Meaning and Responsibility

(forthcoming in *Mind & Language*; preprint, please cite final version)

Ray Buchanan  
*University of Texas at Austin*

Henry Ian Schiller  
*Washington University in St Louis*

**Abstract:** In performing an act of assertion we are sometimes responsible for more than the content of the literal meaning of the words we have used, sometimes less. A recently popular research program seeks to explain certain of the commitments we make in speech in terms of responsiveness to the conversational subject matter (Hoek 2018, Stokke 2016, Yablo 2014). We raise some issues for this view with the aim of providing a more general account of linguistic commitment: one that is grounded in a more general action-theoretic notion of responsibility.

1. Introduction

We cannot reasonably be held accountable for everything that follows from what we say. If, for example, a competent speaker literally produces (1) she need not be asserting, stating, or, in any sense, committing to, (2), nor should her audience take her to have done so:

(1) Marfa is 529 miles from Taos.¹

(2) Marfa is *not* 529.324 miles from Taos.

Even though (2) clearly follows from the strictest literal interpretation of (1), the speaker will not, and should not, be blamed should it turn out false; she is not ‘on the hook’, or responsible for, the truth of (2) on the basis of her utterance.² The same could, of course, be claimed regarding any number of the myriad

---

¹ This example is modeled on an illustrative case in Hoek (2018). As an anonymous referee pointed out to us, it is an open question whether number terms have an exact semantics or a ‘lower-bounded’ semantics, on which an assertion like (1) is given an ‘at least’ reading. Here, however, notice that even if we take the semantics for number terms to be lower-bounded (1) would entail that Marfa is not 528.686 miles from Taos, which seems equally problematic.

² A clarification regarding the operative notion of ‘being on the hook’. Here, and in what follows, we are interested in a notion of commitment that is compatible with (at least some degree of) plausible deniability. Hence, we will allow that a speaker might commit to p by outright asserting it, or (as in Hoek’s cases) by merely conversationally implicating it. Though we might have a harder time plausibly disowning the commitments we undertake in acts of assertion, than we do when we conversationally implicate, we think that this is due to the built in-direction of the latter. A speaker implicates p by saying or making as if to say that q; this indirection leaves open a space for plausible deniability that is closed off when you assert p by using a conventional device that means just that. In the recent literature on lying, some theorists would prefer a more restrictive notion of commitment; one that would differentiate between lying and intentionally misleading (see Stokke 2016 and Viebahn 2021). Such
of other propositions entailed by (1). Likewise, a speaker who utters (3) in response to a question about whether she would like to get breakfast is not liable for the claim that she has never consumed food:

(3) I haven’t eaten yet.

The phenomenon is a familiar one, but why exactly does it arise? That is, why is it that a speaker can, borrowing Daniel Hoek’s (2018) terminology, be *conversationally exculpated* from certain of the prima facie consequences of what she literally says?

As will emerge below, we think that there is a simple, yet fully satisfying answer to this question, one that flows naturally from general considerations pertaining to intentions, intentional action, and the foreseeable consequences thereof. Roughly put, our view is simply that a speaker is ‘on the hook’ for anything that she represents herself as believing by uttering what she does. But in order for us to get our own favored view, we will start with an important recent attempt to provide a ‘pragmatic mechanism’ for generating conversational exculpatures developed in Hoek (2018). One interesting feature of our own account is that it suggests that we should no more expect to find such a *mechanism* for determining what we are ‘on the hook for’ in speech than we should for finding a mechanism for determining what we are responsible for in intentionally acting, more generally. Another feature of our account is that it will help us better appreciate the important, but limited role that *subject matter* and *questions under discussion* play in the generation and recovery of utterance content and commitment.3

2. Aboutness and Exculpature

In recent years, there has been some interest in explaining how a speaker might be excused from certain of the prima facie commitments of what she says in terms of *conversational subject matter*.4 Our intuitions about what commitments someone takes on in a conversation are often responsive to considerations concerning what the conversation she is engaged in is *about*. As Yablo (2014) notes, clarifying subject matter can sometimes be used as an excuse for producing an utterance that might otherwise carry unwanted commitments. For example, a literal utterance of (4) plausibly entails the

---

3 Why focus on the normative notion of what a speaker commits to, rather than the notion of ‘what is said’? We believe that no theoretically useful notion of what is said can be identified apart from what a speaker means, what they’re on the hook for, and the literal meanings of their utterances (cf Borg 2017).

existence of events (the event(s) of Oscar’s toast buttering), properties (the property of being done slowly), and numbers (the number 2):

(4) Oscar buttered the toast slowly, two times.

But even if a philosophically trained speaker knows that these entailments flow from (4), Yablo suggests that she need not worry about taking on such commitments in uttering it if the conversation in which she is engaged is not, in any way, about such metaphysical matters (but instead is about, say, what happened in the kitchen). This is an example of what Hoek calls *conversational exculpature*. Building on Yablo’s initial insight, Hoek argues that the phenomenon of conversational exculpature is essentially tied to conversational subject matter.

Hoek purports to have identified a pragmatic mechanism that generates such conversational exculpatures. Roughly put, Hoek takes conversational exculpature to be a variety of *pragmatic subtraction* wherein a speaker means, or commits to, something “less than” the literal content of her words; a process that helps to excuse the speaker from any would-be commitment of her utterance that is not itself appropriately tied to the subject matter of the conversational context.\(^5\)

Following Mandy Simons (2005), Hoek assumes that assertions often owe their relevance to additional presuppositions the speaker makes in context, and that such contextual presuppositions need not always be part of the antecedent conversational common ground. The “subject matter” of the conversation is a partition imposed upon those worlds by the operative *question under discussion* ("QUD") where the elements of the partition are the relevant alternative answers to that question.\(^6\) Hoek’s basic idea is that in cases of exculpature, the speaker’s “intended message” - that for which she cannot be conversationally exculpated - is the ‘unique, wholly relevant remainder’ that results from intersecting the “literal content \(p\), the underlying contextual presupposition \(q\), and the question under discussion \(S\)” (ibid. 153). As he elaborates:

In cases of exculpature, the speaker’s intended message \(r\) is determined on the basis of the literal content \(p\) of their statement, and two contextual clues: the contextual presupposition \(q\) to which the speaker appeals, and the subject matter \(S\) they address. The speaker’s message \(r\) is the unique proposition that is (A) just about \(S\) and (B) equivalent to \(p\) given \(q\).

---

\(^5\) See Stalnaker 1998, 2002. See Roberts 1996/2012 for the initial discussion of the QUD framework. Following Hoek, we’ll treat the notions of subject matter and QUD as interchangeable. This is to be distinguished from the notion of (positive) subject-matter developed by Kit Fine (2016), which treats it in terms of a fusion of its truthmakers.

\(^6\) See Groenendijk, and Roelofsen 2019 for the relevant semantics for questions, which treats them as sets of alternatives.
In cases of exculpature, Hoek claims, it is “only this relevant remainder \( r \), and not \( p \) or \( q \), is seriously endorsed and added to the conversational common ground” (p. 153; italics ours).

To illustrate, return to one of our initial examples in the Introduction. Suppose a speaker utters (5) in direct response to the question ‘How far is Marfa from Taos?’:

\[
(5) \text{Marfa is 529 miles from Taos.}
\]

The strictest, literal interpretation of (5) entails that Marfa is not 529.324 miles from Taos, yet in this context the speaker who utters it will incur no such commitment. Why so? In this case, assume that (i) the literal content of (5) is that Marfa is exactly 529 miles from Taos (\( p \), below). Note that this is incredibly unlikely, and thus we might insist that (ii) the utterance of (5) presupposes that there is an integer distance between Marfa and Taos (\( q \), below). If we further assume that (iii) the speaker and hearer are only interested in the distance between Marfa and Taos to the nearest mile, we can take the conversational subject matter of the conversation to be \( s \), below.

\[
\begin{align*}
p &= \text{Marfa is exactly 529 miles from Taos.} \\
q &= \text{Marfa is an integer distance from Taos.} \\
s &= \text{The distance between Marfa and Taos to the nearest mile.}
\end{align*}
\]

Here, the Stalnakerian proposition that \( \text{Marfa is 529 miles from Taos to the nearest mile} \) is logically equivalent to \( p \) given \( q \) (the conversational presupposition required to make sense of the utterance) and \( s \) (the conversational subject matter). Moreover, this proposition is as much “just about” the relevant subject matter as it is an answer to the question ‘How far is Marfa from Taos to the nearest mile?’.

Notice that in the foregoing example, the speaker, in some sense, conveys less than the minimal, literal content of her words: what is intuitively conveyed is itself entailed by the relevant minimal content, (p), but not conversely. [That Marfa is exactly 529 miles from Taos entails, but is not entailed by, that to the nearest mile, Marfa is 529 miles from Taos.] One important feature of Hoek’s account is that, if correct, it would also handle cases of exculpature in which the speaker commits to something logically
independent of what she says. Consider an utterance (6) in response to the question ‘What kind of hat did Ellen wear?’:

(6) Ellen wore the same type of hat as Sherlock Holmes.

If a speaker utters (6) in a conversational setting in which both she and her audience presuppose that Holmes wore a deerstalker, she is, and will be taken to be, committed to the proposition that Ellen wore a deerstalker, but not to the proposition that Holmes exists. Notice that in this case, this proposition that the speaker is on the hook for is not itself entailed by the literal content, c. It is, however, entailed by c in tandem with the contextual presuppositions of the speaker and hearer. Notice here that a myriad of other inequivalent propositions of the form Ellen wore a deerstalker and Q are also entailed (where Q is anything whatsoever that we mutually presuppose, for example, that grass is green). On Hoek’s account of these various contextual implications of c and the common ground of the conversation, the speaker only commits herself to those propositions that address the conversational subject matter - i.e., propositions that answer the operative QUD, here, ‘What kind of hat did Ellen wear?’. (We assume, along with Hoek, that a proposition is an answer to a QUD just in case it is equivalent to one of the elements of the partition of the context set that the QUD creates.) Informally, Hoek suggests that we might give the following gloss regarding how Nina’s audience might come to infer the relevant message:

We know that Nina does not seriously believe what she said. She does not believe in Holmes and besides, she is not talking about Victorian detectives. She must in fact be telling us something relevant, that is, something about Ellen. It is clear enough how what she said connects to that topic: Nina is talking as if Holmes were a real detective, who really wore one of those funny hats. Given this assumption, what she said is another way of saying Ellen wore a hat like that. Accordingly, that must be the information about Ellen she intends for us to pick up on. (ibid. 162-163)

In effect, Hoek takes conversational exculpature to be a species of relevance implicature that is essentially routed through considerations of conversational subject matter. In the relevant range of cases, a speaker commits to only that content that is logically equivalent to what she literally said given the conversational common ground. Importantly, Hoek emphasizes that his account “fails to generate alternative readings in most contexts: exculpature is only defined given a suitable configuration of literal content, contextual presupposition, and subject matter.” (p. 159).

7 This gives Hoek’s view an advantage over accounts on which linguistic commitments are determined merely as a matter of contextual implication.
If we could successfully generalize from examples such as the foregoing, we would have a principle-driven story about one important species of pragmatically determined content. Any viable pragmatic theory should explain why in utterances like (5) speakers cannot be held accountable for every entailment that flows from the literal meaning of her words. Hoek offers an elegant, plausible explanation of how this could be.

Hoek’s discussion is complex and we encourage the reader to have a look at the original paper for the formal details of his account. In what follows, however, we want to single out two claims in his presentation that strike us as especially central to the philosophical underpinnings of the account, and that will also help to motivate our own favored story regarding the generation and recovery of speakers' commitments.

Notice that in the official statement of the view above, Hoek makes two important claims, both of which we will take issue with. First, there is a claim regarding how exculpatures are “determined”:

(DET) “In cases of exculpature, the speaker’s intended message r is determined on the basis of the literal content p of their statement, and two contextual clues: the contextual presupposition q to which the speaker appeals, and the subject matter S they address.” (p. 159)

Second, there is a predictive claim regarding the content of the speaker’s message in those cases in which the “configuration of literal content, contextual presupposition, and subject matter” required for exculpature are met:

(PRED) In cases of conversational exculpature, ‘the speaker’s [intended] message r is the unique proposition that is (A) just about S and (B) equivalent to p given q.’ (ibid.)

For ease of exposition in what follows, let’s say that a proposition p is appropriately QUD-responsive in context c just in case it meets conditions (A) and (B) above; i.e., it is (i) logically equivalent to the literal content, q, of the speaker’s utterance (given the contextual presuppositions c), (ii) an answer, or partial answer, to the operative QUD in c. Further, we will call the configuration of ‘literal content, contextual presupposition, and subject matter’, that Hoek claims is required for conversational exculpature, favorable conditions. According to this view, when favorable conditions are met the speaker’s intended

---

8 Hoek claims that uniqueness is guaranteed by a constraint that he calls ‘Independence’ (ibid., p.173). Roughly put, the Independence constraint states that the subtracted contextual presupposition should have “no bearing” on the subject matter. In what follows, our worries regarding Hoek’s account do not turn on objections to this (putative) constraint, or on examples in which it is violated.
message - that which she ‘seriously endorses and is added to the common ground’ (p. 153) - is “determined” (more anon) by considerations of the literal content of her words, the subject matter, and the relevant contextual presuppositions; the content of that message being the unique proposition that is appropriately QUD-responsive.

Though there is much to admire in this elegant picture, we ought to be very cautious about taking it on board. As we are about to see, the problem isn’t simply that there are counterexamples to (DET) and (PRED). We think that the deeper problem concerns the motivation for looking for such a “pragmatic mechanism” for conversational exculpature in the first place.

3. Intentions, Commitment, and Subject Matter

In order to evaluate the foregoing account of exculpature, it is crucial to first resolve an important ambiguity in (DET) and with that a correlative unclarity in (PRED). In exactly what sense might QUD-responsiveness “determine” anything concerning the speaker’s ‘intended message’ (in favorable conditions or not)? ‘Determines’ as it occurs in (DET) is ambiguous between a (a) metaphysical reading and (b) an epistemic reading. On the former metaphysical reading, (DET) amounts to the following claim:

(DET-M) In cases of conversational exculpature, the speaker’s intended message is grounded in facts regarding QUD-responsiveness.

According to the latter epistemic reading, (DET) amounts to the claim that the speaker’s intended message is “figured out”, or recovered, on the basis of considerations pertaining to QUD-responsiveness:

(DET-E) In cases of conversational exculpature, the speaker’s intended message is recoverable on the basis of considerations pertaining to QUD-responsiveness.

As we will argue momentarily, (DET-M) is false and a proper appreciation of why this is so will be a first step towards better understanding the underlying metaphysics of speaker commitment in discourse. Moreover, seeing the problem with (DET-M) will also allow us to see how we might more plausibly locate Hoek’s important insights regarding exculpature and QUD-responsiveness in the epistemological story about how we successfully recover speaker commitments in certain cases.

However the notion of ‘intended message’ is to be cashed out (more anon), it is implausible that any facts concerning subject matter/QUD could “determine” it in the sense required for (DET-M); i.e., as
an account of what generates, or gives rise to, conversational exculpatures. What a speaker intends - and, more generally, what she is intentionally doing by uttering what she does - is itself fully grounded in the facts concerning her mental states at that time, including what she intends, expects, desires, etc. The claim that the mutually understood QUD in a context can reach back, as it were, and make it the case that a speaker had this or that specific communicative intention in uttering what she did seems to us implausible. Even more generally, we cannot see how facts regarding the QUD could metaphysically determine anything that the speaker is intentionally doing in producing her utterance.

To make this point vivid, consider a speaker, Rosa, and hold her audience-directed intentions in a particular conversational setting fixed. Suppose, for example, that Rosa utters (7) intending for Oscar to entertain that his Amazon Prime package has arrived as a result of his recognition of her (communicative) efforts:

(7) It’s here.

Hold fixed this intention and her other relevant psychological states, vary the mutually understood QUD in any way you like. Could that variation by itself plausibly make a difference to her “intended message”? No. Of course, if Oscar had just explicitly asked ‘When is the car picking us up’ and she utters (7) with the same intentions as she did in the original set up, she will likely be misunderstood. But, that fact does not itself change her intentions or her intended message. What your audience takes you to have meant - or is in a position to take you to have meant - does not, and cannot determine, what you actually intended to get across.

This last point suggests that the more plausible interpretation of (DET) is the epistemic one, (DET-M) above. After all, there is no doubt that the fact that a particular proposition is QUD-responsive is sometimes an important epistemic consideration in favor of taking it to have been intended or meant by a cooperative, helpful speaker. In the case above, for example, the reason that Rosa’s audience would have been justified in taking her intended message to be that the car had arrived is precisely because that

---

9 For some reflections on these features of intentions more broadly, see Davidson 1963, Bratman 1987. There is the possibility that the content of a speaker’s intended message is itself somehow determined by the meaning of the words she uses in a context, a view that Lepore & Stone (2015) call direct intentionalism. This is a controversial position, and we are skeptical that any theoretically justifiable notion of speaker intention can be cashed out in these terms.

10 There is a reading of what we’re claiming here that might seem question begging (i.e., hold fixed what the speaker intends, and then interrogate your intuitions on whether there could be some variation in what she intends). What we mean, however, is that by holding fixed a speaker’s intentions and varying other contextually relevant prospects we manage to hold fixed what we could plausibly be interested in the speaker having aimed at getting across, not just her intentions tout court (Grice 1957). What we don’t hold fixed, however, is the rationality of her having that intention.

11 Davis (1998) emphasizes this point to great effect in his critical discussion of the standard Gricean account of conversational implicature.
proposition is appropriately QUD-responsive. It is here - on the epistemological side of utterance interpretation - that we can best locate Hoek’s insights about QUD-responsiveness and speaker commitment. Moreover, we think that this shift in the interpretation of (DET) suggests the need for a correlative shift with regards to (PRED). Rather than taking Hoek’s apparatus as issuing directly in predictions regarding what the speaker did or did not intend in a particular communicative setting, better to take it as issuing in predictions regarding what one’s audience *ought (in the epistemic sense)* to interpret you as having intended/committed to:

(Pred-E) In cases of conversational exculpature, you ought to interpret the speaker as intending the unique proposition, r, that is (A) just about S and (B) equivalent to p given q.

This understanding of Hoek’s framework is supported by at least some of the surrounding commentary on the formal apparatus. For example, at one point Hoek suggests ‘the central predictive claim of the account is … that [the QUD-responsive reading] is available in any context where it is well defined.’ (ibid. P. 171). If an ‘available reading’ of a speaker’s utterance is simply something that her audience might *reasonably take* her to have intended/meant, (PRED-E) offers a plausible construal of what Hoek is hoping for his account to predict. In effect, if Hoek’s framework is understood in this way it offers an account of how we ought (in the epistemic sense) to interpret speakers in cases of conversational exculpature.

(PRED-E) and (DET-E) are considerably more plausible than the metaphysical/grounding reading of Hoek’s claims. Indeed, we think that these claims successfully capture something of genuine epistemic importance regarding how our hearers successfully recover our intended message in a certain range of cases (See Section 5). That being said, however, we think that there are counterexamples to (DET-E). More importantly, however, we will argue that these claims do not successfully predict all instances of the target phenomenon. A more general account is needed.

4. Some Reservations

In assessing whether Hoek’s framework offers a plausible general account of the recovery of speaker commitment (i.e., an account of what we ought to take the speaker to have intended/be committed to), it is important to remember its limited scope. Hoek claims that his account ‘only has application in those cases in which there is “a suitable configuration of literal content, contextual presupposition, and subject

---

12 Thanks to an anonymous referee for encouraging us to pursue this interpretation
matter.” (p. 159) (i.e., cases in which there are favorable conditions). As best we can see, however, cases of conversational exculpature are not so limited.

Reconsider one of Hoek’s own examples from earlier:

(8) Ellen has the same kind of hat that Sherlock Holmes does.

According to Hoek, a speaker uttering (8) in response to a question about the kind of hat Ellen wears (i.e., in a conversation whose subject matter involves Ellen’s sartorial choices) need not be understood as committing herself to the existence of Sherlock Holmes. Rather she will naturally (and correctly) be understood as only committing to the claim that Ellen wears a deerstalker. But as best we can see, this interpretation is not itself essentially tied to the relevant QUD. After all, wouldn’t this interpretation be equally justified, even if, say, Ellen had uttered (8) as an interjection in a conversation about the vote recount in the Senate Race? Or, if she’d uttered (8) in a discourse initial setting (for example, as a passing remark to a stranger on the bus)? 13 That is, isn’t it plausible that a speaker uttering (8) could be plausibly seen as undertaking exactly the same commitments even in contexts where there is a different, un-related QUD, or none at all? (PRED-E) does not, however, have clear application in such cases, so a more general story is needed.

Of course, a proponent of (PRED-E) might try to respond to the worry about discourse initial utterances by claiming that there must always be some QUD or other - for example, ‘How is it with the world?’ 14 We are skeptical that this is right. But notice that even if it were correct, it would be doubtful that such a general, all purpose QUD could do much by way of helping to determine the (oftentimes) fine-grained commitments we can take on in such cases. In the case of (8) above, such a general all purpose QUD would wildly overgenerate “available” readings as it would fail to discriminate between indefinitely many other propositions that are themselves logically entailed by the strict literal interpretation of (8). Likewise, we are equally skeptical of any response according to which a speaker makes a conversational interjection she must always be understood as raising and answering a new QUD thereby. Not only does this strike us as ad hoc; it also helps to confirm a more general suspicion of ours regarding appeals to QUDs: oftentimes a QUD is itself made manifest via a recognition of the speaker’s intended message and not vice versa (cf Grindrod & Borg 2019). If you haven’t seen your friend, Carla in several weeks and she approaches you and utters (9):

---

13 Roberts (1996) identifies discourse initial utterances with addresses of a question about how the world is. We might think that in situations where speaker and hearer share a sufficient amount of background information and interests, discourse initial utterances may address a more specific subject matter. The point is that even if the speaker addresses an incredibly general question, their utterance will oftentimes have the same normative features.

14 See Stalnaker 2002, and Roberts 2012, who takes up this suggestion. See Camp 2018 for some recent comments.
Perhaps there is some sense in which Carla might be taken as raising, and thereby answering the question ‘Where has Carla been?’ by uttering (9), but this is clearly derivative from the successful recovery of what she intended to convey in the first place. Oftentimes, the QUD, rather than enabling a hearer to determine the speaker’s intended meaning, is established only once that meaning has been identified.

Moreover, there are other cases in which the strict, literal interpretation of a speaker’s utterance delivers something more (i.e., entails something stronger) than the intuitively communicated content that the account under consideration cannot plausibly accommodate. Consider a case originally due to Bach (1994). Suppose a speaker utters (10) in direct response to the question ‘How severe is Oscar’s paper cut?’:

(10) Oscar isn’t going to die.

As Bach notes, the strictest, literal interpretation of (10) entails that Oscar is immortal, yet in this context the speaker who utters it will incur no such commitment; in this conversation, she is excused from this entailment of what she (strictly speaking) says. But why so? In this case, assume that the literal content, $p$, of (10) is that Oscar is immortal and that it is common ground that Oscar has a paper cut ($q_1$). How, exactly, do we get from that literal content of (10) and $q_1$ to the intuitively correct interpretation of the speaker’s remarks - i.e., (r) that Oscar will not die from the paper cut. If we were to try to apply the Hoek-inspired account to this case, we would have to find some contextual presuppositions in the context such that (b) comes out logically equivalent to the literal content of (10) given the mutually understood QUD and $q_1$:

QUD = How severe is Oscar’s paper cut?

\[ p = \text{Oscar is immortal.} \]
\[ q_1 = ? \]
\[ q_2 = \text{Oscar has a paper cut.} \]
\[ r = \text{Oscar won’t die from that paper cut.} \]

But what could this additional presupposition plausibly be? It can’t be simply that Oscar is mortal, since that is logically inconsistent with both $p$, as well as the target proposition, $r$. If we try to simply reverse engineer the needed presupposition to arrive at the intuitively correct prediction (r), we’d need something
along the lines of the following: *that either Oscar dies from that paper cut or he does not die at all*. We very much doubt that the speaker or her audience could plausibly be said to be presupposing any such thing (even in the most attenuated sense of merely *acting as if* they accept it for the purpose of the conversation). More generally, we doubt that the speaker/hearer can be plausibly taken as presupposing anything that would render the target interpretation appropriately QUD-responsive in this case. At best, [PRED-E] does not have application in this case.

The foregoing examples cases are ones in which speakers incur linguistic commitments in the absence of any QUD or in the presence of an unrelated one (as in the example of the interjection), as well as cases in which speaker’s do not seem to be making the presuppositions that the theory requires (i.e., these are cases in which there is not an appropriate configuration of ‘literal content, contextual presupposition, and subject matter’ to apply the theory). But, as we are about to argue, even when conditions are favorable for the application of the [PRED-E], sometimes speakers are on the hook for more than the account would suggest, and sometimes less. Let’s start with an example of the former.

Even when there *is* a mutually recognized question under discussion, we worry that (PRED-E) does not always issue correct predictions. Suppose that we ask you (11), to which you respond by (12):

(11) Will Rick Perry be at the party tonight?
(12) Every former governor in Trump’s cabinet will be at the party.

Further, let’s assume that we mutually presuppose (11a):

(11a) Rick Perry is the only former governor serving in Trump’s cabinet

Notice that according to the proponent of (PRED-E), the proposition expressed by (11b) is predicted to exhaust your intended message given (a) and (11) as the QUD:

---

15 Thanks to an anonymous referee for correcting an earlier presentation of this example, and for helping us better see the challenge it poses for the account under consideration.

16 As an anonymous referee pointed out, a proponent of [PRED-E] might claim that the foregoing example is not a worry for their view, if by “conversational exculpature” they only mean to be discussing cases in which the intended meaning of an utterance is “determined” by subtracting a candidate contextual presupposition; i.e., cases unlike (10). Since we claim that in this case there is no realistic contextual presupposition, this would indeed not count as a case of “conversational exculpature” in this stipulative sense. We are, of course, happy to follow suit and talk in this way, limiting the intended applicability of the [PRED-E], accordingly. Moreover, we are also open to the possibility that the case of (5) might be explained in some other way (perhaps, in terms of lexical modulation or some variety of contextual enrichment). Our goal in what follows, however, is to give a general account of how (and why) speakers take on the commitments that they do in speech (and can sometimes be excused for obvious entailments of what they say). Insofar as the examples in the text above successfully illustrate such speaker commitments (or the lack thereof), [PRED-E] cannot then plausibly be taken as a general theory of speaker commitment and responsibility that we seek to provide in this paper.
(11b) Rick Perry will be at the party.

Why so? In this conversational setting, (11b) is logically equivalent to the literal content of (12) given the contextual presupposition (11a). That is, given the contextual presuppositions at play, the set of worlds at which Perry is at the party = the set of worlds at which every former governor in Trump’s cabinet will be at the party. Moreover, (11b) is also, given the context, the answer to the explicit QUD raised in (11). Hence, (11b) is predicted to exhaust that which you committed to in uttering what you did. On the proposal under discussion, in cases of exculpature, you ought to understand the speaker’s intended message to be the unique proposition, r, that is (A) just about S and (B) equivalent to p given q and here that is (11b) (recall the claim that it is “only this relevant remainder r, and not p or q, is seriously endorsed and added to the conversational common ground” (p. 153; italics ours)). But Is (11b) all that you are on the hook for in uttering (10)? No. Plausibly, you are also on the hook for the literal content of what you said: namely, the proposition that every former governor in Trump’s cabinet will be at the party. To appreciate why this is so, suppose it was revealed to us that our presupposition (11a) is in fact mistaken and that Perry is not the only former governor in Trump’s cabinet. As it turns out, Trump’s Secretary of Agriculture, Sonny Perdue, was also once a governor. Notice that if Sonny were not going to be at the party, you would have intuitively asserted/committed to something false by uttering (12). We claim that the reason for this is that you are indeed responsible for the literal content of (12) even though that content is not the one predicted by (PRED-E). In short, there are cases where (PRED-E) undergenerates commitments.

In other cases, however, it seems that the account overgenerates commitments. To illustrate what we have in mind, suppose that we mutually presuppose that numerous descriptions are uniquely co-instantiated with being the 44th US president. For example, suppose we all presuppose that the 44th president is one and the same as the former US senator whose middle name is ‘Hussein’ and is one and the same as the husband of Michelle Obama, 2012 Time’s 2012 Person of the Year, and so on. Given these numerous mutual presuppositions, we ask you (13), to which you respond by (14):

(13) Who is coming to Albuquerque next week?
(14) The 44th President of the US is coming to Albuquerque next week.

17 One suggestion Hoek makes about how to understand his view is that the conversational subject matter plays a role in determining what is contextually presupposed, such that we presuppose whatever makes the utterance most relevant to the QUD (page number). But if this suggestion is correct, then there actually seems to be a demand on us to presuppose that Rick Perry is the only former governor in the cabinet, because this presupposition is what makes (12) most relevant to (11)).
What exactly have you committed yourself to thereby? If (PRED-E) were correct, we should understand you to have committed to much more than you might have thought you were signing on for by uttering (14). Assuming a Russellian analysis of definite descriptions, notice that, for example, (*) the proposition that the husband of Michelle Obama is coming to Albuquerque next week is logically equivalent to the literal content of (14) given the contextual presuppositions. Moreover, since (*) also counts as an answer to the explicit QUD given in (13), it meets the conditions of being appropriately QUD-responsive. But is it plausible that you asserted or committed to that proposition in uttering (14)? Though you indeed believe (*), you don’t make yourself thereby liable for it by uttering what you did. Notice that if, for example, it turned out that Barack and Michelle never officially filed their marriage papers at the courthouse and they were not actually legally married, you would have no grounds whatsoever for retracting your utterance of (14). We think that the best explanation for why you’d have no reason to retract is simply that in uttering (14) you did not commit yourself to the truth of (*) or, for that matter, that the Time’s 2012 Person of the Year is coming to Albuquerque, or any number of the other contextually relevant propositions of the same form. If this is correct, (PRED-E) also overgenerates speaker commitments.

In some cases, [PRED-E] undergenerates speaker commitment; in other cases, it overgenerates. If this is correct, it shows that whatever its other merits, (PRED-E) simply cannot be the full story regarding conversational exculpature.

6. Intention, Responsibility, and Speaker Commitment

In the foregoing discussion, we distinguished between two distinct questions one might ask about speaker commitment and conversational exculpature; one metaphysical, the other epistemological. On the one hand, we would like a story of what metaphysically grounds, or makes it the case, that a speaker takes on a particular commitment in uttering what she does. On the other hand, we want an epistemological account of how it is that the speaker’s audience successfully recovers those commitments from her utterance. In what follows, we offer our favored answer to these questions with a focus on the former.

What determines - in the metaphysical sense - speaker commitment? When we remember that speech acts are just that - intentional acts performed by speakers with certain purposes in mind - the shape of an answer to this question begins to emerge. Very roughly put, we think that a speaker is ‘on the hook’ for all, and only, the contents of those beliefs that she intentionally provides evidence for by uttering what she does.

In a typical literal utterance of a declarative sentence, a speaker will intentionally provide her audience evidence of certain of her beliefs. Our view is that in an assertoric act, a speaker commits to
some content $p$ just in case she is responsible for having given her audience evidence that she believes it. Here the notion of responsibility should be understood in the normal action-theoretic way: an agent can be held responsible for anything whatsoever she does of her own volition.\footnote{Though as many have noted (cf. Bratman 1987) which of an agent’s volitional activities we rightly hold her responsible for will depend on a variety of factors, including the stakes of her decision, her overall plan, her alternative choices, etc.} When it comes to the issue of what a speaker commits to in an act of assertion, the relevant responsibility is for her having intentionally provided her audience evidence that will, by her lights, put them in a position to infer that she believes that $p$. For short, we will say that when a speaker meets these conditions she has \textit{represented herself as believing that $p$ by her utterance}.\footnote{We think of utterance interpretation on an inferentialist model. If you prefer a perceptual model, just rephrase the foregoing as ‘...evidence that will put their audience in a position to see...’} Put in these terms, we can state our favored general account of speaker commitment as follows:

\begin{quote}
\textit{[Commitment]:} S commits to $p$ by uttering $u$ iff S represents herself as believing that $p$ by uttering $u$.
\end{quote}

When a speaker’s utterance satisfies the conditions of [Commitment] she makes herself \textit{responsible} for having given evidence of the belief that $p$ and it is for this reason that she can be held liable should it turn out that not $p$.

In assessing [Commitment] it is important to keep three points in mind. First, for present purposes, we will remain officially agnostic about the precise relationship between our favored notion of linguistic commitment and Grice’s notion of speaker-meaning.\footnote{Davis (2002) makes an important case against traditional Gricean accounts of speaker-meaning and in favor of a novel, simpler analysis couched solely in terms of the intentional expression of belief, like [Commitment] above. Here, we don’t want to adjudicate whether Davis has successfully analyzed speaker-meaning (see Buchanan (2012), Davis (2013), and Zeman (2014) for discussion). As should be clear from the text, however, we are more than sympathetic to Davis’s picture if it is offered as an account of speaker-commitment in the sense that we have been discussing.} There are, however, some initially compelling reasons for resisting identifying the former with the latter. For one thing, the conditions in [Commitment] are considerably less demanding than those standardly required by Griceans for speaker meaning, including the full range of higher order, or reflexive, intentions that Griceans typically appeal to in their accounts. For now, we will note just one potential disanalogy between speaker commitment and speaker meaning. Some authors write as though speaker meaning is conceptually tied to understanding in the following sense: understanding an utterance requires successful recognition of what the speaker intended in producing it. That is, if S meant that $p$ by $u$ then you must entertain that $p$ in order to understand $u$. Interestingly, this condition does not seem to hold for speaker commitment. Suppose that
during a spirited conversation about classic rock over Thanksgiving dinner, your least favorite uncle utters (15):

(15) Come on! None of the Beatles could even read music.

Your uncle has plausibly represented himself as believing (*) that (among other things) Ringo could not read music. But did he also mean it in the Gricean sense? Not necessarily. Even if, for whatever reason, you did not entertain anything specifically about Ringo on the basis of his utterance, did you thereby fail to understand (15)? Insofar as the answer is ‘No’, we should allow that speaker meaning and speaker commitment can sometimes come apart. While we strongly suspect that any plausible analysis of speaker meaning will provide conditions that are themselves sufficient for speaker commitment in the sense at hand, we are unsure whether those conditions will also be necessary. For now, however, it suffices to say that we are offering [Commitment] solely as an analysis of what determines the speaker’s liabilities. While we have some doubts about whether a speaker must mean p in the Gricean sense to be liable for it, we won’t try to fully substantiate this suspicion here.\footnote{It is worth noting that our account of commitment also departs in a significant way from accounts in the so-called neo-pragmatist tradition (Brandom 1983, 1994; Kukla & Lance 2009) which treat a speaker’s commitments as arising from conventional properties of the words uttered.}

Second, and relatedly, we claim that a speaker commits to the contents of those beliefs she intentionally represents herself as having by uttering what she did. For example, consider a nervous employee who asks their boss for a raise. The employee is aware that their trembling voice will give evidence of their fear/belief that the request will be denied. While she intentionally produces her utterance and it is - even by her own lights - good evidence of her fear/belief, she will not satisfy the conditions required by [Commitment] thereby since she does not produce her utterance as evidence of her fear/belief. Compare: when you utter the word ‘Oscar’ you will be intentionally providing evidence of the fact that you believe that it begins with the letter ‘O’, even though you will not, in the normal course of events, be uttering ‘Oscar’ \textit{in order to} provide evidence of this belief. Likewise, you might know that (normally) in putting on your jacket you are intentionally doing something that is evidence of the belief that it is cold, but this is quite different from putting on your jacket \textit{in order to} get your host to recognize your belief that the room is too cold. We think that this distinction is the key for understanding cases such as earlier example:

(16a) Who is coming to Albuquerque next week?
(16b) The 44th President of the US is coming to Albuquerque next week.
Given the relevant conversational presuppositions, your utterance of (16b) will provide your audience evidence of your belief that the husband of Michelle Obama is coming to Albuquerque next week, that Time’s 2012 Person of the Year is coming… and a host of other propositions. But the reason that you are not liable for these further claims is that you did not produce (16b) specifically in order to provide evidence of beliefs with those contents; your utterance is indeed evidence of those beliefs, but it wasn’t produced with the goal of being such evidence. In a normal conversational setting, a sincere, literal utterance of a declarative sentence will be evidence of numerous beliefs on the part of the speaker, not all of which she can be held accountable for simply in virtue of saying what she did. In order to ascertain what she has committed to by her utterance we must uncover certain of the reasons she had for uttering what she did; in particular, reasons pertaining to those beliefs she intentionally produced her utterance as evidence of.

Third, notice that if [Commitment] is correct, our liability in speech comes from what we intentionally provide evidence for. As best we can see, the things that we intentionally do in using language are subject to all of the rational constraints characteristic of intentional action, more generally. For example, if it is generally true that a rational agent can intentionally \( \varphi \) only if they have the ability to \( \varphi \), then so too for what they can intentionally provide evidence for by an utterance. Likewise, insofar as the intentions that lie behind our intentional actions are generally subject to consistency requirements of the kind advocated for by Bratman, so too with those intentions that lie behind our communicative endeavors. For example, consider a (plausible) means-end constraint on rationally formed intentions:

\[
\text{Means-End: if a rational agent forms the intention to } \varphi \text{ but believes that } \Upsilon \text{ is a necessary means to do so, they will form the intention to } \Upsilon \text{ if they had not already done so.}
\]

We think that a linguistic analogue of this constraint can be found in certain cases of presupposition accommodation. When a speaker ultimately seeks to provide her audience evidence of the belief that \( p \) by \( u \), but thinks that they must first take her to believe that \( q \) in order to do so, she will (if rational) also intend that \( u \) be evidence of the belief that \( q \). For example, suppose that your friend has overheard you talking about a pet named ‘Carla’ and she asks you (17), to which you respond by uttering (18):

(17) Can Carla lay an egg?
(18) Carla is a pug.

Now in this case we think it is clear that you have, by way of uttering (17), committed to (19).
(19) Carla can’t lay an egg.

But what about (20)?:

(20) Pugs can’t lay eggs.

Insofar as you think that your audience must take you to believe (20) in order to appreciate that you believe (19), you must (if rational) also be uttering (18) in order to provide evidence of a belief whose content is (20). On our account, she is then predicted to be liable for (19) (as well as (20)). This strikes us as exactly the right verdict.

Returning to the examples with which we began, we ask why is it, for example, that an average speaker uttering (1) will not be on the hook for the truth of (2) even though it is clearly entailed?

(1) Marfa is 529 miles from Taos.
(2) Marfa is not 529.324 miles from Taos.

The explanation is simple: a normal speaker does not produce (1) as evidence of their believing the proposition expressed by (2). For one thing, your average speaker won’t believe anything so granular as (2) to begin with. But, even if she did have such a belief, what grounds could she have for possibly expecting that her utterance of (1) will give her audience evidence that she believes (2)? Perhaps if she were addressing the Society for Cartographical Precision, she could reasonably form such an intention, but not during your average coffee shop small talk. Likewise, unless Nina has special reason for thinking that her audience is a fervent believer in the reality of Sherlock Holmes (he actually lived on Baker St. etc.), how could she reasonably expect to be taken as speaking literally in uttering (6)?

(6) Ellen has the same kind of hat that Sherlock Holmes does.

That is, how could she, by uttering (6) reasonably be expecting to provide her audience evidence that will put them in a position to infer she believes in Holmes? Such an expectation would only be justified if she thought that - somehow - the fact that she uttered a sentence whose literal content entails Holmes’ existence could outweigh the audience’s reasonable background assumption that she, like virtually everyone else, believes the Holmes stories to be fiction.\(^\text{22}\) Compare: why, exactly, is it that Romeo is

\(^{22}\) In a case in which we have enough background knowledge concerning the speaker’s beliefs, these expectations (Holmes is fictional, etc.) might be outweighed. Imagine, the speaker has told you on many previous occasions that her next door neighbor, “Sherlock” is a famous British detective and the subject of some non-fiction crime books by
justified in his expectation that by uttering ‘Juliet is the Sun’ he will not thereby put his audience in a position to infer that he believes that Juliet is a burning celestial object at the center of our solar system?

We see no point in postulating a special proprietary notion of linguistic responsibility and commitment. If [Commitment] is correct, the variety of responsibility we incur in speech is fundamentally no different in kind from the variety of responsibility we incur from intentionally doing what we do. As such, we should no more expect there to be precise boundaries for what we make ourselves liable for in speech, than we should for what we make ourselves responsible for by our intentional actions, generally. It is oftentimes vague whether some particular consequence of our action was, or was not, not foreseeable in advance. The case of assertion is no different since it will oftentimes be vague whether the speaker indeed thinks that her audience will be in a position to take her to believe that p as a result of her utterance, and hence, correlativey vague whether she committed to p. We do not think that such vagueness is worrisome for our view. Insofar as our intuitive judgements regarding speaker commitment go lock and step with the vagueness predicted by [Commitment], we have further confirmation of our favored view.

The foregoing picture of speaker commitment is a thoroughly intentionalist one; you are responsible for anything you represent yourself as believing by uttering what you do. Let’s now return now to the epistemological question of how we successfully recover speaker commitments to see how we might best accommodate Hoek’s important insights about the connection between subject matter/questions under discussion and conversational exculpature and commitment.

When a speaker engages with an audience that wants to know whether Q - perhaps this was already clear, or perhaps her audience explicitly asked her - and the speaker has information that could be

Doyle. With this background, we might indeed (plausibly) take her to be committing to the existence of Sherlock Holmes by her utterance of (8). Likewise, if a speaker volunteers (6a) we would be forced to weigh the relative plausibility of her believing in the existence of Sherlock Holmes versus that of her believing that one can have tea, or talk about football, with a fictional entity.

(6a) I had tea with Sherlock Holmes this morning, and we talked about football.

In this case, a literal interpretation of her utterance - and all of the commitments that come along with it - might be the best that we can do. The holistic and context-sensitive nature of IBE might seem worrisome for anyone who was hoping for a systematic, predictive algorithm for linguistic interpretation. We, however, see these aspects of our favored view as a feature, rather than a bug. Thanks to an anonymous referee for a healthy dose of skepticism regarding the unconstrained nature of IBE-based pragmatics and for the invitation to discuss examples such as (6b) above.

23 Of course, in other cases it will not be vague. Imagine, for example, a competent speaker who uses a well-known slur. Even if they do not specifically intend to represent themselves as believing that the relevant target of the slur is worthy of denigration, etc., they will almost certainly be on the hook for slurring, nevertheless. Here, as is more generally the case, if C is an obviously foreseeable consequence of your performing a certain action, you can be held to account for bringing about C whether or not you specifically intended to do just that (this is especially so when C is evaluated as having negative moral consequences). Thanks to an anonymous referee for the suggestion to discuss slurs in the context of [Commitment]. See Chapter 13 of Bratman (1986) for a discussion of intention, intentional action, and responsibility that we are sympathetic to.
of help in settling the issue, and it costs her nothing to provide it, she might, if feeling helpful, take that into consideration in planning her conversational contribution. The speaker’s beliefs regarding her audience’s interests might then play a motivational role in her choosing to utter what she does. Likewise, in some cases, a speaker’s communicative plan might be indirectly shaped by what she thinks her audience would like to discuss. For example, S might sometimes explicitly intend for her audience A to come to recognize what she meant, and/or committed to, by A’s reasoning specifically from the assumptions that (i) S understands that they want to know whether Q, (ii) that S can answer the question whether Q and (iii) S is feeling helpful. In such a case, the speaker’s beliefs about issues pertaining to conversational subject matter might play a role in her assessment of the means by which she intends to achieve her communicative goals. Analogously, an agent’s beliefs about whether the window is locked can figure in her formation of a rational plan to let more air into the room. But even in cases such as this, it is still the speaker’s intentions so formed (and other relevant mental states) that determine what she commits to rather than anything extrinsic to those states such as facts regarding what A is actually interested in or what S and A were just talking about five minutes ago (to bring this out imagine that, for e.g., that S’s beliefs about what A is interested in discussing are wrong). The speaker’s assessment of the conversational interests of her interlocutor is important, but is not determinative of what we intentionally do in speech, including what we intentionally represent ourselves as believing in acts of assertion.

On the epistemological front of utterance interpretation, our assessment of what question the speaker is seeking to address - if such there be - can sometimes importantly figure in our inference to what she was committed to in uttering what she did. But, in our view, figuring out a speaker’s commitments is a species of the general enterprise of trying to figure out what someone else is intentionally doing. Our knowledge of the intentional actions of another is always based on an inference to the best explanation given our total evidence regarding her (including the mental states we take her to have), the specific context, and her publicly observable behavior. As best we can see, the epistemological story of conversational exculpature and commitment is no different. For example, the fact that you just asked the speaker about where to buy a burrito, or that you took the conversation to be about vintage Levi’s, or… are considerations that could favor a particular hypothesis regarding what the speaker was committing to, but they must always be weighed against the overall evidence you have available in the context. If our favored view, [Commitment], is on the right track, we should be no more optimistic about finding a general algorithm, or mechanism, for determining what a speaker has committed to by her utterance than we are for recognizing what she is intentionally doing, more generally.  

---

24 Our skepticism here is based, in part, on the abductive, holistic, and highly-context sensitive nature of IBE-based reasoning. One should no more expect to find an algorithm for making such inferences than one for calculating what makes a hypothesis, as Lipton (1991) once put it, ‘loveliest’, See, for e.g., Chomsky (1986) and Fodor (2001) for some important worries regarding the possibility of a rule-based account of pragmatic interpretation.
Conclusion

As Hoek points out, in an act of assertion, we can sometimes be held liable for more than what we strictly speaking said, and sometimes less. We have argued that these liabilities are essentially tied to facts about what we represent ourselves as believing in saying what we do. When we represent ourselves as believing a particular proposition, we give our audience evidence not just of the fact that we believe it; we also indirectly give them evidence that they ought to believe it, too. If, on the basis of your say so, your audience comes to treat information that you have given them as a premise in their own practical and theoretical deliberation, you might bear some responsibility for what they go on to think or do. We think our favored account sheds some light on how, and why, we incur these responsibilities in speech. If we are correct, the full story about how, and why, we incur the responsibilities and liabilities that we do in speech, is simply one small part of the more general account of responsibility that our colleagues in action theory and ethics are seeking to uncover.  

References


---

25 We would like to thank Josh Dever, Daniel Drucker, Jon Litland, the members of Ray and Josh's course “Language, Action, and Decision” (UT Austin, Spring 2020), and the participants in the Words Workshop, especially Sarah Broadie, Jonathan Jenkins Ichikawa, Devin Morse, and Alex Davies for helpful discussion of this material. Additionally, special thanks are due to Gary Ostertag, as well as two anonymous referees for Mind and Language, for written comments on earlier drafts that greatly helped to improve the paper.


