on the word in the transcript that they heard as most prominent. The questionnaire did not allow participants to advance without both listening to the recording and selecting a word.

In addition to the 14 target sentence utterances, two constructed "filler" utterances were included within each questionnaire to ensure that all participants were paying attention. In each filler utterance, one word was produced with an exaggerated pitch accent. Participants were

---

[4] The 42 discourses used in the projection annotation task but not RPT were ones in which the transcript and audio recording differed substantially from each other.

[5] The annotations were performed with IRB approval; distribution was restricted to IP addresses in the United States. Annotators were paid $0.85 for completing the questionnaire.

Mark exactly one prominence.

Play

I didn't even know they had anything like that.

Continue

Figure 2: Sample trial in the prominence annotation task.

expected to select this word as the most prominent. The annotation task was completed by 190 unique participants (some participants completed multiple questionnaires). We excluded data from participants who responded incorrectly to one or both fillers (N=36) or did not report American English as their native language (N=5). The final dataset included annotations from 162 unique participants.

The annotations were classified into three categories. Items for which more than 65% of participants identified the prominent word in the matrix clause were labeled "matrix". Items for which more than 70% of participants identified the prominent word in the complement clause were labeled "complement".[6] Items that met neither of these criteria were labeled as having "broad" focus, i.e., the entire sentence was assumed to be in focus. The distribution of labels by the factivity of the embedding predicate is shown in Table 1.[7] Examples of the three categories are given in (16)-(18), with the percentage of annotators who selected each word indicated below it (words that were not selected by any annotators are left blank). The information in parentheses following each item includes the item ID in the CommitmentBank, the total number of annotators for that item in the prominence annotation task, and the mean speaker certainty rating from the projection annotation task.

	Matrix	Complement	Broad	Total
Factive	12	50	18	80
Non-factive	27	181	62	270
Total	39	231	80	350

Table 1: Distribution of prominence annotations by predicate factivity.

---

[6]We increased the threshold for labeling an item "complement" relative to the "matrix" threshold because the complement clause is almost always longer than the matrix clause.

[7]The factive predicates were *see, find, know, realize, bother, recognize, understand* and *notice*. The non-factive predicates were *believe, bet, convince, feel, foresee, guarantee, guess, hear, hypothesize, imagine, mean, occur, say, seem, swear, take, tell* and *think*.

(16) Matrix Focus
    a.   if they know some practice is wrong
               100   (SWBD-004, N=9, μ=1.63)
    b.   I   didn't think that was right
       28.6 14.3   28.6 28.6   (SWBD-263, N=7, μ=-0.63)

(17) Complement Focus
    a.   I think it went to Lockheed, didn't it?
               100   (SWBD-428, N=9, μ=1.25)
    b.   Do you feel that yours is invaded?
       11.1   44.4   44.4   (SWBD-417, N=9, μ=-0.33)

(18) Broad Focus
    a.   I   don't think we   did it
       14.3 14.3 14   28.6 28.6   (SWBD-348, N=7, μ=-2.33)

Prior research on projection of CCs has mostly been restricted to utterances in which either the complement or the matrix clause has a single narrowly focused constituent. However, the prominence annotations suggest that these sorts of utterances are relatively infrequent. In our data, annotators rarely identified the same word as prominent: for only 26% (92) of the utterances more than 70% of the annotators agreed on the prominent word. Moreover, 23% of the utterances were categorized as exhibiting "broad" focus.

4. Results

Participants' mean certainty ratings by predicate factivity and collapsed prominence category are shown in Figure 3. The data were modeled using a mixed-effects linear regression predicting participants' certainty ratings from predicate factivity (factive vs. non-factive), focus (matrix vs. complement vs. broad), and their interaction. The random effects structure included by-participant and by-item random intercepts. A log-likelihood comparison between this model and a model without the interaction fixed effect between focus location and predicate factivity revealed that the interaction was significant ($\chi^2(2) = 8.66$, $p < 0.02$), such that focus only had an effect on the CCs of factive predicates. The CCs of factive predicates were less projective with complement focus than with matrix focus ($\beta = -1.52$, $SE = .46$, $t = -.3.28$), and those with complement focus were less projective than those with broad focus ($\beta = -.83$, $SE = .39$, $t = -2.11$). Factive CCs with broad focus and factive CCs with matrix focus were equally projective ($\beta = -0.68$, $SE = .53$, $t = -1.27$).

For the purposes of evaluating the predictions of QUD-based accounts of projection and to compare with prior experimental research on laboratory speech, we restrict our discussion to the items in the categories of "matrix" and "complement". However, we note that a full picture of the relation between prosody and projection behavior of CCs will involve investigating these other sorts of prosodic patterns, in particular distinguishing between cases in which a single constituent is prominent vs. multiple constituents within a clause as well as instances in which the entire utterance is focused, i.e., our "broad" focus cases.
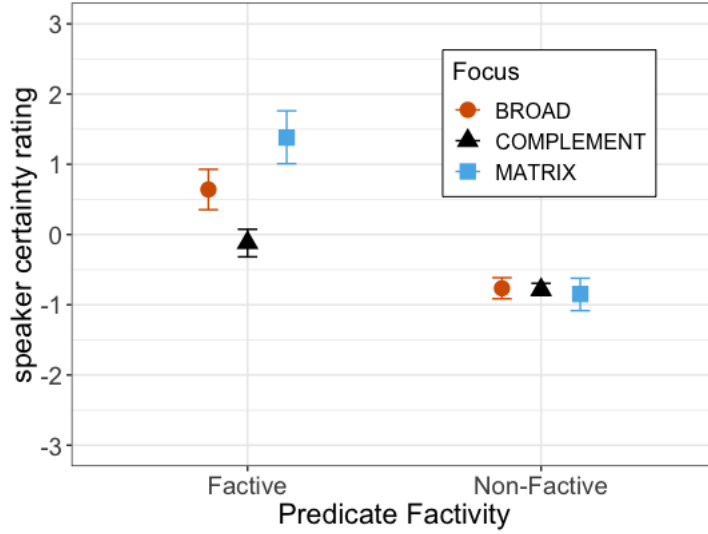
Figure 3: Mean speaker certainty ratings by predicate factivity and focus (with 95% confidence intervals). The y-axis values correspond to the 7-point Likert scale labeled at 3/"The speaker is certain that it is true", 0/"The speaker is not certain whether it is true or false", and -3/"The speaker is certain that it is false".

5. Discussion

Our two main findings are (i) that the CCs of factive predicates are less projective when a complement constituent is focused compared to when a matrix constituent is focused, and (ii) prosodically-marked focus does not influence the projection of non-factive CCs. In this section, we discuss the extent to which these findings are consistent with prior findings on laboratory speech. We then outline a preliminary analysis of our findings that reframes insights from QUD-based accounts of projection to capture the behavior of non-factive CCs.

For factive CC projection, our findings are consistent with those from prior laboratory speech studies, confirming that speakers and hearers do attend to information structure in drawing factive inferences: factive CCs are more projective when the matrix is focused compared to when the complement is focused. This finding is predicted by existing QUD-based analyses of factive presupposition projection (Beaver et al., 2017; Simons et al., 2017).

Our findings for non-factives, however, are not consistent with findings from the single lab-speech based experiment that included non-factive CCs. As in Djärv and Bacovcin (2017), we found that prosodically-marked information structure affects non-factive CC and factive CC projection differently. However, whereas Djärv and Bacovcin (2017) found that non-factive CCs were *more* projective with complement than matrix focus, we found that information structure had no effect on non-factive CC projection. One difference between our data and the stimuli used in Djärv and Bacovcin (2017) that might account for this difference in findings relates to the person of the matrix subject. Whereas Djärv and Bacovcin's (2017) stimuli were restricted to sentences with third person matrix subjects, our data included sentences with first and second person matrix subjects as well.

Another difference between Djärv and Bacovcin's (2017) study and ours has to do with the location of the focused constituent. Whereas the subject of the complement was focused in all

their complement-focus stimuli, utterances in our dataset were much more diverse: in some cases, the complement subject was focused as in (19a) and (27b), but other constituents were often focused as well as in (19b), and there were also cases in which multiple constituents were focused as in (19c). Each of the non-factive complement focus cases in (19) received negative mean certainty ratings, indicating that the CC did not project in these cases.

(19)   a.   I really did not feel that I   would buy   a Dodge
                                88.9   11.1   (SWBD-221, N=9, μ = -2.44)
        b.   I can't say I really enjoyed it all that much
                              85.7   14.3   (SWBD-046, N=7, μ = -2.80)
        c.   I don't think the teachers could stand the stress all   year long frankly
                        44.4   44.4   11.1
                                                       (SWBD-298, N=9, μ = -0.50)

As discussed in detail in section 2.3, Djärv and Bacovcin (2017) account for their finding that prosody influences projection in terms of potentially conflicting inferences between the lexical semantics of the embedding verb and the prosodically-evoked QUD. We propose an alternative analysis to account for the differing projection behavior of factive and non-factive CCs with respect to prosodically-marked focus. Our analysis starts from the observation that there is more than one way to background information in a discourse. On the one hand, speakers can signal that information is in the conversational background by acting as if it is already in the common ground. This is Stalnaker's (1974) notion of "speaker presupposition". On the other hand, speakers can signal that information is backgrounded in the sense that it is not currently being updated at that moment in the discourse; rather, that content remains constant with respect to the current state of the common ground. This is the type of backgrounding accomplished by information structure (Vallduví, 1993). However, in information structure frameworks such as Vallduví (1993), backgrounding does not always indicate that backgrounded content is an accepted part of the common ground. This distinction between information structural backgrounding and common ground status is further supported by Djärv and Bacovcin's (2017) finding that non-factive CCs are more projective when the complement is backgrounded.

The observation that factive predicates often, but not always, lead to the inference that the CC is in the common ground can be accounted for following a QUD-based analysis such as that outlined by Beaver et al. (2017). That work claims that in an out-of-the-blue, overhearing context of an utterance with a factive predicate, listeners infer a QUD in which each alternative proposition is true, i.e., in which the alternatives are veridical. For example, in an out-of-the-blue situation, the utterance in (20a) (cf. Djärv and Bacovcin, 2017) generates the alternatives in (20b). However, a listener who overhears (20a) is unlikely to take the QUD to include all possible propositions. Rather, a reasonable way to restrict to the alternatives in (20b) is to assume that they only include propositions that are potentially discoverable, i.e., those propositions that are true. In other words, the listener can reasonably infer that the QUD is (20c). This QUD entails that p is true, so *Anna left town* is not-at-issue, and backgrounded in the common ground sense.

(20)   a.   John might have discovered that Anna left town.
        b.   {q: John discovered p}
        c.   {q: John discovered p | p is true}

In contrast, the use of a non-factive predicate like *believe* does not generally lead to an inference that the CC is in the common ground, because, as Beaver et al. (2017) point out, such predicates are non-veridical. The semantics of *believe* alone do not allow the listener to restrict the alternatives in (21b) to those in which p is true: p need not be true for John to believe that it is. Unless the listener has some external reason to do so (e.g., p has high prior probability, or the context suggests that the speaker is committed to its truth), p is unlikely to be taken as common ground.

(21)   a.   John might have believed that Anna left town.
       b.   {q: John believed p}

Focus on the matrix verb in (20a) and (21a) evokes the alternatives in (22):

(22)   {q: John R that Anna left town | R is a cognitive attitude}

Whether the QUD constructed from the alternatives in (22) entails p, i.e., *that Anna left town*, depends on the set of attitudes that instantiate R. If they are all veridical, then the QUD will entail p, and p will project. However, if at least one alternative involves a non-veridical attitude, the QUD will not entail p, and the CC will not project. While the precise identity of the alternatives is context-dependent, the semantics of non-factive predicates obligatorily introduce this non-veridicality into the alternative set, yielding the QUD in (23).[8] For example, the QUD evoked by (21a) with focus on *believe* inevitably includes the alternative *John believes Anna left town*, which does not entail that *Anna left town*. The fact that a non-factive verb like *believe* was used in the first place produces a scalar implicature that stronger veridical attitudes like *know* could not have been felicitously used. That is, it must be possible that the CC is false, otherwise the stronger predicate would have been used. So, even though the CC is information-structurally backgrounded by the prosodic realization of the utterance, it will still be considered at-issue, and hence will not project.

(23)   {q: John R that Anna left town | R is a cognitive attitude and there is at least 1 non-veridical R}

For factive embedding verbs, there is no such expectation that the QUD contains an alternative with a non-veridical attitude. When (20a) is uttered with focus on the embedding verb, the alternative set (22) cannot be constrained to a QUD in which there are non-veridical attitudes among the alternative propositions. Rather, the listener is likely to infer the QUD in (24):

(24)   {q: John R Anna that left town | R is a cognitive attitude there are 0 or more non-veridical R}

However, the discourse context or other speech cues can still provide other evidence to the contrary. For example, speaker A's utterance in (25) obtains a low projection score (-1.25) as it seems to express some uncertainty towards B's previous assertion, by asserting that it wasn't already in A's beliefs.

(25)   B: I think we could have made a different agreement back in the late forties era...
       A: I didn't realize that we'd signed some type of deal like that.
            43.0   57.0   (SWBD-363, N=7, μ=-1.25)

---

[8]There are non-factive predicates that are considered veridical, such as *be right* and *demonstrate* (Anand and Hacquard, 2014). However, the non-factive predicates in our data are exclusively non-veridical.

This analysis extends to complement focus cases for factives and non-factives as follows. When a constituent within the complement clause is focused, the QUD is constructed with the embedding verb fixed in the information structural background. The alternatives evoked with focus on *Anna* for (20a) and (21a) are shown in (26a) and (26b), respectively. Here, the domain restriction that generates the QUD applies to the individuals who are under consideration as town-leavers. Neither the factive (26a) nor the non-factive (26b) entail the CC *Anna left town*.

(26)   a.   {*q:* John discovered *x* left town}
       b.   {*q:* John believed *x* left town}

Under this analysis, projection of the CC will vary depending on whether other contextual information indicates that it is actually resolved as a speaker commitment. This variability in projection for individual items is attested in our data. For example, unlike the other non-factive complement focus examples shown in (19), (27a) received a high mean speaker certainty rating (+2.13), as it is reasonable to believe the speaker is committed to their descriptions of their own past. Similarly, (27b) shows a factive with narrow focus on the embedded subject, which obtained a relatively high positive mean rating (+1.85). In this case the speaker was talking about a test she took and the salient alternatives are the facts on the test that she now knows.

(27)   a.   I can't believe I was so brazen before
                  10.0   90.0   (SWBD-62, N=10, μ=2.13)
       b.   I didn't realize that our garbage isn't being decomposed
                           100   (SWBD-303, N=9, μ=1.85)

Our analysis is similar to Djärv and Bacovcin's (2017) in the sense that the lexical semantics of the embedding predicates interacts with QUD-based inferences. However, unlike Djärv and Bacovcin (2017) who claim that (the absence of) conflicting inferences from these two sources predict differences in projection for factives and non-factives, we suggest that this difference is due to the extent to which the entailments of the QUD are constrained by the semantics of the embedding predicate. This is supported by our experimental results which show that the effect of prosodic variation is much stronger on factives than non-factives. At the very least it supports the case that factive presuppositions are very easily overridden and hence are like the soft presupposition triggers described in Abusch (2010), rather than hard lexical triggers. Moreover, a lexical trigger based analysis alone does not explain our results for the non-factives: we did not see significant differences in projection ratings based on focus placement conditions, but we do see variation in projection for individual items, e.g., (19a) and (27a).

However, these results can be explained by the fact that, for non-factives, information-structural backgrounding does not constrain the QUD alternatives enough to signal that the CC is not-at-issue. Thus, in most cases non-factives do not project. Nevertheless, our study indicates that many factors beyond utterance information structure affect whether the CC is considered at-issue and hence affect perception of speaker certainty, e.g., the a priori probability of the CC given the speaker's public beliefs and the previous discourse. This can produce projection behavior for non-factives and suppress it for factives. This is in line with Simons et al. (2017) who argue that higher level questions in a discourse structure can affect what is considered at-issue, beyond the immediate prosodically-evoked QUD. Overall, analyzing projection in terms of what is at-issue in a discourse –rather than in terms of conflicts between lexical triggers and

the QUD– allows for a more parsimonious analysis of both factives and non-factives.

6. Conclusion

Prior comprehension experiments on prosodically-marked focus and CC projection have exclusively employed carefully manipulated laboratory speech. Our research using naturally-occurring utterances provides crucial evidence about the extent to which previous observed effects of prosodically-marked focus on CC projection extend to the communication of projective meaning in spontaneous speech. Consistent with what has been found for utterances with factive predicates, we found that the CCs of factive prediates are sensitive to prosodically-marked focus: the CC is less likely to project when the complement receives narrow focus compared to when the matrix clause receives narrow focus. However, for utterances with non-factives, prosody was not found to influence the projection of the CC. We proposed an analysis of these findings along the lines of Beaver et al. (2017), emphasizing that speakers can background information by acting as if it is already in the common ground, or by signaling via information-structure that it is not currently being updated. For utterances with factive predicates, information-structural backgrounding can constrain the QUD alternatives such that the CC is not-at-issue, leading to to a projection interpretation. However, non-factive predicates do not easily allow the QUD to be restricted to true alternative propositions; thus, the CC remains at-issue regardless of information-structure, leading to a non-projecting interpretation in the absence of evidence to the contrary.

To provide more rigorous empirical support for this analysis, future work will look more specifically at the types of alternative sets people construct when doing projection ratings. We would also like to know what sort of entailments are accessible from different factive/non-factive driven QUDs. For example, are existential presuppositions actually inferrable from complements with subject focus? Beyond this, our findings suggest that future work ought to investigate how other prosodic patterns relate to projection.[9] This study only looked at the location of the most prosodically prominent word in an utterance. We need to better understand how utterances with multiple prosodic prominences are interpreted in terms of information structure, and how this relates to our 'broad focus' data and contrastive topic constructions. Full prominence annotation of our data using the RPT method may help shed light on this.

References

Abrusán, M. (2011). Predicting the presuppositions of soft triggers. *Linguistics and philosophy 34*(6), 491–535.

Abrusán, M. (2016, June). Presupposition cancellation: Explaining the 'soft–hard' trigger distinction. *Natural Language Semantics 24*(2), 165–202.

Abusch, D. (2010). Presupposition triggering from alternatives. *Journal of Semantics 27*(1), 37–80.

Anand, P. and V. Hacquard (2014). Factivity, belief, and discourse. In *The Art and Craft of Semantics: A Festschrift for Irene Heim*, pp. 69–90. MIT Working Papers in Linguistics.

Beaver, D. (2010). Have you noticed that your belly button lint colour is related to the colour of your clothing? In R. Bäuerle, U. Reyle, and E. Zimmerman (Eds.), *Presuppositions and Discourse: Essays Offered to Hans Kamp*, pp. 65–99. Elsevier.

---

[9]Vaikšnoraitė et al. (2019) have taken a promising first step in this direction by eliciting productions of constructed factive sentences in contexts that bias either projecting or non-projecting interpretations.

Beaver, D., C. Roberts, M. Simons, and J. Tonhauser (2017). Questions under discussion: Where information structure meets projective content. *Annual Review of Linguistics 3*, 265–284.

Beckman, M. E. and G. Ayers (1997). *Guidelines for ToBI Labelling, version 3.0*. The Ohio State University.

Büring, D. (2003). On d-trees, beans, and b-accents. *Linguistics and philosophy 26*(5), 511–545.

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language*, 1–42.

Chierchia, G. and S. McConnell-Ginet (1990). *An Introduction to Semantics*. Cambridge: MIT Press.

Cole, J., T. Mahrt, and J. Roy (2017). Crowd-sourcing prosodic annotation. *Computer Speech and Language 45*, 300–325.

Cummins, C. and H. Rohde (2015). Evoking context with contrastive stress: Effects on pragmatic enrichment. *Frontiers in Psychology 6*.

de Marneffe, M.-C., M. Simons, and J. Tonhauser (2019). The CommitmentBank: Investigating projection in naturally occurring discourse. In *Proceedings of Sinn und Bedeutung (SUB) 23*, pp. 107–124.

Djärv, K. and H. A. Bacovcin (2017). Prosodic effects on factive presupposition projection. In *Proceedings of Semantics and Linguistic Theory (SALT) XVII*, pp. 116–133.

Heim, I. (1983). On the projection problem for presuppositions. In *Proceedings of the Second West Coast Conference on Formal Linguistics*, pp. 114–25.

Karttunen, L. (1971). Some observations on factivity. *Papers in Linguistics 4*(1), 55–69.

Kiparsky, P. and C. Kiparsky (1970). Fact. In M. Bierwisch and K. Heidolph (Eds.), *Progress in Linguistics*.

Mahrt, T. (2018). *Acoustic cues for the perception of the information status of words in speech*. Ph. D. thesis, University of Illinois at Urbana-Champaign.

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics 5*, 1– 69.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics 1*, 75–116.

Simons, M., D. Beaver, C. Roberts, and J. Tonhauser (2017, April). The best question: Explaining the projection behavior of factives. *Discourse Processes 54*(3), 187–206.

Stalnaker, R. (1974). Pragmatic presuppositions. In M. K. Munitz and P. K. Unger (Eds.), *Semantics and Philosophy*, pp. 197–213. NYU Press.

Tonhauser, J. (2016). Prosodic cues to presupposition projection. In *Proceedings of Semantics and Linguistic Theory (SALT) XXVI*, pp. 934–960. Ithaca, NY: CLC Publications.

Tonhauser, J., D. Beaver, and J. Degen (2018). How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics 35*(3), 495–542.

Vaikšnoraitė, E., M.-C. de Marneffe, and T. Tonhauser (2019). On the prosody of presupposition projection: A production study. In *Proceedings of the 49th Meeting of the North-Eastern Linguistics Society*, Volume 3, pp. 241–249. University of Massachusetts, Amherst.

Vallduví, E. (1993). The informational component. *IRCS Technical Reports Series*, 188.

van der Sandt, R. A. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics 9*(4), 333–377.