Non-Catastrophic Presupposition Failure[1]

S. Yablo

1. Background

I will be talking in this paper about the problem of presupposition failure. The claim will be (exaggerating some for effect) that there is no such problem -- more like an opportunity of which natural language takes extensive advantage.

---

[1] Papers sort of like this one were presented at Indiana, UC Davis, UC Berkeley, UC San Diego, Yale, Brown, Penn, Kentucky, Oxford (as the 2005 Gareth Evans lecture), ANU, Monash, the Chapel Hill Colloquium, an APA session on Metaontology, and graduate student conferences at Pittsburgh and Boulder. I am grateful to Richard Holton, Sally Haslanger, Agustin Rayo, Caspar Hare, Kai von Fintel, Danny Fox, Irene Heim, Karen Bennett (who commented at the APA), Anne Bezuidenhout (who commented at Chapel Hill), Larry Horn, Sarah Moss, John Hawthorne, Lloyd Humberstone, and especially Bob Stalnaker for questions and advice. I learned too late of the literature on "logical subtraction" (see Humberstone 2000 and references there); it holds out hope of a different and perhaps more straightforward route to incremental content.

The last two sentences are a case in point. The first was, "I am going to talk about the F"; the second was, "there is no F." If the second sentence is true -- there is no F -- then the first sentence, which presupposes that there is an F, suffers from presupposition failure. In theory, then, it should strike us as somehow compromised or undermined. Yet it doesn't. So here is one case at least where presupposition failure is not a problem.

The title is meant to be understood compositionally. Presuppositions are propositions assumed to be true when a sentence is uttered, against the background of which the sentence is to be understood. Presupposition failure occurs when the proposition assumed to be true is in fact false.[2] Failure is catastrophic if it prevents a thing from performing its primary task – in this case making an (evaluable) claim. Non-catastrophic presupposition failure then

[2] Really I should say "untrue" rather than "false," to allow for presuppositions that lack truth-value because they themselves suffer from presupposition failure. Looking ahead to the Donnellan examples, "The man drinking a martini is that guy" is (so it seems) not false but undefined if no one is drinking a martini.

becomes the phenomenon of a sentence still making an evaluable claim despite presupposing a falsehood.

I said that presuppositions were propositions taken for granted when a sentence is uttered, against the background of which the sentence is to be understood. [3] It would be good to have some tests for this. Here are three, loosely adapted from

---

[3] Are we to think of presupposition as a relation that sentences bear to propositions (Strawson), or a relation that speakers bear to propositions (Stalnaker)? There may be less of a difference here than meets the eye. The first relatum for Strawson is utterances or tokens of S, from which it is a short step to speakers presupposing this or that in uttering S. Stalnaker for his part appreciates that certain sentences S should not be uttered unless this or that is (or will be as a result of the utterance) pragmatically presupposed. It does little violence to either's position to treat "S presupposes π" as short for "All (or most, or contextually salient) utterances of S presuppose π," and that in turn as short for "Speakers in making those utterances always (often, etc.) presuppose that π." (Von Fintel ms and Simons 2003 are illuminating discussions.) Semantic presupposition would be the special case of this where S presupposes π as a matter of meaning, that is, S-users presuppose π not for conversational reasons but because semantic rules require it.

the paper that got me thinking about these issues (von Fintel 2004 -- don't miss it!).[4]

One is the "hey, wait a minute" test.[5] If π is presupposed by S, then it makes sense for an audience previously unaware of π to respond to an utterance of S by saying "hey, wait a minute, I didn't know that π." If π is asserted, that response would be silly; of course you didn't know, the point of uttering S was to tell you. Suppose you say, "I'm picking my guru up at the airport." I can reply, "Hey, I

---

[4] Strawson noticed that while some King-of-France sentences strike us as unevaluable ("The KoF is bald"), others seem false ("The KoF visited the Exhibition yesterday"). Von Fintel criticizes earlier accounts of this contrast (by Strawson and Peter Lasersohn) and proposes an interesting new account. He does not address himself to a third possibility noted by Strawson, that a sentence with false presuppositions should strike us as true. This paper agrees with von Fintel's basic idea: some KoF-sentences "are rejected as false…because they misdescribe the world in two ways: their presupposition is false, but in addition there is another untruth, which is independent of the failure of their presupposition" (2004, 325). But it implements the idea differently.

[5] Taken apparently from Shanon 1976.

didn't know you had a guru," but not, "hey, I didn't know you were going to the airport." This suggests that your having a guru was presupposed while your going to the airport was asserted. A likelier response to what is asserted is, "is that so, thanks for telling me."[6] [7]

---

[6] This test seems to work best for semantic presuppositions (see note 3). Looking ahead a bit, "The man drinking a martini is a philosopher" does not invite the reply, "Hey, I didn't know that guy was the one drinking a martini." One can, however, say, "Hey, I didn't know that guy was drinking a martini." So perhaps a version or variant of the test applies to (some) non-semantic presuppositions as well.

[7] Von Fintel attributes to Percus a test that is in some ways similar. "R, and what's more, S" sounds fine if S asserts more than R, but wrong if S only presupposes more. So, "John thinks Judy is a chiropractor" can be followed by "And what's more, he is right to think Judy is a chiropractor," but not "And what's more, he realizes Judy is a chiropractor." This seems to indicate that "He realizes that BLAH" presupposes what "He is right to think that BLAH" asserts, viz. that BLAH, and asserts what it presupposes, viz. that he believes that BLAH.

Second is the attitude attributed when we say that someone denies that S, or hopes or regrets that S; the presupposition π is exempted from the content of that attitude. Hoping you will pick up your guru at the airport may be in part hoping your guru will be picked up, but it is not hoping that you have a guru in the first place. Denying that you are going to pick up your guru at the airport is not denying the conjunction of you have a guru with you are going to pick your guru up at the airport.[8] So a second mark of presuppositions is that π does not figure in what you hope or deny or regret in hoping or denying or regretting that S (Stalnaker 1999, 39).

[8] This observation goes essentially back to Frege. Frege considers the sentence "whoever discovered the elliptic form of the planetary orbits died in misery." He notes that its negation is "whoever etc. did not die in misery" rather than "Either whoever discovered the elliptic form of the planetary orbits did not die in misery or there was nobody who discovered the elliptic form of the planetary orbits" (1872, 162-3). If we assume (as he did) that denial is assertion of the negation, this amounts to the claim that "somebody discovered the elliptic form of the planetary orbits" is no part of what is denied when we deny that "whoever etc. died in misery."

A third test is that presuppositions within limits project, that is, π continues to be presupposed by more complex sentences with S as a part. If you say, "I don't have to pick up my guru after all," or, "it could be I will have to pick my guru up," these statements still intuitively take it for granted that you have a guru. Our earlier tests confirm this intuition. One can still reply, "hold on a minute, you have a guru?" And to hope that you don't have to pick your guru up is not to hope that you have a guru.[9]

Note that one test sometimes used to identify presuppositions is missing from this list: π is presupposed iff unless π holds, S says nothing true or false. That test is useless in the present context because it makes NCPF impossible; π is not classified as a presupposition unless its failure would be catastrophic.

A sentence suffers from catastrophic presupposition failure only if, as Strawson puts it, "the whole assertive enterprise is wrecked by the failure of [S's] presupposition" (1964, 84). There is also the phenomenon of what might be

[9] Related to this, presuppositions fail to project in certain contexts, such as conditionals with π as antecedent. "I don't remember if I have a guru, but if I do, it could be I am supposed to pick my guru up at the airport" does not presuppose that I have a guru.

called disruptive presupposition failure. This occurs when π's failure does not wreck the assertive enterprise so much as reveal it to have been ill advised. It could be, for instance, that π was an important part of the speaker's evidence for S. It could be that π was part of what made S relevant to the rest of the conversation. It could even be that S entails π so that π's falsity guarantees that S is false too.[10]

Disruption is bad, but it is not (in our sense) a catastrophe. On the contrary, a remark is implausible or irrelevant or false because of what it says, and that something was said suggests that the assertive enterprise has not been wrecked after all. I mention this because Stalnaker, who has written the most about these

[10] This relates to a passage in "Pragmatic Presuppositions": "using the pragmatic account [of presupposition], one may say that sometimes when a presupposition is required by the making of a statement, what is presupposed is also entailed, and sometimes it is not. One can say that "Sam realizes that P" entails that P – the claim is false unless P is true. "Sam does not realize that P," however, does not entail that P. That proposition may be true even when P is false. All this is compatible with the claim that one is required to presuppose that P whenever one assers or denies that Sam realizes it" (1999, 54).

topics, is addressing himself more often to the disruptive/non-disruptive distinction than the catastrophic/non-catastrophic distinction.[11] This paper is meant to be entirely about the latter.

---

[11] For instance here:

> Where [presuppositions] turn out to be false, sometimes the whole point of the inquiry, deliberation, lecture, debate, command, or promise is destroyed, but at other times it does not matter much at all…Suppose…we are discussing whether we ought to vote for Daniels or O'Leary for President, presupposing that they are the Democratic and Republican candidates respectively. If our real interest is in coming to a decision about who to vote for…, then the debate will seem a waste of time when we discover that in reality, the candidates are Nixon and Muskie. However if our real concern is with the relative merits of the character and executive ability of Daniels and O'Leary, then our false presupposition makes little difference (1999, 39).

2. Relevance to philosophy

Why should we care about non-catastrophic presupposition failure? There are reasons from the philosophy of language, from epistemology, and from metaphysics.

The philosophy of language reason is simple. All of the best-known theories of presupposition (among philosophers, anyway) suggest that failures are or ought to be catastrophic. This is clearest for Frege's and Strawson's theories -- for those theories more or less define a sentence's presuppositions as preconditions of its making an evaluable claim. Assuming as before that a sentence's primary task is to offer a true or false account of how things are, presuppositions on Frege's and Strawson's theories are automatically propositions whose failure has catastrophic results.

Next consider Stalnaker's theory of presupposition. Stalnaker-presupposition is in the first instance a relation between speakers and propositions; one presupposes π in uttering S if one thinks that π is (or will be, as a result of the utterance) common ground between relevant parties. A sentence presupposes π

only to the extent that S is not appropriately uttered unless the speaker presupposes that π.

Why on this account should presupposition failure be problematic? Well, the point of uttering S is to draw a line through the set of worlds still in play at a particular point in the conversation -- one is saying that our world is on the S-true side of the line rather than the side where S is false. Since the worlds still in play are the ones satisfying all operative presuppositions, the speaker by presupposing π is arranging things so that her remark draws a line through the π-worlds only.

But then what happens when π is false? Because the actual world is outside the region through which the line is drawn, it is hard to see how in drawing this line the speaker is saying anything about actuality. It's as though I tried to locate Sicily for you by saying that as between North and South Dakota, it's in the North, although truth be told it's not in either Dakota. Similarly it is not clear how I can locate actuality for you by saying that as between the π-worlds where S is true and the ones where it is false, it's in the first group, although truth be told it's not a π-world at all.[12]

[12] See Beaver 2001 for theories of the kind favored by many linguists. These seem at least as unaccommodating of NCPF as the ones philosophers like,

That was the philosophy of language reason for caring about NCPF; the standard theories seem to rule it out. A much briefer word now on the epistemological and metaphysical reasons.

The epistemological reason has to do with testimony, or learning from others. Someone who utters a sentence S with truth-conditions C (S is true if and only if C obtains) might seem to be telling us that C does obtain. But if we bear in mind that π is one of the conditions of S's truth, we see that that cannot be right. For it makes two false predictions about the phenomenon of NCPF. The first is that all presupposition failure is non-catastrophic; if π is false, then the speaker is telling us something false, hence the assertive enterprise has not been wrecked. The second is that what the speaker is telling us can never be true. The fact is that some presupposition failure is catastrophic and some isn't; and the claim made can be either true or false. To suppose that speakers are saying inter alia that π in uttering S collapses the first two categories – catastrophic, non-catastrophically true -- into the last – non-catastrophically false.

---

for a reason noted by Simons: "Dynamic theories of presupposition claim that presupposition failure results in undefinedness of the context update function – the dynamic correlate of truth valuelessness" (2003, 273).

So here is the epistemological relevance of NCPF. It reminds us that speakers are not in general vouching for everything the truth of their sentence requires; they vouch for the asserted part but not (in general) for the presupposed part.

This leads to the metaphysical reason for caring about NCPF. Quine famously argues like so: "scientists tell us that the number of planets is 9; that can't be true unless there are numbers; so scientists tell us inter alia that there are numbers; so unless we consider ourselves smarter than scientists, we should believe in numbers." This assumes that speakers are vouching for all the truth-conditions of the sentences coming out of their mouths. But there being a thing that numbers the planets is no part of what Clyde Tombaugh (the discoverer of Pluto) was telling us – no part of what he was giving his professional opinion about -- when he spoke the words, "the number of planets is 9." A different metaphysical upshot will be mentioned briefly at the end.

3. Frege and Strawson

I said that the best-known theories suggest that all presupposition failure ought to be catastrophic, and that the suggestion is implausible. I did not say that the best-known theorists are unaware of this problem. Well, Frege might have been unaware of it. Even he, though, gives an example that might be taken as a case in point:

> Somebody using the sentence "Alfred has still not come" actually says "Alfred has not come," and at the same time hints - but only hints - that Alfred's arrival is expected. Nobody can say: "since Alfred's arrival is not expected, the sense of the sentence is false" (1918, 331)

Frege's use of hint makes it sound as though we are dealing with an implicature. But "still" is by the usual tests a presupposition trigger. ("Hang on, I didn't know Alfred was supposed to be here!") Suppose for argument's sake that the tests are right.

Frege says that the thought is not automatically false if Alfred was unexpected. By this he presumably means that the thought's truth-value depends not on how expected Alfred was but on whether he has indeed come. Even if the

presupposition fails -- he was not expected -- a claim is still made that can be evaluated as true or false.

So the Alfred example looks like a case of non-catastrophic presupposition failure. Of course, Frege would not see it that way, because the presuppositions that he (and later Strawson) has mainly in mind are existential presuppositions:

> If anything is asserted there is always an obvious presupposition that the simple or compound proper names used have reference (1872, 162)

The sentence

> Whoever discovered the elliptic from of the planetary orbits died in misery

is said to lack truth-value unless someone did indeed make the indicated discovery (1872, 162). Strawson in similar fashion says that if someone produced the words

> The King of France is bald,

we would be apt to say that "the question of whether his statement was true or false simply did not arise, because there was no such person as the King of France" (1950, 12).

But, and here he goes beyond Frege, Strawson notices that failure even of a sentence's existential presuppositions does not prevent it from making an evaluable claim:

> Suppose, for example, that I am trying to sell something and say to a prospective purchaser The lodger next door has offered me twice that sum, when there is no lodger next door and I know this. It would seem perfectly correct for the prospective purchaser to reply That's false, and to give as his reason that there was no lodger next door. And it would indeed be a lame defense for me to say, Well, it's not actually false, because, you see, since there's no such person, he question of truth and falsity doesn't arise (1954, 225).

This is an example of what Strawson calls "radical failure of the existence presupposition" (1964, 81) -- radical in that "there just is no such particular item at

all" as the speaker purports to be talking about. It shows that for the existence presupposition to fail radically is not necessarily for it to fail catastrophically.

Now, if the existence presupposition can fail radically -- there is no such item as the speaker purports to be talking about -- one expects that the uniqueness presupposition could fail radically too -- there are several items of the type the speaker purports to be talking about. Consider another example of Strawson's:

> if, in Oxford, I declared, "The Waynflete Professor of Logic is older than I am" it would be natural to describe the situation by saying that I had confused the titles of two Oxford professors [Waynflete Professor of Metaphysics and Wykeham Professor of Logic], but whichever one I meant, what I said about him was true (1954, 227)

This becomes radical failure of the uniqueness presupposition if we suppose that in confusing the titles had confused the individuals too, so that his remark was no more directed at the one than the other. Does the failure thus reconstrued remain non-catastrophic? I think it does. The remark strikes us as false if the Waynflete

and Wykeham Professors are both younger than Strawson, and true (or anyway truer) if he is younger than them.[13]

What about non-radical failure of the existential and uniqueness presuppositions? By a non-radical failure I mean that although the description used is not uniquely satisfied, the subject does have a particular item in mind as the intended referent. The uniqueness presupposition fails non-radically when one says, "The square root of N is irrational," meaning to refer to the positive square root, forgetting or ignoring that N has a negative root too. This kind of remark does not court

[13] Suppose Strawson had said, "The Philosophy Professor at St Andrews is older than me," not realizing that St Andrews had two Professors. Such a statement again seems correct if both are older and incorrect if both are younger -- indeed (arguably) if either is younger. Stalnaker in conversation suggests treating this as a case of pragmatic ambiguity; the utterance seems true when it is true on both disambiguations, false when it is false on both (or perhaps false on either). I do not see how to extend this treatment to superficially similar cases. "All eight solar planets are inhabited" seems false, but it is presumably not ambiguous between nine attributions of inhabitedness, each to all solar planets but one.

catastrophe since it strikes us as correct if both roots are irrational, and incorrect if both are rational, and no other outcome is possible.

That was my example of non-radical failure of the uniqueness presupposition, not Strawson's; his would be the Oxford mix-up, assuming that the intended referent was, say, Gilbert Ryle, then Waynflete Professor of Metaphysics. Strawson also gives an example where it is the existential presupposition that non-radically fails:

> perhaps, if I say, "The United States Chamber of Deputies contains representatives of two major parties," I shall be allowed to have said something true even if I have used the wrong title, a title, in fact, which applies to nothing (1954, 227)[14]

[14] This example is important in Strawson's debate with Russell. Some empty-description sentences strike us as false, as Russell's semantics predicts. But others are such that "if forced to choose between calling what was said true or false, we shall be more inclined to call it true" (Strawson 1954, 227). Russell cannot claim too much credit for plugging truth-value gaps, if he sometimes plugs in the wrong value. (I ignore the wide-scope

So although Strawson doesn't put it this way, his discussion suggests a four-fold classification along the following lines:[15]

	uniqueness presupposition	existential presupposition
radical failure of the	Waynflete Prof of Logic[16]	lodger next door
non-radical failure of the	square root of N	Chamber of Deputies

The fourth of Strawson's categories – non-radical failure of the existential presupposition -- proved the most influential, as we shall see.

4. Donnellan and Stalnaker

---

negation strategy as irrelevant to the examples Strawson is concerned with here.)

[15] This classification is not meant to be exhaustive; perhaps, e.g., the description applies to exactly one thing, but that thing is not the intended referent.

[16] Understood so that the speaker is thinking confusedly of both professors at once.

Strawson appreciates, of course, that the judgments just noted seem at odds with his official theory, particularly with the principle that "If someone asserts that the ϕ is ψ he has not made a true or false statement if there is no ϕ" (Donnellan 1966, 294). Donnellan's famous counterexample to that principle would thus not have come as a surprise to him:

> Suppose one is at a cocktail party and, seeing an interesting-looking person holding a martini glass, one asks, "Who is the man drinking a martini?" If it should turn out that there is only water in the glass, one has nevertheless asked a question about a particular person, a question it is possible for someone to answer (1966, 287).

Given that

> Strawson admits that we do not always refuse to ascribe truth to what a person says when the definite description he uses fails to fit anything (or fits more than one thing) (1966, 294),

what does Donnellan think he is adding to Strawson's own self-criticism? Donnellan is not very explicit about this but here is my best guess as to his reply.

What Strawson admits is that the person has said something true. He does not (according to Donnellan) admit that the statement originally at issue, viz. "the man drinking a martini is a famous philosopher," is true. One might wonder, of course, what we are doing if not "awarding a truth value ….to the original statement." The answer is that we "amend the statement in accordance with [the speaker's] guessed intentions and assess the amended statement for truth or falsity" (Strawson 1954, 230). The statement Strawson is willing to call true, then, is not the one suffering from presupposition failure, and the one suffering from presupposition failure he is not willing to call true. (Elsewhere Strawson says the original statement is true only in a secondary sense.) Donnellan is bolder: he thinks that the unamended, original statement "The ϕ is ψ" can be true in the absence of ϕs, if the description is used referentially.

A second difference between Donnellan and Strawson is this. Strawson paints a mixed picture featuring on the one hand a presupposition that the description is uniquely satisfied, and on the other hand an intention to refer with that

description to a certain object. Donnellan simplifies matters by turning the referential intention into an additional presupposition:

> when a definite description is used referentially, not only is there in some sense a presupposition…that someone or something fits the description,…but there is also a quite different presupposition; the speaker presupposes of some <u>particular</u> someone or something that he or it fits the description. In asking, for example, "Who is the man drinking a martini?" where we mean to ask a question about that man over there, we are presupposing that that man over there is drinking a martini – not just that <u>someone</u> is a man drinking a martini (1966, 289).

This may not seem like progress; before we had one failed presupposition to deal with, now we have two. But, and this is the third difference between Donnellan and Strawson, the "new" failed presupposition, rather than being an obstacle to evaluation, is what <u>enables</u> evaluation, by pointing the way to an evaluable hypothesis: that man is a famous philosopher.

Stalnaker attempts to put all this on a firmer theoretical foundation. Imagine O'Leary saying, "The man in the purple turtleneck is bald," where it is understood that the man in question is <u>that</u> man (Daniels). The propositional content of

O'Leary's statement is that Daniels is bald. The fixation of content here is along lines more or less familiar from Kaplan. Just as the character of an expression like "you" determines its denotation as a function of context,

> there are relatively systematic rules for matching up [referential] definite descriptions with their denotations in a context (1999, 41).

The rule for "you" is that it contributes the addressee; the rule for a referential description is that it contributes

> the one and only one member of the appropriate domain who is presupposed to have the property expressed in the description (1999, 41).

Crucially from our perspective,

> it makes no difference whether that presupposition is true or false. The presupposition helps to determine the proposition expressed, but once that proposition is determined, it can stand alone. The fact that Daniels is bald in no way depends on the color of his shirt (1999, 43).

So we see that Stalnaker does have an account to offer of <u>some</u> cases of NCPF. NCPF occurs (in these cases) for basically Kaplanian reasons. A conventional meaning is given by a systematic character function mapping contexts (= sets of

worlds) to propositions. And there is nothing to stop a set of worlds from being mapped to a proposition defined on worlds outside of the set.

This is fine as far as it goes. But NCPF is ubiquitous, and character as Kaplan understands it is reserved to a few special terms. Stalnaker knows this better than anyone, of course; he was one of the first to charge two-dimensionalists with an undue optimism about the project of extending Kaplan-style semantics from demonstratives to the larger language. Some NCPF may be a matter of characters mapping contexts to propositions defined outside those contexts, but not much. An example of Kripke's brings out the extent of the difficulty:

> Two people see Smith in the distance and mistake him for Jones. They have a brief colloquy: "What is Jones doing?" "Raking the leaves." "Jones," in the common language of both, is a name of Jones; it <u>never</u> names Smith. Yet, in some sense, on this occasion, clearly both participants in this dialogue have referred to Smith, and the second participant has said something true about the man he referred to if and only if Smith was raking the leaves (Kripke 1977, 14)

Assuming Smith was raking the leaves, the second participant says something true with the words, "Jones is raking the leaves," despite (or because of) the false presupposition that it is Jones they see off in the distance. The example has a Donnellan-like flavor, but the explanation will have to be different; a proper name like "Jones" does not have a reading on which it denotes whoever is presupposed to be Jones in the relevant context. This is why I say there is no general account of NCPF in Stalnaker.[17] I will be suggesting, however, that he does provide the materials for such an account.

So, to review. A sentence's presuppositions are (generally) no part of what it says. Presuppositions can however function as <u>determinants</u> of what is said. The suggestion is that they can influence what is said equally well even if false.[18] It

---

[17] This is not to say he doesn't have particular explanations to offer in particular cases. Often he appeals to a device like Strawson's (see above). The original statement – "Jones is raking the leaves" – suffers from presupposition failure, so is not evaluable. Had the speaker been better informed, she would have made a statement – "Smith is raking the leaves" -- whose presuppositions are true. Our evaluation of the second statement is then projected back onto the first.

remains to explain how exactly the trick is pulled off. Explaining this will be difficult without an account of the mechanism by which presuppositions exert their influence. Because we are really asking about that mechanism. Does it ever in the course of its π-induced operations find itself wondering whether π is true?

There are hints in the literature of three strategies for making π (not a part of but) a guide to asserted content. The first tries to get at what S says by ignoring the possibility that π fails. The second tries to get at what S says by restoring π when it does fail. The third tries to get at what S says by asking what more than π needs to be true for S to be true. I will be arguing against IGNORE and RESTORE and defending SAY-MORE.

5. IGNORE

Asserted content as conceived by the first strategy addresses itself only to π-worlds. It just ignores worlds where π fails. Thinking of contents as functions from worlds to truth-values, ignoring a world is being undefined on that world. S's asserted content is thus a partial function mapping π&S-worlds to truth, π&~S-worlds to falsity, and worlds where π fails to nothing at all.

> [1] S's asserted content **S** is the proposition that is true (false) in a π-world w iff S is true (false) in w, and is otherwise undefined.[19]

There might seem to be support for this in a passage from Stalnaker:

> in a context where we both know that my neighbor is an adult male, I say, "My neighbor is a bachelor," which, let us suppose, entails he is adult and a male. I might just as well have said "my neighbor is unmarried." The same information would have been conveyed (1999, 49).

The same information would indeed have been conveyed if by "information conveyed" we have in mind assertive content in the sense of [1], for (ignoring

---

[19] So far this says nothing about S's truth-value in worlds where π fails. Let S be "The KoF is so and so." Russellians will call S false in worlds where France lacks a king. Strawsonians will say it is undefined. They agree, however, on S's truth-value in worlds where France has a unique King, and those are the only worlds that [1] cares about. Later I will be stipulating that S's truth-value in a world goes with the truth-value of the IGNORE proposition, the one defined by [1].

worlds where my neighbor fails to be an adult male), my neighbor is a bachelor if and only if he is unmarried.[20]

Never mind whether the IGNORE strategy can be attributed to Stalnaker; does it succeed in making π not a part of S's asserted content but a determinant of that content? It does. π influences what S says by marking out the set of worlds on which **S** is defined. But π is not a part of what S says, for [1] makes **S** undefined in worlds where π is false, and it would be false in those worlds if S said in part that π.

The IGNORE proposition has some of the features we wanted. But what we mainly wanted was an **S** that could still be evaluated in worlds where π failed. And here [1] does not deliver at all. "The King of France is sitting in this chair"

[20] Stalnaker would not identify what is said with a proposition defined only on π-worlds. Such an identification would make nonsense of passages like the following: "To make an assertion is to reduce the context set in a particular way…all of the possible situations incompatible with what is said are eliminated" (1999, 86). It is not clear to me how closely his notion of what is said -- he sometimes calls it "the proposition expressed" -- lines up with my assertive content, but certainly the correspondence is not exact.

sounds to most people just false. But there is nothing in the IGNORE proposition to support this judgment, for the IGNORE proposition is undefined on worlds where France lacks a King.

Methodological digression: I said that "The KoF is sitting in this chair" sounds to most people just false. Why not go further and declare that it really <u>is</u> false? Strawson for his part is reluctant to take this further step. "The KoF is sitting in this chair" is <u>not</u> false in what he considers the term's primary sense:

> <u>sometimes</u> [however] the word "false" may acquire a <u>secondary</u> use, which collides with the primary one (1954, 230)

One option is to follow Strawson in calling sentences like "The KoF is behind that door" false only on a secondary use of that term, and sentences like "The US Chamber of Deputies has representatives from two major parties" true only on a secondary use of "true." The task is then to explain why some gappy sentences <u>count</u> as false, while others count as true. Another option would be to follow Russell and call both of the above sentences false in the <u>primary</u> sense. The task would then be to explain why some primarily false sentences ("The man with the

martini is a philosopher") count as true, while others ("The King of France is bald") count as neither true nor false.

Given that both theories (Russell's and Strawson's) need an analogous sort of supplement to deal with intuitive appearances of truth and falsity, either could serve as our jumping-off point; the choice is really between two styles of theoretical bookkeeping. That having been said, let's consider ourselves Strawsonians for purposes of this paper. S's semantic content -- what in context it means -- will be a proposition defined only on π-worlds; it is semantic content that determines S's truth-value. [21] Truth-value intuitions are driven not by what a sentence means, however, but by what it says: its asserted content.

---

[21] Von Fintel 2004 and Beaver and Krahmer (ms) also take this option. Because sentences and their semantic contents have the same truth-value (if any) in all worlds, we can be casual (sloppy) about the distinction between them. So, for instance, it makes no difference to an argument's validity whether we think of it as made up of (i) sentences, (ii) the propositions that are those sentences' semantic contents, or (iii) sentences and propositions combined.

So, “The KoF sits in this chair” strikes us as false because it says in part that that someone sits in this chair. “The US Chamber of Deputies has representatives from two major parties” strikes us as true because it says that the House of Representatives has representatives from two major parties. Both of our remaining strategies are aimed at carving out a notion of asserted content that predicts truth-value intuitions in a way that semantic content is prevented from doing by the fact that it is undefined on worlds where π fails.

6. RESTORE

Let S be "The KoF is sitting in this chair.” Even if we agree with Strawson that S is lacking in truth-value, there is still the feeling that it escapes on a technicality. The chair’s emptiness is all set to falsify it, if France’s lack of a king would just get out of the way. One response to this obstructionism is to say, fine, let's give France a king; then S’s deserved truth-value will shine through. This is the idea behind RESTORE. Instead of ignoring worlds where π fails, we attempt to rehabilitate them, in the sense of bringing them back into line with π. Of course one can’t literally turn a non-π world into a π-world, so in practice this means looking at S’s truth-value in the closest π-worlds to w.

Now, for S to be true (false) in the π-worlds closest to w is, on standard theories of conditionals, precisely what it takes for a conditional π → S to be true (false) in w. So we can let the idea be this:

[2] **S** is true (false) in w iff π → S is true (false) in w.[22]

Why does "The KoF is sitting in this chair" strike us as false? Even if France is supplied with a king, still he is not to be found in this chair. Why does "The KoF is bald" strike us as lacking in truth-value? Supplying France with a king leaves the issue still unresolved; in some closest worlds the added King is bald, in others not.[23] So the RESTORE strategy has prima facie a lot going for it.

I don't doubt that for some similarity relation and some associated similarity-based conditional, [2] gives the right results. But if we confine ourselves to the conditionals we know of and have intuitions about -- the indicative and the subjunctive -- the strategy fails. Let me give some examples before attempting a diagnosis.

---

[22] I assume that π→S is false iff π→~S is true.

[23] See Lasersohn 1993 and von Fintel 2004.

Bertrand Russell, invited to imagine what he could possibly say to God if his atheism proved incorrect, replied (not an exact quote), "I would ask him why he did not provide more evidence of his existence." I infer from this that Russell accepted a certain indicative conditional

G. If God exists, he is doing a good job of hiding it.

Now the consequent of this conditional presupposes what its antecedent affirms; so G is of the form $\pi \rightarrow S$, read as if it is the case that π, then it is the case that S. This according to [2-ind] is the condition under which what S says is true. But then it would seem that S ought to count for Russell as true, given that he accepts G. And something tells me that it does not strike Russell as true that God is doing a good job of hiding his existence.

So this remark of Russell's shows that [2] in its indicative version does not give a correct account of asserted content. Now consider a different Russell remark: "If there were a God, I think it very unlikely that he would have such an uneasy vanity as to be offended by those who doubt his existence." From this it seems that Russell would have accepted

H. If there were a God, he would be generous to doubters.

H is of the form $\pi \rightarrow S$, read as if it were the case that π, it would be the case that S,. This according to [2-sub] is the condition under which what S says is true. So it would seem that S ought to count for Russell as true, given that he accepts H. But Russell is not at all inclined to think that God is generous to doubters.

Non-theological example: Is the King of France at this moment somewhere in Europe, Africa, Australasia, or the Americas? Not a chance. But the corresponding conditionals are plausibly correct; those are the places he would be, if he existed, and the places he is, if he exists. [2-sub] does get "The King of France sitting in this chair" right, for the King if he existed would not be in this chair. But imagine for a moment that this chair is the long lost French throne; the King of France would (let's say) be sitting in this chair if France had a king. [2-sub] predicts that our intuitions should shift. But it does not make it any more plausible to suppose that the King of France is sitting in this chair to be told that he would be sitting in it if France had a king. Imagine now that this chair is the long lost French throne and French kings if any are master illusionists; if France

has a king, he is sitting in this chair. This does not affect our truth-value intuitions at all. It is enough for them that the chair is empty.

The problem we are finding with [2] (I will focus for simplicity on [2-sub]) is an instance of what used to be called the "conditional fallacy." According to Shope (1978, 402 -- I have taken some liberties), the conditional fallacy is

> A mistake one makes in analyzing a statement p by presenting its truth as dependent upon the truth of a conditional of the form: 'If a were to occur, then b would occur', when one has overlooked the fact that although statement p is actually true, if a were to occur, it would undermine p and so make b fail to occur.

Philosophers have tried, for instance, to analyze dispositions in counterfactual terms:

> x is fragile = if x were to be struck, it would shatter

But x would not shatter if the molecular properties M making it fragile go away the moment that x is struck. What we meant to say, it seems, is that

x is fragile = if x were struck and retained M, it would shatter.

[2-sub] tries to analyze false-seemingness in counterfactual terms:

S counts as false = if π, S would be false.[24]

But suppose S counts as false in virtue of certain facts F, and restoring S's presuppositions chases those facts away. (Europe would not have been King-of-France-free if France had had a king.) What we should have said, it seems, is that

S counts as false = if π&F, S would be false

This is essentially what we do say in the next few sections. I mention this now because the motivation to be offered below is different, and we won't be stopping to connect the dots.

---

[24] Perhaps the fallacy comes in an indicative version too. One is tempted to analyze "Jones is totally reliable" as: if Jones says X, X is true. But if Jones says 0=1, that means not that 0=1 but that Jones is unreliable.

7. SAY-MORE

A passage discussed earlier deserves a second look. Stalnaker had us choosing between "my neighbor is a bachelor" and "my neighbor is unmarried," it being understood that my neighbor is an adult male. He says that the same information would be conveyed whichever sentence we chose. But in a part of the passage we didn't get to, he puts the word "increment" before "information":

> the increment of information, or of content, conveyed by the first statement is the same as that conveyed by the second (1999, 49).

The word increment suggests that we are to ask what more it takes for S to be true, supposing the requirement of π's truth is waived or assumed to be already met. This is the idea behind SAY-MORE. What S says, its assertive content, is identified with what more S asks of a world than that it should verify π.[25]

[25] Suppose we use prop(π) for the properties a world needs to verify π, prop(S) for the properties a world needs to verify S, and prop(S\π) for the additional properties π-worlds must have to be worlds where S is true. It is

Determining these additional requirements may sound like a tricky business; but it is not so different from something we do every day, when we look for the missing premises in an enthymematic argument. To ask what further conditions (beyond π) a world has to meet to be S is essentially to ask what premises should be added to π to obtain a valid argument for S:

π

???

S

So we can put the SAY-MORE strategy like this:

[3] **S** is whatever bridges the logical gap between π and S.

---

not in general the case that prop(S\π) = prop(S)-prop(π). An analogy might be this. A rich man can get into heaven only by giving millions to charity; so prop(heaven-goers\rich) includes giving millions to charity. But giving millions away is not in prop(heaven-goers)\prop(rich), because lots of people who don't give millions away still get into heaven, e.g., the deserving poor.

Of course, the gap might be bridgeable in more than one way. I propose to postpone to finesse this issue for now by letting **S** be the result of lumping all otherwise qualified gap-bridgers together. So, for instance,

France has exactly one king.

???

The King of France is sitting in this chair.

becomes valid if for ??? we put either "Some French king is sitting in this chair" or "All French kings are sitting in this chair." Assuming both statements bridge the gap equally well (see below), the assertive content is "Some and all French kings are sitting in this chair."

A lot more needs to be said, obviously, and some of it will be said in the next section. Right now though I want to try [3] out on a series of examples, one from Strawson, two adapted from Strawson, one from Donnellan, one from Kripke, and one from Langendoen.

A The lodger next door offered me twice that sum

B The author of Principia Mathematica also wrote Principia Ethica.

C All ten solar planets are inhabited.

D The man drinking a martini is a philosopher.

E Jones is burning the leaves.

F My cousin is not a boy anymore.

All six sentences are meant to strike us as false -- the first because there is no lodger next door; the second because neither PM author wrote PE; the third because most solar planets (they number nine, not ten) are uninhabited; the fourth because Daniels (who is in fact drinking water) is not a philosopher but an engineer; the fifth because that man (it's really Smith) is not burning but raking the leaves; and the sixth because my cousin (whether a boy or not) is only eight years old.

How would Stalnaker explain the appearance of falsity in these cases? This is to some extent speculative, but here is what I suspect he would say. A seems false for Russellian reasons: it is equivalent to a conjunction one of whose conjuncts is "There is a lodger next door." B seems false for supervaluational reasons: it is false on all admissible disambiguations. C and E seem false for the sort of reason Strawson offered (section 4): we amend them to "All nine solar planets are

inhabited" and "Smith is burning the leaves" before assigning a truth-value. D seems false for Kaplanian reasons: its character applied to the context of utterance issues in a falsehood, viz. Daniels is a philosopher. F seems false because I use it to make an assertion not about my cousin's sex (that's presupposed) but my cousin's age.

The hope is that we can replace these various explanations with one, perhaps closest in spirit to Stalnaker's undeveloped proposal about F: the sentences seem false because what they assert is false. [3] tells us how to find the propositions asserted**;** we ask what assumptions have to be added to

π_A There is exactly one lodger next door,

π_B Principia Mathematica has exactly one author.

π_C There are exactly ten solar planets.

π_D That man [pointing] is the man drinking a martini.

π_E That man [pointing] is Jones.

π_F My cousin is a male human being.

for it to follow that

A The lodger next door offered me twice that sum

B The author of Principia Mathematica also wrote Principia Ethica.

C All ten solar planets are inhabited.

D The man drinking a martini is a philosopher.

E Jones is burning the leaves.

F My cousin is not a boy anymore.

The needed assumptions would seem to be

A Some and all lodgers next door offered me twice that sum.

B Some and all Principia Mathematica authors also wrote Principia Ethica.

C All solar planets are inhabited.

D That man is a philosopher.

E That man is burning the leaves

F My cousin is an adult.

A-**F**, the asserted contents of A–F, really are what A-F only appear to be, namely false.[26] The suggestion (once again) is that this is not a coincidence. A – F appear false because of the genuine falsity of what they assert or say.

I have been stressing the role asserted content plays in explaining felt truth-value, but it is also relevant to judgments about what is said, contributing in this second way even where truth-value intuitions are lacking. This is the application that matters to Stalnaker:

> it is possible for … presuppositions to vary from context to context, or with changes in stress or shifts in word order, without those changed requiring variation in the semantic interpretation of what is said. This should make possible a simpler semantic theory… (1999, 53).

There is that much less need to multiply meanings if "one [can] use the same sentence" against the background of different assumptions to assert different

[26] The term "asserted content" might be in some cases misleading, since one does not hear "The lodger next door offered me twice that sum" as asserting that some and all lodgers next door offered me twice that sum. Other terms sometimes used are "allegational," "proffered" or "at-issue" content.

things. (Grice of course makes similar claims on behalf of implicature.[27]) [3] shows why asserted content would fluctuate in this way: the shape of the logical gap between π and S is clearly going to depend in part on π. To illustrate with the Donnellan case, the gap between π_D and D is filled by a proposition about Daniels because that is who we presume to be drinking a martini; if we decide it is really O'Leary then the gap-filler changes accordingly. Or consider Stalnaker's elaboration of Langendoen's example:

> normally, if one said "my cousin isn't a boy anymore" he would be asserting that his cousin had grown up, presupposing that he is male." But one might, in a less common context, use the same sentence to assert that one's cousin had changed sexes, presupposing that she is young (1999, 53-4)

The first proposition he mentions (my cousin has grown up) corresponds to the missing premise in

> My cousin is and always has been a male human being.
> ???

[27] See "Further Notes on Logic and Conversation" in Grice 1989, especially the discussion of "modified Occam's razor" on pp 47-9.

My cousin is not a boy any more.

while the second (my cousin has changed sexes) is the premise one needs to make a valid argument of

My cousin is and has always been a human child.

???

My cousin is not a boy any more.

A final example concerns the cognitive content of proper names. How are we to reconcile Mill's idea that names mean their referents with Frege's observation that a name's contribution to cognitive content is not predictable from its meaning so defined? Stalnaker's candidate for the role of (unpredicted) cognitive content is the diagonal proposition, but assertive content can be helpful in this regard as well. Suppose that "n" is a name and that n is presumed to be the so and so. Then "n is F" asserts that the so and so is F (that is what it takes to get from the stated presumption to the conclusion that n is F). This is why "A meteor is about to destroy the Earth" is experienced as saying that a meteor is about to destroy this very planet.

8. Definitions

There is more to bridging the gap between π and S than combining with π to imply S. It is crucial that **S** not be a "bridge too far" at either end. Let me explain what I mean by that.

X is a bridge too far at the S end if it combines with π to imply more than S. So, to take again the Donnellan example, the proposition we want is That man is a philosopher, not That man is a philosopher & snow is white. The latter proposition combines with π to yield "The man with the martini is a philosopher & snow is white," which is a stronger conclusion than we were aiming for.

X would be a bridge too far at the π end if it made for redundancy in the premises -- if it repeated material already present in π. The proposition we want is That man is a philosopher, not That man is a martini-drinking philosopher. It is stated already in π that he is drinking a martini, and there is no need to repeat what is already stated.

How do we enforce the requirement of not being a bridge too far at the S end? Suppose **X** and π imply a stronger statement than S = a statement that S does not imply. Then S does not imply **X**&π (or it would imply the stronger statement). Turning this around, if we stipulate that S does imply **X**&π, that will prevent **X** from being a bridge too far at the S end,.[28]

How do we enforce the requirement of not being a bridge too far at the π end? Suppose we have our hands on the reason why **X** is true (false) – its truth-maker or falsity-maker. **X** has a trace of π in it iff **X** could not be true (false) for that reason unless π too were true (false). So

[4] **X** is π-free iff

> **X** is true (false) and could be true (false) for the same reason, that is, with the same truth-(falsity-)maker, even if π were false (true) .

[28] There is no question of S not implying π -- we have stipulated that S has truth-value only if π is true – so the requirement is really just that S should imply **X.**

I will leave to an appendix my attempt at a theory of truth- and falsity-makers. But the idea is this. Truth-preservation across all worlds w can be called global implication.[29] Local implication (in some particular w) is truth-preservation across all situations in w. A fact in w is a proposition true in w. Given all this,

[29] Implication should preserve definedness as well as truth. The reason is this. "The KoF is bald" should not π-free imply the falsehood, "Among the bald people is a French king." The latter is false because there are no French kings among the bald people, which is compatible with France's having a unique King. This would be a case of π-free implication, if it were a case of implication. But it is not a case of implication, because although truth is preserved, definedness is not. (There could be a world lacking in bald people where France had a unique king.) On an intuitive level, requiring **X** to be defined wherever S is defined is requiring that **X**'s presuppositions be no stronger than S's. This fits with the idea that S counts as false because it implies something whose weaker presuppositions allow it to be false where S is undefined.

[5] A truth-maker for **X** in w is fact that implies **X** globally and is implied by **X** locally.[30]

Now we can explain what is involved in bridging the gap between π and S. Certainly **X** should combine with π to imply S. Also though S should return the favor; it should imply **X**&π. So much is basically to say that although **X** may well be defined on additional worlds, **X** restricted to the π-worlds is true/false in the very same worlds as S. Since **X** extends S beyond the π-worlds. let's call it an extension of S. Finally **X** should be π-free. All in all, then,

[6] **X** bridges the gap between π and S iff **X** is a π-free extension of S.

Asserted content was to be the sum or conjunction of gap-bridgers, so

[7] **S** is the conjunction of S's π-free extensions.

[30] Falsity-makers are similar. It is not assumed that **X** has only one truth- or falsity-maker. A disjunction might be made true either via its left disjunct or its right.

This can be simplified, however. Extensions are just maximum-strength implications; they are implications false in as many π-worlds as possible, compatibly with still being implied by S. Like any maximum-strength implications, they are conjunctions of regular implications. But then the conjoining extensions is conjoining conjunctions of implications. which by the associative law for conjunction is just conjoining implications. So we can get (roughly) the same results as [7] in a more digestible form if we switch to

[8] **S** is the conjunction of S's π-free implications.[31]

This way of putting it further clarifies why an S that is undefined due to presupposition failure might nevertheless strike us as false. S does not seem false merely because it implies a falsehood, for all the sentences we are talking about do that much, just by virtue of implying π. (E.g., "The KoF is bald" implies that

[31] This should be understood to mean "natural, intelligible implications" (and above, "natural, intelligible extensions"). Otherwise asserted content becomes hard to distinguish from semantic content, as the intersection of all π-free propositions implied is liable to be defined only on the π-worlds and the actual world.

France has a unique king.) S seems false because it implies a falsehood, the reasons for whose truth-value have nothing to do with π.

"The KoF is sitting in this chair" implies "Someone is in sitting in this chair." "Someone is sitting in this chair" doesn't suffer from presupposition failure, so we can evaluate it; it is false. Moreover, and this is crucial, "Someone is sitting in this chair" is false for a reason that could still obtain even if France had a unique king, viz. that the chair is empty.

Compare "The KoF is bald." It implies "France has a King." "France has a King" doesn't suffer from presupposition failure, so we can evaluate it; it is false. So "The KoF is bald" counts as false, if "France has a king" is π-free. Could "France has a bald king" have been false for the same reason even if France had a king? No, it could not. The reason "France has a bald king" is false is that France has no king. Clearly it could not have been false for that reason in a world where π was true -- a world where France had a unique king.

9. Claims

A theory of NCPF should address itself to three questions. First, how is it possible, that is, how can presupposition failure ever fail to be catastrophic? Second, why does catastrophe strike in some cases but not others? Third, limiting ourselves to the "good" cases, why do some of these sentences strike us as true and others as false?

Stalnaker gave the outlines of an answer to our first question: how is NCPF so much as possible? π helps to determine the proposition expressed; π's truth-value is not directly relevant to its role as proposition-determiner; and the proposition determined is not limited to worlds where π is true. I call this the outline of an answer because Stalnaker doesn't really explain how false πs can play the same determinative role as true ones.[32]

Now that we have some idea of the mechanism by which π exerts its influence, we can see why truth-value doesn't come into it. Those of us who use stacks of books as bookends know that false books perform just as well in this capacity as true ones. It's the same with presuppositions. The shape of the logical gap

[32] Leaving aside special cases like the referential use of definite descriptions..

between π and S is defined by π's content; whether that content obtains doesn't much matter.[33]

So that's our explanation of how NCPF is possible. It may seem that we have succeeded too well. If π's falsity is irrelevant to its role in determining asserted content, why is presupposition failure ever catastrophic?

This would be a good question if catastrophic presupposition failure had been characterized as presupposition failure resulting in the loss of asserted content. But that is not how we explained it. Presupposition failure is catastrophic, we said, if it has the result that S makes no claim. And having an asserted content does not suffice for making a claim.

The reason is this. Part of what is involved in S's making a claim is that for matters to stand as S says is for them NOT to stand as ~S says. It can happen that S is so tainted by its association with π that **S** and **~S** cease to disagree when π

[33] It can matter a little. If π is true, then any false implication **X** of S is automatically π-free; **X**'s falsity-makers coexists with π in this world, so obviously the two are compatible. Likewise if π is false then S's true implications are automatically π-free.

fails. A sentence whose negation is (counts as) just as true as itself takes no risks and cannot be used to convey real information.

[9] S makes a claim iff: if **S** is true, then **~S** is false.

S makes no claim: **S** is true and **~S** is also true

Now we can define the central notion of this paper:

[10] S is a case of catastrophic presupposition failure iff

π's falsity has the result that S makes no claim.

The poster child here is "The KoF is bald." To find its assertive content we ask, what beyond France's having a unique king is required for the KoF to be bald? The candidates are, let's suppose, (i) France has a bald King, and (ii) any French Kings are bald. We have already seen that (i) is not π-free, because it is false for a reason incompatible with π, namely France's lack of a king. What about (ii)? Its truth-maker (again, France's lack of a king) is consistent with π being false, so (ii) is π-free. It appears that there is nothing for "The KoF is bald" to say but Any French Kings are bald, and nothing for its negation to say but Any French Kings

are non-bald, that is, France has no bald Kings.. Both propositions are true; so according to [9], "The KoF is bald" makes no claim, and according to [10], "The KoF is bald" suffers from catastrophic presupposition failure.

10. "True" and "false"

This leaves the question of what distinguishes the presuppositionally challenged sentences that count as true from the ones that count as false. First a stipulation: the claim S makes -- when it makes a claim – is its asserted content. A sentence's felt truth-value goes with the truth-value of the claim it makes:

[11] S counts as true (false) iff it makes a true (false) claim.

Take "The KoF is sitting in this chair." What π-free implications does it have? One obvious implication is The chair contains a French king. This is false because No one is sitting in the chair, or perhaps because No king is …., or No

French king is ….[34] All of these are compatible with France's having a unique King. It looks then like "The KoF is sitting in this chair" claims in part that The

[34] Note that that France lacks a king does not make "This chair contains a French King" false, since it does not imply the latter's presupposition that there is such a thing as this chair. What about France lacks a King and there is such a thing as this chair? This is formally eligible but a better -- more proportional -- candidate for the role is This chair is empty (see Yablo 2003 and the Appendix). Objection: If "The KoF sits in this chair" π-free implies the false "A French King sits in this chair," shouldn't "The KoF is heavier than this chair" π-free imply the false "A French King is heavier than this chair"? Yet "The KoF weighs more than this chair" does not strike us as false. I reply that "A French king is heavier than this chair" is not π-free, because in this case there is no better candidate for falsity-maker than France lacks a king and there is such a thing as this chair. No simple fact about the chair itself falsifies "A French king is heavier than this chair" as the chair's emptiness falsifies "A French king sits in this chair." More generally, "The KoF bears R to x" does not strike us as false when R is an "internal relation" like taller-than or heavier-than -- a relation that obtains in virtue of intrinsic properties of the relata. The proposed explanation is that facts purely about x do not suffice for the falsity of "A French king bears R to x"; France's lack

chair contains a French king. That claim is false, so the sentence that makes the claim ("The KoF is sitting in this chair") counts as false.

A couple of examples finally of counting as true despite presupposition failure. The first will use a concrete term, the second, to move us back to the ontological relevance of NCPF, will use a term that's abstract.

A long time ago there were two popes, one in Rome, the other in Avignon. Imagine that we check into a monastery one fine night in 1400, and scrawled on the bedpost we read, "The pope slept here." A bit of research reveals that both popes in fact slept in the bed in question. Then it seems to me that the inscription strikes us as true; for although it was making a stronger claim than apparently, this stronger claim was correct. To check this against the theory, we must hunt around for π-free implications of "The pope slept here." One such implication is that some pope slept here; another is that all popes slept here. So far then it looks like the assertive content is that some and all popes slept here. It is because this is true in the imagined circumstance that "The pope slept here" counts for us as true.

---

of a king has to be brought in, which makes the implication no longer π-free. (See in this connection Donnellan 1981 and von Fintel 2004.)

Looking into my back yard last Monday, I saw one cat; look into the yard last Tuesday, I saw two cats, and so on. I now form the hypothesis that this pattern will continue forever [35] -- to state the hypothesis more explicitly,

"For all n, the number of cats in my yard on the nth day = n."

I presuppose in saying this that no matter what day it is, there is a unique thing that numbers the cats in my yard on that day (and more generally that whenever there are finitely many Fs, there is a unique thing that numbers them). Now consider the statements on this list:

on the first day there is one cat in my yard
on the second day there are two cats in my yard
on the third day there are three cats etc.
etc.

All of them are implied by my hypothesis, and it is easy to see that each is a π-free implication. A typical falsity-maker is the fact that on the third day there

[35] This example is from Burgess and Rosen (ms).

are no cats in my yard. That there are no cats in my yard can happen just as easily in platonistic worlds (where π holds) as in wholly concrete worlds. A typical truth-maker is the fact that on the third day the cats in my yard are Zora, Teasel, and Yossele. For the cats in my yard to be Zora, Teasel, and Yossele can happen just as easily in concrete worlds (where π fails) as in platonistic worlds.

Now I haven't argued that these are all the π-free implications. But if they are, then what my hypothesis says is that on the first day there is one cat, on the second there are two, and so on. This fits with our intuitive sense that the hypothesis counts as true or false according to how many cats my yard contains on which days; the existence of numbers plays no role whatsoever. Note the analogy with the King of France; just as "The KoF is sitting in this chair" counts as false because of the chair's material contents -- nothing to do with French royalty – and "The number of cats on the nth day = n" counts as true, if it does, because of my yard's material contents -- nothing to do with numbers.

11. Parting thoughts on abstract ontology

Nominalists maintain that abstract terms do not refer. They will find it suggestive, then, that (supposedly empty) abstract terms make in some cases the same sort of contribution to felt truth-value as definitely empty concrete terms do.

Platonists will complain that empty concrete terms make a negative contribution; simple King-of-France sentences almost all count as false, to the extent that they make a claim at all. One would expect the emptiness of abstract terms, if they were empty, to manifest itself the same way. But the sentences we construct with abstract terms very often strike us as true.

I agree that the failure of a concrete term to refer prevents it from exercising positive semantic influence. But there is a reason for this. "The King of France"'s semantic contribution goes way beyond our notions of what a French king would have to be like. He is (or would be, if he existed) an original source of information of the type that makes simple King-of-France sentences count as true. Numbers by contrast are not (would not be) an original source of information on any topic of interest; their contribution is exhausted by what they are supposed to be like. This makes the presupposition that numbers exist "fail-safe" in the sense that its failure makes (or would make) no difference whatever

to which applied arithmetical sentences count as true. I am tempted to conclude that nothing in the felt truth-values of those sentences has any bearing on the issue of whether numbers exist. s

Appendix

Situations are parts of worlds; worlds are maximal situations; truth-in-a-world is a case of truth-in-a-situation. Suppose that A is a sentence and situation s is part of world w. Then s verifies (falsifies) A only if it contains everything potentially relevant to A's truth-value in w. Thus for s to verify "All swans are white," it is not enough that all the swans in s are white; s should also contain all of w's swans. The formal upshot is a condition called persistence: if A is true (false) in s and s is part of s*, then A is true (false) in s*.[36] Now we introduce two notions of implication (the first is basically familiar, the second not):

[36] See Kratzer for an enlightening discussion of persistence.

A implies B globally iff for all possible worlds w, A is true in w only if B is true in w.[37]

A implies B locally in w iff for all s ≤ w, A is true in s only if B is true in s.[38]

Here is our first stab at a definition of truth-maker. A truth-maker for X in w is a T that implies X across all possible worlds, and that holds in every w-situation where X holds:

T makes X true in w iff

(a) T is true in w,

(b) T implies X globally,

(c) X implies T locally in w.

(F makes X false in w iff it makes ~X true.) This runs into a problem, however.[39]

---

[37] I ignore that implication should preserve definedness as well. See note 29.

[38] Compare the definition of lumping in Kratzer 1989.

It could happen that $\underline{X}$'s truth in $\underline{w}$ is overdetermined by $\underline{T_1}$ and $\underline{T_2}$ (e.g., $\underline{X}$ might be a disjunction with unrelated disjuncts both of which are true). Then $\underline{X}$ may well lack a truth-maker. It won't imply $\underline{T_1}$ in $\underline{w}$ because there are $\underline{w}$-situations where $\underline{X}$ holds thanks instead to $\underline{T_2}$, and it won't imply $\underline{T_2}$ because there are $\underline{w}$-situations where $\underline{X}$ holds thanks instead to $\underline{T_1}$. Still, there ought to be a $\underline{v_1} \leq \underline{w}$ such that $\underline{X}$ implies $\underline{T_1}$ in $\underline{v_1}$ and a $\underline{v_2} \leq \underline{w}$ such that $\underline{X}$ implies $\underline{T_2}$ in $\underline{v_2}$. So rather than asking $\underline{X}$ to imply $\underline{T}$ in $\underline{w}$, we should ask it to imply $\underline{T}$ in some subsituation of $\underline{w}$.

$\underline{T}$ makes $\underline{X}$ true in $\underline{w}$ iff for some $\underline{v} \leq \underline{w}$

(a) $\underline{T}$ is true in $\underline{v}$,

(b) $\underline{T}$ implies $\underline{X}$

(c) $\underline{X}$ implies $\underline{T}$ in $\underline{v}$,

Now that we are explicitly contemplating multiple truth- and falsity-makers, we need to adjust the definition of π-freedom:

$\underline{X}$ is $\underline{\pi\text{-free}}$ in $\underline{w}$ iff

---

[39] Originally raised by Heim against Kratzer.

> X is true in w and it has a truth-maker that holds also in worlds where π is false, or X is false in w and it has a falsity-maker that holds also in worlds where π is true.

This runs into a different problem. "France has a bald king" had better not be a π-free consequence of "The King of France is bald," or the latter will count as false which it shouldn't. The obvious falsity-maker is France has no king, which is indeed not compatible with π. But another technically eligible falsity-maker is France has no bald kings. This is compatible with France's having a unique king, and since our definition requires only that some falsity-maker be transportable to π-worlds, it seems we are sunk. The solution I suggest is to impose a proportionality requirement along roughly the lines of Yablo 2003. France has no bald kings is needlessly complicated in that a strictly simpler condition (France lacks a king) still satisfies conditions (a)-(c). Thus we should think of (a)-(c) as defining candidacy for the role of truth- (falsity-) maker. and add that the successful candidate should not involve gratuitous complications in whose absence (a)-(c) would still be satisfied.

Bibliography

Beaney, Michael. 1997. The Frege Reader (Oxford: Blackwell)

Beaver, David. 2001. Presupposition and Assertion in Dynamic Semantics (Palo Alto: CSLI Press)

Beaver and Krahmer, "Presupposition and Partiality" (ms)

Bezuidenhout and Reimer. 2004. Descriptions and Beyond (Oxford: OUP)

Burgess, John, and Gideon Rosen. ms. "Nominalism Reconsidered," to appear in Stewart Shapiro, ed., Handbook of Philosophy of Mathematics and Logic

Burton-Roberts, Noel. 1989. The Limits to Debate: A Revised Theory of Semantic Presupposition (Cambridge: Cambridge U Press)

Cohen, Ariel. 2000. "The King of France Is, In Fact, Bald," Natural Language Semantics 8: 291-295

Cole, P. (ed.) 1981.Radical Pragmatics (New York: Academic Press),

Donnellan, Keith. 1966. "Reference and Definite Descriptions," Philosophical Review 75: 281-304

Donnellan, Keith. 1981. "Intuitions and Presuppositions," in Cole 1981, 129-142

von Fintel, Kai. "What is Presupposition Accommodation? (ms)

von Fintel, Kai. 2004. "Would You Believe It? The King of France is Back," in Bezuidenhout and Reimer 2004, 315-341

Frege, Gottlob. 1872. "On Sense and Meaning," page references to Beaney 1997

Frege, Gottlob. 1918. "The Thought," page references to Beaney 1997

Glanzberg, Michael. "Expression Failure and Presupposition Failure" (ms)

Glanzberg, Michael. 2002."Context and Discourse," Mind and Language 17: 333-375

Grice, H. Paul. 1989. Studies in the Way of Words (Cambridge, MA: Harvard University Press)

Heim, Irene. 1983. "On the Projection Problem for Presuppositions," WCCFL 2

Humberstone, Lloyd. 2000. "Parts and Partitions," Theoria 66: 41-82

Kratzer, Angelika. 1989. "An Investigation into the Lumps of Thought," Linguistics and Philosophy 12: 607-653

Kripke, Saul. 1977. "Speaker's Reference and Semantic Reference," originally in Midwest Studies in Philosophy II, as reprinted in French, Uehling, and Wettstein (eds.), Contemporary Perspectives in the Philosophy of Language (U of Minnesota Press: Minneapolis), 6-27

Lasersohn, Peter. 1993. "Existence Presuppositions and Background Knowledge," Journal of Semantics 10: 113-122

Van der Sandt, Rob. 1989. "Anaphora and Accommodation," in Renate Bartsch, Johann van Benthem and Peter van Emde Boas (eds.) Semantics and Contextual Expression (Foris: Dordrecht)

Shanon, Benny. 1976. "On the Two Kinds of Presupposition in Natural Language," Foundations of Language 14: 247-249

Shope, Robert. 1978. "The Conditional Fallacy in Contemporary Philosophy", Journal of Philosophy 75: 397-413

Simons, Mandy. 2003. "Presupposition and Accommodation: Understanding the Stalnakerian Picture", Philosophical Studies 112: 251-278

Stalnaker, Robert. 1970. "Pragmatics," Synthese 22, page references to the reprint in Stalnaker 1999

Stalnaker, Robert. 1974. "Pragmatic Presuppositions," in Munitz and Unger, eds. Semantics and Philosophy (New York: NYU Press), page references to the reprint in Stalnaker 1999.

Stalnaker, Robert. 1978. "Assertion," in Cole, ed., Syntax and Semantics, vol. 9, page references to the reprint in Stalnaker 1999

Stalnaker, Robert. 1999. Content and Context (Oxford: OUP)

Stalnaker, Robert. 2002. "Common Ground." Linguistics and Philosophy 25: 701-721

Strawson, P.F. 1950, "On Referring," Mind 59; page references to the reprint in Strawson 1971

Strawson, P.F. 1954. "A Reply to Mr. Sellars." Philosophical Review 63:216-231

Strawson, P.F. 1964, "Identifying Reference and Truth-Values," Theoria 30; page references to the reprint in Strawson 1971

Strawson P.F. 1971. Logico-Linguistic Papers (London: Methuen)

Yablo, S. 2003. "Causal Relevance," Philosophical Issues 13

Zeevat, Hank. 1992. "Presupposition and Accommodation in Update Semantics," Journal of Semantics 9: 379-412