Appendix I
Three Theories

In this appendix, we summarize three important earlier proposals (briefly mentioned in Section 5) and discuss challenges that they (and we) encounter.[1]

♦ Abusch (2002, 2010)

One can take Abusch's (2002, 2010) starting point to be the observation that focus can yield presupposition-like phenomena.[2] For instance, (1) may evoke the set of alternatives {You will vote for Biden in November, you will vote for Trump in November}. With the assumption that at least one of these alternatives is true, we obtain the inference that you will in fact vote in November.

(1) Will you vote for $[\text{Biden}]_F$ in November?
=> you will vote in November

By assuming that the items in (2) lexically activate appropriately-defined scales, Abusch can derive their presuppositional behavior.

(2) a. stop: {stop, continue}
b. win: {win, lose}
c. be right: {be right, be wrong}
d. be aware: {be aware, be unaware}

Here all expressions are taken to have a bivalent meaning, but it is presupposed that one of the alternatives is true, hence the desired presuppositional effect. This is particularly easy to see with the semantic decomposition offered in (3)a,b for *x stops smoking* and *x continues smoking*: it is immediate that their disjunction that is equivalent to *x smoked*, i.e. their joint presupposition.

(3) a. x stops smoking = x smoked and x doesn't smoke
b. x continues smoking = x smoked and x smokes
c. (x stops smoking or x continues smoking) $\Leftrightarrow$ x smoked

The immediate difficulty is that with different alternatives, different results are obtained. In (4), *continue* is given the set of alternatives {continue, start}, and as a result *x continues smoking* presupposes their disjunction, namely that *x smokes*. While this reading may be accessed in special contexts, one needs to explain why in general it is presupposed that *x smoked*, while it is *asserted* that *x still smokes*.[3]

(4) Deviant alternative set: continue: {continue, start}
a. x starts smoking = x didn't smoke and x smokes
b. x continues smoking = x smoked and x smokes
c. (x starts smoking or x continues smoking) $\Leftrightarrow$ x smokes

In Abusch's defense, there seems to be genuine arbitrariness in some cases. She mentions the exquisite contrast between *x is aware that p* and *x is right that p:* both have a believing/saying component, and a veridical component, but *x is aware that p* presupposes that *p* is true and asserts that *x believes p*, whereas *x is right that p* presupposes that *x said/believes that p* and asserts that *p* is true. The contrast naturally follows from Abusch's alternatives in (2)c,d: *aware* and *unaware* both have a veridical component, hence their disjunction does too; *be right* and *be wrong* both have a 'believing/saying' component, hence it too is preserved by their disjunction. But the lexical nature of the account seems to be essential to derive this fine-grained contrast, since there is no obvious difference

---

[1] See Schlenker (to appear) for a detailed discussion of how these theories can or cannot account for presuppositional data pertaining to iconic pro-speech gestures and ASL classifier predicates.

[2] See for instance Büring (2012) for a far more detailed discussion of the inferential effects of focus and givenness, which we do not discuss in the present paper.

[3] Special contexts involve contrastive focus, as in *Did Robin start smoking or did Robin continue to smoke?* The fact that alternative can change presuppositions when manipulated is of course consonant with Abusch's theory, but it also highlights the need for certain lexical stipulations – as Abusch herself makes transparently clear.

between the global meaning of *be aware* and *be right*.

As a result, this contrast is a direct challenge to our theory (and to any theory that triggers presuppositions on the basis of bivalent meanings). Still, there might be a useful difference between the implications of these two verbs: *x was right that p* usually conveys that *p* was controversial at the relevant time, whereas *x was aware that p* doesn't. Thus (5)b suggests, a bit oddly, that it was controversial that Obama should abide by the Constitution (by contrast, *Obama is right that Trump has to abide by the Constitution* might be more natural if one thinks that the embedded clause is now a matter of controversy); this controversiality inference is absent from (5)a,c,d.[4]

(5) a. Obama was aware that he had to abide by the Constitution.
b. ?Obama was right that he had to abide by the Constitution.
c. Obama correctly believed that he had to abide by the Constitution.
d. Obama went by his correct belief that he had to abide by the Constitution.

No matter what the precise source of this controversiality inference (and we grant that this point requires further investigation), it may affect the triggering rule. In (6), we have made explicit the controversiality component of *x is right that p* and its absence in *x is aware that p*. *F* stands for *p* (to evaluate the factive presupposition) or *x believes that p* (to evaluate the presupposition that the agent believes the complement).

(6) a. Be aware: $P(believe_{t-1}\ (F) \mid believe_{t-1}\ C\ \&\ acquire_{t}$ (p and x believes that p))
b. Be right: $P(believe_{t-1}\ (F) \mid believe_{t-1}\ C\ \&\ acquire_{t}$ (p **and p is controversial** and x believes that p)

The boldfaced conjunct in (6)b will make it less likely that, upon learning that *x is right that p*, one antecedently knows that *p* is true, with the effect that for *F = p* we should get a lower probability than in the case of (6)a. The latter is similar to the case of *x knows that p*: chances are that upon learning that both *p* and the belief that *p* hold, one had an antecedent knowledge of the state of the world. For *F = x believes that p*, there is no reason to assume a high value in (6)a. Can a high value be justified in (6)b? In case *be right* is interpreted as a speech act, such a high value makes sense: upon learning that x is right in a disagreement about p, one will often have a pre-existing belief about the disagreement and thus about x's beliefs. It's less clear that this extends to silent disagreements. But the contrast between *be aware* and *be right* as analyzed here seems to go in the right direction, although a full analysis has yet to be developed.

♦ **Simons et al. (2010)**

Simon et al. (2010) start from the (very convincing) observation that 'Questions under Discussion' (QUDs) may affect whether a presupposition arises or not. For instance, in (7) the QUD is whether France has a king, and as a result the definite description fails to presuppose that there is king, as it would be bad conversational practice to presuppose what's under discussion.

(7) –Does France have a king?
–Well, the king of France didn't attend the opening of Parliament.

This fact alone can be derived by every theory that allows for local accommodation. Simons et

---

[4] Two remarks should be added.
1. (5)b might improve if replaced with: *Obama said that he had to abide by the Constitution, and he was right*. But the underlined conjunct yields an inference that the point was non-obvious at the time, which might satisfy the controversiality inference of *be right*.
2. A natural question is whether the deviance of (5)b is just due to the at-issue nature of the veridical implication, whence it should be non-trivial *at the time of speech* that Obama had to abide by the Constitution (when he was President). Our impression is that (5)b implies that it was controversial *at the time* that Obama had to abide by the Constitution. This implication seems to us to be absent from (i), where the non-triviality requirement holds at most of the time of speech.

(i) Obama had to abide by the Constitution and he was aware of it.

The control in (5)d is intended to show that the presupposition of *x is right that p*, to the effect that *x believes or says that p*, is not responsible for the deviance of (5)b (note also that the presupposition may just be about a belief, as in: *Obama believed that he had to abide by the Constitution, and he was of course right (that he had to))*.

al. go much further in suggesting that entailments may 'project' beyond operators, and may thus behave like presuppositions, in case they are not relevant to the QUD, thanks to the definitions in (8):

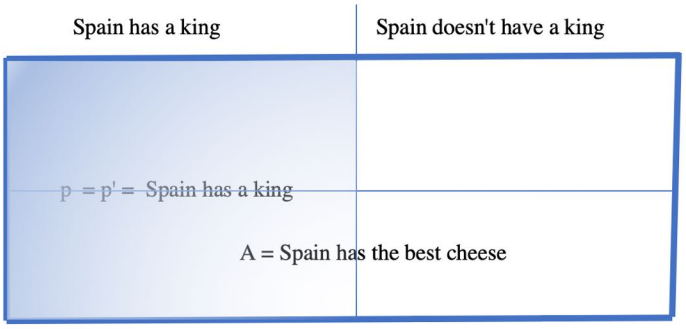
(8) Definition of relevance and at-issueness in Simons et al. (2010)
a. An assertion is relevant to a QUD iff it contextually entails a partial or complete answer to the QUD.
b. A question is relevant to a QUD iff it has an answer that contextually entails a partial/complete answer to the QUD.
c. A proposition p is at-issue relative to a question Q iff the question ?p is relevant to Q.

To illustrate, in (9) we initially treat *know* as bivalent and veridical (it entails the truth of its complement) and note that the entailment *you can eat raw vegetables* is not relevant to the QUD, and thus it is not at-issue and projects; things are different with the entailment that France has a king in (7).

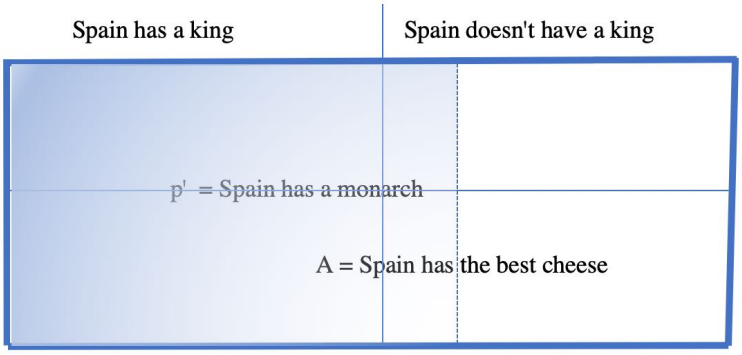
(9) Background scenario: a nutritionist has been visiting first grade classrooms to talk to the children about healthy eating
Q: What most surprised you about the first graders?
A: They didn't know that you can eat raw vegetables.

The difficulty, noted in Chemla (2006), is that this account is insufficiently predictive and/or makes impossible predictions. Suppose *p* provides an answer to Q, and take *p'* to be an entailment of *p* (*p'* could be identical to *p*, as in the case illustrated in (10)). Find a proposition *A* that partly overlaps with every cell of Q (so that *not A* also overlaps with every cell of Q). For instance, take a 2-cell QUD similar to (7), *Does Spain have a king?*, with a simple answer, *p = Spain has a king* (which does not trigger a presupposition). Take *A* to be: *Spain has the best cheese in the world*. This certainly overlaps with each of the two cells. Take an entailment *p'* of *p*, either *p* itself, or something weaker, for instance *Spain has a monarch*. Now *p'* has *(p' or A)* as an entailment, and it is not relevant to Q, so *(p' or A)* can project. Similarly, *(p' or not A)* can project. If both do, then we get a presupposition that *p'*: it now becomes presupposed that Spain has a monarch. The reasoning that leads to the prediction that *p* should project is illustrated in (10). (11) is similar but pertains to a mere entailment *p'* of *p*. Since the reasoning is general, every entailment *p'* of *p* is predicted to project – which couldn't be correct, and calls, at a minimum, for some constraints.

(10) QUD corresponding to *Does Spain have a king?* The answer is p = *Spain has a king*.
With A = *Spain has the best cheese*, *(p or A)* as well as *(p or not A)* are entailments that fail to address the QUD and should thus project, hence *p* should project as well.



(11) QUD corresponding to *Does Spain have a king?* The answer is now p = *Spain has a king*. We consider an entailment *p'* = *Spain has a monarch*. With A = *Spain has the best cheese*, *(p' or A)* as well as *(p' or not A)* are entailments that fail to address the QUD and should thus project, hence *p'* should project as well.



On the assumption that this account can be constrained, it still encounters another problem, discussed by Abrusán (2011) (and noted by the authors themselves). Abrusán notes that the account incorrectly predicts that presuppositions should fail to be generated in (12) and (13), since the fact that the first graders failed the exam, or that John used to smoke, are relevant to the QUD; but intuitively the presuppositions are generated in this case.

(12) Q: What most surprised you about the first graders?
A: They didn't know that they have failed the exam.

(13) Q: What do you know about John?
B: He still didn't quit smoking.

♦ **Abrusán (2011)**

❑ *Insights*

After criticizing Abusch's and Simons et al.'s accounts, Abrusán (2011) proposes a triggering algorithm for the verbal case. For her, entailments that are not 'about' the event time of the verb get presupposed (this is motivated by considerations on attention, as only entailments that are about the event time are taken to be the main point of such constructions). We won't be concerned with how 'aboutness' is defined.[5] In the case of change of state verbs, which will be of some importance below, Abrusán's theory works as in (14).

(14) a. John stopped smoking at t_1.
b. Entailment 1: John does not smoke at t_1
c. Entailment 2: John smoked at t_2 (where t_2 is some contextually given interval before t_1)

The simplified representation in (14)a, with event time t_1, comes with several entailments, two of which are stated in (14)b,c. Entailment 1 is about event time t_1 and thus it does not get presupposed. By contrast, Entailment 2 is not about event time t_1 and thus it gets presupposed.

Our analysis agrees with Abrusán's in some but not in all cases. Take for instance change of state verbs. For Abrusán, *John stopped smoking at* t_1 triggers a presupposition that John smoked before t_1 because this entailment is not about the matrix event time. For us, the intuition is a bit different: upon learning that John doesn't smoke at t_1 but did before, one would typically antecedently know the part about the past, namely that John smoked before t_1. We discuss two types of counterexamples to Abrusán's theory. The main problem is that her analysis overgenerates because it treats as presupposed some entailments which are not. From our perspective, this happens when an entailment *p* of an expression *pp'* is not about the matrix event time (hence Abrusán predicts a presupposition), and yet

[5] In a nutshell, a sentence S is *not* about an individual [named by] c just in case for every interpretation M, S is true in M iff for every interpretation M' which is just like M for atomic sentences, except possibly ones that contain c, S is true in M'. Abrusán (2011) discusses various necessary refinements to obtain a theory of presupposition generation. In particular, she needs to treat the matrix and the embedded time arguments in *John knows (at* t_1*) that it is raining (at* t_1*)* as counting as different so as to differentiate between entailments that should vs. shouldn't become presupposed).

when one learns that *pp'*, one doesn't typically antecedently know that *p* (so the present analysis predicts no presupposition). In other cases, Abrusán's theory undergenerates, because an entailment *p* of an expression *pp'* is about the matrix event time (hence it is at-issue for Abrusán), and yet when one learns that *pp'*, one typically antecedently knows that *p*.

❑ *An overgeneration dilemma*

As mentioned by N. Klinedinst (p.c.), a class of verbs discussed in Anand and Hacquard (2014) presents systematic exceptions to Abrusán's analysis, as illustrated in (15) (*establish* could be added to the class, as noted by N. Klinedinst):

(15) a. The bloody gloves {demonstrate, imply, prove, show} that Mary committed the murder.
=> Mary committed the murder.
b. Do the bloody gloves {demonstrate, imply, prove, show} that Mary committed the murder?
≠> Mary committed the murder.
(modified from Anand and Hacquard 2014)

It is clear that the entailment about the past murder is not about the matrix event time of the construction and is thus predicted by Abrusán to be presupposed, contrary to fact.

Abrusán's theory can posit that these constructions do not entail their propositional arguments to begin with, as shown by examples such as (16), where the embedded clause is not presented as true (Abrusán, p.c.). Since Abrusán's theory only intends to turn *lexical* entailments into presuppositions, it does not predict presuppositions in this case.

(16) That article {demonstrates, implies, proves, shows} that Obama was born in Africa.
≠> Obama was born in Africa

But in view of the cases of contextual triggering discussed in Section 4, these data present a dilemma for Abrusán's theory: if it sticks to the view that presuppositions are triggered from lexical rather than contextual entailments, it fails to account for the diverse examples we discussed where contextual information is crucial. If it allows itself to take as input contextual entailments, it will overgenerate in the case of (15).[6]

From the present perspective, when one learns that *x demonstrates* (or: *implies, proves, shows, establishes*) *that p, p* is often hard/non-trivial, and one often does *not* antecedently know that *p*; one might learn *p by* learning that x demonstrates it. In other words, it makes sense to assume that the probability in (17) is relatively low, and if so we do not expect *p* to count as an epistemic precondition

---

[6] In addition, an algorithm that works on top of lexical entailments alone might encounter difficulties in other cases, since it is possible to use *know* without a factive entailment, with the meaning of *be absolutely certain that*, as in the real life examples in (i); or Abrusán's (ii) (*Very Serious People* is a mocking title used to refer to pundits who are certain of facts that are false); the challenge will thus be to draw a principled distinction between *know* and the examples in (25) in the main text.

(i) a. Warsh is a Very Serious Person, and all Very Serious People know that deficits are bad. (Tim Duy)
(retrieved on December 24, 2019 at https://economistsview.typepad.com/timduy/2017/10/kevin-warsh-very-serious-person.html)
b. But the belief that America suffers from a severe "skills gap" is one of those things that everyone important knows must be true, because everyone they know says it's true. (Paul Krugman)
(retrieved online on December 24, 2019 at https://www.nytimes.com/2014/03/31/opinion/krugman-jobs-and-skills-and-zombies.html?ref=opinion)

(ii) The keys were not in the drawer but she knew that they were there, so she foolishly kept on searching. (Abrusán 2011).

A proponent of Abrusán's theory might argue that these examples involve a variety of context shift, be it Free Indirect Discourse or 'protagonist projection', investigated in Abrusán (to appear). But in both varieties, shifting is blocked by unshifted readings of *yesterday* (i.e. readings from the speaker's rather than from the agent's perspective). Still, it seems to us that sentences like (iii) are compatible with a non-factive reading of *know*, which casts doubt on this explanation.

(iii) Yesterday evening, Trump just knew that yesterday morning/the day before yesterday Biden had plotted to overthrow him.

of *x demonstrates that p*.

(17) $P(believe_{t-1}(p) \mid believe_{t-1} C \,\&\, acquire_t (p \text{ and } x \text{ demonstrates that } p))$

The same analysis extends to (3)b in the main text (= PULL-GUN-SHOOT), which is another case of overgeneration: here the initial state (= the gun's presence under the jacket) is predicted to be presupposed but isn't, because upon learning that someone pulls a gun from his jacket and shoots, one typically doesn't antecedently know that the person has a gun.

Further problems arise when we consider presuppositions that are clearly triggered from contextual entailments. Take *x is pregnant* and *x is a doctor*: in both cases, common sense knowledge yields an entailment that something relevant happened before the matrix event time: *x was impregnated at least 5 days ago, x went to medical school*. If we adopted a context-sensitive version of Abrusán's algorithm, these entailments would be predicted to be presupposed, but in fact neither is.[7] From our perspective, this is because neither is an epistemic precondition of the target proposition: upon learning that *x is pregnant*, one typically does not antecedently know that *x was impregnated at least 5 days ago;* similarly, when one learns that *x is a doctor,* one typically does not antecedently know that *x went to medical school;* rather, it is *by* learning that *x is a doctor* that one acquires the belief that *x went to medical school*.

(18) a. Mary is pregnant. => Mary was impregnated at least 5 days ago
b. Is Mary pregnant? ≠> Mary was impregnated at least 5 days ago

(19) a. Mary is a doctor. => Mary went to medical school
b. Is Mary a doctor? ≠> Mary went to medical school

The case of *be pregnant* is straightforward: it is clear that in general upon learning that someone is pregnant, one does not have antecedent knowledge that this person was impregnated at least 5 days before: the probability in (20) is plausibly low.

(20) $P(believe_{t-1}(x \text{ was impregnated at least 5 days ago}) \mid believe_{t-1} C \,\&\, acquire_t (x \text{ is pregnant}))$

Abrusán's theory correctly predicts that in change of state constructions, the initial state gets presupposed, as we illustrated at the outset in (8) in the main text. But what about the final state? She argues that such inferences are in fact pragmatic rather than lexical.[8] But once the possibility of contextually triggered presuppositions is taken into account, it becomes harder to explain why entailments about the final state are not presupposed.

Take the sentence *Will the police blow up this suspicious package?* Certainly, *x blows up y* contextually entails that, at a later point, y is destroyed. Furthermore, such a presupposition wouldn't make the question idle – it would in effect be asking about the manner of destruction. But this is not how the question is understood: the final state just isn't presupposed.[9] From the present perspective, the asymmetry between initial and final states is relatively expected: upon learning that x blows up y, one typically has greater antecedent knowledge of the past than of the future.

---

[7] Abrusán (2011) did not claim to have a general triggering algorithm, and thus we are considering here potential extensions of her proposal.

[8] Abrusán argues that the contrast between (ia) and (ib) suggests that *smoked before* is entailed but *didn't smoke after* isn't. But (ib) leaves open the possibility that there was a small interval between the stop and the new start. (ii) is thus a better test: Abrusán expects (maybe correctly) that (ii)b should be more acceptable than (ii)a.

(i) a. #John stopped smoking, but he has never smoked before.
b. John stopped smoking, but then he started again.

(ii) a. #(At 12:05pm sharp,) the temperature stopped being negative, but was never negative before.
b. (At 12:05pm sharp,) the temperature stopped being negative, but was never non-negative after.

[9] The same point could be made with the verbal gesture TAKE-OFF-ROTATING, used in Schlenker (to appear) to represent a helicopter take-off. There is certainly a contextual entailment that the helicopter will end up in the air, but it is not presupposed: *At 12:05, will the company's helicopter* TAKE-OFF-ROTATING? definitely doesn't presuppose that shortly after 12:05 the helicopter will be in the air.

❑ *Undergeneration*

There are also cases in which Abrusán's theory arguably undergenerates, and treats as at-issue some entailments that are in fact presupposed. One such case is discussed by Abrusán (2011): it pertains to *regret*. (21)a triggers the inference that John believes that he worked as a linguist (this might come in addition to a factive presupposition to the effect that John in fact worked as a linguist[10]). But this inference pertains to the matrix event time, so without a refinement, Abrusán's theory predicts that it should not be presupposed, contrary to fact.

(21) a. Does John regret that he worked as a linguist?
b. => John believes that he worked as a linguist

Abrusán (2011) proposes two potential solutions. One is that *regret* is a normal factive verb (presupposing the truth of its complement), but that it allows for Free Indirect Discourse uses in which "it reports an attitude of a subject towards facts as perceived by him". However the mere *possibility* of a Free Indirect Discourse reading does not explain the 'belief' entailment that is in fact obtained.[11] More relevantly, Abrusán considers the possibility that there is indeed a belief entailment, but that it is not about the matrix event time. Rather, it is about a time right before the matrix event time, so that this entailment can be presupposed; pragmatic reasoning then leads to the inference that the entailment still held at the matrix event time. This raises the same problem as before about the contextual entailments: why does the output of pragmatic reasoning not feed the triggering mechanism in this case although it does in other cases?

In addition, an analysis based on pragmatic enrichment does not seem right: in a situation in which a person's beliefs keep changing because of memory problems, *x regrets that p* still comes with the presupposition that *x believes that p*, not with the (contextually plausible) presupposition that *x believed that p:*

(22) John has Alzheimer's and spends his time forgetting about his past career in linguistics.
Does John regret that he worked as a semanticist?
=> John believes that he worked as a semanticist

In (22), the lexical part of Abrusán's analysis of *regret* predicts that right before the time of utterance John believed that he had been a semanticist. There is no pragmatic reason to strengthen this inference to encompass the time of utterance, since John keeps forgetting. But the inference still goes through, which suggests that it is lexically triggered by *regret*.

From the present perspective, can we explain why *x regrets that p* doesn't just entail but in fact presupposes that *x believes that p*? The question is as follows: on the assumption that one has learned that *x believes that p and x regrets that p,* is it usually the case that one antecedently knows that *x believes that p*? This seems plausible on the assumption that one usually knows more about a person's beliefs than about their desires. This makes sense in view of the fact that a person's beliefs can be inferred from the information they have access to, in particular the state of the world, whereas their desires are more complex to infer. This would need to be explored in greater detail, but in any event there is nothing in the present theory to block presupposition generation in this case, unlike in Abrusán's

---

[10] We write that the factive presupposition *might* come in addition to the belief presupposition because an alternative possibility is that it derives from it. In a nutshell, many modalized presuppositions become factive when their modal character is not explicitly justified (this is known as the 'Proviso Problem', discussed for instance in Lassiter 2012). This arises for instance with *want*: out of the blue, *Ann doesn't want me to stop smoking* triggers the inference that I smoked before; but when the target sentence is preceded by *Ann wrongly believes that I am a smoker, and __* , the factive inference disappears.

[11] As mentioned in fn. 6, Abrusán (to appear) adds a further possibility, namely what she calls 'protagonist projection', which has slightly different properties from Free Indirect Discourse. This doesn't seem to change the dialectical situation: the mere *possibility* of protagonist projection doesn't explain why the inference in (21)b is nearly obligatory. In addition, our impression is that the existence of the 'belief' inference isn't affected by the addition of expressions that block protagonist projection/Free Indirect Discourse, such as an unshifted uses of *yesterday* or of an expressive, as in (i).

(i) Up to this point, Sam hasn't regretted praising his idiotic assistant during yesterday morning's meeting.

theory.

We believe that there are further cases of undergeneration in Abrusán's theory. *At t, x deters y from q-ing* presupposes that at t there is a risk that y harms x through q (the details don't matter as long as there is a non-trivial inference *about t*). For instance, *In 2016, did the Obama administration deter Russia from influencing the election?* presupposes that in 2016 Russia tried, or at least had a possible desire, to influence the election; it isn't clear how the strategies deployed by Abrusán for *regret* will work in this case.[12]

Within our theory, the analysis of *deter* requires that we assess the probability in (23). Now it is plausible that one usually knows about threats before learning about actions take to deter them, which might explain why this probability is high and the entailment gets presupposed.

(23) $P(believe_{t-1}$ (there is a risk that y harms x through q) $|$ $believe_{t-1}$ C & $acquire_t$ (x deters y from q-ing)

In sum, there are systematic cases in which Abrusán's system overgenerates, and these can in principle be avoided by the present theory; there also seem to be cases of undergeneration, which seem easier to avoid within the present framework.[13]

---

[12] See Schlenker (to appear) for further potential problems pertaining to ASL classifier predicates.

[13] Two remarks should be added about earlier accounts.

1. Anand and Hacquard 2014 seek to explain why verbs of the *inform* class often presuppose the truth of their complement. They note that these verbs refer to "discourse moves which lead to the acceptance of the complement *p* into the common ground of the reported discourse", and they propose that "this acceptance of *p* can easily bleed into the actual common ground (...): an illusion of factivity arises whenever a reported context is taken to faithfully represent the conversational community in the world of evaluation." This line of analysis raises two questions. First, why do verbs of this class yield a presupposition that that the complement is true relative to the original conversation? Second, why does that presupposition get 'imported' into the reporting conversation? No general account is offered yet. In addition, this line of explanation won't extend to presuppositions that are triggered outside of speech reports, unlike the one we develop in this piece.

2. There might be a connection between the analysis developed in this piece and the theory of presupposition as abduction sketched in Thomason et al. (2006) (see also Hobbs et al. 1993). As an anonymous reviewer once noted, their idea is to find "the most probable explanation for what was said". "Typically if the utterance contains an expression that does not make sense unless p is true, there is a good case for adding p to the speaker's beliefs". Thomason et al. 2006 frame their proposal within a dynamic theory and thus (we believe) they too can relativize triggering to a local context. Furthermore, theirs is clearly a communicative theory. The challenge, of course, is to find a non-circular way to define what it means for an expression to "not make sense" as opposed to "be false". Concretely, in (3)a (*PICK-UP-GUN-SHOOT*) and (3)b (*PULL-GUN-SHOOT*) alike (in the main text), the presence of a gun can be inferred, but only in the former case does it seem to be clearly presupposed. In (2) in the main text, the gesture *UNSCREW-ceiling* provides iconic information about an action of unscrewing a bulb from the ceiling, but only one component (the presence of the bulb on the ceiling) is presupposed. Thomason et al. explicitly state that "utterances, like other actions, may depend for their effects on preconditions about the situation in which they are performed"; as we understand things, the challenge is to define in a general fashion what these preconditions are.

Appendix II
Formal Illustration

In this appendix, we illustrate the presupposition triggering rule developed in Section 6.2 (copied in (24)) by providing a possible worlds interpretation in which an individual x believes a proposition p just in case each of the worlds compatible with what x believes satisfies p; in other words: x's epistemic base entails p (a standard assumption in possible world semantics).[14]

(24) **Presupposition triggering relative to a context (propositional case)**
For a (contextually) given probability threshold *a*, for a propositional expression *E* in context c', for random time variables t and t', trigger a presupposition *p* if:
(i) c' |= E => p, and
(ii) P(believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**) ≥ a

where P (• | _) is the subjective conditional probability of • given _ and **E** and **p** are the semantic values of *E* and *p* respectively (when no confusion arises, we will forego boldfacing).[15]

In words: trigger a presupposition *p* from an expression *E* in context c' if:
(i) *E* contextually entails *p* relative to c', and
(ii) if one antecedently believed c' and acquires the believe that *E*, there is a high enough chance (above threshold *a*) that one antecedently believed p.

We must thus reason probabilistically about certain complex events, cases in which one antecedently believed at t-1 a proposition c' (corresponding to the local context of the target expression), and one acquired at t a belief that **E**, the proposition corresponding to the target construction. Take the sentence *Will Sally TURN-WHEEL?*, which presupposes that Sally is in front of a wheel (in a car situation, she is in the driver's seat). For simplicity, we focus on just one epistemic agent x, and on the times t and t-1. We need to consider all possible events in which (i) x initially believed c' and then (ii) x acquired the belief that Sally turned a wheel in front of her. Within our possible worlds interpretation of belief, (i) means that the agent's epistemic base at t-1 was included in the set of worlds corresponding to c', represented in all our illustrations below by the outermost rectangle. (ii) means that the agent's epistemic base at t-1 is not included in **E**, but that it becomes included in **E** at t. **E** will be represented as a colored rectangle in the upper left corner of c'. We assume that information grows monotonically, and thus beliefs at t are more specific than beliefs at t-1. Thus we consider all cases in which the agent's epistemic base was initially within c' but was not included in **E** at t-1, and then became included in **E** at t. For simplicity, we will assume throughout that at t the agent's epistemic base is identical to (rather than just included in) **E**, and hence everything will hinge on the agent's initial epistemic base at t-1.

---

[14] A reviewer asks whether our analysis leads to the prediction that truths of arithmetic are presupposed. If one adopts a possible worlds framework in the analysis of *believe$_{t\text{-}1}$ **p***, as is done for concreteness, the answer is 'yes', at least if truths of arithmetic are true in every possible world. But this problem is entirely due to the lack of fine-grainedness of possible world semantics, and it already arises in all standard treatments of presuppositions. For instance, in the analysis of Heim (1983), the context set is a set of possible worlds, and whatever follows from the context set is presupposed. As a result, all truths of arithmetic are presupposed.

[15] As hinted by a reviewer, we could also state that a presupposition is triggered if the probability that one antecedently believed p is 'much greater' than the probability that this wasn't the case, as stated in (i*), where *b* is the parameter giving the difference between the two probabilities (the crucial parts are boxed)

(i*) Alternative version of (24)(ii): P(believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**) ≥ P(not believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**) + b

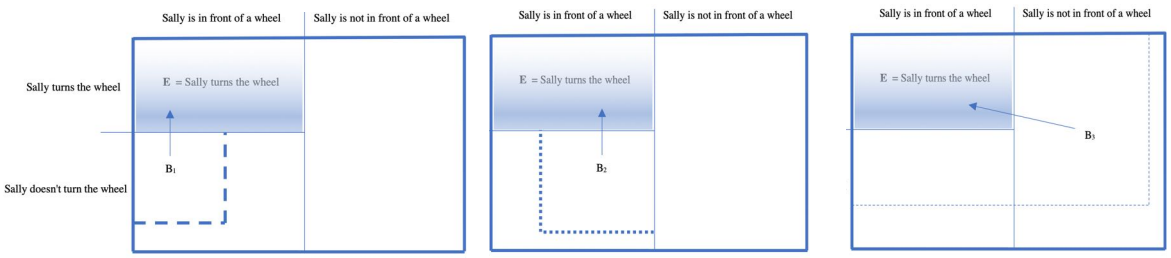
Since P(believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**) + P(not believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**) = 1, this alternative version can be rewritten as (ii*) below:

(ii*) P(believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**) ≥ (1 - P(believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**)) + b
i.e. P(believe$_{t\text{-}1}$ **p** | believe$_{t\text{-}1}$ c' & acquire$_t$ **E**) ≥ (1+b)/2

In other words, setting a = (1+b)/2, (i*) above is equivalent to (24)(ii).

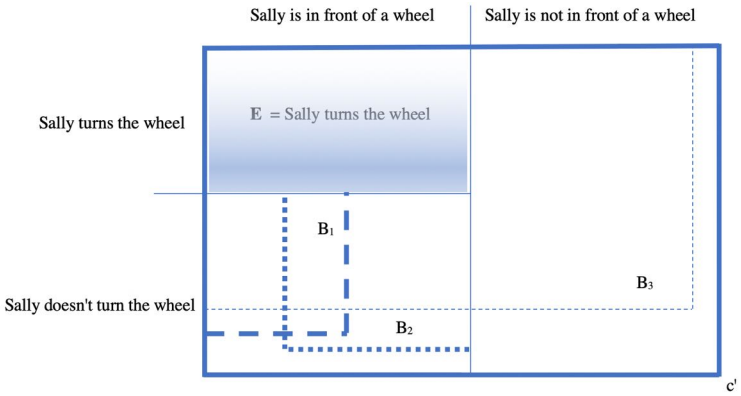
To take a concrete case, three (disjoint) events of belief change are depicted in (25). **E** corresponds to the set of possible worlds of c' in which Sally turns a wheel in front of her. The left-most part of c' corresponds to worlds in which Sally is in front of a wheel (**E** is a strict subset of that area), while the right-most part of c' corresponds to worlds in which Sally isn't in front of a wheel. The initial belief states are $\mathbf{E}\cup B_1$ (in event e_1), $\mathbf{E}\cup B_2$ (in event e_2), and $\mathbf{E}\cup B_3$ (in event e_3).

(25) 3 events in which one (i) initially believes c', and (ii) acquires the belief that Sally turned a wheel in front of her: in e_1, the agent's epistemic base contracts from $\mathbf{E}\cup B_1$ to **E**; in e_2, from $\mathbf{E}\cup B_2$ to **E**; in e_3, from $\mathbf{E}\cup B_3$ to **E**



The three initial belief states $\mathbf{E}\cup B_1$, $\mathbf{E}\cup B_2$ and $\mathbf{E}\cup B_3$ are represented in one and the same figure in (26).

(26) 3 events in which one (i) initially believes c', and (ii) acquires the belief that Sally turned a wheel in front of her



Our three events of belief change thus differ only with respect to the initial epistemic state (the posterior epistemic state always ends up being **E**): for $i\in\{1, 2, 3\}$, in event e_i, the agent initially had belief $\mathbf{E}\cup B_i$ (at t-1), and then acquired belief **E** (at t).

The key in (26) is that two of the initial belief states, namely $\mathbf{E}\cup B_1$ and $\mathbf{E}\cup B_2$, are included in the set of worlds in which Sally is in front of a wheel (on the left-hand side of the rectangle): these are cases in which the agent initially knew that Sally was in front of a wheel. But there is also one initial belief state, namely $\mathbf{E}\cup B_3$, where this is not the case: the corresponding area is not entirely within the left-hand side of the rectangle.

To apply our analysis, we must establish a probability threshold for *a* in (24), and assign probability weights to the three events e_1, e_2 and e_3. We will initially take the threshold *a* to be at 80%, and we will discuss two assignments of probabilities, one (= Scenario 1) in which the desirable result is obtained (*Will Sally TURN-WHEEL?* presupposes that Sally was initially in front of a wheel), and one (= Scenario 2) in which it is not.

All that matters when computing the conditional probability in (24)(ii) is the relative probability weight of e_1, e_2, e_3 relative to the disjunction of these events. For instance, $P(e_1 \mid e_1 \text{ or } e_2 \text{ or } e_3)$ is just $P(e_1) / (P(e_1 \text{ or } e_2 \text{ or } e_3)$, and since e_1, e_2, e_3 are mutually exclusive, this is just $P(e_1) / (P(e_1) + P(e_2) + P(e_3))$. We simplify the discussion by taking $P(e_1) + P(e_2) + P(e_3) = 1$, i.e. by taking e_1, e_2 and e_3 to be exhaustive (nothing hinges on this as long as the proportions $P(e_i)/((P(e_1) + P(e_2) + P(e_3))$ for $i = 1, 2, 3$ are as we specify). In Scenario 1, the probabilities of our three events of belief change are specified as in (27).

(27) Scenario 1: it is presupposed that Sally was in front a wheel (probability threshold: $a = 80\%$)
e_1 (= the agent initially had belief $\mathbf{E}\cup B_1$ and acquired belief $\mathbf{E}$) has probability 40%
e_2 (= the agent initially had belief $\mathbf{E}\cup B_2$ and acquired belief $\mathbf{E}$) has probability 40%
e_3 (= the agent initially had belief $\mathbf{E}\cup B_2$ and acquired belief $\mathbf{E}$) has probability 20%

In this scenario, upon acquiring the belief that $\mathbf{E}$, there is a 40% chance that one antecedently believed $\mathbf{E}\cup B_1$, but this doesn't rise above the 80% threshold that would trigger a presupposition, and similarly for $\mathbf{E}\cup B_2$. However, there is an 80% chance that one antecedently believed $\mathbf{E}\cup B_1\cup B_2$, since both antecedent belief states $\mathbf{E}\cup B_1$ and $\mathbf{E}\cup B_2$ are included within $\mathbf{E}\cup B_1\cup B_2$. As a result, $\mathbf{E}\cup B_1\cup B_2$ ought to be presupposed, and since $\mathbf{E}\cup B_1\cup B_2$ is included in the set of possible worlds in which Sally is in front of a wheel, this should be presupposed.

There are of course weaker propositions that are marked to be presupposed as well: c' itself, and also $\mathbf{E}\cup B_1\cup B_2\cup B_3$, since (upon acquiring the belief that $\mathbf{E}$) there is a 100% chance that one antecedently believed that proposition. But since $\mathbf{E}\cup B_1\cup B_2$ is strictly stronger than $\mathbf{E}\cup B_1\cup B_2\cup B_3$ (or c', for that matter), only the stronger presupposition makes itself felt.

It can be noticed that $\mathbf{E}\cup B_1\cup B_2$ is a bit stronger than just *Sally is in front of a wheel*. This need not be a problem since typically further information is presupposed as well in this case, e.g. pertaining to the size and position of the wheel.

In Scenario 2, defined as in (28), it should not be presupposed that Sally is in front of a wheel.

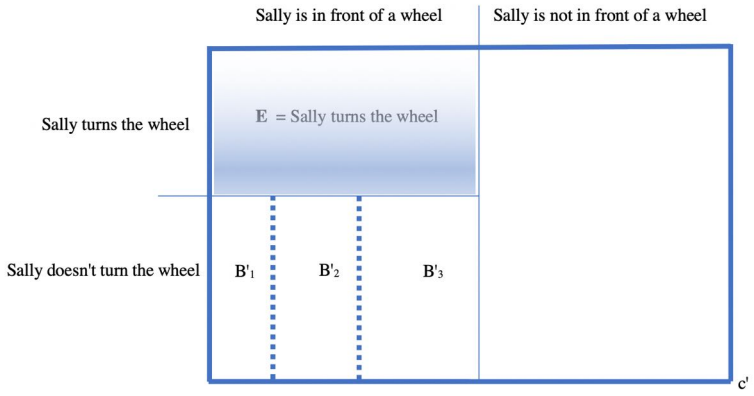
(28) Scenario 2: it is not presupposed that Sally was in front a wheel (probability threshold: $a = 80\%$)
e_1 (= the agent initially had belief $\mathbf{E}\cup B_1$ and acquired belief $\mathbf{E}$) has probability 30%
e_2 (= the agent initially had belief $\mathbf{E}\cup B_2$ and acquired belief $\mathbf{E}$) has probability 30%
e_3 (= the agent initially had belief $\mathbf{E}\cup B_3$ and acquired belief $\mathbf{E}$) has probability 40%

Crucially, the chance that (upon acquiring the belief that $\mathbf{E}$) one antecedently believed $\mathbf{E}\cup B_1\cup B_2$ is only of 30%+30% = 60%, which does not rise above the 80% threshold. Similarly, the probability that one antecedently believed $\mathbf{E}\cup B_1\cup B_3$ is only of 70%, and similarly for $\mathbf{E}\cup B_2\cup B_3$. There is a 100% chance that one antecedently believed $\mathbf{E}\cup B_1\cup B_2\cup B_3$, but this area contains worlds in which Sally isn't next to a wheel, so it is not presupposed that Sally is in front of a wheel.

The claim that *Sally* TURN-WHEEL presupposes that Sally is next to a wheel should thus be translated into a claim that, in view of general knowledge, standard discourse situations are like Scenario 1 and not like Scenario 2.

Many more cases would need to be discussed in future research. Some cause problems for the present theory, such as Scenario 3, illustrated in (29)-(30):

(29) A problematic case with three antecedent belief states



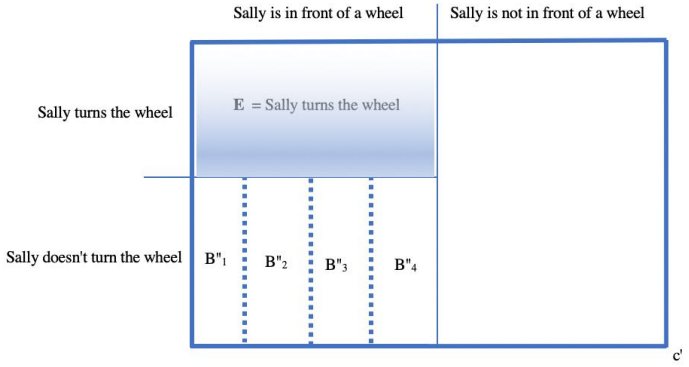
(30) Scenario 3, with B'_1, B'_2 and B'_3 as in (29), and probability threshold $a = 60\%$
e'_1 (= the agent initially had belief $\mathbf{E}\cup B'_1$ and then acquired belief $\mathbf{E}$) has probability 30%
e'_2 (= the agent initially had belief $\mathbf{E}\cup B'_2$ and then acquired belief $\mathbf{E}$) has probability 30%
e'_3 (= the agent initially had belief $\mathbf{E}\cup B'_3$ and then acquired belief $\mathbf{E}$) has probability 40%

Since we are assuming a threshold of $a = 60\%$, we have that $\mathbf{E}\cup B'_1\cup B'_2$, $\mathbf{E}\cup B'_1\cup B'_3$, and

$\mathbf{E}\cup B'_2\cup B'_3$ are all selected to be presupposed. The reason is this: there is a 60% chance that one antecedently believed $\mathbf{E}\cup B'_1\cup B'_2$, since this is the case both in event e'_1 and e'_2, whose probability weights add up to 60%. Similarly, there is a 70% chance that one antecedently believed $\mathbf{E}\cup B'_1\cup B'_3$, and also a 70% chance that one antecedently believed $\mathbf{E}\cup B'_2\cup B'_3$. As a result, the following propositions are all selected to be presupposed: $\mathbf{E}\cup B'_1\cup B'_2$, $\mathbf{E}\cup B'_1\cup B'_3$, and $\mathbf{E}\cup B'_2\cup B'_3$. Therefore their intersection must be presupposed as well. But their intersection is nothing other than **E**.[16] This cannot be, as this would go against a pragmatic principle that mandates that every expression is non-trivial in its local context. But then we are caught in a dilemma: either no presupposition is predicted at all, despite the fact that it is in this case certain that upon acquiring the belief that **E** one antecedently believed that Sally is in front of a wheel; or we predict a presupposition that is overly strong, making the target expression itself trivially true.

We can easily generalize the problem for cases in which the threshold is higher, e.g. $a = 75\%$, as in (31)-(32).

(31) Another problematic case with four antecedent belief states



(32) Scenario 4, with $B"_1$, $B"_2$ and $B"_3$ and $B"_4$ as in (31), and probability threshold $a = 75\%$
$e"_1$ (= the agent initially had belief $\mathbf{E}\cup B"_1$ and acquired belief **E**) has probability 25%
$e"_2$ (= the agent initially had belief $\mathbf{E}\cup B"_2$ and acquired belief **E**) has probability 25%
$e"_3$ (= the agent initially had belief $\mathbf{E}\cup B"_2$ and acquired belief **E**) has probability 25%
$e"_4$ (= the agent initially had belief $\mathbf{E}\cup B"_2$ and acquired belief **E**) has probability 25%

In Scenario 4, there is a 75% that (if one believed c' and acquired the belief that **E**), one antecedently believed $\mathbf{E}\cup B"_2\cup B"_3\cup B"_4$, and similarly for $\mathbf{E}\cup B"_1\cup B"_3\cup B"_4$, $\mathbf{E}\cup B"_1\cup B"_2\cup B"_4$, and $\mathbf{E}\cup B"_1\cup B"_2\cup B"_3$, so each of those is selected to be presupposed, hence their intersection should be presupposed as well. But their intersection is none other than **E**.

We do not know of any principled way to block this state of affairs, or to argue that it cannot arise.[17]

---

[16] Note that $((\mathbf{E}\cup B'_1\cup B'_2)\cap(\mathbf{E}\cup B'_1\cup B'_3))\cap(\mathbf{E}\cup B'_2\cup B'_3) = \mathbf{E}\cup B'_1\cap(\mathbf{E}\cup B'_2\cup B'_3) = \mathbf{E}$.

[17] E. Chemla (p.c.) suggests that in situations in which **E** is (absurdly) predicted to be presupposed, the threshold a must be raised. We leave this possibility for future research.

One separate issue that could be discussed in the future pertains to the relation between the triggering rule in (24) and presupposition accommodation. The reason is this: when we apply the rule in Scenario 1 (= (27)), we trigger a further constraint on c', namely that it should be strengthened (i.e. set-theoretically reduced) to a context c'^{+} so as to entail the presupposition **p**, i.e. that Sally is in front a wheel. This result could be achieved by global accommodation, i.e. by strengthening the global context C, or possibly by local accommodation, i.e. by replacing c' with c'^{+} without modifying C. In either case, one could stop here, or one could further require that the triggering rule still yield an appropriate result when applied in c'^{+}. But this further requirement would be automatically satisfied: since by construction c'^{+} entails **p**, it will be the case that: $P(\text{believe}_{t-1}\ \mathbf{p} \mid \text{believe}_{t-1}\ c'^{+}\ \&\ \text{acquire}_t\ \mathbf{E}) = 1 \geq a$.

Appendix III
Exploitation

In the main text (in (29)), we posited in essence that **p** must be treated as presupposed relative to its local context c' as soon as P(believe$_{t-1}$ **p** | believe$_{t-1}$ c' & acquire$_t$ **pp'**) ≥ a. But one could start instead from an analysis in which P(believe$_{t-1}$ **p** | believe$_{t-1}$ c' & acquire$_t$ **pp'**) ≥ a only yields a requirement that, relative to its local context c', *p* has a probability ≥ a of being entailed by that context. Still, following the general spirit (but not the letter) of "rational speech act" approaches to communication (e.g. Goodman and Frank 2016), this rule might in some cases be exploited to give rise to a near-certainty that *p* is entailed by its local context. In other words, if *stops smoking* comes with a requirement that there is probability ≥ a that *used to smoke* is locally entailed, strategic reasoning could lead to a near-certainty that *used to smoke* is in fact locally entailed.

We assume that there can be uncertainty about what is taken for granted in the conversation. At the initial state of the reasoning, then, the requirement is that listener L_0 should have probability ≥ a = (1-ε) of treating *used to smoke* as being presupposed, by way of the mechanisms we outlined above. But there is also a probability ε that L_0 treats *used to smoke* as being at-issue. The speaker S_1, for her part, is in the following situation. She could utter *pp'*, with a small risk of *p* being treated as at-issue. But if this is her goal, she could be explicit about it, for instance by saying *Robin used to smoke and stopped,* i.e. *p and pp'*.[18] So there are two cases to consider depending on whether S_1 means for *p* to be treated as presupposed, or as being at-issue. We will write as U_1(pp' | p at-issue) the expected utility for S_1 of an utterance of *pp'* on the assumption that she wants to treat *p* as being at-issue, and more generally U_1(utterance | intended meaning) will be the utility obtained by a certain utterance given a certain intended meaning. We will take utility to be 1 if the meaning is conveyed, 0 if it's not, and the cost of an expression E will be written as c(E). We consider each case in turn.

Case 1: S_1 wants *p* to be treated as at-issue

U_1(pp' | p at-issue) = (probability of communicative success * 1) - cost
= ε - c(pp')

U_1(p and pp' | p at-issue) = (probability of communicative success * 1) - cost
= 1 - c(p and pp')

S_1 being rational, S_1 will produce *p and pp'* just in case

U_1(p and pp' | p at-issue) > U_1(pp' | p at-issue)

or in other words just in case

1 - c(p and pp') > ε - c(pp'), i.e. 1 - ε > c(p and pp') - c(pp')

If we write δ for the cost difference between the expression *(p and pp')* and the expression *pp'* (i.e. δ = c(p and pp') - c(pp')), this condition will be satisfied whenever ε + δ < 1.

Case 2: S_1 wants *p* to be treated as presupposed

The equations are different when *p* is intended to be presupposed.

U_1(pp' | p presupposed) = (probability of communicative success * 1) - cost
= (1- ε) - c(pp')

U_1(p and pp' | p presupposed) = (probability of communicative success * 1) - cost
= 0 - c(pp')
= -c(pp')

[18] See Schlenker 2008 for a theory of presupposition projection that is based on the competition between … *pp'*… and … *(p and pp')*… .

If these are the only alternatives, S_1 can only produce *pp'* if she intends for *p* to be presupposed, since uttering *(p and pp')* would yield a certainty of miscommunication and (because of the cost) a negative utility. So to summarize:

(33) Behavior of S_1
In all cases, if S_1 intends *p* to be presupposed, S_1 utters *pp'*; in other words: P(pp' | p presupposed) = 1 (and thus P(p and pp' | p presupposed) = 0).
Whenever $\varepsilon + \delta < 1$: if S_1 intends *p* to be at-issue, S_1 utters *(p and pp'):* P(p and pp' | p at-issue) = 1 (and thus P(pp' | p at-issue) = 0).

Now the strategic listener, L_1, reasons on the basis of S_1's utility maximization behavior, as summarized in (33). On the assumption that L_1 wants to recover the meaning intended by S_1, it is clear that this leads to the conclusions in (34).[19]

(34) Behavior of L_1
In call cases, if S_1 utters *(p and pp')*, always treat *p* as at-issue.
Whenever $\varepsilon + \delta < 1$: if S_1 utters *pp'*, always treat *p* as presupposed

In other words, despite the presence of a small chance ε that *p* was not treated as presupposed by the naive listener, strategic reasoning may lead to the conclusion that a more sophisticated listener will treat *p* as presupposed with probability 1.

[19] In greater detail: L_1 needs to determine the probability that S_1 intended *p* to be treated as at-issue or as presupposed. Using Bayes's rule,
P(p at-issue | pp') = [P(pp' | p at-issue)/P(pp')] * P(p at-issue) = 0 since P(pp' | p at-issue) = 0. Therefore P(p presupposed | pp') = 1. And in all cases, P(p at-issue | p and pp') = 1.

Appendix IV
Epistemic Preconditions vs. Counterfactual Stability

Schlenker (to appear) sketches a triggering rule according to which *counterfactually stable local entailments are presupposed*. As stated in (50) in the main text, the test asks that one consider a contextual entailment *p* of *E*, and that one assume, relative to the assumptions of the context, that *E* holds true. Then it assesses the counterfactual stability of the entailment *p* by asking whether, on the counterfactual assumption that *E* had not been the case, *p* would still have held, as stated in (35)a.

(35) Assume C |= E => p (i.e. *p* is a contextual entailment of *E*). If $\rightarrow$ represents the counterfactual conditional, the test is whether:
C, E |= (not E) $\rightarrow$ p
Yes: treat *p* as a presupposition.
No: do not treat *p* as a presupposition.

We will prove a very limited equivalence between a version of the 'counterfactual stability' idea and the present analysis based on epistemic preconditions. The key is to provide the non-standard semantics for conditionals in (36). It is a 'variably strict' conditional in the domain of worlds considered by a strict conditional is given by a belief state Bel_t, but is expanded to an earlier and weaker belief state if the antecedent of the conditional is inconsistent with Bel_t.

(36) **Semantics of $\rightarrow$**
We assume that information growth is monotonic, and thus that one believes at t+1 everything one believed at t, but not conversely, and that beliefs can be arbitrarily weakened if one is willing to go far enough in the past. We write as => the material implication.
If w belongs to a belief state Bel_t, then for non-modal expressions *F* and *G*, Bel_t, w |= F $\rightarrow$ G has the following truth conditions:[20]
(i) indicative case: if Bel_t |≠ not F, Bel_t, w |= if F, G iff Bel_t |= F => G;
(ii) counterfactual case: if Bel_t |= not F, and if Bel_{t-1} = the most recent belief state $Bel_{t'}$ such that $\mathbf{Bel_{t'}}$ **|≠ not F**, then
Bel_t, w |= if F, G iff Bel_{t-1} |= F => G.

It can now be observed that a version of the analysis with counterfactual stability is equivalent to a version of the analysis with epistemic preconditions.

(37) We write Bel_t |= F for: for each w in Bel_t, Bel_t, w |= F.
Using the notations of the main text, and on the assumption that C |= E => p, the following two conditions are equivalent.

a. Modified Counterfactual Stability
if $believe_{t-1}$ C and $acquire_t$ **E**, Bel_t |= (not E) $\rightarrow$ p

b. Modified Epistemic Precondition
P($believe_{t-1}$ **p** | $believe_{t-1}$ C & $acquire_t$ **E**) = 1

Note that the conditions in (37)a imply $believe_{t-1}$ C & **E**, which connects the condition Bel_t |= (not E) $\rightarrow$ p to the condition C, E |= (not E) $\rightarrow$ p in (35).

The proof is as follows.

(i) To prove (a) => (b), assume Modified Counterfactual Stability, and assume $believe_{t-1}$ C and $acquire_t$ **E**. By (a), Bel_t |= (not E) $\rightarrow$ p. Since $believe_t$ **E**, Bel_t |= E and the conditional is counterfactual, so using (36)(ii) we get $\underline{Bel_{t-1} |= (not\ E) => p}$. Since $believe_{t-1}$ C and (by assumption) C |= E => p, we also have that $\underline{Bel_{t-1} |= E => p}$. Taken together, the two underlined condition imply that $believe_{t-1}$ **p**. In other words, if $believe_{t-1}$ C and $acquire_t$ **E**, $believe_{t-1}$ **p**, and hence: P($believe_{t-1}$ **p** | $believe_{t-1}$ C & $acquire_t$ **E**) = 1.

[20] Note that both cases can be unified under clause (ii) if 'closest earlier belief state' is replaced with: 'closest earlier or identical belief state'.

(ii) To prove (b) => (a), assume P(believe$_{t-1}$ **p** | believe$_{t-1}$ C & acquire$_t$ **E**) = 1 and believe$_{t-1}$ C and acquire$_t$ **E**; it immediately follows that believe$_{t-1}$ **p**. Since Bel$_t$ |= E, to evaluate Bel$_t$ |= (not E) $\rightarrow$ p we must apply the counterfactual rule in (36)(ii). Since acquire$_t$ **E**, Bel$_{t-1}$ is the most recent belief state that doesn't entail *E*, and the condition is that Bel$_{t-1}$ |= (not E) => p. It is fulfilled because believe$_{t-1}$ **p**, ie. Bel$_{t-1}$ |= p.[21]

[21] Without fully comparing the two frameworks, we note that the analysis based on counterfactual stability has one potential advantage over the one advocated in this piece, and one empirical drawback. The potential advantage is that judgments about the value of counterfactual statements might be easier to elicit than ones about the intuitive probability of certain conditions. For instance, for an example involving the verb *take off*, we could state the counterfactual stability test as follows:

(i) Suppose at 12:05 the company's helicopter was on the ground and took off. If this entire thing hadn't happened, would it still be the case that right before 12:05 the company's helicopter was on the ground?

A drawback of an approach based on counterfactual stability is that it probably makes incorrect predictions for the *prove* class. *Did Smith prove that Jones was the murderer?* does not presuppose that Jones was the murderer. But the test based on counterfactual stability might well make the wrong prediction in this case: on the assumption that Jones was the murderer and Smith proved it, if this entire thing hadn't happened, it is likely that Jones would still have been the murderer.

Supplementary Materials:
Raw ASL data

Raw ASL data can be downloaded at the following URL:[22]

https://drive.google.com/file/d/1Xl-u_FizMm5S2G7CQAQYh9HpiQLtOylB/view?usp=sharing

[22] As can be seen there, the consultant explicitly noticed that in the first judgment task he adopted the *viewer's* rather than the *signer's* perspective when interpreting the question, contrary to what the question asked; this is easily corrected for, and when this is done the first judgment task gives rise to the same pattern as the other three judgment tasks. Taking this correction into account, the raw inferential scores for the questions in (16) in the main text were as follows: a: 5 5 5 6; b: 6 6 6 6; c: 6 5 6 6; d: 4 5 5 5.