Presupposition: What went wrong? *

Lauri Karttunen
Stanford University

Abstract When the first generation of generative linguists discovered presuppositions in the late 1960s and early 1970s, the initial set of examples was quite small. Aspectual verbs like *stop* were discussed already by Greek philosophers, proper names, *Kepler*, and definite descriptions, *the present king of France*, go back to Gottlob Frege and Bertrand Russell by the turn of the century.

Just in the span of a few years my generation of semanticists assembled a veritable zoo of ‘presupposition triggers’ under the assumption that they were all of the same species. Generations of students have learned about presuppositions from Stephen Levinson’s 1983 book on *Pragmatics* that contains a list of 13 types of presupposition triggers, an excerpt of an even longer unpublished list attributed to a certain Lauri Karttunen. My task in this presentation is to come clean and show why the items on Levinson’s list should not have been lumped together.

In retrospect it is strange that the early writings about presupposition by linguists and even by philosophers like Robert Stalnaker or Scott Soames do not make any reference to the rich palette of semantic relations they could have learned from Frege and later from Paul Grice. If we had known Frege’s concepts of *Andeutung* – Grice’s *conventional implicature* – and *Nebengedanke*, it would have been easy to see that there are types of author commitment that are neither entailments nor presuppositions.

Keywords: presupposition, projection, conventional implicatures, question-under-discussion, at-issue-meaning, factives, implicatives, invited inferences

1 Levinson’s list

Fifteen years ago, in the summer of 2001 at the European Summer School in Logic, Language, and Information in Helsinki there was a special event to mark *Twenty Years of Two-Level Morphology*. My presentation was titled “A Short History of Two-Level Morphology.” It recounted the story of how finite-state transducers came

* Warm thanks to my friends and colleagues in the Language and Natural Reasoning group at the Center for the Study of Language and Information at Stanford, especially to Cleo Condoravdi, Stanley Peters, and Annie Zaenen, for all the discussions we have had on this topic.

©2016 Karttunen
to be used for morphological analysis and generation: they were compact, efficient, and applicable to all languages.

After the talk and the reception that followed, I was walking back to my hotel. I heard quick steps approaching me from behind. It turned out to be a young woman, out-of-breath because she had been running to catch up with me. She started our conversation with a breathless question:

- Are you the SAME Lauri Karttunen who wrote *Presuppositions of Compound Sentences*?
- Yes, I am.
- I DIDN’T KNOW THAT YOU WERE STILL ALIVE!

She introduced herself as Jennifer Spenader, a graduate student at the University of Stockholm. Her surprise was real. I had not published anything on semantics in the previous decade, the paper on the presuppositions of compound sentences was nearly twenty years old by then. Our dialogue continued:

- Levinson says you have a list of 31 presupposition triggers but he only mentions 13. Please send the full list to me. I am writing my dissertation on presuppositions.¹
- I am not sure that I still have that list. I will send it to you if I find it.

Fifteen years later by now, I am still the same person, happy to be alive, but I have never found a copy of that old class handout. I don’t know how Levinson got hold of one; I never had any contact with him. That attribution has haunted me ever since.

Here is the list of PRESUPPOSITION TRIGGERS as it appears in Levinson 1983: 181–184:

i. Definite descriptions (Strawson1950, 1952)
ii. Factive verbs (Kiparsky & Kiparsky, 1971)
iii. Implicative verbs (Karttunen, 1971)
iv. Change of state verbs (Sellars, 1954; Karttunen, 1973)
v. Iteratives
vi. Verbs of judging (Fillmore, 1971)
vii. Temporal clauses (Frege,1892 (1952); Heinämäki, 1972)
ix. Implicit clefts with stressed constituents (Chomsky 1972; Wilson & Sperber, 1979)
x. Comparisons and contrasts (Lakoff, 1971)
xi. Non-restrictive relative clauses
xii. Counterfactual conditionals
xiii. Questions (Katz 1972; Lyons 1977)

The error, from which the field still has not completely recovered, was the idea that the items on this list exemplify the same phenomenon. The zoo of presupposition

¹ Completed in the following year (Spenader 2002).
Presupposition: What went wrong?

triggers should have been constructed with separate cages for different species. The quest for a unified theory of presupposition, pragmatic or semantic, has been a failure.\(^2\)

By the time linguists discovered presuppositions in the late 1960s (Horn 1969; Morgan 1969; Langendoen & Savin 1971), the topic had been discussed by logicians and philosophers for a long time (Frege 1892; Russell 1905; Strawson 1950).\(^3\)

In hindsight it is a pity that the philosophers and linguists engaged in the early discussions about presuppositions in the 1970s only referenced Frege’s 1892 paper on *Über Sinn und Bedeutung*. They all seem to have been unaware of the relevance of Frege (1918), a paper called *Der Gedanke* (The Thought).\(^4\) I discovered this work only a couple of years ago preparing my talk for SALT 24. Larry Horn (2007) had taken notice of it much earlier. Although it does not contain all the distinctions that should be made, Frege (1918) would have been a good starting point.

2 Frege’s *Voraussetzung*, *Andeutung*, and *Nebengedanke*

Frege’s paper on *Sinn und Bedeutung* (Sense and Reference) discusses of proper names like *Kepler* and definite descriptions like *the one who discovered the elliptic form of the orbits of the planets*. In Frege’s view main sentences denote a truth value. Any sentence containing a term referring to something that does not exist is fundamentally defective, it does not have a truth value. The negation of such a sentence is also neither true nor false. *Kepler died in misery* and *Kepler didn’t die in misery* both presuppose the existence of Kepler. The word *Voraussetzung* is Frege’s term for presupposition.\(^5\) Frege had a cumulative view of presuppositions. The temporal clause in *After the separation of Schleswig-Holstein from Denmark*,...
Prussia and Austria quarreled presupposes that Schleswig-Holstein was separated from Denmark. If that had not happened, the sentence would not denote anything. Frege does not discuss the presuppositions of compound sentences formed with if... then, and, and or, which would have falsified the cumulative hypothesis (Karttunen 1973).

The 1892 article remarks that a connective like although, but, and the untranslatable doch doesn’t affect the sense of the sentence that follows but illuminates it in a peculiar fashion. Replacing the sentence with another one with the same truth value would not affect the truth value but it might be seen as inappropriate like singing a sad song in a lively way (Frege 1892: 226). The later Der Gedanke paper spells out this idea in more detail:

Much of language serves the purpose of aiding the hearer’s understanding, for instance the stressing of part of a sentence by accentuation or word-order. One should remember words like "still" and "already" too. With the sentence “Alfred has still not come” one really says "Alfred has not come" and, at the same time, hints that his arrival is expected, but it is only hinted. It cannot be said that, since Alfred’s arrival is not expected, the sense of the sentence is therefore false. The word "but" differs from "and" in that with it one intimates that what follows is in contrast with what would be expected from what preceded it. Such suggestions in speech make no difference to the thought.

... But what is essential depends on one’s purpose. To a mind concerned with what is beautiful in language what is indifferent to the logician can appear as just what is important. (Frege 1918: 295–296).

As Horn (2007) points out, this Frege’s idea of “hinting” or “intimating” surfaces half a century later as Paul Grice’s notion of conventional implicature as the following convoluted passage indicates:

If I say (smugly) He is an Englishman; he is, therefore, brave, I have certainly committed myself, by the virtue of of the meaning of my words, to its being the case that his being brave is a consequence (follows from) his being an Englishman. But while I have said that he is an Englishman, and said that he is brave, I do not want to say that I have said (in the favored sense) that it follows from his being an Englishman that he is brave, though I have certainly indicated, ad so implicated that this is so. I do not want to say that my utterance of this sentence would be, strictly speaking false should the consequence in question fail to hold. (Grice 1975: 44–58).

The word translated as “hinting” in one sentence and as “intimating” in the other is the same German word in Frege’s text: andeuten. As Horn suggests, in the best of
Presupposition: What went wrong?

worlds we could adopt its nominalization, Andeutung, as the name for the semantic relation that we now clunkily refer to as conventional implicature. It may be too late for that. Grice himself doesn’t make much of the notion, he is mainly interested in CONVERSATIONAL IMPLICATURES and his conversational maxims.

Included under Frege’s concept of Andeutung are the affective meaning of words like bastard and jerk in Christopher Potts’s (2005) work on conventional implicature. The choice between the neutral horse or the pejorative nag makes no difference to the thought. The assertive force does not extend over that in which these words differ (Frege 1918: 295).

The third semantic relation that appears in Frege’s writings is Nebengedanke. Max Black translates it as subsidiary thought. In Frege 1892: 44,47 it appears in his discussion of non-restrictive relative clauses. According to Frege,

1. Napoleon, who recognized the danger to his right flank, himself led his guards agains the enemy position.

expresses two thoughts:

2. a. Napoleon recognized the danger to his right flank.
   b. Napoleon himself led his guards agains the enemy position.

If one or the other of the clauses in (2) is false, the whole sentence (1) is false. There may also be a subsidiary thought that the knowledge of the danger was the reason for Napoleon’s action. But one may in fact doubt whether this thought is merely lightly suggested or really expressed (Frege 1892: 47). This characterization of subsidiary thought sounds very similar to what Michael Geis and Arnold Zwicky (1971) call an INVITED INFERENCE.

If the linguists and philosophers at the time of the first boom of presupposition studies around 1970 had been in command of Frege’s palette of semantic relations ranging from Voraussetzung (presupposition) to Andeutung (conventional implicatures) to Nebengedanke (invited inference), the picture that emerged might have

6 The best English translation for Frege’s andeuten might be allude, with allusion as the nominalization.
7 Kent Bach (1999) argues that Frege and Grice are wrong about still and but. For him the existence of any non-truthconditional meaning is a myth. Potts (2005) agrees with Bach about therefore and but.
8 Like many 20th century logicians, Frege rejects the idea that universal quantifiers all and every have existential import, contrary to Aristotle and the recent developments in Natural Logic (Valencia 1991). In a letter to Husserl (Frege 1997) he argues that in expressions like all M are N do not presuppose that there are any M, nor do they commit the speaker such a view as an Andeutung would do, although it is a subsidiary thought. In Horn’s estimate, Frege’s treatment of the existential import of universals as a pragmatic aspect of communicated meaning rather than a semantic implication actually allows for a subtler and more insightful account of quantified statements than does the presuppositional story (Horn 2007: 45).
been more nuanced, not limited to just two semantic relations: entailment and presupposition.

3 Levinson’s list: a second look

Looking at Levinson’s list in the light of Frege’s distinctions we can recognize some clear cases of Fregean presuppositions, conventional implicatures, and invited inferences. Let us first provisionally winnow the list presuppositions to the following easy cases.

3.1 Easy cases

i. Definite descriptions
v. Iteratives
vii. Temporal clauses
viii. Cleft sentences
ix. Implicit clefts with stressed constituents
x. Comparisons and contrasts
xiii. Questions

In making this selection we of course need to get away from the narrow notion of logical presupposition that is well-defined only for expressions that could have a truth value. We recognize that questions and commands can have presuppositions. Although Frege’s concept of presupposition does not apply in cases such as

(3) a. When did Kepler die?
    b. Tell me about Kepler!

Frege probably would have agreed with us that, if Kepler had never existed, the expressions in (3) would be flawed. Questions like (3a) would have no answer, commands like (3b) could not be followed. They fail as speech acts. As many commentators have noted, e.g. Soames (1976: 169), we have intuitions about whether statements are true or false but we have no intuitions about whether something is truth-valueless rather than false. That latter distinction is a technical construct, it is misguided to take it as basic concept, we need a pragmatic notion. Let us say for now that presupposition failure results in a serious failure in communication. More about that later in Section 5.

3.2 Factives

The item that has generated more controversy over the years than anything else on Levinson’s list is Factives. The verb and adjective constructions listed in Kiparsky &
Presupposition: What went wrong?

Kiparsky (1970) are a heterogeneous collection. The article was written and the height of the Generative Semantics movement in the late 1960s that sought to represent semantic relations in the underlying syntactic representation in Deep Structure, a notion that is completely out of fashion now. In other respects, the article is still a good survey of primary data but one should have distinguished at the very beginning at least the following five types of expressions:

ii. Factive verbs
   a. Certain predicates with that-clause subjects:
      that $S$ be odd/tragic (as opposed to likely)
      that $S$ count/matter/suffice (as opposed to happen)
   b. Certain emotive adjectives with complements:
      $NP$ be happy/glad/furious that $S$ (as opposed to hopeful)
      $NP$ be sad/delighted/disappointed to $VP$ (as opposed to willing)
   c. Certain propositional attitude verbs:
      $NP$ know/regret/forget/remember that $S$ (as opposed to believe)
   d. Verbs of discovery:
      $NP$ discover/find out/notice/observe that $S$ (as opposed to suspect)
      $NP$ be discovered/found out/noticed/observed to $VP$ (as opposed to suspected)
   e. Certain verbs of communication: acknowledge, admit, confess (as opposed to say)

The subgroup (ii.a) belongs to the easy cases listed in (3.1). These constructions clearly pass the classical tests for presuppositions (Chierchia & McConnell-Ginet 2000: 350): the truth of a presupposition is preserved under negation, questioning and in the antecedents of conditionals. Examples such as

(4) a. Isn’t it odd that desire lasts so much longer than the ability to perform?
   b. Does it count that I celebrated every inch of you?

commit the speaker to the truth of the that-clause.

These easy cases involve just one person, the author, and the world that she inhabits as it is already known to be, no future contingencies, not dream worlds, or worlds as imagined by other people. 9

The most commonly cited examples of factive verbs, know, regret, remember etc. in subgroup (ii.c), are more complicated because in addition to the author there is another person involved, the attitude holder which we will call the protagonist. Even when the author and the protagonist are in some sense the same person they may be in different worlds as Morgan (1969) demonstrated with examples like:

(5) I dreamt that I was a German and nobody knew it.

9 In literary works matters can become more complicated. In a style called free indirect discourse the narrator gives the illusion of being one or several of her characters and saying what they would say.
It is clear that we can have false beliefs about the actual world and true regrets based on those false beliefs:

(6)  
   a. We misheard that the dinner was free and were glad to not to have to pay.  
   b. Sally misremembered not having left a tip and regretted it.  

The question is what happens if we replace regretted in (6b) by knew or were aware? Is there a context where it is OK to say that someone ‘knows’ or ‘is aware of’ something that is false in the actual world? That is controversial. In the case of Morgan’s dream world in (5) know is fine because the complement is true in that world. My impression is that the question about the actual world is more a philosophical issue than a question about the proper use of English. Depending on the answer the (ii.c) class might have to be split in two.

In contexts where the difference between the beliefs of author and those of the protagonist have not been spelled out, the default assumption is that they are aligned. If there is no reason to think otherwise,

(7)  Sally regretted not having left a tip.

does commit the author to the proposition that in the actual world Sally did not leave a tip and regretted. But the inference that it is the actual world where that is the case rests on the assumption that, in the absence of any indication to the contrary, the beliefs of the author are aligned with the beliefs of the protagonist. Grice probably would have called this a generalized conversational implicature.

The ‘coming-to-know’ verbs in (ii.d), discover, find out, etc. commit the author to the truth of the complement in affirmative assertions but it has been known from the beginning that these ‘semifactives’ are different from verbs like regret. Stalnaker (1974), Gerald Gazdar (1979), Rob van der Sandt (1992) gave rather elaborate pragmatic accounts of how the presuppositions of these particular verbs could get cancelled in situations where truth of the complement is an open issue. The presence of negative polarity items in the complement clause in examples such as (8a) and (8b) indicates that the author is not sure of the truth of the complement. The author of (8c) seems not to take any stand on the final result of the police investigation.

(8)  
   a. The police did not discover that any cars had been tampered with.  
   b. If the police did not discover that any cars had been tampered with, the suspects should let go.  
   c. Perhaps the police will discover that some cars had been tampered with but so far they haven’t.

The pragmatic accounts of the presupposition cancellation were formulated for first and second person subjects. As David Beaver (2010) pointed out, they don’t
Presupposition: What went wrong?

carry over to the type of examples in (8). Beaver gives an abundance of examples similar to (8) from the internet of cases where it is clear from the context that the author is not committed to the truth of the complement clause.

Instead of making the assumption that the ‘coming-to-know’ verbs are presupposition triggers only to have to tie yourself into knots trying to explain how that presupposition sometimes get cancelled or goes away, and doesn’t project, it is better to admit that it was a mistake to assume that these verbs have presuppositions in the classical sense. They are in a class by themselves. Affirmative assertions definitely commit the speaker to the truth of the complement but negative sentences, questions, and conditionals with coming-to-know verbs are in principle noncommittal. Most of the time there are enough clues in the context to indicate whether the author takes the complement as true or as yet to be established but it is not part of the lexical meaning of these verbs. What is part of it, however, is that discovering, realizing, noting, etc. lead to facts. There is a real distinction between coming to know that \( p \) and coming to think that \( p \); discovering that \( p \) and assuming that \( p \). In retrospect, calling the discovery verbs ‘semifactives’ was not a bad choice of a term once you decouple it from the notion of presupposition.

Finally, the last group of factives (i.e., acknowledge, admit, confess, etc. do not commit the speaker to the truth of anything other than the protagonist having communicated something that the protagonist wished to present as a fact. The different flavors of these verbs involve conventional implicatures, Frege’s Andeutungen: affirming what is expected (acknowledge), what was denied before (admit), or revealing some new personal information (confess).

With these verbs it is often difficult to figure out the relationship between the protagonist’s and the author’s world view. In an interview with ex-Vice president Dick Cheney, if the reporter asks Was the Iraq war a mistake? and Cheney says No, the interviewer could report the answer as

\[(9) \quad \text{Cheney did not acknowledge that the Iraq war was a mistake.}\]

The audience on Fox News of (9), most likely assumes that the world view of the reporter aligns with that of Cheney, hence the war was not a mistake contrary to some contrary opinions. The audience of MSNBC would have a different interpretation of (9), the war was a mistake, Cheney is living in some parallel universe.

3.3 Implicatives

Karttunen (1971) claimed that the characteristic feature of implicative verbs such as manage is that they presuppose some sufficient and necessary condition for the event described by their complement clause. For example, (10a) and (10b) would
presuppose that solving the problem was difficult for Ann or that it was unexpected that she would succeed.\textsuperscript{10}

(10) \begin{enumerate}
\item a. Ann managed to solve the problem.
\item b. Ann didn’t manage to solve the problem.
\end{enumerate}

Rebekah Baglini and Frances Itamar (2016) have more complicated theory of what is presupposed. According to them the presupposition of manage is that there is some causal element that is necessary for the truth of the complement but that is insufficient to bring it about without an additional, situation-dependent component, a ‘catalyst’ that determines whether or not Ann actually solved the problem. For them assertion of (10a) is something like

(11) The catalyst, which was necessary but not sufficient for Ann to solve the problem, actually caused Ann having solved the problem.

One advantage that this analysis has is that it doesn’t attempt to pin down whether the catalyst involves trying, difficulty, unexpectedness, likelihood, or something else. Of all the two-way implicative verbs of English manage is the most bleached, least specific in characterizing the impediment that the protagonist has to overcome. However, it seems a mistake to me now to think of that as a presupposition. In retrospect it seems that the two-way implicatives in (12) have all the hallmarks of Frege’s Andeutungen (Grice’s conventional implicatures).

(12) \begin{enumerate}
\item a. Ann remembered/bothered/dared/deigned/happened to open the door. 
\[ \models \text{Ann opened the door.} \]
\item b. Ann didn’t remember/bother/dare/deign/happen to open the door. 
\[ \models \neg \text{Ann didn’t open the door.} \]
\end{enumerate}

The positive examples in (12a) entail that Ann opened the door, the negative ones in (12b) entail that she didn’t. By using one of the implicative verbs the author commits herself to some proposition over and beyond what is asserted. It is not always clear what that additional piece of meaning is. By saying Ann remembered to open the door the author may imply that it was Ann’s prior intention to do so or she might mean that Ann did what was expected of her. This implication is conventional, it is part of the lexical meaning of remember in the remember to VP construction. But just as in the case of Frege’s yet and Grice’s therefore, the implicated proposition has no effect on the truth of the author’s statement. Even if it turns out that Ann just

\textsuperscript{10} Bach (1999: 334) argues unconvincingly that whatever manage contributes to the meaning of Bill managed to finish his homework is not implicated. The content “includes both the finishing and the entailed difficulty.” But whoever asks Did Bill manage to finish his homework? is asking whether Bill finished his homework, not about the combination of finishing and difficulty.
happened to open the door, without intending or anybody expecting it, the assertion *Ann remembered to open the door* is, strictly speaking, true. The answer to a yes-no question like *Did Ann remember to open the door?* is the same as the answer to *Did Ann open the door?* A command like *Remember to open the door!* can be understood and followed even if the implication that the addressee is supposed to open the door is conveyed as part of the command itself.

The fact that the *happen to* VP is one of the two-way implicative constructions illustrated in (12) is difficult to reconcile with the Baglini-Frances idea that two-way implicatives are associated with a presupposition involving some necessary causal catalyst. The author of *Ann happened to open the door* alludes that Ann’s opening or not opening the door was accidental. The negation *Ann didn’t happen to open the door* is just the same: Ann didn’t open the door but she might have, and whether she did or not was due to pure chance as far as the author cares or knows about.

The verbs in (12) have the entailment signature +|- meaning that they yield a positive entailment in positive contexts and a negative entailment in negative contexts. The two-way implicatives in (13) have the opposite signature -|+, negative entailment in the positive, positive entailment in the negative.\(^\text{11}\)

(13) a. Ann failed/forgot/neglected to open the door.
   \[\vdash \text{Ann didn’t open the door.} \quad -|+\]

b. Ann didn’t fail/forget/neglect to open the door.
   \[\vdash \text{Ann opened the door.} \quad -|+\]

The two-way implicatives yield an entailment in both positive and negative contexts. The four types of one-way implicatives yield an entailment only under one polarity but a suggestion (an invited inference) under the other.

(14) a. Ann was not able to speak up.
   \[\vdash \text{Ann didn’t speak up.} \quad o|-\]

b. Ann refused to speak up.
   \[\vdash \text{Ann didn’t speak up.} \quad -|o\]

c. Ann was forced to speak up.
   \[\vdash \text{Ann spoke up.} \quad +|o\]

d. Ann didn’t hesitate to speak up.
   \[\vdash \text{Ann spoke up.} \quad o|+\]

### 3.3.1 Invited inferences

If the polarity in (15) is reversed, there is no entailment. It is not a contradiction to say *I was able to speak up but chose not to do so.* However, if I say *I was able...\(^{11}\)

---

\(^{11}\) I am adopting here the notation used by Bill MacCartney and Cristopher Manning (2009), simpler than that in Karttunen 2012. For the two-way implicatives there are two classes, +|- (*manage*) and -|+ (*fail*). For the one-way implicatives, to be discussed shortly, there are four classes: +|o (*forced*), -|o (*refuse*), o|- (*able*), and o|+ (*hesitate*), where + indicates positive entailment, - negative entailment, and o no entailment under the given polarity: positive|negative.
Ann wasn’t able to speak up | Ann was able to speak up | Universe

<table>
<thead>
<tr>
<th>Ann wasn’t able to speak up</th>
<th>able</th>
<th>Ann was able to speak up</th>
</tr>
</thead>
<tbody>
<tr>
<td>≡</td>
<td>∧</td>
<td>≫</td>
</tr>
<tr>
<td>Ann didn’t speak up</td>
<td>¬ speak up</td>
<td>Ann spoke up</td>
</tr>
</tbody>
</table>

Ann didn’t speak up | Ann spoke up

Figure 1 A cover relation: Ann was able to speak up ⊾ Ann didn’t speak up

to read your paper last night and it turns out that I didn’t actually read it, even if I haven’t said an outright lie, my utterance invites the false inference that I read the paper. It would be perceived as very misleading in a situation where my having read the paper is a topic of interest for the addressee and whether I did or not is not yet known to her. This is not true of just the be able construction, all the one-way implicates exhibit to varying degrees the same phenomenon. This is what Frege called Nebengedanke, Grice generalized conversational implicature and Geis and Zwicky invited inference.

(15) a. Ann was able to speak up. ⊾ Ann spoke up. o |-  
b. Ann didn’t refuse to speak up. ⊾ Ann spoke up. -|o  
c. Ann was not forced to speak up. ⊾ Ann didn’t speak up. +|o  
d. Ann hesitated to speak up. ⊾ Ann didn’t speak up. o|+

That the effect is systematic is probably related to the fact that the main sentence with a one-way implicative verb and the complement clause are semantically in MacCartney’s COVER relation, ⊾, with negation on one member of the pair. For example, Ann was able to speak up ⊾ Ann didn’t speak up are in a cover relation because, in any universe of discourse, at every world one or the other is true and there are some worlds where both are true (MacCartney & Manning 2009: 3). This relation is pictured in Figure 1 where Universe is the set of all possible worlds split in two ways into four subsets: the worlds where Ann was/wasn’t able to speak up, the worlds where Ann spoke/didn’t speak up. The intersection of two of the four subsets in the middle consists of worlds where Ann was able to speak up but didn’t. To specify that the actual world is in that set, the speaker would have to say explicitly Ann was able to speak up but didn’t. If the speaker only says Ann was able to speak

716
Presupposition: What went wrong?

_up_ without the additional clause of Ann not speaking up, by Gricean maxims she can be interpreted as implicating that Ann did speak up, a scalar implicature of sorts. This explanation works for all the cases in (15). The cover relations are:  

(16) a. Ann was able to speak up ∼ Ann did not speak up
b. Ann did not refuse to speak up ∼ Ann did not speak up
c. Ann was not forced to speak up ∼ Ann spoke up
d. Ann hesitated to speak up ∼ Ann spoke up

The strength of invited inferences varies from construction to construction. If you look at the uses of _NP was able to VP_ on Google, you see that in an overwhelming number of cases the author meant to convey that the protagonist not only was able to but in fact did VP. With _NP hesitated to VP_ the effect, the suggestion that NP did not VP, is very weak but still noticeable. The phrasal one-way implicative _NP had a chance to VP_ the effect is strong but if the verb is _get_ rather than _have_ it suggests even more strongly that not only did the protagonist _get a chance to VP_ but he used it.  

There is no invited inference in cases such as (17) because the falsity of the complement proposition is explicitly indicated, as in (17a), or common knowledge, as in (17b):

(17) a. She had a chance to leave but chose to stay.
b. Bill Clinton had a chance to kill Bin Laden before September 11 attacks.

Like two-way implicatives, many one-way implicatives are triggers for conventional implicatures. The cover relation for _hesitate_ gives us two putative entailments: _Ann didn’t hesitate to speak up_ ⊨ _Ann spoke up_; _Ann didn’t speak up_ ⊨ _Ann hesitated to speak up_. We accept the first one as valid but not the second one. Hesitating to speak up implies considering whether to speak up. That implication does not follow from _Ann didn’t speak up_. In contrast, the two entailments in Figure 1 both are valid because, unlike _hesitate, be able_ is not encumbered by any _Andeutung_. It is

12 MacCartney & Manning (2009) discuss the cover relation only in connection with _hesitate_ pointing out that in MacCartney’s NatLog system, an entailment such as (14d) arises from the join, _⋈_, of two relations, negation and cover: _¬_ ν _∼_ ⊨ ⊥, where join is defined as _R_ ν _S_ \[def\] = _\{<x,z>: \exists y (<x,y> \in _R_ \land <y,z> \in _S_)>\} _⊥_ is MacCartney’s symbol for _forward entailment_. See MacCartney 2009: §5.

13 In an ongoing experiment with George Supaniratisai, Ignacio Cases and myself at Stanford, we are training an RNN (recurrent neural network) to match the inferences of human speakers of English. We assign to the _be able to VP_ not the logically correct _o|-_ signature but the _0.9|-_ signature reflecting our estimation that there is a _0.9_ probability that the author is using the construction as if it was a two-way _+|-_ implicative construction. We encode _get a chance to VP_ as _0.7|-_ and _have a chance to VP_ as _0.4|-_. These are our seat-of-the-pants estimates that sorely need empirical validation.
self-evident that whatever you do, you are able to do. It is not the case that for any action you did not do, you hesitated or ever considered doing it.  

3.3.2 Uninvited inferences

In addition to the non-logical inferences discussed in the preceding section there is another class of inferences that does not involve any Gricean principles. They arise from our common sense understanding of what kind of situation would prompt a speaker to make an utterance with a particular kind of semantic content. Some examples are given in (18).

(18)  
- a. I meant to answer your email right away. $\sim$ I didn’t answer your email right away.  
- b. I didn’t mean to hurt your feelings. $\sim$ I have hurt your feelings.  
- c. I intended to call on you last night, but it rained. $\sim$ I didn’t call on you last night.  
- d. I didn’t intend to kill my father. $\sim$ I killed my father.

The speaker of (18a) probably hasn’t quickly answered the addressee’s email. The speaker of (18b) probably thinks that she has hurt the addressee’s feelings. The construction \textit{intend/mean to VP} is of course not semantically of the same type as \textit{fail}, $-|+$. There is a pragmatic reason why we are inclined to draw such inferences. If we feel responsible for some bad outcome, a standard way of excusing ourselves is to assert that what happened was not what we intended, as in (18b) and (18d), or to say that we intended to something that we actually failed to do, as in (18a) and (18c). When someone talks about her intentions in the past tense, it is probably because the supposed intention was not in fact realized.

3.4 Aspectual verbs

Levinson’s class (iv), state of change verbs, includes items like \textit{stop} that are the oldest presupposition triggers in the literature. The Greeks regarded questions like \textit{Have you stopped beating your father?} as paradoxical because whichever way the addressee answers it, \textit{Yes} or \textit{No}, he ends up acknowledging that he has been beating his father. In affirmative statements \textit{stop} clearly commits the speaker to having beaten his father and it also passes the negation test for presupposition triggers.

14 The \textit{NatLog} system (2009), does not deal with presuppositions, conventional implicatures, or invited inferences. In \textit{NatLog}, \textit{NP managed to VP} and \textit{NP VPed} are equivalent: \textit{Ann managed to speak up} $\equiv$ \textit{Ann spoke up}. 

718
Presupposition: What went wrong?

But the paradoxical second person questions are a special case. In third person questions and in conditionals the judgements are less clear. As Dorit Abusch (2012) and others have shown, the if-test does not always lead to the expected result.

(19) a. If John stops smoking, Mary will buy him a camera.
   b. In a brochure addressed to resident:
      If you stopped smoking in 2001, you are eligible for a payment from Tobacco Indemnity Fund. (from Kadmon 2001)
   d. I have no idea whether Jane ever smoked, but she hasn’t stopped smoking. (from Simons 2001)

(19a) is unproblematic but the agency sending out (19b) presumably is not presupposing that a particular recipient is or has been a smoker. In this case stopped is understood as a shorthand for were smoking and stopped. Abusch speculates that the difference has to with how familiar the author is with the protagonist. Examples like (19c) show that by indicating that the question is based on speculation, the author escapes any commitment to the proposition that the addressee ever smoked. In this respect the aspectual verbs are similar to the semifactuals in (ii.d). In the case of (19d), as Mandy Simons (2001: 433) points out, the author may know some special symptom displayed by a person who has stopped smoking that Jane does not display. 15

What remains is that a command like Stop smoking does commit the author to the addressee being a smoker. Otherwise the command could not be followed.

3.5 Verbs of judging

Levinson’s class (vi) got onto the list by mistake. Charles Fillmore’s (1971) paper distinguishes three roles: the author, a judge, and a defendant. If the author is assuming the role of the judge and says It was very bad of you not to answer Harry’s letter, then I have criticized you for having done something I consider a bad thing. But if the author and the judge are different as in John criticized Bill for not answering Harry’s letter, the author is not taking any stand as to whether not answering Harry’s letter was or would have been bad or what Bill in fact has done or not done. It is the author’s interpretation of something that John said to Bill.

Since the verbs of judging do commit the author to the opinions she attributes to the protagonist, this class of verbs should not have made it to Levinson’s list.

15 This reasoning brings to mind Alan Ross Anderson’s (1951) example showing that subjunctive conditionals are not always counterfactual: If Jones had taken arsenic, he would have shown exactly those symptoms which he does in fact show.
4 Projection

As Beaver & Geurts (2014) observe in their SEP article:

For nearly four decades, the Holy Grail of presupposition research has been to explain the behavior of presuppositional expressions occurring in embedded positions.

There have been many proposals to explain the projection phenomenon. The gamut ranges from formidably complicated formal accounts such as Schlenker 2007 to a delightfully simple idea developed in Simons, Tonhauser, Beaver & Roberts 2010; Simons, Beaver, Roberts & Tonhauser 2016. I will not be able to tackle Schlenker’s theory but I will comment briefly in Section 4.2 on the Simons et al.’s proposals.

But first to the origin of the problem.

4.1 Attitude verbs

In his ‘plugs-holes-and-filters’ paper Karttunen (1973) found it difficult to decide whether a verb like believe was a ‘hole’ or a ‘plug’, that is, whether the presumed presupposition triggered by believe in

(20) Bill believes that Fred has stopped beating Zelda.

projects or gets stopped by the matrix verb (Karttunen 1973: 188-190). “It has proved difficult to get speakers to agree on what the right answer should be, which may indicate that the question has been put in a wrong way.” The final conclusion was that believe and other non-factive attitude verbs are plugs.

Karttunen (1974) considered that decision an error. His new view was that verbs like believe are neither plugs or holes but belong to a new class of verbs that require that the presuppositions of the complement are ‘satisfied’ by the beliefs of the protagonist. In other words, (20) presupposes that Bill believes that Fred has been beating Zelda. Irene Heim (1992) assumed that this new view was basically correct and incorporated it in a more ambitious and comprehensive theory:

The enterprise is carried out in a framework of context change semantics, which incorporates Stalnaker’s suggestion that presupposition projection results from the stepwise fashion in which information is updated in response to complex utterances. The empirical focus is on predicates of desire and on the contribution of counterfactual mood. (Heim 1992: 183)

It now seems self-evident to me that the view that Karttunen (1974) and Heim (1992) had on belief reports overlooked the fact the author has a considerable leeway in describing someone’s beliefs. Here is a simple example.
Suppose my friend Laura gets a ride in Kris’s new car to Ken’s house. Laura does not know much about cars but she is so impressed by the experience that she says to Ken, *I want to buy a car like Kris’s.* Later on Ken says to me *I think Laura wants to buy the kind of car that Kris has.* As Kris and Ken and Laura are all my friends, I know that Kris has a Tesla and that Ken and Laura are clueless about cars. They don’t know that the car Laura is enthused about is called *Tesla.* It would be quite OK for me in these circumstances to say to someone who knows about electric cars but doesn’t know anything about Kris:

(21)  
   a. Laura wants to buy an electric car.  
   b. Laura wants to buy a Tesla.  
   c. Ken thinks that Laura wants to buy a Tesla.

Heim (1992: 207–212) invokes the well-known *de re*/*de dicto* ambiguity to explain cases where someone’s desires and beliefs are described in language that the protagonist herself could not have used. For (21b), the two readings are crudely paraphrased as in (22):

(22)  
   a. There is a Tesla that Laura wants to buy.  
   b. Laura wants there to be a Tesla that she buys.

(22a) is false because there is no particular car that Laura wants to buy, but so is (22b) because, having no idea of what Teslas are, Laura cannot possibly have on a conscious level a desire to acquire a Tesla.  

In (21) there must be at least one person involved with the knowledge to say a *Tesla* to explicate what Laura meant by saying *I want to buy a car like Kris’s.* In principle it could be any one of the three, the author, Ken, or Laura herself if she knew about electric cars.  

It is not necessary or useful to view this in terms of presupposition satisfaction. In the course of describing other people’s beliefs and wishes speakers are sometimes required to bring in supplementary information as in (21), and free to put their own slant on the content. Here is an example.

(23)  
   a. Sanders:  
      I want to quadruple the tax on billionaires.  
   b. WSJ reporter:  
      Sanders wants radical changes in the US Tax Code.

---

16 It might be shown experimentally that Laura’s brain responds to an image of a Tesla in some significantly different way than to the image of any other car but there it is unclear of whether that by itself would be a confirmation of (22b) unless these brain responses correlate with wants and wishes. All that is outside of formal semantics for now at least.
c. Romney:

I believe Sanders wants to destroy the US economy.

The WSJ reporter’s description of what Sanders wants to do is a fair journalistic caption of what Sanders said. It is not a de re rendering of what Sanders said he wants but it is consistent with it. Sanders himself might well be happy with this headline provided that his specific idea of quadrupling the taxes on billionaires gets mentioned somewhere in the article.

In Romney’s world, it may well be the case that given his model of how the US economy works, plugging in a quadruple tax increase for people like himself has catastrophic consequences for the US economy as a whole. In the Romney world this statement may well correctly reflect the consequences of what Sanders wants to do. Opinionated and partisan, yes, but you cannot call Romney’s belief false on its own terms.

4.2 Question under discussion

In a series of papers and talks and in an ongoing research project Simons et al. have been developing a novel account of presupposition projection. Instead of viewing projection as a result of the inherent properties of particular lexical items and constructions, the aim is to give a pragmatic theory that accounts for the observed judgements on the basis of to basic concepts: FOCUS and QUESTION UNDER DISCUSSION (QUD). It is assumed that in every sentence some constituent, or the sentence itself, is in focus. The focus generates a set of alternatives for the focus position. The alternatives are formalized as questions (Simons et al. 2016: 2). The CONGRUENT QUESTION (CQ) is closely related to the structure of a sentence or a clause. The DISCOURSE QUESTION (DQ) relates to the questions that the ongoing discourse can be taken to resolve. For example, in a case like

(24)

a. James just found out that Harry’s having a graduation party.

b. He didn’t find out that [Ị-HARRY]’s having a graduation party.

the speaker of (24b) contradicts the speaker of (24a) preceding assertion. The focus on the name Harry indicates that this is an instance of METALINGUISTIC NEGATION (Horn 1985). There is no difficulty explaining why the complement does not project but one needs an explanation for why (24b) presupposes that someone in the set of alternatives to Harry is having a graduation party. Because find out is lexically marked as entailing the truth of its complement, all the alternatives entail that someone is having a graduation party, hence it projects. This explanation is said to be more parsimonious than the account that assumes conventional triggering.17

17 Simons et al. do not mention metalinguistic negation. They assume that the complement of find out “is not conventionally marked as projective” (Simons et al. 2016: 11–12). But since find out must be
Presupposition: What went wrong?

The claim that presupposition projection can explained without “conventional triggering” rests on two theses:

i. Every utterance has a some constituent, or the sentence itself, in focus.
ii. Every utterance relates to some QUD, either CQ or DQ.

Both of these assumptions seem implausible. The referred previous works, Selkirk 1984 for (i.) and Roberts 1996, 2012 for (ii.), present these as theses, not as empirical findings.

As for the complements of factive predicates, Simons et al. 2016 predict that under the right condition

Content which projects is either not-at-issue because entailed by the CQ; or not-at-issue because required to be invariant across all alternatives (true in the assumed common ground) in order to render the CQ of an utterance conversationally coherent. In both cases, the content which projects is content which does not address any current conversational question, either the CQ or the DQ. (Simons et al. 2016: 22)

It is unclear whether this theory makes any prediction at all for examples such as (4a) and (4b), repeated in (25) for convenience,

(25) a. Isn’t it odd that desire lasts so much longer than the ability to perform?
   b. Does it count that I celebrated every inch of you?

where the complement of the factive predicate appears in a question. Reading these examples in their original context (Shakespeare 1597; Putney 2010) suggests that there is no QUD that is being addressed by these particular utterances. Yet, the complement of a factive construction projects in such cases.

Simons et al. 2016, the latter of the two papers by the same set of authors, contains a surprising claim: under some circumstances the difference between know and believe disappears. Even the complement of believe projects, that is, it appears that the author is committed to it being true (Simons et al. 2016: 21). This finding does not appear in the earlier Simons et al. 2010 paper but it could have already been made there. The relevant example is simpler (Simons et al. 2010: 321):

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.

marked as entailing the truth of the complement, their explanation for why it projects relies just as much on lexical markings as the account they claim is less parsimonious.
Simon et al. go through an argument to show that, since the proposition you can eat raw vegetables does not provide an answer to Q, it is not at-issue and hence projects. Although the authors do not comment on this, switching from know to believe makes no difference in this case if everything else is kept the same. Given the same background scenario, the two examples in (26) seem equivalent as far as projection is concerned.

(26) A: They didn’t know that you can eat raw vegetables.
   A’: They didn’t believe that you can eat raw vegetables.

A similar observation is only made in the 2015 paper with a more complicated case that involves the same fallacy as this simple example. If we change the A’ answer in (26) to be (27) there is no projection of you cannot eat raw vegetables.

(27) A”: They believed that you cannot eat raw vegetables.

Why is that? The difference is in the background scenario. The speaker in (26) and (27) is a nutritionists. For all we know, nutritionists believe that eating raw vegetables is not only possible but a good thing. The fact (26) commits the speaker to raw vegetables being edible should not be seen as arising by way of projection from the complement, it is already a part of the background scenario. If the complement clashes with the built-in assumptions about nutritionists as in (27) there is no illusion of projection.

The idea that the only kind of content that projects is what is not at-issue with respect to the QUD seems not to work at all for cases like (28):

(28) Q: When did Finland become independent?
   A: It must have been after the Bolsheviks came to power in Russia but before Lenin died in 1924.

Here the content of the after and the before clauses jointly provide a partial answer to the question under discussion. They are both presupposed and projected. Temporal clauses are among the oldest examples of presupposition, first discussed by Frege (1892: 224).

At this stage the QUD approach to projection is a research program. There are many questions still under discussion.
Presupposition: What went wrong?

5 Presupposition or Conventional Implicature?

Going back to Voraussetzung and Andeutung, that distinction, presumably clear to Frege, has become difficult for us to make. There is an overlap between phenomena that have been called presuppositions by some semanticists and conventional implicatures by others. Some linguists have changed their minds even more than once on this issue (Karttunen 1971; Karttunen & Peters 1979; Karttunen 2012). Are these concepts distinct, or opposite ends of some continuum?

Ironically, one of the clearest definitions of conventional implicature comes from Bach (1999: 331), the lone prophet who preaches that they do not exist:

(29) CI: A proposition is a conventional implicature of an utterance just in case
a. the speaker (speaking seriously) is committed to the truth of the proposition,
  b. which proposition that is depends upon the (or a) conventional meaning of some particular linguistic device in the utterance,
   c. but the falsity of that proposition is compatible with the truth of the utterance.

Frege and Grice would have been agreement with Bach about the definition in (29) although they believed that conventional implicatures are real, as most of us do.

The conditions (a) and (b) also apply to presuppositions. But the tradition has it that (c) makes a fundamental difference. False implicatures can be ignored in ordinary conversation, false presuppositions have serious consequences, for a logician, a truthvalueless statement, for a pragmatist, a failure of communicational purpose of some sort. We will try to tease out that difference later on in one concrete case.

There are other differences. Presuppositions are expressible as propositions, typically about something existing or having happened. Conventional implicatures can be quite ineffable, hard to express in words. What do exactly does jerk mean as opposed to bloke? But some expressions classified as conventional implicatures by Potts (2005), such as appositives, have propositional content. Amaral, Roberts & Smith (2007) give many examples that show that conventional implicatures interact with the at-issue meaning of the text. In such cases, as Karttunen & Zaenen (2005) show, conventional implicatures behave like presuppositions. For example, what they implicate can be filtered away by an if-clause:

(30) a. The heiress married Bill and the lucky man became a millionaire.
  b. If the heiress married Bill, the lucky man became a millionaire.

In (30b) the author’s characterization of Bill as the lucky man may be understood as conditional on whether the heiress actually married him.
There are many other linguistic means besides the *if*-clause to introduce a hypothetical context. The author of (31) is not committed to the appositive *a third time offender* if the protagonist is found innocent but he is committed to Roberts having been convicted at least twice before.\(^{18}\)

(31) The jury is very likely to convict Roberts. In that case, as a third time offender, he will never get out of jail.

Under California’s harsh ‘three-strikes-and-you-are-out’ law, a person convicted of three felonies can be locked away for a life.

In trying to find cases to test (29c) with students, I have found that people tend to have much better intuitions about money than truth-value gaps. Here is a case to consider.

i. You and I think that Bill has a new girl friend named Sally. We both think that Bill intends or ought to send her a Valentine card.

ii. You think that Bill often forgets to do things he intends to do. I agree that Bill is forgetful in general but not about his romances.

iii. You say: *I bet you $10 that Bill will forget to send a Valentine card to Sally.* I accept the bet.

After February 14, we go to Sally to find out what happened. Let’s assume we hear either A1 or A2:

A1. No. I didn’t get a card from Bill but he sent me a really sweet tweet.

A2. No. Why would Bill send me anything? We are not romantically involved.

Who won the bet? In the case of A1, I have to agree that a tweet, sweet as it may be, doesn’t really count as a Valentine’s card. Reluctantly I pay you $10. What we presumed to be the relationship between Bill and Sally was correct but my bet on what would happen was wrong.

In the case of A2, we both agree that our bet was invalid. Why? (32a) implies (32b).

(32) a. Bill will forget to send a Valentine to Sally.

b. Bill intends or ought to send a Valentine to Sally.

By (29c) the falsity of the (32b) should not matter for the truth or falsity of (32a) itself. That seems correct if (32a) is taken as a simple statement. But in making

---

\(^{18}\) If the author believes that Roberts has exactly two previous convictions, in the second sentence of (31), Roberts could be referred to as a *two-time offender* (his present status) or as a *three-time offender* (his status, if convicted again).
the bet we both assumed that (32b) was true, elevating that aspect of the semantic content of (32a) to a presupposition that our bet was based on.

This is as close as I have been able to come to convince students that there is a difference between conventional implicature and presupposition. I am not aware of any experimental studies on this issue. It may well turn out that all triggers of conventional implicatures are not all alike in this respect. It seems unlikely that a false implicature triggered by the word bother could invalidate a bet like I bet that Bill won’t bother to send a Valentine to Sally.

Given our inability to find a clear-cut difference between conventional implicature and presupposition, the most we can conclude is that the two species of not-at-issue speaker commitment fall under the genus of PROJECTIVE MEANING (Potts 2005). In some cases the assignation to one or the other species is determined by a lexical item and a syntactic construction, in other cases pragmatic factors such as the purpose of a speech act, like making a bet, also play a role.

Even if presuppositions, conventional implicatures and entailments are of different species, their semantics have to interact in ways that we still have not been able to resolve, either formally or computationally. Assuming that again involves a presupposition and fail a conventional implicature, I suspect that in spite of the precise work of Kamp (2001) in the DRS framework on the meaning of again and all of Karttunen & Peters (1979), we don’t have an account of the entailments in (33):

\[
\begin{align*}
(33) & \quad \text{a. She has never failed to return my calls.} \\
& \quad \Downarrow \quad \text{She has always returned my calls.} \\
& \quad \text{b. She has never failed to return my calls again.} \\
& \quad \Downarrow \quad \text{She once didn’t return my call but never after that.}
\end{align*}
\]

I would love to be shown wrong on that conjecture.

In future years, the means of collecting large amounts of judgements and the advances in DEEP LEARNING technology (Goodfellow, Bengio & Courville 2016) will enable us to model human judgements with a high degree of accuracy but without delivering any understanding of the how a trained RNN, a recursive neural network, actually works. Modeling human performance by a black box that simulates what a human does is not the end goal of what we hope to achieve. Nevertheless, it would be progress.

6 Conclusion

The end of the 1960s and the early 1970s was an exciting time for people like myself who had a chance to stake a claim on syntactic and lexical semantic content that had been ignored by centuries of works by linguists and philosophers. Carol and Paul
Kiparsky discovered factives, a vast expansion beyond the class of definite descriptions and aspectual verbs discussed in the literature up to that point. I discovered implicatives with their associated presuppositions/conventional implicatures, never discussed anywhere before, and the problem with the Frege/Langendoen&Savin cumulative view of the presuppositions of compound sentences. I benefited from the work of Larry Horn, Jim McCawley, George Lakoff, Jerry Morgan, and many others in that period. Joyfully we created a mess that we left to the current generation to clean up.

Bart Geurts sums up the early developments in presupposition studies as follows:

> An especially stark illustration of the disparity of the field, at least in its early days, is the work of a Karttunen, who within the span of six years published three theories that were mutually inconsistent, technically as well as conceptually. (Geurts 1999: 5)

I don’t disagree with his assessment. But the ‘disparity of the field’ still persists. We will never have a unified account of the phenomena on Levinson’s list of presupposition triggers because they are not all of the same species.

References


Presupposition: What went wrong?


Presupposition: What went wrong?

Reprint of Roberts 1966.

Lauri Karttunen
CSLI, 210 Panama St.
Stanford, CA 94305
laurik@stanford.edu