Lexically triggered veridicality inferences*

Aaron Steven White
University of Rochester

Abstract
The use of certain lexical items, such as know and manage, tends to trigger inferences about the veridicality of the eventualities or propositions on which those lexical items operate; while the use of others, such as think and want, does not. I review the empirical landscape pertaining to such lexically triggered veridicality inferences—focusing in particular on different forms of factive and implicative predicates—and I discuss various conventionalist and conversationalist approaches to explaining them.

1 Introduction

Taken literally, the exchange in (1) seems an exercise in solipsism.

(1)  
Bo: I thought you were heading out to get groceries.
   Jo: I remembered that I went a few days ago, but I’m now realizing I forgot to grab beer.

Imbuing (1) with coherence requires not a few steps of inference. These inferences involve not only Bo’s beliefs and Jo’s memories, but also what linguistic expressions of those mental states convey about the world—or at least, speakers’ commitments thereof. For instance, beyond having a memory of a particular sort, we infer from (1) that Jo is committed to actually having gone to get groceries—and, if we take her (and her memory) to be reliable, that she actually went to get groceries. And beyond having a lack of memory of a particular sort, we infer from (1) that Jo is committed to not actually having gotten beer—and again, if we take her to be reliable, that she didn’t actually get beer.

I refer to such inferences as veridicality inferences. In this article, I focus specifically on lexically triggered veridicality inferences, by which I mean inferences (i) whose premise includes (the meaning of)

---

* This article is currently under second review for Handbook of Pragmatics (John Benjamins). I would like to thank Fabienne Martin, Tom Grano, and an anonymous reviewer for useful comments on an earlier version of this article. This work was supported in part by a National Science Foundation grant BCS-1748969/BCS-1749025 The MegaAttitude Project: Investigating selection and polysemy at the scale of the lexicon.
some open class lexical item (e.g. a noun, verb, or adjective)—potentially in interaction with (the meaning of) its syntactic context—and (ii) whose content is explicitly linguistically realized. For example, returning to (1), the inference about Jo’s (commitment to) actually having gone to get groceries seems to be underwritten—at least in part—by her use of the verb remember; and its content is linguistically realized in remember’s subordinate clause (or clausal complement): that I went a few days ago. Similarly, the inference about Jo’s (commitment to) not actually having gotten beer seems to be underwritten—again, at least in part—by her use of the verb forget; and its content is linguistically realized in forget’s subordinate clause: to grab beer.¹

Importantly, not all veridicality inferences are lexically triggered under this definition. For instance, most standard cases of presupposition triggers involve some veridicality inference or other: in (2), we infer from Bo’s question (in conjunction with the cleft in Jo’s response) that something actually bit Bo; we infer from the possessive in Jo’s response that Mo actually owns (or is otherwise related to) a dog; we infer from the iterative again in Mo’s response (in conjunction with Jo’s assertion) that Mo’s dog was actually the biter; and from the feminine pronoun in Mo’s response, we infer that the dog is actually female. But these veridicality inferences are triggered by plausibly closed class items, such as pronouns (she) and iteratives/additives (again), or bits of constructions composed of closed class items, such as interrogatives, clefts (it was ___ that ___), and possessives (__’s ___); thus, they are not lexically triggered in the relevant sense.

(2)  Bo: What just bit me?
    Jo: It was Mo’s dog that bit you.
    Mo: And she’ll do it again if you don’t shut up.

Further, not all lexically triggered inferences are unambiguously veridicality inferences in the way Jo’s having gone to get groceries is in (1). For example, in (3), Bo is in some sense committed to there actually

¹ In the latter case, the content of the inference is of course not exhaustively realized in the subordinate clause to grab beer: At the very least, the negative component of the inference is plausibly encoded as knowledge (broadly construed) about the trigger forget itself; and whether or not the grabber is linguistically realized in the subordinate clause, separately from the realization of the forgetter, depends on one’s syntactic theory. I include these cases under the definition given above, with the caveat that not all inferences involving such enriched versions of some linguistically realized contents are veridicality inferences, at least relative to that content. For instance, from (1), we also infer that Jo was supposed to (or had some plan to) grab beer, but having an obligation or plan to do something does not mean that thing was actually done.
being beer in the fridge, but assuming that Jo and Bo are similarly reliable, our inference from Bo’s utterance in (3) is more tentative than our inferences from Jo’s utterance in (1).²

(3) In the context of (1)
    Bo: I think there’s beer in the fridge.

While it may turn out that there is a unified explanation for all of the inferences so far discussed, I take the current approach to provide a useful cross-section of phenomena, whether one’s goals are descriptive or explanatory in nature. With both types of goal in mind, I begin in Section 2 by discussing known axes of variation in when lexically triggered veridicality inferences arise. In Section 3, I turn to theoretical accounts that attempt to explain this variation within either a conventionalist framework—relying more heavily on knowledge about particular lexical items and how they compose to explain inferences—or a conversationalist framework—relying more heavily on knowledge of general communicative principles in conjunction with additional (non-linguistic) knowledge of the communicated content to explain inferences. In Section 4, I conclude with what I take to be the major open questions in this literature.

2 Variation in lexically triggered veridicality inferences

Standard methods for investigating variation in lexically triggered veridicality inferences across lexical items tend to focus on projection tests, which use entailment cancelling operators (Section 2.1). In addition to these tests, I discuss a variety of other known sources of variation (Section 2.2), including other semantic operators (Section 2.2.1) and extra-sentential contextual effects (Section 2.2.2).

Throughout this section, I attempt to remain as theory-agnostic as possible, deferring discussion of the exact nature and source of the inferences involved—e.g. whether a particular inference is implied by the content of a presupposition, assertion, etc.—to Section 3. This tack is slightly different from standard approaches to describing the data of interest here—where the boundary between a descriptive category, defined in terms of some set of empirical results, and an explanatory category, constructed as part of a system for explaining such theoretical results, is often not sharply delineated. My aim in avoiding such categories in this section is to delineate, as sharply as possible, the data to be explained from the theoretical

² See Urmson 1952; Hooper 1975; Asher 2000; Simons 2007; Hunter 2016; Koev 2018 on parenthetical uses of clause-embedding predicates, like those in (3), as well as their closely related s(entence)-lifted (Ross 1973) counterparts (see also Glass under review for a possibly related phenomenon in Chinese).
apparatus used to explain it, with the hope that this better brings the major distinctions among different theoretical approaches (discussed in Section 3) into relief.

2.1 Variation relative to entailment cancelling operators

Much of the modern literature on lexically triggered veridicality inferences grows from earlier work on presupposition (Frege 1892/1997; Russell 1905; Strawson 1950). As such, much of what is known about variation in this domain has focused on entailment cancelling operators—such as negation, conditionals, questions, and various kinds of modal elements. As the name suggests, insertion of an entailment cancelling operator scoping over the content of the original sentence tends to result in a sentence whose use does not trigger all of the (necessary) inferences triggered by the original sentence. For example, the content of (5) is an inference triggered by the use of (4) in all contexts.

(4) Jo ate a sandwich.
(5) Jo ate something.

But negating (6), questioning (7), conditionalizing (8), or modalizing (9) example (4) results in sentences whose use no longer triggers the inference that (5).

(6) Jo didn’t eat a sandwich.
(7) Did Jo eat a sandwich?
(8) If Jo ate a sandwich, she won’t be hungry.
(9) Jo may have eaten a sandwich.

When the insertion of such an operator does not nullify (or cancel) some inference, the inference is often said to project through that operator (Morgan 1969; Langendoen and Savin 1971). As I discuss below, cancellation and projection are only two of multiple possibilities for how entailment cancelling operators affect inferences of a sentence not containing that operator.

2.1.1 Matrix negation

Uses of both (10) and its modification in (11) are commonly judged to trigger the inference that the speaker actually went to the grocery store.

(10) I remembered that I went to the grocery store.
(11) I didn’t remember that I went to the grocery store.

This contrasts with the effect of inserting a negation into the embedded clause: from (12), we infer that the speaker didn’t go to the grocery store.

(12) I remembered that I didn’t go to the grocery store.

This pattern of inferences can be captured by saying that, from any use (by a trustworthy speaker) of a sentence $X$ remember $S$ or $X not$ remember $S$, one should infer $S$. This description simultaneously captures the fact that uses of both (10) and (11) trigger the inference that the speaker went to the grocery store as well as the fact that uses of (12) trigger the inference that the speaker did not. Following Kiparsky and Kiparsky 1970, any lexical item that shows this pattern—e.g. know, discover, regret, etc.—is termed a factive.

One difficulty that immediately arises for such classifications—even ones based just on the contrast between sentence pairs like that in (10) and (11)—is that at least some lexical items do not show stable inferences across syntactic contexts. For instance, we clearly do not infer from (13) that the speaker went to the grocery store; indeed, we tend to infer that they did not.

(13) I didn’t remember to go to the grocery store.

This is initially surprising, since we draw the same inference from (14) that we do from (10).

(14) I remembered to go to the grocery store.

Related patterns can be seen cross-linguistically. For instance, in Modern Greek (Varlokosta 1994; Giannakidou 1998; 1999; 2009; Roussou 2010), Persian (Farudi 2007), Hungarian (Abrušán 2011), Hebrew (Kastner 2015), and Turkish (Ozyildiz 2017), veridicality inferences are sensitive to the presence of nominal(izing) morphology attached to verbs or their clausal complements—often taking the form of a pronoun, definite determiner, or demonstrative in the language.

These observations suggest the need to specify an inference signature, not for lexical items themselves, but for lexical items in some linguistic context (P. Schulz 2003; Kratzer 2006; Moulton 2009; White 2014; Bogal-Allbritten 2016)—e.g. one for remember with a tensed clausal complement (remember that), as in (10) and (11), and one for remember with an infinitival complement (remember to), as in (13) and (14). I henceforth refer to such combinations of lexical items with their linguistic context as
constructions, though this terminological choice should be taken only as descriptive shorthand, not as a commitment to a constructionist view (Goldberg 2013).

Thus, the remember that construction is factive, while the remember to construction is not. Following Karttunen 1971a, nonfactive construction that nonetheless trigger veridicality inferences—e.g. remember to—are termed implicatives. Because this contrastive definition encompasses many logically possible inference patterns—depending not only on whether the lexical item is negated (i.e. has negative matrix polarity) or not (i.e. has positive matrix polarity), but also whether the inference that is triggered involves negation or not—it is useful to define a shorthand for possible patterns.

Such a shorthand is proposed by Nairn, Condoravdi, and Karttunen (2006) in the form of implication signatures (see also MacCartney 2009; Karttunen 2012). Implication signatures aim to capture the behavior of a lexical item (in a particular syntactic context) by specifying two trits (trinary digits) of information. The first trit specifies whether a particular lexical item triggers an inference when found in a positive polarity matrix clause, and if so, whether that inference is positive (+) or negative (−); and the second specifies whether a particular lexical item triggers an inference when found with negative matrix polarity, and if so, whether that inference is positive (+) or negative (−). Thus, factives, like remember that, have the implication signature ++; and implicatives like remember to have the implication signature +−.

This system implies $3^2 = 9$ possible implication signatures, all of which are attested (more or less robustly)—at least in English. For example, the inverse of the signature of remember to (−+)—the signature of fail to, neglect to, and forget to—is exemplified in (15) and (16). From (15), we infer that the speaker did not grab beer; but from (16), we infer they did.

(15) I {failed, neglected, forgot} to grab beer.
(16) I didn’t {fail, neglect, forget} to grab beer.

Constructions that have either of the two signatures (+−, −+) are together called the two-way implicatives (Karttunen 1971a). These are contrasted with one-way implicatives, which only trigger a positive (+) or negative (−) inference with one matrix polarity—the lack of an inference in a particular direction being marked by ◦. There are four such combinations (+◦, −◦, ◦+, ◦−), wherein lexical items with + for their first trit are sometimes termed (objective) veridical (Giannakidou 1994; 1998; 1999; 2001; Zwarts 1995; Egré 2008; Anand and Hacquard 2014; Spector and Egré 2015) and those with – for their first trit are sometimes termed (objective) antiveridical (Giannakidou 1998).

(17) +◦ structures: prove that, show that
(18) −◦ structures: pretend that, hallucinate that
(19) **o+ structures:** hesitate to

(20) **o– structures:** think to, know to

The last two signatures implied by this system are sometimes referred to as the *antifactives/counterfactive* (– –) and the *nonveridicals/nonfactives* (○○). The former class is relatively sparsely populated in English—with *pretend to* being one of the few candidates (Givón 1973)—but the latter is quite large, encompassing many doxastic (belief) and communicative predicates, such as *believe that, think that, say that/to, claim that/to, assume that, suppose that*, etc. (White and Rawlins 2018; White et al. 2018).³

2.1.2 **Conditionals and interrogatives**

It has long been known that lexically triggered veridicality inferences do not respond uniformly to all entailment cancelling operators. For instance, Karttunen (1971b) observes that some constructions that appear to be factive under the negation test (21) do not trigger veridicality inferences when within the scope of questions (22) and conditionals (23).

(21) John didn’t {regret, realize, discover} that he has not told the truth.
(22) Did you {regret, realize, discover} that you had not told the truth?
(23) If I {regret, realize, discover} later that I have not told the truth, I will confess it.

According to Karttunen, constructions like *discover that, find out that, and see that* do not trigger veridicality inferences within the scope of a question, even though they do under negation. This pattern contrasts with the one observed for *regret and realize*, which trigger veridicality inferences in both contexts.

Karttunen furthermore suggests that, while *realize that* patterns with *regret that* relative to negation and questions, it patterns with *discover that* relative to conditionals. Insofar as such judgments are robust, this variability potentially suggests a need to specify two further trits in an implication signature—one for questions and another for conditionals—resulting in $3^4 = 81$ logically possible signatures. To my knowledge, it is not known whether all possible signatures in this expanded set are attested—and depending

---

³ See Nairn, Condoravdi, and Karttunen 2006; Karttunen 2012; Lotan, Stern, and Dagan 2013; White and Rawlins 2018; White et al. 2018 for attempts at an exhaustive classification of the implication signatures of English lexical items, and see Falk and Martin 2017 for a similar attempt in French. See Pavlick and Callison-Burch 2016; Rudinger, White, and Van Durme 2018 for methods to automatically extract predicates that fall under these signature from text corpora.
on one’s theoretical position, this may not be a particularly useful question to ask (see discussion of conversationalist approaches in Section 3).

In practice, it is common to follow Karttunen 1971b in only drawing a distinction among two subclasses of factives: the true factives—e.g. regret that, be happy that, be a bummer that—which tend to involve an emotive component and purportedly do not show variability in whether they trigger veridicality inferences under different entailment cancelling operators;\(^4\) and the semifactives—e.g. discover that, realize that, find out that—which tend not to involve an emotive component and do show such variability (Hooper and Thompson 1973; Hooper 1975; Simons 2007; Egré 2008; Beaver 2010; Djärv, Zehr, and Schwarz 2018).\(^5\) True factives are often assimilated with the more general class of hard triggers, whose inferences are relatively stable across environments; and semifactives are often assimilated with the more general class of soft triggers, whose inferences are more variable (Abusch 2002, 2010; Abbott 2006; Romoli 2011, 2014; cf. Abrusán 2011, 2016; Tonhauser, Beaver, and Degen 2018).

2.1.3 Modal elements
Lexically triggered veridicality inferences can also vary with the kinds of modal element that scope over them. For instance, note that (24) triggers the veridicality inference that Bo smokes, and this does not change with the insertion of the modal auxiliary may in (25)—i.e. the veridicality inference triggered by acknowledge projects through may.

(24) Bo acknowledged that he smokes.
(25) Bo may acknowledge that he smokes.

Similar to may, the inference apparently projects through the subordinate clause-taking predicate force in (26); but unlike force, the predicate order in (27)—even in the same syntactic context—appears to block the inference (Karttunen 1973).

(26) Jo forced Bo to acknowledge that he smokes.

\(^4\) A small minority of authors deny that emotive predicates trigger veridicality inferences, instead arguing that they only trigger inferences about the beliefs of the emotion’s experiencer (Schlenker 2005; Egré 2008). This is often attributed to Klein 1975 (Abrusán 2011; Anand and Hacquard 2014). See Djärv, Zehr, and Schwarz 2018 for an experimental investigation of this question.

\(^5\) The term semifactive tends to cover both constructions like realize that and discover that, even though those constructions purportedly differ with respect to which entailment cancelling operators they trigger veridicality inferences under.
Jo ordered Bo to acknowledge that he smokes.

This may be related to the fact that order does not trigger veridicality inferences itself, while force does; but this cannot be the entire explanation, since may does not block the projection of some veridicality inferences, acting similarly to force. Such patterns lead Karttunen (1973) to classify lexical items not only by their veridicality inferences—as factives, implicatives, etc.—but also by whether they are holes, plugs, or filters.

A hole is any construction that, like force NP to, does not alter an inference triggered by the content it scopes over; a plug is any construction that, like order NP to, that blocks an inference that would be so triggered; and a filter is a construction that acts like a hole under certain conditions and a plug under others. Karttunen characterizes holes as generally being factive, implicative, or aspectual predicates and plugs as generally being communicative predicates and (possibly) all other propositional attitude verbs that do not fall into the class of holes. Filters for him only include a small set of functions items, including conditionals, conjunction, and disjunction, and so I do not discuss them here.

2.2 Other conditions on variation

While the behavior of lexically triggered veridicality inferences with entailment cancelling operators is by far the most well-studied, these inferences are also known to be sensitive to other semantic operators (Section 2.2.1), and there is a growing body of work investigating the role of contextual factors in determining whether or not veridicality inferences are triggered (Section 2.2.2).

2.2.1 Other semantic operators

Since at least Karttunen 1971a, it has been known that certain kinds of modal predicates sometimes trigger veridicality inferences. For instance, be able in (28) triggers the inference that the quarterback completed exactly two passes; and in (29), be able triggers the inference that the quarterback did not complete any passes.

(28) The quarterback was able to complete exactly two passes. (compare Karttunen’s Ex. 46)
(29) The quarterback wasn’t able to complete any passes.

Thus, be able appears to be some sort of implicative. But what kind of implicative is it? At least the inference triggered by (28) is felt to be weaker, in some sense, since being able to do something does not imply doing it. Karttunen (2016) suggests that be able is not alone in this pattern, citing some other
implicatives that follow it: *refuse to* in (30), tends to trigger the inference that Ann spoke up, while *not force to* in (31) and *hesitate to* in (32) tend to trigger the inference that Ann didn’t speak up.

(30) Ann didn’t refuse to speak up. (Karttunen’s 15b)
(31) Ann was not forced to speak up. (Karttunen’s 15c)
(32) Ann hesitated to speak up. (Karttunen’s 15d)

Interestingly, the weakness of one-way implicatives’ inferences in a particular direction is not true in many Indo-European languages with a morphologically realized grammatical aspect distinction: when the ability modal (Bhatt 1999) along with other non-epistemic modals (Hacquard 2006; 2009) and propositional attitude predicates (Hacquard 2008) are combined with perfective aspect, the veridicality inference of sentences like (28) is always triggered; but when they are combined with imperfective aspect, they are not (cf. Mari and Martin 2007; Homer 2010; see also Hacquard’s (to appear) rebuttal). This phenomenon—which is often referred to as *actuality entailment* (Bhatt 1999)—is attested in at least Modern Greek (Bhatt 1999; Giannakidou and Staraki 2013), Hindi (Bhatt 1999), French (Hacquard 2006; Mari and Martin 2007; Homer 2010), Italian (Hacquard 2006), and Spanish (Borgonovo and Cummins 2007). Further, this interaction with aspect appears to extend beyond modal predicates and Indo-European: (Spector and Egré 2015) present a contrast in Hungarian (attributed to Mártă Abrusán) between *tell* with perfective aspect, which always triggers the inference that the content of its embedded clause is true, and *tell* with imperfective aspect, which does not.

In addition to such morphologically realized conditions, prosodic factors also affect whether certain veridicality inferences are triggered. For instance, Beaver (2010) notes that if intonation signals focus on a factive, as in (33)b and (33)d, the veridicality inference is generally triggered; but if intonation signals focus on the clause carrying the content of the veridicality inference, as in (33)a and (33)c, it generally is not (see also Beaver and Clark 2008 and see Tonhauser 2016 for experimental corroboration; cf. Djärv and Bacovcin 2017).

(33) Beaver’s (2010) Ex. 73
a. If I discover that your work is [plagiarized]_, I will be [forced to notify the Dean]._
b. If I [discover]_ that your work is plagiarized, I will be [forced to notify the Dean]._
c. If the T.A. discovers that your work is [plagiarized]_, I will be [forced to notify the Dean]._

---

Bhatt lists additional languages where similar patterns purportedly hold (Ch. 5, fn. 4).
d. If the T.A. [discovers] that your work is plagiarized, I will be [forced to notify the Dean].

As Beaver (p. 93) puts it:

Example [a] suggests the professor has a suspicious mind, but that the student may be innocent. In contrast, [b] suggests that the student is innocent, but that the professor is prepared to act as if she had not made the discovery[...] or else that the professor is mad. [c] does not imply that the student is guilty. And finally, [d] conjures up an image of complicity between the all-knowing professor and the guilty student.

2.2.2 Contextual conditions on variation

While standard classifications rely mainly on diagnostic tests based on entailment cancelling operators, there are now a menagerie of other contextual factors that have been noted as conditioning whether a veridicality inference is triggered by a particular construction. One prominent case of this is noted by Stalnaker (1974): a change in the subject of a predicate can modulate veridicality inferences. For instance, while the content of is the same as that of (34)b—when (34)b is spoken by Pavarotti—(34)a seems to trigger the inference that Loren is in New York, while (34)b does not.

(34) Chierchia and McConnell-Ginet’s (1990) Ex. 73
   a. If Pavarotti discovers that Loren is now in New York, he will be angry.
   b. If I discover that Loren is now in New York, I’ll be angry. (Pavarotti is speaker.)

This pattern is plausibly connected to the more general observation that veridicality inferences are not triggered when the speaker is assumed to not know whether the content of the would-be inference is true (Karttunen 1974; Gazdar 1979; Levinson 1983; Geurts 1994; Simons 2001). For instance, aspectual predicates like start, begin, stop, finish, and continue are often judged to give rise to inferences about a state (not) holding before the event denoted by the predicate (Sellars 1954; Newmeyer 1969)—e.g. the use of (35) tends to trigger the inference that the speaker used to smoke but no longer does.7

7 They also trigger inferences about a state (not) holding after the event denoted by the predicate. This post-state inference appears to be unlike the pre-state inference in being more sensitive to entailment cancelling operators. I return to this point in Section 3.
(35) I recently stopped smoking.

Further, these predicates initially appear to trigger these inferences under the entailment cancelling operators already discussed—e.g. (36) and (37) tend to trigger the same inference as (35); and (37) additionally triggers the inference that the addressee currently smokes.

(36) I can’t stop smoking.
(37) If I stop smoking, will you stop smoking?

But when the context is varied only slightly—e.g. if the speaker is unfamiliar with the addressee as implicated in (38)—these inferences are not triggered: as Simons (2001) notes, the interrogative in (38) can be roughly paraphrased as (39).

(38) I notice that you keep chewing on your pencil. Have you recently stopped smoking? (Simons’ Ex. 1)
(39) Is it the case that you have recently been a smoker and have recently ceased to be one? (Simons’ Ex. 2)

This pattern suggests that ignorance about the truth or falsity of an inference that would otherwise be triggered by some construction is enough to not trigger that inference—at least for some constructions (e.g. semifactives and aspectuals).

In addition to properties of a speaker’s epistemic or doxastic state, certain aspects of the content of a veridicality inference have also been argued to be relevant. For instance, Karttunen et al. (2014) suggest that the emotive valence of evaluative adjectives interacts with the contextually determined valence of the content of a clause to result in distinct inferences (see also Karttunen 2013): if the valence of an evaluative adjective complement is consonant with the assumed valence of the property described by its complement (40), they suggest that +− (implicative) inferences are triggered; and if not (41), they suggest, ++ (factive) inferences are triggered.

(40) I am not saying that I was not stupid to have trusted someone because they were family but it doesn’t mean that they should get away with it. (Karttunen et al.’s Ex. 7a)
(41) I was not fortunate to be born with long and beautiful eyelashes like many women. (Karttunen et al.’s Ex. 8b)
It remains an open question to what extent this observation is related to other contextual factors, such as those discussed earlier in this section.

3 Explanatory approaches

Explanatory approaches to lexically triggered veridicality inferences lie along a continuum with extreme *conventionalism* on the one pole and extreme *conversationalism* on the other. Under approaches on the conventionalist extreme of this continuum (Katz and Langendoen 1976), all lexically triggered veridicality inferences arise from knowledge about particular lexical items and how they compose. Under approaches on the conversationalist extreme (Stalnaker 1973; 1974), all lexically triggered veridicality inferences arise from knowledge of general communicative principles in conjunction with additional (non-linguistic) knowledge of the communicated content.

In practice, most extant proposals lie somewhere in between these two extremes. Indeed, extreme conventionalism, in particular, has largely been abandoned since the 1970s because of its inability to handle the sorts of variability discussed in Section 2 (see Levinson 1983, Section 4.4.1 and references therein). Nonetheless, this continuum remains useful for structuring discussion, and so I partition approaches into conventionalist (Section 3.1) and conversationalist (Section 3.2), based roughly on the relative reliance on lexical and compositional knowledge for triggering veridicality inferences.8

3.1 Conventionalist approaches

The key conventionalist conceit is that it is knowledge *about* lexical items *qua* linguistic objects that drives lexically triggered veridicality inferences. And since (Stalnaker 1973; 1974; Karttunen 1974), it has been recognized that a *sine qua non* of any conventionalist account is that it provide a theory that links such knowledge with the communicative acts a speaker performs (see also Wilson 1975; Kempson 1975; Boër and Lycan 1976; cf. Katz and Langendoen 1976). In general, this explanation takes something like the following form: lexically triggered veridicality inferences arise via reasoning about what the speaker must

---

8 Due to space constraints, I do not discuss in detail different formal syntactic and semantic accounts of how linguistic expressions of interest here—generally, predicates that combine with subordinate clauses—come to have the particular semantic properties that they do. See Grano in prep for a recent overview of relevant literature on such subordinate clause-taking predicates.
commit themselves to—and in many cases, furthermore taken as entailed by (or at least consistent with (Gazdar 1979; Levinson 1983)), what is commonly agreed on—in using a particular expression.

This position naturally captures inferences that are invariantly triggered by an expression across contexts of use—i.e. entailments—but the explanation gets somewhat trickier for inferences that only appear to be triggered in particular syntactic structures or contexts of utterance—i.e. that are defeasible—as discussed in Section 2. A major commitment of a conventionalist approach is that, unlike conversational implicatures (Grice 1989), which are triggered any time the same content is expressed under particular contextual conditions, at least some lexically triggered veridicality inferences are detachable—i.e. in principle, synonymous constructions could differ in their lexically triggered veridicality inferences.

This literature has largely been split between investigations of factive constructions and investigations of implicative constructions. This split arises in part from the fact that at least some of the veridicality inferences triggered by factive constructions—henceforth factive inferences—are backgrounded (and potentially furthermore presupposed) in some intuitive sense, while veridicality inferences triggered by implicative constructions—henceforth implicative inferences—are (generally) foregrounded in some intuitive sense. I thus discuss explanatory approaches to factive inferences (Section 3.1.1) and implicative inferences (Section 3.1.2) separately, attempting (where possible) to draw connections between these approaches.

3.1.1 Factive inferences

In prominent early accounts, factives’ lexically triggered veridicality inferences are explained by assigning a particular construal to meaning postulates, which associate lexical items (or constructions thereof) directly with the desired implications (Karttunen and Peters 1979; Gazdar 1979). For example, Karttunen (1971b) gives (42) as a meaning postulate for discover that is generalizable to other semifactives—i.e. those factives whose veridicality inferences are more variable across semantic operators—and (43) as a meaning postulate for regret that is generalizable to other true factives—i.e. those factives whose veridicality inferences remain relatively stable across semantic operators. Thus, insofar as some sentence containing a true factive V implies it is possible that x (not) V S for some x and S, that sentence is predicted to trigger the inference that S. In contrast, given just these postulates, semifactives are predicted to only trigger veridicality inferences when unembedded or when scoping under negation (and no other operator).

(42) Karttunen’s Ex. 11
   a. ∀x: ∀s: x discover s implies s
   b. ∀x: ∀s: x not discover s implies s

(43) Karttunen’s Ex. 11’
Karttunen and Peters’s (1979) solution is to posit a new dimension on which expressions can compose: in addition to an extension expression, which determines an expression’s truth conditional content, they posit an implicature expression. Meaning postulates such as (42)a are then directly imported into implicature expressions, with implies being converted to conventionally implicates (terminology due to Grice 1989; see Potts 2005 for an updated version of this sort of approach). A crucial component of Karttunen and Peters’ system is that defeasibility of a veridicality inference is strictly a product of the conventional implicature dimension, and so the system necessarily gets more complicated when importing meaning postulates like (42)b, since Karttunen and Peters must posit (i) that negation is ambiguous between a form that supports projection and one that doesn’t (see Horn 1972, 1989); and (ii) that this form only occurs with some predicates but not others. Importing the meaning postulates in (43) gets even hairier, since it requires similar ambiguities for it is possible that and a corresponding specification of which it is possible that cooccurs with which negation. Nonetheless, it is technically possible to get this part of the system working. Where this system falters most is in incorporating defeasibility based on contextual conditions, like those discussed in Section 2.

Gazdar (1979) provides a potential solution to this problem that is similarly multi-dimensional but attempts to explain contextual conditions via general principles determining when an inference is not triggered when it otherwise might have been—i.e. when it is cancelled. In short, Gazdar posits a strict priority ordering on (potential) inferences: conventional, defeasible inferences are only triggered if they are consistent with (i) the mutually agreed upon commitments of the interlocutors (the context); (ii) the entailments of the utterance; (iii) the clausal implicatures of the utterance; and (iv) the conversational implicatures of the utterance (relative to the current context). If any inference that comes later in the ordering is inconsistent with the accumulation of inferences coming earlier, that inference is not triggered.

Gazdar’s proposal portends a shift toward blending aspects of the conventionalism and conversationalism: though it posits that all lexically triggered veridicality inferences are at root a product of lexical knowledge, contextual factors may manipulate whether that lexical knowledge is activated. This shift continues in two influential proposals by Heim (1983b, 1992) and van der Sandt (1992), which both aim to refactor the conventional source of factive inferences.9

---

9 I do not discuss the distinction between Heim and van der Sandt’s dynamic theories and transparency-based theories (Schlenker 2008; 2009; 2010), which present a strongly conventionalist alternative based in classical logic, instead referring the interested reader to Schlenker 2011a, 2011b.
3.1.1.1 Common ground approaches Heim (1983b) argues that associating expressions with implications that are, in principle, independent of their content—as Karttunen and Peters (1979) and Gazdar (1979) do—misses generalizations about the relationship between an expression’s content and those inferences. To remedy this, she proposes to bind these two kinds of knowledge together under a single form of conventional knowledge within a dynamic semantics framework: context change potentials. Within this framework, lexical items are associated with partial functions that determine how their use updates a context—a formalization of Lewis’ (1969) notion of a common ground in terms of possible worlds. Specifically, contexts are formalized as sets of possible worlds compatible with what is commonly accepted among conversational participants, and updating a context with a particular expression entails applying that expression’s context change potential to the context in order to yield a new one. This update will generally result in further constraining the set of possible worlds that are compatible with what has been accepted as common ground by the conversational participants.

Heim (1992) explores two different ways that this system may give rise to lexically triggered veridicality inferences, focusing mainly on cases of projection through entailment cancelling operators. The first is to encode a definedness condition as a part of a lexical item’s context change potential that requires that any inference that that lexical item triggers must be entailed by the current context. Another way of saying this in her system is that updating the current context with the content of a particular inference results in the same context. In contexts where such entailment does not hold, accommodation (Lewis 1979; see also Beaver and Zeevat 2007) of the content of the inference is then necessary (Heim 1983b)—i.e. the context must first be updated with that content before applying the context change potential of the sentence to the result. Under her system, this accommodation is what gives rise to at least some lexically triggered veridicality inferences—specifically, those of factive predicates.

Heim (1992) also suggests an alternative route to generating lexically triggered veridicality inferences. Following Karttunen 1974, she notes that, prima facie, (44) seems to trigger both the inference that Patrick has a cello and the inference that Patrick believes he owns a cello.

---

10 For Heim 1982; 1983a; 1983b, they are in fact sets of pairs of worlds and assignments of indices to individuals. This more complicated system is necessary for capturing presuppositions of quantified statements—among other phenomena—and is not strictly necessary for understanding her approach to lexically triggered veridicality inferences.

11 In fact, she suggests two alternative routes. I do not discuss her other alternative, which relies on smuggling veridicality inferences in via de re interpretations (see Section 5.1), largely because it is generally viewed as a nonstarter (Geurts 1998).
(44) Patrick wants to sell his cello. (Heim’s Ex. 1)

The former inference—which seems likely to be triggered by the possessive *his*—does not arise in certain contexts, such as (45); only the weaker inference that Patrick believes he owns a cello does.

(45) Patrick is under the misconception that he owns a cello, and he wants to sell his cello. (Heim’s Ex. 2)

This behavior raises the question where the stronger inference of (44) comes from when not in contexts like that in (45). Elaborating on an idea of Karttunen’s (see also Kay 1992), Heim suggests that there is “spill-over” from beliefs (*local contexts*) to the *global context*—i.e. the representation of the conversational participants shared commitments. To illustrate this, she notes that one tends to infer from (46) that the speaker is committed to it in fact having been raining, even though strictly speaking, they are only committed to John having such a belief.

(46) John believes that it stopped raining. (Heim’s Ex. 66)

She then suggests (p. 212) that…

…when we hear (66) out of the blue, we know two things: first, as a matter of the semantics of this sentence, we know that it requires the presupposition that John believes it was raining. Second, we know that the speaker takes this to be uncontroversial and unsurprising. Now why would it be unsurprising that John has such a belief? The most natural guess is that it would be unsurprising because it was in fact raining and John was in an appropriate position to find out.

Anand and Hacquard (2014) give a similar account of the inferences triggered by communicative predicates as part of their generalization that only doxastic predicates—e.g. *know, be aware, see*—but not communicative predicates—e.g. *say, tell, demand*—are factive. In particular, they argue that predicates describing response-stance discourse moves (Cattell 1978), such as *admit, acknowledge, and confirm*—which they report as the only attested cases of factive communicative predicates—do not have veridicality inferences that are encoded as constraints on the global context. They claim that apparent veridicality inferences from uses of such sentences arise via spill-over from the discourse that a communicative predicate describes a particular discourse move in and the discourse in which that move is reported.
(47) Does the book {acknowledge, admit, confirm} that Mary is the murderer? (Anand and Hacquard’s Ex. 14)

Anand and Hacquard’s proposal is abstractly similar to Heim’s proposal in that both assume that some sort of default reasoning process causes aspects of some local context—e.g. a belief context or reported common ground—to be merged with aspects of the global context.

3.1.1.2 Anaphoric approaches Under van der Sandt’s (1992) proposal all “[p]resuppositions are simply anaphors. They only differ from pronouns or other kinds of semantically less loaded anaphors in that they contain enough descriptive content to establish a reference marker in case discourse does not provide one” (p. 345). Thus, in the particular case of factives, their use involves reference to some familiar contentful abstract object (see Asher 1993; Spenader 2003). And when such an object cannot be found, factive inferences come about as a consequence of constructing that object via a process of accommodation, which is abstractly similar to Heim’s (though necessarily different in detail), using a notion of recursively embedded discourse structures formalized within Discourse Representation Theory (Kamp and Reyle 1993; see also Zeevat 1992; Beaver 1995). And like Heim, van der Sandt explains variability in whether a factive inference is triggered via a process of local accommodation within some embedded discourse structure.

This proposal has clear affinity with the proposal put forth by Kiparsky and Kiparsky (1970), wherein the subordinate clause in a factive construction is derived from a complex noun phrase headed by fact—i.e. (48) is derived from something with the same structure as (49).

(48) Jo knows that Bo left.
(49) Jo knows the fact that Bo left.

One (though not the only) way to construe van der Sandt’s approach for factive constructions in particular thus reduces to saying that factive inferences are triggered not directly by open class lexical items like know, but by some component of the construction that those open class lexical items cooccur with. Conversely, under this proposal, open class lexical items like think would not be able to cooccur with this component.

One question that arises is what this component might be? In Kiparsky and Kiparsky 1970, it is some other special open-class item—a silent version of the subordinate clause-taking noun fact. Other approaches in this vein instead associate the complementizer that (or some null counterpart) with the inference (Kratter 2006; Moulton 2009; Sheehan and Hinzen 2011; Bogal-Allbritten 2016), while others assume that, in addition to a complementizer, subordinate clauses in factive constructions can have (sometimes covert) definite determiners that trigger the veridicality inference (Takahashi 2010; Kastner
This last approach has initial support from the cross-linguistic evidence discussed in Section 2.1, wherein veridicality inference correlated with the presence of a definite determiner or demonstrative near the complementizer in a subordinate clause.

3.1.2 Implicative inferences

Unlike early approaches to factive inferences, early approaches to implicative inferences do not assume direct encoding of the veridicality inference in a meaning postulate—rather relying on deductive reasoning to generate the inference. For instance, Karttunen (1971a) gives the following template for +– implicatives—where \( v \) might be replaced by, e.g., *manage*.

(50) Karttunen’s Ex. 37

a. PRESUPPOSITION: \( v(S) \) is a necessary and sufficient condition for \( S \).

b. PROPOSITION: \( v(S) \).

This ensures that…

…if the main sentence is an affirmative assertion, it states, according to the speaker's supposition, that a sufficient condition for the truth of the complement sentence is fulfilled. Thereby the speaker indirectly asserts that the complement is also true. A negative assertion claims that a necessary condition for the truth of the complement is not fulfilled; therefore it must be false. If the main clause is questioned, the speaker must be ignorant of whether the complement sentence by itself would make a true assertion. (p. 352)

A similar pattern can be used to capture any subtype of implicative by manipulating the nature of the presupposition—e.g. (51) for –+ constructions, (52) for o–, and (53) for +o.

(51) Karttunen’s Ex. 41

a. PRESUPPOSITION: \( v(S) \) is a necessary and sufficient condition for \( \neg S \).

b. PROPOSITION: \( v(S) \).

(52) Karttunen’s Ex. 54

a. PRESUPPOSITION: \( v(S) \) is a necessary condition for \( S \).

b. PROPOSITION: \( v(S) \).

(53) Karttunen’s Ex. 59

a. PRESUPPOSITION: \( v(S) \) is a sufficient condition for \( S \).
b. **PROPOSITION:** \( v(S) \).

One problem with this account is that it is a bit *just so* in that the proposed necessity/sufficiency inferences are not independently justified. For instance, the inferences triggered by a use of *manage or fail* seem to involve a range of contextually determined modalities—including effort and intention (Karttunen 1971a), but also difficulty or unlikelihood (Coleman 1975)—none of which are necessary or sufficient conditions.

Aiming to further delve into the relationship between the presuppositional and propositional components of implicative predicates, Baglini and Francez (2016) present a modification of Karttunen’s proposal. Their proposal has two components, which interact to give rise to implicative entailments. The first component is similar to (50)a: *manage* \( p \) triggers an inference that there is some lexically underspecified (but familiar) causally necessary but causally insufficient *catalyst* for bringing about \( p \)—e.g. effort on the part of the manager. The important difference from (50)a is that the catalyst need not be some event of managing itself, but rather some piece of the causal puzzle—e.g. effort, intention, etc.—for how \( p \) might come about.\(^{12}\) The second component is also similar to (50)b: a sentence containing *manage* entails that the aforementioned catalyst did in fact cause \( p \) to be true. Therefore, *manage* \( p \) entails \( p \). This proposal also explains why *not manage* \( p \) entails *not p* since the catalyst is causally necessary, if it does not bring \( p \) about, then \( p \) is not true.

As it stands Baglini and Francez's proposal only handles two-way implicatives like *manage*, whose implicative direction and matrix polarity match. Nadathur (2016) modifies it to capture inverted two-way implicatives \( (\neg \rightarrow) \), like *fail* and *neglect*, and one-way implicatives. Her basic idea is a natural integration of Karttunen’s and Baglini and Francez's. For Nadathur, all implicatives at least presuppose causal necessity of some catalyst, but in contrast to Baglini and Francez's *manage*, two-way implicatives furthermore presuppose causal sufficiency of that catalyst—i.e. that that catalyst is the “only open prerequisite” (p. 1012). The addition of the causal sufficiency presupposition for two-way implicatives would leave Baglini and Francez's explanation for the entailments of *manage* intact, since that proposal relies only on *manage* presupposing causal necessity. But since Baglini and Francez's proposal explains the entailments of *manage* with only the causal necessity presupposition, Nadathur’s approach must explain why one-way implicatives, like *be able*, are not two-way, like *manage*.

To do this, Nadathur proposes that, rather than being used to assert that a contextually available catalyst actually caused \( p \), implicatives are only used to assert that those catalysts hold. But because two-

---

\(^{12}\) Baglini and Francez (2016) develop a rich model of causal dynamics based on K. Schulz 2011, which I do not discuss here, since this intuitive description is sufficient for current purposes. For alternative models of such dynamics and their relation to language, see Copley 2019 and references therein.
way implicatives make reference to causally necessary and sufficient catalysts, existence of the catalyst is enough to capture the positive and negative implicative directions of manage (and potentially fail).\footnote{The specifics of Nadathur’s account of two-way implicatives are more complicated. She does not employ a notion of sufficiency directly, rather relying on a form of circumscription reasoning (McCarthy 1980), implemented using exhaustification (see also K. Schulz and Van Rooij 2006) and resulting in a form of conditional perfection (Geis and Zwicky 1971).} It furthermore predicts that, if a predicate only presupposes causally necessary catalysts, it only has a single implicative direction. For instance, the negative implicative direction of be able under Nadathur’s account is negative, since be able asserts that a particular causally necessary catalyst—i.e. ability—does not exist. And it has no positive implicative direction, since we cannot conclude from the existence of a single causally necessary catalyst, such as ability, that \( p \) is true.

Inverted \((\rightarrow)\) implicatives—e.g. fail—are captured by saying that a verb can presuppose causal necessity and sufficiency for not \( p \). Thus, Nadathur’s system straightforwardly predicts inverted two-way implicatives, like fail, as well as inverted one-way implicatives, like hesitate, whose negative implicative direction is positive: \((54)\)b entails \((55)\)a but \((54)\)a does not entail \((55)\)b.\footnote{One place this proposal fails is that it cannot capture one-way implicative with positive implicative direction. Indeed, unmodified, it predicts such verbs should not exist. This prediction is incorrect for verbs like refuse, which have no negative implicative direction but whose positive implicative direction is negative.}

\[(54)\]
\[
\begin{align*}
\text{a. } & \text{John hesitated to join the fray.} \\
\text{b. } & \text{John didn't hesitate to join the fray.}
\end{align*}
\]

\[(55)\]
\[
\begin{align*}
\text{a. } & \text{John joined the fray.} \\
\text{b. } & \text{John didn't join the fray.}
\end{align*}
\]

Combined, these proposals also potentially provide traction on a phenomenon that was noted by Bhatt (1999) for predicates like be able but which appears general to one-way implicatives: though one-way implicatives only have a single implicative direction, they tend to trigger sometimes quite strong inferences in the other direction. Karttunen (2012) likens this inference to the one seen in conditional perfection (Geis and Zwicky 1971), and Nadathur (2016) suggests that this is a consequence of pragmatically strengthening a necessary condition to a sufficient condition (see Karttunen 2016, Sec. 3.3.1 for a closely related account).
It is somewhat unclear, however, how this syncs with the observation that, in at least some other languages, grammatical aspect modulates whether one gets a veridicality inference with root (non-epistemic) modals, like be able (Bhatt 1999) and have to (Hacquard 2006), as well as some propositional attitude predicates, such as want (Hacquard 2008). These inferences are often referred to as actuality entailments because, in contrast to those triggered by English be able with negative polarity, they are nondefeasible—in contrast to conditional perfection inferences. There is a large literature on the syntactic and semantic conditions that give rise to actuality entailments that I do not discuss here, instead referring interested readers to Hacquard to appear and Giannakidou and Mari to appear, Ch. 6.

In addition to actuality entailments, it is also unclear how the Baglini and Francez-Nadathur proposal captures predicates that evoke coincidence, like happen to and turn out, or propositional attitudes, like remember and forget (see Karttunen 2016 for further critiques). The coincidental predicates are challenging to explain on a causal necessity/sufficiency story like the Baglini and Francez-Nadathur account because they do not seem to be associated with any particular necessary and sufficient conditions, while the propositional attitude verb predicates are challenging to explain because the inferences that they trigger do not stay constant across syntactic contexts (see Section 2.1.1): (56) seems to trigger the inference that Jo has some (possibly weak) obligation or goal to make the bed—plausibly analogous to the effort inferences for manage and fail (van Leusen 2012)—whereas (57) triggers the inference that she actually did (Karttunen 1971a).

(56) Jo did(n’t) {remember, forget} to make the bed.
(57) Jo did(n’t) {remember, forget} that she made the bed.

One explanation for the propositional attitude predicates in particular, suggested by White (2014), is that—unlike manage and fail, which plausibly trigger the necessity/sufficiency inference themselves—remember and forget do not trigger the apparently analogous obligation inference. Rather, he suggests, the subordinate clause itself contributes the relevant obligation inference and because remember and forget are factive, they trigger a veridicality inference about the content of the infinitival clause—similar to (58).

(58) John did(n’t) {remember, forget} that he had to make the bed.

The main difference between (56) and (58) is that the veridicality inferences triggered by (56) are intuitively much stronger than those triggered by (58). He addresses this difference by positing that the stronger

---

15 Thanks go to an anonymous reviewer for pointing this out.
inferences of (56) arise in much the same way as actuality entailments in languages with overt grammatical aspect, relying on independent tests for semantic restructuring (Hacquard 2008; Grano 2012) that suggest different forms of semantic composition are active in (56) and (58).

3.2 Conversationalist approaches

In contrast to conventionalist accounts, which are generally happy to associate lexical items with fairly rich pragmatics effects, conversationalist accounts attempt to reduce the amount of lexical stipulation as much as possible—rather relying heavily on general conversational principles to derive inferences. Such an approach holds promise not only because it is potentially more formally elegant—requiring fewer stipulations—but also because it places less burden on a language learner who would otherwise need to learn such stipulations (see Dudley 2017 and references therein for discussion of how factive inferences are acquired).

In practice, nearly all of these accounts, must posit some amount of conventional lexical knowledge—some more than others. For instance, nearly all conversationalist approaches of semifactives start from the assumption that know $P$ entails $P$—i.e. that know is veridical. But in general, the goal is to squeeze as much as possible out of the interaction of such knowledge with general communicative principles. In line with this, it behooves these approaches to find cases of lexically triggered inferences that appear to be not only defeasible, but also relatively nondetachable from an expression’s content—i.e. independent of the particular linguistic expression of some content.

The inferences triggered by semifactives and aspectuals (but generally not true factives or implicatives) are common targets, since at least some of them have long been known to be defeasible (as discussed in Section 2). For instance, Simons (2001) argues for a conversationalist approach to some of the veridicality inferences associated with semifactive and change-of-state constructions—in particular, those that project through negation—on the basis that, in addition to being defeasible, at least some of those inferences are like conversational implicatures (Grice 1989) in also being nondetachable. This is evidenced in (59)—where different paraphrases of stop also trigger the inference that Jane was laughing—and (60)—where different paraphrases of realize also trigger the inference that Harry was a fool.

(59) Jane didn’t {quit, cease, discontinue} laughing (Simons’ Ex. 14)
(60) Harry didn’t {realize, come to know, become aware} that he was a fool. (Simons’ Ex. 16)

Her explanation elaborates on an idea proposed by Stalnaker (1974) and has two components: (i) the use of a sentence that expresses some proposition $P$—e.g. that Jane stopped laughing—raises the question of
whether or not a proposition P is true—e.g. either that Jane stopped laughing or that she didn’t; and (ii) insofar as proposition P entails proposition Q (but not vice versa)—e.g. that Jane stopped laughing entails that Jane smoked—"[a] speaker who raises the question whether P indicates a belief that Q is true” (p. 13). Similar reasoning can be used to derive factive inferences under the assumption that know p implies p (but not vice versa). Thus, while this account requires that lexically triggered veridicality inferences are derived at least in part from conventional knowledge about lexical items, the relevant conventional knowledge is relatively minimal.

Simons (2001) notes that this proposal massively overgenerates veridicality inferences. For instance, for simple negated sentences involving a transitive verb, like (61), this proposal predicts that the inference that Jo ate something is triggered, contrary to fact: under her proposal, (61) should raise the question whether Jo ate the sandwich; but because Jo ate the sandwich entails that Jo ate something (but not vice versa), the inference that Jo ate something should be triggered.

(61) Jo didn’t eat the sandwich.

In the same vein, for change-of-state constructions, like stop, this proposal predicts not only the attested inference that Jane smoked before the reference time but also the unattested inference that she didn’t smoke after the reference time. To fix this latter case, she proposes to appeal to a conceptual distinction: in the particular case of change-of-state constructions, she posits that there is an important difference between pre-conditions for an event and its post-conditions.16

To address this overgeneration (among other issues), Abusch (2002, 2010) proposes an alternative approach: (i) lexical items are associated with alternative lexical items; (ii) in using a lexical item associated with such a set, a speaker signals that they are presupposing the disjunction of those alternatives. This approach can be exemplified using the factive know that, for which she posits be unaware that as an alternative. Abusch assumes that, as a matter of lexical knowledge one infers from x knows P that P and x believes P and from x is unaware P that P and x does not believe P. Thus, in using x knows P (or x is unaware P or x does not know P and so on), a speaker presupposes (P and x believes P) or (P and x does not believe P), from which P (the veridicality inference) can be inferred. (A similar derivation can be given for stop under the assumption that it has continue as an alternative.) To capture variability in whether a lexical items triggers a particular inference, Abusch appeals to mechanisms drawn from Heim 1983b.

16 A similar fix would be necessary for factives like know: x knows P assymmetrically entails x believes P and thus Simons’ account predicts that one should infer x believes P from x does not know P, contrary to fact.
Simons’ and Abusch’s approaches are abstractly related in the sense that they both posit that lexically triggered veridicality inferences are derived by reasoning over alternatives; the main difference is in the source of those alternatives: a general principle for Simons and conventional lexical knowledge for Abusch. This allows Abusch to tune the alternative sets to fit the attested veridicality inferences better than a highly general approach like Simons’; but it does so at the cost of potentially being less explanatory, since she does not propose a general rule for determining the alternative set a lexical item is associated with. Much following work in the conversationalist vein can be seen as attempting to interpolate between these two approaches.

Some such approaches attempt to further constrain the alternatives lexical items can be associated with. For instance, Romoli (2011, 2014), following Chemla 2009, 2010, does this by associating lexical items like know and stop with alternative sets (like Abusch) in which they are the strongest element—i.e. wherein they asymmetrically entail all other alternatives (like Simons). For instance, he associates know \( P \) with the alternative \( P \) and stop \( P \) with the alternative used to \( P \)—like other accounts, assuming that know \( P \) entails \( P \) and that stop \( P \) entails used to \( P \) as a matter of convention.

Other approaches similarly focus on the notion of entailment for computing lexically triggered veridicality inferences but do not make the notion of lexically associated alternatives central—instead relying on a mix of lexical knowledge and informational structural notions (Abbott 2000, 2006; Simons et al. 2010; Beaver et al. 2017). For instance, Abrusán (2011, 2016) proposes a system that relies crucially on the notion of sentences and their uses (utterances) having a set of inferences that are part of their main point (or at-issue). For her, an utterance’s main point is determined by bottom-up influences, such as conventional knowledge about the meaning of a construction, and top-down influences coming from the context; and anything that is not part of the main point of an utterance will project through entailment cancelling operators (see also Simons et al. 2010). The bottom-up influences determine the basic set of inferences that must be part of the main point (cf. Wilson and Sperber 1979)—effectively, all and only the inferences that are about the event described by the matrix clause predicate (or more precisely the time span over which it occurred)—and the top-down influences—e.g. focus indicated by intonation (see Section 2.2.1)—may add to these inferences. Variability comes about as a produce of which inferences these top-down influences affect.

Simons et al. (2010, 2017) propose an alternative that attempts to reduce reliance on bottom-up influences, like the ones posited by Abrusán, moving back in the direction of (Simons 2001)’s original question-based proposal. They do this by deriving the question any particular use of a sentence raises (or is used to answer) via its focus alternatives (Rooth 1992). Specifically, they assume that all sentences have focus on at least one constituent (Selkirk 1984) and that the question raised by the use of a sentence is (roughly) whichever one would be derived by replacing the focused constituent with a WH word and, in
prototypical cases, removing entailment cancelling operators. For instance, prototypically, the question raised by (62) is (62)a.

(62) \[Jo\]_f doesn’t know that Bo left.
   a. Who knows Bo left?
   b. \{X knows Bo left \mid X is animate and contextually relevant\}

Under standard approaches to questions, constituent questions, such as (62)a, correspond to a set of (possible, true, complete, etc.) answers (Hamblin 1973; Karttunen 1977; Groenendijk and Stokhof 1984) wherein the domain of the WH word is contextually restricted, such as those in (62)b. They refer to this set as the current question and posit that “[p]rojection of the content of the complement of an attitude verb occurs if the Current Question for the utterance entails this content” (p. 192), where a question entails some content if the disjunction of its answers entails it.\(^{17}\) Under the assumption that \(X\) know \(P\) entails \(P\) for any \(X\) this account then derives the lexically triggered veridicality inference of (62).

In contrast, the inference is not necessarily triggered by the use of (63).

(63) Jo doesn’t \[know\]_f that Bo left.
   a. How is Jo related to the proposition that Bo left?
   b. \{Jo Vs Bo left \mid V is a contextually relevant relation between entities and propositions\}

For instance, if the contextually relevant relations turn out to include believing, then at least one alternative will not entail that Bo left, and therefore no inference about Bo’s actually leaving will be triggered. But if the context only supports alternative relations that are themselves veridical—e.g. discovering, realizing, etc.—then the inference is triggered by the use of (63).

Simons et al.’s explanation gets a bit trickier for sentences like (64), since the proposition that Bo left does not show up in all alternatives, and therefore, not all alternatives entail that Bo left.

(64) Jo doesn’t know \[that Bo left\]_f.
   a. What does Jo know?
   b. \{Jo knows \(P\) \mid \(P\) is a contextually relevant proposition\}

\(^{17}\) Note the similarity to Abusch 2002, 2010 and dissimilarity from Simons 2001, which crucially requires that only the positive alternative in the polar question raised by the use of the sentence is relevant.
To remedy this, they appeal to constraints on the set of contextually relevant propositions determined by the selection restrictions of the predicate—specifically, that to know something, that thing must be true.

A main contribution of Simons et al.'s theory is that it integrates lexically triggered veridicality inferences into a broader conversationalist theory of discourse in much the same way as Heim 1992 and van der Sandt 1992 within conventionalist approaches. To do this integration, they rely on Roberts' (1996) notion of a Question Under Discussion (QUD; see also Rooth 1992; D. Beaver and Clark 2008) for defining their notion of what is (not) at-issue. The basic idea behind their QUD-based approaches is that (a certain subset of coherent) discourses move forward by attempting to answer a(n explicitly or implicitly determined) discourse question. To answer this discourse question, subinquiries can be raised, subject to relevance constraints on whether the subinquiry will help answer the discourse question; and if they satisfy these constraints, they become the current question. This current question may in turn act as a discourse question for another subinquiry; or it may be answered, in which case the discourse question it was a subinquiry of becomes the current question again. Crucially, to be licitly used in answering a question, a sentence must have a focus structure that is compatible (or congruent) with the current question; or if not compatible with the current question, the focus structure must be associated with a question that might reasonably become the current question because of its relevance to answering the discourse question. Following the hypothesis laid out above, context projects insofar as a current question that gets introduced this way entails that content.

This proposal is reasonably viewed as a recasting of standard conventionalist accounts that rely on (global) accommodation to capture factive inferences into a conversationalist framework. One potential difference is that Simons et al. argue that their system can actually handle further cases that seemingly do not rely on lexical knowledge at all. For instance, believe is a prime case of a nonveridical predicate, but in certain contexts, it can appear factive (or at least implicative). Consider (65) and (66) in a context where Phil is looking for the car, which Amy has parked in the parking garage, contrary to her normal tendency, and which Polly does not know.

(65) Simons et al.'s Ex. 34
Polly: Why is it taking Phil so long to get here?
Amy: He didn’t know that the car’s parked in the parking garage.

(66) Simons et al.'s Ex. 36
Polly: Why is it taking Phil so long to get here?
Amy: He didn’t believe that the car’s parked in the parking garage.
Simons et al. argue that the inference that the car is parked in the parking garage is derived the same way in both cases. Specifically, they hypothesize that “[p]rojection of the content of the complement of an attitude verb occurs if the best explanation for relevance of the [current question] to the [discourse question] requires attribution of acceptance of that content to the speaker” (p. 192), and thus, the inference triggered by Polly’s utterances in both (65) and (66) are derived by general considerations of relevance.

4. Conclusion

The last 50+ years of research on lexically triggered veridicality inferences have provided a wealth of empirical observations, formal tools, and philosophical perspectives on the use of those tools. Nonetheless, there is still much work to be done in understanding the nature and source of these inferences: while different lexically triggered veridicality inferences seem likely to arise in very different ways (Karttunen 2016), there currently exists no single theory that predicts all the extant inference patterns discussed in Section 2 of this paper under a cohesive system of semantic and pragmatic mechanisms (cf. Giannakidou and Mari to appear and references therein). Developing such a cohesive system requires further disentangling the role of lexical and compositional knowledge from that of conversational principles. Such a disentanglement itself requires integrating insights from the vast literature on the semantics of propositional attitude reports and clause-embedding (see Grano in prep for a review) with the equally vast literature on formal pragmatics. But the time is right for such an integration: not only do researchers have access to massive corpora annotated for veridicality (Saurí and Pustejovsky 2009; 2012; de Marneffe, Manning, and Potts 2012; Lee et al. 2015; White et al. 2016; Stanovsky et al. 2017; Rudinger, White, and Van Durme 2018) that can be used to evaluate theoretical coverage, they also have at their fingertips sophisticated, explicit computational models of pragmatic reasoning (Frank and Goodman 2012; Goodman and Stuhlmüller 2013) that have shown initial promise in capturing lexically triggered veridicality inferences (Qing, Goodman, and Lassiter 2016).

References


