Counterfactuals to the Rescue[*]

Cleo Condoravdi
Stanford University

1 Introduction

Edgington (2008) argues that the stakes of philosophical theorizing about conditionals are so high because of the role conditionals play in thought and reasoning. While advocating an analysis of counterfactuals in terms of conditional probability, she singles out a central problem that all theories of counterfactuals must address:

> All theories of subjunctives—Goodman's, Lewis's, mine—share what is essentially the same problem, that of specifying what you hang on to and what you give up when you make a counterfactual supposition: suppose such-and-such had been the case; what do you hold constant? For Goodman this is the problem of cotenability, for Lewis it's the problem of closeness, for me, it's the problem of which context shift is appropriate, which probability distribution is the appropriate one, given that it is not the one that represents your present state of belief. (p. 13–14)

This paper uses grammatical evidence from the phenomenon of polarity reversal in counterfactuals to shed light on the role context and, in particular, pragmatic presuppositions play in determining what you hold constant and what you give up when interpreting counterfactuals. It makes the case that counterfactuals exhibiting polarity reversal settle that the consequent follows not only from the explicitly mentioned antecedent but also from each of a set of evoked alternative antecedents. In order for this to happen, opening up the possibility of the antecedent requires also opening up the possibility of the alternatives. This, in turn, implies that in making the counterfactual assumption, you not only give up propositions you take for granted that are inconsistent with it, but also certain contextually entailed propositions even though they are consistent with the counterfactual assumption. The close study of grammatical systems shows that language endows simple words with the kind of meaning to enable compact reasoning about alternative possibilities without requiring of speakers to spell it all out.

[*] I would like to thank Lee Walters and Sven Lauer for discussion and comments on the paper.

2 Polarity Reversal in Counterfactuals

Positive polarity items (PPIs) are expressions which are, generally, not acceptable within the scope of negation in the same clause. They are a rather heterogeneous lot, including the temporal adverbials *already* and *still*, as well as degree modifiers such as *pretty* and *rather*. (1) – (3) exemplify the regular distribution of three of these positive polarity items.[1]

(1) a. He has already arrived. / He has arrived already.
 b. ?# He has not already arrived. / ?# He has not arrived already.
 NOT > ALREADY

(2) a. He is still at home. / He is at home still.
 b. ?# He is not still at home. / ?# He is not at home still.
 NOT > STILL

(3) a. She is pretty smart.
 b. ?# She is not pretty smart.

The temporal adverbials can syntactically appear over or under negation, and the logical scope mirrors the overt syntactic scope. Specifically, if the negation appears to the left of the adverbial, as in (1-b), (2-b), it takes scope over it; if the negation appears to the right of the adverbial, as in (4), it scopes under the adverbial. Thus, (4-a) and (4-b) are on a par with (1-a) and (2-a), since the PPIs are not in the direct scope of negation.

(4) a. He is already not at his post.
 ALREADY > NOT
 b. She is still not at her post.
 STILL > NOT

Exceptionally, PPIs can appear in the scope of negation in certain environments. Baker (1970a,b), who first studied the phenomenon systematically, introduced the term 'polarity reversal' and identified the antecedent of counterfactual conditionals as an environment in which PPIs can appear in the scope of negation. (5) shows the exceptional acceptability of the three PPIs above within negated antecedents of counterfactuals.

(5) a. If he had not already arrived, we would have postponed the meeting.
 b. If he were not still at home, we would have missed him.
 c. If she weren't pretty smart, she wouldn't have gotten out of this situation.

1 As the variants in examples (1) and (2) illustrate, the temporal adverbials show a degree of freedom as to their syntactic placement.

The counterfactual suppositions in (5) — *if he had not already arrived*, *if he were not still here*, *if she weren't pretty smart* — all include a negation that scopes over a PPI. Therefore, (5-a), (5-b), and (5-c) should be at least as problematic as (1-b), (2-b), and (3-b), but they are not.

A confounding factor that needs to be set aside is that (1-b), (2-b), and (3-b) appear acceptable in certain contexts, especially if *not* is emphatically stressed and the utterance is understood as an instance of denial. According to Horn (1989) and van der Sandt (1991), a denial is a speech act objecting to some aspect of the full informational content of a previous utterance, or even to the particular form of the utterance. Thus, it can reject that utterance for any number of disparate reasons: because the proposition it expresses is false, or because it suffers from presupposition failure, or because one of its implicatures is false, or because it gives rise to objectionable connotations or inferences of any kind. The polarity of a sentence used as a denial may be negative or positive.[2] Consequently, occurrences of (1-b), (2-b), or (3-b) can be acceptable if they engage with a previous utterance by a different speaker. By contrast, whenever a denial construal is disfavored, as in (6) and (7), PPIs resist being in the scope of negation.

(6) a. John said he'd come late but he is already here.
b. # John said he'd come early but he is not already here.

(7) a. John said he'd leave early but he is still here.
b. # John said he'd leave late but he is not still here.

The acceptability of (1-b), (2-b), and (3-b) under certain special conditions should, therefore, be set apart from cases of polarity reversal, which is systematic and, it would seem, grammatically conditioned. In polarity reversal, it is properties of the larger construction within which the negation and the PPIs appear which render the latter under the scope of the former licit.

As a general characterization of the environments that allow polarity reversal, Baker (1970a,b) proposed, roughly, that a sentence S_1 with a polarity sensitive expression exhibiting polarity reversal is acceptable if it entails or pressuposes another sentence S_2 containing the same polarity sensitive expression whose polarity

2 For instance, in the mini-dialogues below, B's utterance, which objects to the truth-conditional content of A's utterance, has positive polarity. In (i) the polarity of the two sentences uttered by A and by B, respectively, is switched, whereas in (ii) it remains the same.

(i) A: John never showed up.
B: What do you mean? John is already here.

(ii) A: John left work early today.
B: What do you mean? John is still in his office.

is the opposite of that of S_1. The idea is that an S_1 containing a PPI within the direct scope of negation would be acceptable if it entails or presupposes some S_2 which contains the PPI in an affirmative context.

Baker (1970b) and, with certain amendments, Karttunen (1971) attributed polarity reversal in counterfactuals to the presupposition of the falsity of the antecedent, and regarded the phenomenon of polarity reversal in counterfactuals as striking evidence for the role of presupposition in accounting for polarity licensing. (8) is one of Baker's (1970b) examples, which, he claims, "is not appropriate unless the speaker is willing to commit himself to the stronger presupposition in [(9)]" (p. 6).[3]

(8) If someone hadn't already succeeded in making radio contact with your husband, the Coast Guard would almost surely be expressing its concern.

(9) Someone has already succeeded in making radio contact with your husband.

Where does the presupposition come from? Subjunctive conditionals[4] do not necessarily presuppose that their antecedent is false (Anderson 1951; Stalnaker 1975; Edgington 1995). In fact, they can be used to reason about either the truth or the falsity of the antecedent given what is known, or taken for granted, about the consequent. Both (10) and (11), for instance, are used in a context which is taken to be consistent with the antecedent. In (10) the conditional is used to argue for the truth of its antecedent given the truth of its consequent. In (11) the conditional is used to argue for the falsity of its antecedent given the falsity of its consequent, and, therefore, the falsity of the antecedent cannot be taken for granted in advance, i.e., it cannot be pragmatically presupposed prior to the utterance of the conditional.

(10) If he had still been in his office then, the lights would have been on. The lights were on. Therefore, he was still in his office then.

(11) If he had still been in his office then, the lights would have been on. The lights were off. Therefore, he was not in his office then.

Subjunctive conditionals with polarity reversal in the antecedent, however, can only have counterfactual uses. Baker observed that a subjunctive conditional with polarity reversal cannot be used to argue for the falsity of its antecedent, as evidenced

3 Baker takes both *someone* and *already* to be positive polarity items. The status of *someone* as a PPI is controversial but that of *already*, which I will focus on here, is not. 'Stronger presupposition' is meant to compare the factual presupposition in (9) with a weaker non-factual presupposition.

4 I use the term 'subjunctive' to make reference to the form of a conditional, and the term 'counterfactual' for conditionals used in a context in which their antecedent is settled to be false. The term 'subjunctive' does not properly characterize the full range of morphosyntactic properties of the relevant conditionals, but I adopt it, as it is a standard term to pick out the particular class of conditionals which can have a counterfactual interpretation.

by the contrast between (13) and (11). We can also observe that a subjunctive conditional with polarity reversal cannot be used to argue for the truth of its antecedent; consider the contrast between (12) and (10).

(12) # If he hadn't still been in his office then, the lights would have been off. The lights were off. Therefore, he was not in his office then.

(13) # If he hadn't still been in his office then, the lights would have been off. The lights were on. Therefore, he was still in his offce then.

According to Baker, these contrasts are due to the fact that the PPI cannot appear in the scope of negation unless the conditional comes with the presupposition that its antecedent is false.

Although Baker's official principle governing when a PPI is acceptable under polarity reversal makes reference to a notion of presupposition as a relation between two sentences,[5] in the discussion about the phenomenon, the notion of presupposition that he refers to is that of speaker presupposition.[6] Karttunen (1971) argues that the polarity rule should not make reference to a relation between two sentences in isolation but rather be further relativized to a set of premises that are taken for granted. In the conclusion of his paper, Karttunen explicitly endorses the pragmatic notion of speaker presupposition.

> [A]s sound as Baker's principle is, it is too narrow in requiring that there be some logical relation (entailment or presupposition) between the sentence which violates the general polarity rule and the corresponding sentence with reversed polarity. It is enough if the latter is regarded by the speaker as a necessary truth in the particular state of affairs that he is considering. One case where the speaker clearly should have such a belief is when the latter sentence stands in a certain logical relation to the former, that is, is either entailed or presupposed by it. It is this subset of polarity reversals that is explained by Baker's principle in its original form. Whatever the correct formulation ultimately turns out to be, it will have to cover a larger class of cases. (p. 294–295)

The Baker–Karttunen generalization can be stated in a way that avoids assuming that subjunctive conditionals come with a presupposition regarding their antecedent: a PPI in the antecedent of a subjunctive conditional allows polarity reversal only if the conditional is counterfactual, i.e., it is uttered in a context which is incompatible with its antecedent.

---

5 See the principle in (5b$''$) in Baker (1970b: 12).

6 This can be seen, for instance, in the quote given above, in connection with (8) and (9).

In later work, Ladusaw (1980), Krifka (1991, 1995), Szabolcsi (2004), and Schwarz & Bhatt (2006) pursued alternative explanations for polarity reversal. Krifka and Szabolsci attributed polarity reversal to the presence of a higher operator taking both the negation and the PPI in its scope; Ladusaw and Schwarz & Bhatt to a special type of negation. In the terminology of Szabolsci, the higher operator, or the special type of negation, 'rescues' the PPI. The aim then of these kinds of approaches is to characterize the types of rescuing operators. But with the exception of Schwarz & Bhatt, this work did not consider the case of counterfactuals.

Schwarz & Bhatt claim that rescuing involves a special negation, dubbed 'light negation'. They independently motivate this assumption based on certain distributional facts regarding the syntactic position of negation and other phrases within a clause in languages like German. They then show that the environments where negation has a special syntactic distribution largely coincide with the environments that support rescuing of PPIs.[7] From a semantic point of view, on Schwarz & Bhatt's account, negation is not just two-way ambiguous between regular negation and light negation, but multiply ambiguous, as the interpretation of light negation has to vary across different environments in which it is supposed to occur. For counterfactuals, they state: "We do not know for sure how light negation comes to enforce counterfactuality. But one possibility that comes to mind is that counterfactual light negation triggers a factive presupposition, that is, the presupposition that its scope is true" (p. 92). For instance, the negation in the antecedent of (8), which, given the PPI in its scope, has to be light negation, presupposes (9). This presupposition is then inherited by the entire conditional, as presuppositions originating in the antecedent do in general, and this is what imposes a counterfactual interpretation on the conditional.

Although it can derive the counterfactual interpretation of subjunctive conditionals in their role as PPI rescuers, the price this account pays is quite high given the multiple ambiguity of negation it has to assume. An analysis that does not postulate ambiguity of negation would seem to be preferable on general methodological grounds. Zeijlstra (2012) offers such an alternative account of Schwarz & Bhatt's distributional facts. But the issue of why counterfactuals act as PPI rescuers remains to be addressed.

A common feature of the approaches outlined above is that the status of an expression as a PPI does not follow from its lexical semantics, and for the particular case of polarity reversal in counterfactuals, that the interpretation of the counterfac-

7 Ladusaw (1980) also appeals to a special negation operator to explain the appearance of PPIs in the scope of negation. That operator, according to Ladusaw, has identical truth conditions with the standard negation operator but carries a conventional implicature that "someone has believed until recently that the proposition in its scope is true". Whereas Ladusaw intended his special negation operator to apply to the denial uses we saw in connection with (1-b), (2-b) and (3-b), Schwarz & Bhatt do not see light negation as covering cases of denial.

tual conditional itself does not play a role. In this paper, I take as my starting point a more fine-grained view of polarity sensitivity, and trace the possibility of polarity reversal in counterfactuals to the meaning of the polarity sensitive expressions and the interpretation of counterfactuals. In my proposal, the negation in the antecedents of the conditionals in (5) is regular truth-functional negation, and the need for a counterfactual interpretation of subjunctive conditionals in their role as PPI rescuers will be derived. I will focus on the two temporal adverbials *already* and *still*, as their lexical semantics is better understood. Inevitably, in order for the analysis to cover other cases, the lexical semantics of the relevant PPIs would have to be settled on first. But the reasoning regarding the interpretation of counterfactuals should carry over.

The structure of the rest of the paper is as follows. Section 3 gives more detail on the theory of PPIs that I am assuming. Section 4 fixes the meaning of *already* and *still*. Section 5 develops the analysis and shows how the rescuing behavior of making counterfactual assumptions comes about.

3 Scalar Assertions

The alternative analysis of PPI rescuing in antecedents of counterfactuals that I explore relies on the idea that, in virtue of their conventional meaning, polarity sensitive items are associated with a set of alternatives in addition to their ordinary truth-conditional content. These alternatives, ultimately, lead to a set of alternative propositions that are ordered by semantic strength with the proposition corresponding to the plain truth-conditional content of the sentence. The polarity of the sentence obviously figures crucially in the direction of the ordering by strength. The effect of the alternatives on the overall meaning is at the core of the informativity-based theories of polarity sensitivity proposed by Krifka (1995) and Chierchia (2006).

The limited distribution of one class of polarity sensitive expressions, including the PPIs of interest here, has to do with the fact that they give rise to *scalar assertions*. Scalar assertions are assertions of sentences that, semantically or pragmatically, evoke alternatives and in which the proposition based on the plain truth-conditional content is informationally ordered with respect to the alternatives. A scalar assertion conveys more information than a plain assertion, as it negates any alternative propositions that are informationally stronger than the proposition based on the plain truth-conditional content.[8]

Plain and scalar assertions update the information state of a context, correspond-

8 Utterances of *John ate some of the cookies*, for instance, constitute scalar assertions whenever *all*, the lexical alternative to *some*, is pragmatically evoked. A scalar assertion of this sentence is more informative than the assertion of its truth-conditional content because it also conveys the information corresponding to the scalar implicature that John did not eat all of the cookies.

ing to what is common ground between the relevant agents in the context, as is familiar from Stalnaker (1973, 1975). If we construe information states as sets of possible worlds, the relation of strength (informativity) between such information states is in terms of the subset relation. With a slight abuse of notation but for ease of exposition, in the following, I use the same symbol for contexts and their information state parameter (the context set, in Stalnaker's terms). The basic effect of a plain assertion of ϕ in a context c can be identified with the update function + defined as in (14). It relies on the notion of felicity, a necessary condition of which is presupposition satisfaction.[9]

(14) $c + \phi = \{w \in c \mid w \in [\![\phi]\!]_c\}$ provided ϕ is felicitous relative to c, else undefined.

(15) ϕ is felicitous relative to context c only if c satisfies ϕ's presuppositions.

(16) Context c satisfies presupposition p of ϕ iff c entails p.

The relation of informational strength between two sentences can then be defined in terms of the update function as in (17).

(17) ϕ_1 is informationally at least as strong as ϕ_2 iff for any context c relative to which ϕ_1 and ϕ_2 are felicitous, $c + \phi_1 \subseteq c + \phi_2$.

Scalar assertions of a sentence with alternatives update the context with the proposition corresponding to the plain truth-conditional content and, in addition, negate the truth of any informationally stronger alternative propositions. Thus, in general, scalar assertions result in a more informative context than plain assertions. The effect of a scalar assertion on a context is defined in (18).

(18) $ScalAssert(\langle \phi, Alt(\phi) \rangle, c) =$
$\{w \in c \mid w \in [\![\phi]\!]_c \wedge \neg(\exists \phi' \in Alt(\phi)\, (w \in [\![\phi']\!]_c \wedge c + \phi' \subset c + \phi))\}$

Polarity sensitive expressions with semantically triggered alternatives of a certain kind give rise to scalar assertions. Their acceptability depends on the relation of informational strength between the plain content and the alternatives. They are acceptable in an environment when the plain truth-conditional content, which, of course, depends on the polarity of that environment, is at least as strong as each one of the alternatives. In that case, $c + \phi = ScalAssert(\langle \phi, Alt(\phi) \rangle, c)$. By contrast, in an environment in which they are not acceptable, the plain truth-conditional content is informationally weaker than any one of its alternatives. In that case, $ScalAssert(\langle \phi, Alt(\phi) \rangle, c) = \emptyset$ even when $c + \phi \neq \emptyset$.

9 The requirement of presupposition satisfaction for successful contextual update thus connects linguistically triggered presuppositions with pragmatic presuppositions.

On this kind of theory, the contrast between (1-a) and (1-b), or that between (2-a) and (2-b), would be explained by the fact that the contribution *already* and *still* make to the meaning of the sentences generally result in consistent scalar assertions of (1-a) and (2-a) in contexts relative to which they are felicitous, whereas (1-b) and (2-b) lead to inconsistency in any such context. I show this in the following section.

If we could show that (5-a) and (5-b) are like (1-a) and (2-a) in this respect, rather than (1-b) and (2-b), the rescuing behavior of counterfactuals would follow. This is what I do in section 5, where I show that with polarity reversal the plain content of the conditional is informationally stronger than each one of the alternatives, at least when the conditional is counterfactual.

4 *Already, still*: content, presupposition and alternatives

The temporal adverbials *already* and *still* have a trivial truth-conditional content but give rise to non-trivial implications. For example, (19-a) and (20-a) have the same truth-conditional content, shared with (21). In contrast to (21), they have an implication regarding his being in the cave prior to the reference time, in this case the time of utterance. The polarity of this implication is flipped for the two adverbials, as seen in (19-b) vs. (20-b). They also have the counter-to-expectation implication in (19-c) and (20-c).

(19) a. He is already in the cave.
b. He was not in the cave earlier.
c. He got into the cave earlier than expected.

(20) a. He is still in the cave.
b. He was in the cave earlier.
c. He has stayed in the cave longer than expected.

(21) He is in the cave (now).

The first kind of implication is actually one aspect of a more general implication about a potential change. This implication has the hallmarks of a linguistically triggered presupposition, as it projects through possibility modals and antecedents of conditionals. For instance, (22-a) and (22-b) imply that the treasure was in the cave at an earlier time but may have been moved away, while (23-a) and (23-b) imply that the treasure was not in the cave at an earlier time but may have been moved there.

(22) a. Maybe the treasure is still in the cave.
b. If the treasure is still in the cave, it will be safe.

(23) a. Maybe the treasure is already in the cave.
b. If the treasure is already in the cave, it will be safe.

With predicates where the requisite presuppositions cannot be satisfied in normal contexts, use of the adverbials can lead to infelicity, as in (24-b) and (25-a).

(24) a. It is already late.
b. #It is still late.

(25) a. #It is already early.
b. It is still early.

Löbner (1989) characterized the meaning of these adverbials in terms of admissible intervals; admissible intervals consist of a positive/negative phase followed by a negative/positive phase but no other phases.[10] As he put it, *already* "pick[s] out a well-defined interval out of the overall time-axis, i.e., the time interval starting with the last negative phase beginning before [the reference time] and ending with the eventually following positive phase. ...Starting from the fact that there is a positive phase of p which began before [the reference time], the question [for *still* is whether this phase continues until [the reference time] or has been succeeded by a negative phase" (p. 174).

Recasting Löbner's proposal somewhat, we can say that both adverbials presuppose that there is a transition within an interval with respect to a given property. The property is determined by the predicate the adverbials modify, in the case at hand, the clausal predicate. Let's call such intervals I periods of transition with respect to the property expressed by the clausal predicate. For *already*, the presupposed transition is from a maximal initial subinterval I_{NEG} of I not satisfying the property to a maximal final subinterval I_{POS} satisfying that property. For *still*, the presupposed transition is from a maximal initial subinterval I_{POS} of I satisfying the property to a maximal final subinterval I_{NEG} not satisfying that property. Therefore, *already* requires periods of transition such that $I_{\text{NEG}} \prec I_{\text{POS}}$, while *still* requires periods of transition such that $I_{\text{POS}} \prec I_{\text{NEG}}$. Both *already* and *still* assert of the reference time that it is within I_{POS} but because of their different presupposition their implications differ.

Let t_0 be the relevant reference time. Utterances of (22-a) and of (23-a) presuppose that t_0 is within periods of transition with respect to the temporal property of the treasure being in the cave. There could well be uncertainty in the common ground about what times on the time axis the periods I, I_{POS}, I_{NEG}, or even t_0, correspond to (different times in different possible worlds), as well as the position of t_0 within any given period of transition. Utterances of (22-a) and of (23-a) convey that it is consistent with the relevant agent's information state that t_0 is within the positive phase I_{POS} of the period of transition, hence that the treasure may be in the cave at t_0.

Already and *still* also activate alternatives. These are alternative times within the periods of transition at which the predicate they modify may hold. For *already*

10 Löbner (1989) studied the behavior of the German equivalents of *already* and *still*, *schon* and *noch*.

the alternative times follow the reference time, while for *still* they precede it.[11] Depending on when the transition occurs within a period of transition, the reference time may be within I_{NEG}, while at least some, if not all, of the alternative times are within I_{POS}. But if the reference time is within I_{POS}, then all the alternative times are also within I_{POS}.

The association with alternatives along with their presuppositions is what makes *already* and *still* polarity sensitive.[12] Relative to a context c satisfying the presuppositions of $Already(\phi(t_0))$, the update with the plain truth-conditional content $c + \phi(t_0)$ is informationally stronger than the update with each one of the alternatives $c + \phi(t)$, where $t_0 \prec t$. Similarly for $Still(\phi(t_0))$, except that the alternatives are based on times $t \prec t_0$. Hence assertions of (1-a), (2-a), (19-a), (20-a), for instance, constitute scalar assertions. The relation of information strength is reversed when the polarity of the environment in which *already* and *still* is negative. Relative to a context c satisfying the presuppositions of $\neg Already(\phi(t_0))$,[13] the update with the plain truth-conditional content $c + \neg\phi(t_0)$ is informationally weaker than the update with each one of the alternatives $c + \phi(t)$, where $t_0 \prec t$. Similarly for $\neg Still(\phi(t_0))$, where the alternatives are based on times $t \prec t_0$. Hence assertions of (1-b) and (2-b) constitute scalar assertions leading to an inconsistent update in any context in which their presuppositions are satisfied.

5 Counterfactuals and Alternatives

The truth-conditional, possible-world semantics for counterfactual conditionals of Stalnaker (1968) and Lewis (1973) relies on the notion of comparative similarity between worlds. For a counterfactual conditional to be true in a world w, the proposition expressed by the consequent of the conditional has to be true throughout the worlds in which the proposition expressed by the antecedent is true and which are otherwise as similar to w as can be.[14] The facts of w determine the similarity to w, but not all facts have equal weight and the relation of similarity, as both authors have emphasized, is vague and context-dependent. Here is, for instance, how Lewis (1981) puts it:

> We may think of factual background as ordering the possible worlds. Given the facts that obtain at a world i, and given the attitudes and understandings and features of context that make some of these facts

11 The alternatives also give rise to the counter-to-expectation implication.

12 See also Krifka (1995).

13 Since presuppositions project through negation, the presuppositions of $\neg Already(\phi(t_0))$ are the same as those $Already(\phi(t_0))$.

14 This informal statement presupposes what Lewis calls the Limit Assumption.

count for more than others, we can say that some worlds fit the facts of *i* better than others do. Some worlds differ less from *i*, are closer to *i*, than others. (p. 218)

5.1 The semantics of counterfactuals and presuppositional antecedents

For purposes of formulating the semantics of conditionals, we can model maximal similarity in terms of selection functions which map a world w and a proposition p to a set of worlds in which p is true. For current purposes, we can assume that the value of a selection function is a non-empty set of worlds. Any selection function S also minimally satisfies the properties in (26) for any w and proposition p (construed as a set of worlds).[15]

(26)
a. $S(w,p) \subseteq p$
b. If $w \in p$, then $S(w,p) = \{w\}$
c. If p entails q and $p \cap S(w,q) \neq \emptyset$, then $S(w,p) = \{w' \in S(w,q) \mid w' \in p\}$

On the traditional Stalnaker-Lewis semantics, the truth-conditional content of a subjunctive conditional *if* ϕ, *would* ψ, relative to a selection function S and a context c, is as in (27).[16]

(27) $[\![\textit{if } \phi, \textit{would } \psi]\!]^S_c = \{w \in W \mid S(w, [\![\phi]\!]_c) \subseteq [\![\psi]\!]_c\}$

The semantics in (27) may be adequate for conditionals whose antecedents bear no presuppositions, but it will not do for conditionals with presupposition-bearing antecedents. According to (27), the proposition that is given as an argument to the selection function corresponds to the (unrestricted) truth-conditional content of the antecedent. Below I show that the presuppositional content of the antecedent affects the content of a conditional and amend the semantics in (27) to incorporate the effect of presuppositions in the hypothetical assumption made by the antecedent.

As is well-known, the presuppositional implications of the antecedent of a conditional "project" and become presuppositions of the entire conditional (Karttunen 1974; Heim 1992). Since the temporal adverbials *still* and *already* are associated with presuppositions, the antecedents of the conditionals in (28) and (29), as well as the conditionals themselves, carry the presuppositions triggered by the temporal adverbials.

(28) If he were still in the cave, we would be getting worried.

15 As Lewis (1973: 58–59) shows, these conditions characterize the selection functions "derived from (centered) systems of spheres satisfying the Limit Assumption".

16 I drop the parameter S for $[\![\phi]\!]$ and $[\![\psi]\!]$. If ϕ and ψ do not contain a counterfactual, S does not play a role in their interpretation.

(29) If he were already in the cave, they would have to rush.

An utterance is felicitous only relative to contexts that satisfy (entail) its presuppositions. (28) and (29) are thus felicitous only in contexts which satisfy their respective presuppositions, entailing minimally, for (28), that he was in the cave earlier, and for (29), that he was not in the cave earlier.[17]

The presuppositional content of the antecedent not only imposes a felicity condition on the context in which the counterfactual is uttered, but it also affects the content of the counterfactual assumption.[18] Compare (28) and (29) with (30) and (31).

(30) If he were in the cave (now), we would be getting worried.

(31) If he were in the cave (now), they would have to rush.

Although the truth-conditional content of the antecedents of all four conditionals is the same, the kind of counterfactual assumption (28) is making is different from that of (30), and, similarly, the one (29) is making is different from that of (31).[19] The counterfactuals in (30) and (31) do not necessarily express the same content as the corresponding ones in (28) and (29), even if they are uttered in the same context (e.g., in a context in which it is taken for granted that he was earlier in the cave but

17 As discussed in section 4, the full presupposition requires that the reference time and the alternative times be within periods of transition with respect to the property of him being in the cave.

18 Heim (1992: 205), in her investigation of presupposition projection from antecedents of conditionals, makes the same observation about the effect of presuppositions on the content of counterfactuals: "Let me close this excursion with a remark on the effect of presuppositional requirements in the antecedent of a counterfactual's truth conditions. Recall the context where Mary is presupposed to be in the phone booth. We noted above that an indicative *if*-clause like *If John is in the phone booth* ... in this context amounts to the supposition that both John and Mary are in the booth. This is otherwise for a minimally different subjunctive *if*-clause: If we say *If John WERE in the phone booth*, then it depends on the actual facts and the selection function whether the hypothetical situations under consideration have both people in the booth or have John there instead of Mary. ... *then Mary would be outside* is a felicitous and possibly true continuation. (As opposed to the deviant indicative variant *If John is in the phone booth, then Mary is outside.* This is acceptable only if we are ready to conclude that Mary's being in the phone booth isn't presupposed after all.) This difference, of course, is predicted by [the context change potential of counterfactual conditionals]. But what is also predicted is that if we add to the subjunctive antecedent a *too*, as in *If John were in the phone booth too* ..., then the meaning is in a certain respect more like that of the indicative again: no matter what the selection function and facts of the world, we only get to consider hypothetical worlds with both people in the booth together. So *If John were in the phone booth TOO, then Mary would be outside* is also deviant".

19 *Still* and *already* restrict the temporal reference of the antecedent to a (contextually determined) reference time, which obviously affects the antecedent's truth-conditonal content. In order to control for that, I have made the temporal reference of the antecedents of (28) – (31) be uniformly the time of utterance (marked by the use of *were* as opposed to *had been*).

no longer, or in a context in which it is taken for granted that he was not in the cave earlier, but that he would be eventually). For (30) or (31), the selection function may just pick worlds in which he is now in the cave, regardless of whether he was there earlier or not, or whether he would be there later on or not. By contrast, for (28) the selection function must pick worlds in which he was in the cave earlier and continues to be in the cave now, and similarly for (29) it must pick worlds in which he was not in the cave earlier but is in the cave now. So even if interpreted with respect to the same selection function S and context c, the interpretations of (28) and (30), or of (29) and (31), can differ. We can conclude that the proposition that is given as an argument to the relevant selection function is different for (28) and (30), and for (29) and (31).[20] If a context c satisfies (entails) the presuppositions of the antecedent ϕ of a counterfactual conditional, it encodes the necessary information, but since it entails $\neg\phi$, we cannot have it directly restrict the proposition that corresponds to the truth-conditional content of the antecedent (since $c \cap [\![\phi]\!]_c = \emptyset$).

To take into account the effect of presuppositions on the hypothetical assumption made by counterfactual conditionals, I use Heim's (1992) notion of a revised context and have the proposition expressed by the antecedent be contextually restricted by the revised context. $rev(c,\phi)$ as defined in (32) designates a particular kind of revision of c that allows for consistency with ϕ. It is the maximal (i.e., most inclusive and hence least informative) revision of c that still satisfies the presuppositions of ϕ.

(32) $rev(c,\phi) = \cup\{X \subseteq W \mid c \subseteq X \text{ and } \phi \text{ is felicitous relative to } X\}$

If c entails $\neg\phi$, moving to $rev(c,\phi)$ opens up the possibility of ϕ. As Heim shows, $rev(c,\phi)$ satisfies the presuppositions of ϕ if and only if c does. If ϕ has no presuppositions, then $rev(c,\phi)$ would just be W.

We can now revise (27) to (33).[21]

(33) $[\![\mathit{if}\ \phi, \mathit{would}\ \psi]\!]_c^S = \{w \in c \mid S(w, rev(c,\phi) \cap [\![\phi]\!]_c) \subseteq [\![\psi]\!]_c\}$

The semantics in (33) is more radically context-dependent than the one in (27), as the proposition expressed by a conditional uttered in a context c is a subset of c's context set, rather than a subset of the set of all possible worlds W.

Let us see more specifically what the revised context would be like for (28) and (29). Recall that the presupposition we have assigned to *still* and *already* is that the reference time and the alternative times are located within periods of transition with

20 This is not the only conceivable conclusion. An alternative would be to have presuppositions constrain the selection function chosen in a given context. Thanks to Lee Walters for discussion on this point.

21 To take into account presuppositions of the consequent the semantics would have to be further revised but I will ignore this here.

respect to the property of times expressed by the clausal predicate. Let us designate the property of times expressed by the clausal predicate in (28)–(31) as $\mathsf{he\text{-}in\text{-}cave}$. While for the presuppositionless (30) and (31) uttered in a context c with reference time t_0 the proposition that is given as an argument to the relevant selection function S can be $[\![\mathsf{he\text{-}in\text{-}cave}(t_0)]\!]_c$, it has to be a more restricted proposition for (28) and (29) so as to encode the presuppositional content.

(28) would be felicitous relative to context c only if c entails that the reference time t_0 and the alternative times, which are all times $t \prec t_0$, are located within a period of transition with respect to $\mathsf{he\text{-}in\text{-}cave}$, with $I_{\text{POS}} \prec I_{\text{NEG}}$. In other words, c has to entail that he exits the cave at some point within the relevant period. Given that c entails that he is not in the cave at the reference time, then c is comprised solely of worlds in which the reference time t_0 is within I_{NEG}. For any alternative time t, c allows for possibilities in which t is within I_{NEG} and possibilities in which t is within I_{POS}. $rev(c, Still(\mathsf{he\text{-}in\text{-}cave}(t_0)))$ opens up the possibility for $\mathsf{he\text{-}in\text{-}cave}(t_0)$, so it includes in addition worlds in which the reference time t_0 is within I_{POS}. Those will be worlds in which all the alternative times are within I_{POS}.

Similarly, (29) would be felicitous relative to context c only if c entails that the reference time t_0 and the alternative times, which are times $t_0 \prec t$, are located within a period of transition with respect to $\mathsf{he\text{-}in\text{-}cave}$, with $I_{\text{NEG}} \prec I_{\text{POS}}$. In other words, c has to entail that he enters the cave at some point within the relevant period. c is comprised solely of worlds in which the reference time t_0 is within I_{NEG}. For any alternative time t, c allows for possibilities in which t is within I_{NEG} and possibilities in which t is within I_{POS}. $rev(c, Already(\mathsf{he\text{-}in\text{-}cave}(t_0)))$ opens up the possibility for $\mathsf{he\text{-}in\text{-}cave}(t_0)$, so it includes in addition worlds in which the reference time t_0 and consequently all the alternative times are within I_{POS}.

The respective propositions that are given as arguments to the selection function are as in (34).

(34) a. $rev(c, Still(\mathsf{he\text{-}in\text{-}cave}(t_0))) \cap [\![\mathsf{he\text{-}in\text{-}cave}(t_0)]\!]_c$
b. $rev(c, Already(\mathsf{he\text{-}in\text{-}cave}(t_0))) \cap [\![\mathsf{he\text{-}in\text{-}cave}(t_0)]\!]_c$

$rev(c, Still(\mathsf{he\text{-}in\text{-}cave}(t_0)))$ preserves the information in c that the reference time t_0 and the alternative times are within a period of transition, and similarly for $rev(c, Already(\mathsf{he\text{-}in\text{-}cave}(t_0)))$. What changes is the location of the transition between I_{POS} and I_{NEG} (for *still*), or between I_{NEG} and I_{POS} (for *already*), and, therefore, whether t_0 and the alternative times fall within the positive phase or the negative phase of the relevant period. For instance, for (28) the information that he exits the cave is preserved; what changes is the time of his exit from the cave, so as to allow for the reference time t_0 to be a time at which he is in the cave. For (29) the information that he enters the cave is preserved; what changes is the time of his entering the cave, so as to allow for the reference time t_0 to be a time at which he is in the cave.

For (30) and (31), by contrast, the propositions given as arguments to the selection function are less restrictive, so relative to the same world its values can be outside of $rev(c, Still(\text{he-in-cave}(t_0)))$ or of $rev(c, Already(\text{he-in-cave}(t_0)))$.

5.2 Antecedents with positive polarity

Let us first consider the role of the alternatives for counterfactuals which, like (28) and (29), have antecedents with positive polarity.[22] Krifka (1995) and Chierchia (2006) propose that alternatives need not be exploited only at the level of the matrix sentence. This implies that within a complex expression the relative strength of the plain content vis-à-vis the alternatives can be assessed in an embedded position. In conditionals, specifically, this can happen at the level of the antecedent.

The set of alternatives for the antecedent of (28) is given in (35-a), and that for the antecedent of (29) in (35-b), where t_0 is the reference time.

(35) a. $Alt(Still(\text{he-in-cave}(t_0))) = \{\text{he-in-cave}(t) \mid t \prec t_0\}$
b. $Alt(Already(\text{he-in-cave}(t_0))) = \{\text{he-in-cave}(t) \mid t_0 \prec t\}$

As we saw in the previous section, the propositional argument of the selection function is a more restricted proposition than the proposition corresponding to the truth-conditional content of the antecedent. Therefore, in checking the relative strength of the plain content and the content of the alternatives we need to look at propositions restricted by the revised context. The sets of alternative propositions corresponding to (35) are as in (36).

(36) a. $\{rev(c, Still(\text{he-in-cave}(t_0))) \cap [\![\text{he-in-cave}(t)]\!]_c \mid t \prec t_0\}$
b. $\{rev(c, Already(\text{he-in-cave}(t_0))) \cap [\![\text{he-in-cave}(t)]\!]_c \mid t_0 \prec t\}$

None of the alternatives in (35-a) yields a more restricted proposition than (34-a), and similarly, none of the alternatives in (35-b) yields a more restricted proposition than (34-b). In fact, the proposition in (34-a) is stronger than each proposition in (36-a), and the proposition in (34-b) is stronger than each proposition in (36-b).[23] Therefore, making the counterfactual assumption that he is in the cave at t_0 is the strongest relevant assumption that can be made.[24] Consequently, the PPIs *still* and

22 In the brief discussion in section 4 on the polarity sensitivity of the adverbials, we considered contexts which satisfied their presuppositions and which were compatible with the plain content and the alternatives. When we move to antecedents of counterfactuals, the latter property, of course, no longer holds.

23 On the reasonable assumption that the alternatives do not carry presuppositions of their own.

24 Given condition (26-c) on selection functions, this relation of strength gets reversed for the entire

already are acceptable in antecedents with positive polarity.

5.3 Antecedents with polarity reversal

When the antecedent of a subjunctive conditional exhibits polarity reversal, as in (5-a) and (5-b), repeated in (37), the relative assessment of strength between plain and alternative content can obviously not be done at the level of the antecedent, since negation reverses the strength relations.[25] Therefore, it has to be done at the level of the matrix sentence.

(37) a. If he had not already arrived, we would have postponed the meeting.
b. If he were not still at home, we would have missed him.

We thus have to establish that assertions like (37-a) and (37-b) and, more generally, assertions of the schematic form in (38-a) and (38-b) constitute scalar assertions.

(38) a. $\textit{if}\ \neg(\textit{Already}(\phi)),\ \textit{would}\ \psi$
b. $\textit{if}\ \neg(\textit{Still}(\phi)),\textit{would}\ \psi$

This means that the alternatives are themselves conditionals and we have to assess their relative strength vis à vis the plain content of the conditional uttered. The plain content of the conditionals in (38) is that of the conditional in (39-a) and their set of alternatives are as in (39-b).

(39) a. $\textit{if}\ \neg\phi,\ \textit{would}\ \psi$
b. $\{\textit{if}\ \neg\phi',\ \textit{would}\ \psi \mid \phi' \in \textit{Alt}(\phi)\}$

As discussed earlier, the content of both $\textit{Already}(\phi)$ and $\textit{Still}(\phi)$ is that the reference time t_0 is within the positive phase I_{POS} of the period of transition I determined by ϕ. The content of each alternative ϕ' based on alternative time t' is that t' is within the positive phase I_{POS} of I, where for $\textit{Already}(\phi)$ every $t' \in \textit{Alt}(t_0)$ is such that $t_0 \prec t'$, and for $\textit{Still}(\phi)$ every $t' \in \textit{Alt}(t_0)$ is such that $t' \prec t_0$.

Let us focus attention here on uses of (38-a) and (38-b) in contexts which are incompatible with the antecedent, that is contexts which entail ϕ.[26] In such cases the conditional corresponding to the plain content and the conditionals in the set of alternatives are all counterfactual, since ϕ contextually entails ϕ' for each $\phi' \in \textit{Alt}(\phi)$.[27] For example, when (37-a) is interpreted counterfactually, the context

conditional.

25 This mirrors the contrast between (1-a), (2-a) and (1-b), (2-b).

26 In the next section we consider non-counterfactual interpretations of subjunctive conditionals and why they do not allow polarity reversal.

27 Conversely, $\neg\phi'$ contextually entails $\neg\phi$, for any $\phi' \in \textit{Alt}(\phi)$.

entails that he had arrived by the reference time, and, therefore, that he had arrived by any time later than the reference time.

Let us abbreviate the revision of context c with $\neg(Already(\phi))$, or with $\neg(Still(\phi))$, as c_rev.[28] Given the semantics for subjunctive conditionals in (33), for the plain content to be true in a world $w \in c$, it has to be the case that $S(w, c\_rev \cap [\![\neg\phi]\!]_c) \subseteq [\![\psi]\!]_c$. Similarly, for any alternative to be true in w, it has to be the case that $S(w, c\_rev \cap [\![\neg\phi']\!]_c) \subseteq [\![\psi]\!]_c$. Now, c_rev includes, for any $\phi' \in Alt(\phi)$, both possibilities in which $\neg\phi$ and ϕ' hold and possibilities in which $\neg\phi$ and $\neg\phi'$ hold. For instance, the c_rev corresponding to (37-a) includes worlds in which he has not arrived by the reference time t_0 but has arrived by some later time t, as well as worlds in which he has not arrived either by the reference time t_0 or by the later time t. The c_rev corresponding to (37-b) includes worlds in which he is not at home at the reference time t_0 but is at home at some earlier time t, as well as worlds in which he is at home neither at the reference time t_0 nor at the earlier time t. The fact that c_rev contains possibilities of both types guarantees that the propositions $c\_rev \cap [\![\neg\phi']\!]_c$ given as arguments to the selection function are not inconsistent. The question is, are the selection functions relative to which the conditionals are interpreted subject to any constraints such that the sets in (40) are related in a systematic way for every $w \in c$ and every $\phi' \in Alt(\phi)$?

(40) a. $S(w, c\_rev \cap [\![\neg\phi]\!]_c)$
b. $S(w, c\_rev \cap [\![\neg\phi']\!]_c)$

Given that the way the worlds are selected is based on the facts of w, the question could be formulated as follows: should the fact of a world $w \in c$[29] that some alternative ϕ' holds be preserved under the hypothetical assumption that $\neg\phi$? $c\_rev \cap [\![\neg\phi]\!]_c$ is consistent with the fact that ϕ', since making the hypothetical assumption that the reference time is within the negative phase is compatible with an alternative time remaining within the positive phase of the period of transition determined by ϕ. If the worlds delivered by the selection function in (40-a) are exclusively $\neg\phi$, ϕ' worlds, then the sets in (40) would be disjoint and no connection could be established between the plain and the alternative content of counterfactuals with polarity reversal as in (38). But the fact that ϕ', though consistent with the counterfactual assumption, may be given up and, in fact, may have to be given up.

Lewis (1979), in his explication of overall similarity, postulated that similarity of particular fact is of least or no importance. Analyses of counterfactuals based on premise semantics, such as Kratzer (1981, 1989) and Veltman (2005), have

28 Thus, depending on the particular case under consideration, c_rev is either $rev(c, \neg(Already(\phi))$ or $rev(c, \neg(Still(\phi))$.

29 I am using 'fact of a world w' in the familiar way to refer to a proposition true in w.

emphasized the role that dependence between facts plays in counterfactual reasoning. As Veltman (2005: 164) puts it: "similarity of particular fact is important, but only for facts that do not depend on other facts. Facts stand and fall together. In making a counterfactual assumption, we are prepared to give up everything that depends on something that we must give up to maintain consistency".

Now, for any given ϕ' and $w \in c$, the fact of w that ϕ and the fact that ϕ' are not independent from one another. Given what is presupposed, once ϕ is true throughout the worlds of c, so is any alternative ϕ'. These two types of facts are dependent, though not dependent due to a law-like generalization, or due to some causal connection, or due to a feature of the world that matters for overall similarity independently of the information in the context in which the conditional is used.[30] That ϕ' is true in w, for any $\phi' \in Alt(\phi)$, is a fact of w but it is true based on the fact that the transition from a positive to a negative phase (for *still*), or from a negative to a positive phase (for *already*), occurred when it did and the presupposition of a unique relevant transition within the period of transition. When we make the counterfactual assumption that the reference time is within the negative phase, this necessarily changes the time of the transition and it opens up the possibility that it could have occurred anywhere within the period of transition, including after the latest, or before the earliest, alternative time.

Therefore, the selection function can be assumed to be inclusive and not to distinguish between worlds in which the phase transition occurs after t_0 but before an alternative time t (these would be $\neg\phi$, ϕ'_t worlds) and worlds in which the phase transition occurs after t (these would be $\neg\phi$, $\neg\phi'_t$ worlds) in terms of their similarity to w. Given the constraint on selection functions in (26-c) and the fact that the proposition $c\_rev \cap [\![\neg\phi']\!]_c$ entails the proposition $c\_rev \cap [\![\neg\phi]\!]_c$, it follows that (41) holds.

(41) $$S(w, c\_rev \cap [\![\neg\phi']\!]_c) \subseteq S(w, c\_rev \cap [\![\neg\phi]\!]_c)$$

With this property of the selection function, the plain content contextually entails each element in the set of alternatives, and therefore (38-a) and (38-b) constitute scalar assertions.

The more inclusive the value of the selection function, the less information it encodes about similarity of worlds. Polarity reversal is possible for counterfactuals uttered in a context with no information that would discriminate between those worlds in which the phase transition, if it hasn't occurred by t_0, would occur earlier rather than later. That is arguably part of the presupposition carried by the polarity adverbials—not just the factual constraints on the context but also lack of conditional

30 This is arguably structurally analogous to Tichý's (1976) 'man with a hat' case, except for the different source of the dependence.

information about alternative timing of the transition within the period of transition, encoded via inclusive selection functions. Even if the internal structure of a world forces his being here sooner rather than later, or his staying here longer rather than shorter, use of the adverbials communicates and forces a kind of agnosticism about the time of the transition.

5.4 The role of counterfactuality

As discussed in section 2, Baker observed that polarity reversal in subjunctive conditionals is accompanied by a counterfactual implication. Baker's generalization, reflected in his analysis, was that conditionals of the form in (38) are acceptable *if and only if* the conditionals presuppose $\mathit{Already}(\phi)$ and $\mathit{Still}(\phi)$, respectively.[31] Schwartz & Bhatt took that generalization for granted and postulated the existence of a special negation that presupposes its prejacent, in the case of the conditionals of the form in (38), $\mathit{Already}(\phi)$ and $\mathit{Still}(\phi)$, respectively.[32] This presupposition originating in the antecedent projects to become a presupposition of the conditional. So conditionals of the type in (38) are predicted by both analyses to be acceptable only if uttered in contexts that entail $\mathit{Already}(\phi)$ or $\mathit{Still}(\phi)$, that is only in contexts satisfying the presuppositions of *Already* or *Still* and entailing ϕ.

But this turns out to be the wrong generalization. A conditional with polarity reversal in the antecedent can be counterfactual without the prejacent of the negation being entailed by the context. A case in point are conditionals with conjoined antecedents where the prejacent of the negation is not entailed by the context in which the conditional is used. PPIs in the scope of negation are acceptable as long as the other conjunct renders the conditional counterfactual. (42-a) is an instance where the negated part of the antecedent is *consistent* with the context in which the counterfactual is uttered, given the information conveyed by the first sentence. (42-b) is an instance where the negated conjunct is *entailed* by the context in which the counterfactual is uttered, given that the first sentence conveys (possibly even entails) that John is not here.[33]

---

31 Since Baker relied on a semantic notion of presupposition (see his note 2 on page 15), he must have assumed that subjunctive conditionals are ambiguous between those that presuppose the falsity of their antecedent and those that do not.

32 The conditionals have the logical form in (i), where $\neg_L$ is the special light negation.

(i) a. $\mathit{if}\ \neg_L(\mathit{Already}(\phi)),\ \mathit{would}\ \psi$
    b. $\mathit{if}\ \neg_L(\mathit{Still}(\phi)), \mathit{would}\ \psi$

33 In this respect, the examples in (42) are different from the cases of conjoined antecedents considered by Baker (1970b) and Karttunen (1971). In their examples, the negated conjunct with the PPI carried a presupposition triggered by an expression other than the PPI and satisfied by the first conjunct,

(42) a. I am not sure if John has shown up but, luckily, we can do without him. If, however, he had said he'd come help us and were not already here, I'd be annoyed with him.

b. I don't mind that John hasn't shown up. If, however, he had said he'd come early and were not already here, I'd be annoyed with him.

The implication accompanying (42-a) is that John did not say that he'd come help us and that accompanying (42-b) that John did not say that he'd come early. So the conditionals are counterfactual, albeit not because of the negated conjunct containing the PPI.

The correct generalization then is that a PPI under the scope of negation is acceptable in the antecedent of a subjunctive conditional as long as the entire antecedent is counterfactual. In the case of conjunction it suffices for one of the two conjuncts to be incompatible with the context. In the examples in (42) the first conjunct's being counterfactual suffices for the full antecedent to be incompatible with the context of utterance and hence for the PPI in the scope of negation in the second conjunct to be 'rescued'. This suggests that the 'rescuing' of PPIs within negated antecedents of counterfactuals does not so much require that the prejacent of negation is presupposed to be true, but rather that the conditional itself is interpreted against a context such that the relevant possibilities in which the antecedent is true—what the selection function delivers— are wholly outside the context.

The hypothetical assumptions made in (42) open up the issue of the timing of the transition and allow for possibilities in which the transition occurs anywhere within the relevant period of transition. As with the standard cases discussed in section 5.3, for any world w in the context, the selection function does not distinguish between worlds in which the first conjunct is true and the transition occurs after the reference time t_0 but before an alternative time t and worlds in which the first conjunct is true and the transition occurs after both t_0 and t, regardless of the particular facts of the timing of the transition in w. Therefore, the selection function satisfies (41) and the conditionals in (42) can constitute scalar assertions.

As far as the question of why a subjunctive conditional needs to be counterfactual in order to be able to rescue PPIs, in the analysis pursued here, the answer will be found in how asserting the conditional in a context which is compatible with its antecedent influences the relation between the plain content of the subjunctive conditional and the content of the alternatives. Below I sketch out an argument that this is what happens in the two types of cases discussed in section 2.

itself incompatible with the context, e.g., *If John had a brother and his bother had not come to his help already*, where *his brother* presupposes that John had a brother. Unlike (42), where the negated conjunct is consistent with or even entailed by the context, in Baker's and Karttunen's examples the negated conjunct is simply infelicitous relative to the context.

First let us see how the form of a conditional affects its interpretation. Stalnaker (1975) assumes that the information of a context in which a conditional is asserted restricts the selection functions used in its interpretation, stating:[34]

> [I]f the conditional is being evaluated at a world in the context set, then the world selected must, if possible, be within the context set as well . . . In other words, all worlds within the context set are closer to each other than any worlds outside it. The idea is that when a speaker says If A, then everything he is presupposing to hold in the actual situation is presupposed to hold in the hypothetical situation in which A is true. (p. 275–276)

This amounts to a pragmatic constraint that all else being equal, any conditional χ asserted in context c has to be interpreted relative to selection functions S satisfying (43).

(43) For any $w \in c$, $S(w, p) \subseteq c$, where p is the proposition determined by the antecedent of χ and by c as per the semantics of χ.[35]

Stalnaker (1975), moreover, took the subjunctive marking in English conditionals to be a conventional means for signaling "that the selection function is one that may reach outside of the context set" (p. 276).[36] This means that if a speaker wants to signal that (43) is not in force in the intended interpretation of his assertion of a conditional, he has to use the subjunctive form of the conditional, *if* ϕ, *would* ψ. When ϕ is inconsistent with c, any selection function that satisfies (26-a) cannot simultaneously satisfy (43). In that case, therefore, any selection function S would be such that, for any $w \in c$, its value is fully outside the context set, i.e., $S(w, rev(c,\phi) \cap [\![\phi]\!]_c) \subseteq rev(c,\phi) \setminus c$. When ϕ is consistent with c, the worlds in c can vary in whether they verify ϕ. In those worlds $w \in c$ in which ϕ is true, the selection function would return $\{w\}$ as its value, given the centering condition in (26-b), and the conditional would be true only if ψ is true in w as well. So it is among the non-ϕ-worlds within c that the selection function can reach outside of c.

34 Unlike (27) and (33), on Stalnaker's semantics, selection functions pick a unique world rather than a set of worlds. Given the semantics in (33), a conditional asserted in a given context can only be true in worlds within its context set.

35 As we have seen already, the context not only helps determine the truth-conditional content of the antecedent of a conditional but also the proposition that is given as an argument to the selection function. As specified in (33), for subjunctive conditionals with antecedent ϕ, the proposition that is given as an argument to the selection function is $rev(c,\phi) \cap [\![\phi]\!]_c$. As we will see in section 5.5, for indicative conditionals it is $c \cap [\![\phi]\!]_c$.

36 See also von Fintel (1998).

(44) and (45) exemplify the two types of non-counterfactual uses of subjunctive conditionals and the corresponding contrasts with antecedents with polarity reversal (repeated from (10), (12) and (11), (13) in section 2).

(44) a. If he had (still) been in his office then, the lights would have been on. The lights were indeed on. Therefore, he was (still) in his office then.
b. # If he hadn't still been in his office then, the lights would have been off. The lights were indeed off. Therefore, he was not in his office then.

(45) a. If he had (still) been in his office then, the lights would have been on. The lights were off. Therefore, he was not in his office then.
b. # If he hadn't still been in his office then, the lights would have been off. The lights were on. Therefore, he was (still) in his office then.

Let us first consider what happens in the admissible cases in (44-a) and (45-a). In the first case the conditional is used to argue for the truth of the antecedent after the truth of the consequent is established. In the second case, the conditional is used to argue for the falsity of the antecedent after the falsity of the consequent is established. In both cases, we can assume that the initial context is compatible with both the antecedent ϕ and the consequent ψ of the conditional. In both cases, only those ϕ-worlds in the original context which are also ψ-worlds survive the update with the conditional.[37] In the first case, when the second premise comes in, any non-ψ-worlds (which would have to be non-ϕ-worlds) are removed. Finally, the conclusion, by stating that all remaining worlds are ϕ-worlds, implies that all the non-ϕ-worlds in the original context were non-ψ-worlds. In the second case, when the second premise comes in, all ψ-worlds (some of which could be ϕ-worlds) are removed. Finally, the conclusion, by stating that all remaining worlds are non-ϕ-worlds, implies that there were non-ϕ-non-ψ-worlds in the original context.[38]

In order to explain why polarity reversal is not possible in subjunctive conditionals that are not counterfactual, we have to show that the plain content, $\textit{if}\ \neg\phi, \textit{would}\ \psi$, does not necessarily contextually entail $\textit{if}\ \neg\phi', \textit{would}\ \psi$ for every $\phi' \in Alt(\phi)$. Let us abbreviate the proposition $rev(c,\chi) \cap [\![\chi]\!]_c$ as $\mathcal{A}(c,\chi)$. Consider a context c which does not settle when the transition happened[39] and with some $w \in c$ such that $w \in [\![\neg\phi]\!]_c$, $w \in [\![\psi]\!]_c$ and $w \notin [\![\neg\phi']\!]_c$ for some $\phi' \in Alt(\phi)$. This means that in w one of the alternative times is within the positive phase of the period of transition

37 If there are no ϕ-ψ-worlds in the original context, the conditional is false at all the ϕ-worlds in that context, which would defeat the purpose of the first type of argument.

38 Both types of argument thus establish a non-accidental connection between ϕ and ψ and between $\neg\phi$ and $\neg\psi$, e.g., in the examples above between his being in/away from the office and the lights being on/off.

39 So the $\neg\phi$-worlds in c are comprised of ϕ'-worlds and $\neg\phi'$-worlds, for every $\phi' \in Alt(\phi)$.

determined by ϕ. For such worlds w, $S(w,\mathcal{A}(c,\neg\phi)) = \{w\}$ but $w \notin S(w,\mathcal{A}(c,\neg\phi'))$. So not only do the two sets not stand in the subset relation, they are disjoint. This disrupts the informational ordering between the plain semantic value of the conditional and the semantic value of at least one of its alternatives. The plain content is true in w but the value of $S(w,\mathcal{A}(c,\neg\phi'))$ may well be constituted of worlds which are not uniformly ψ-worlds. For instance, to the extent that the conditional is informative relative to c due to there being ϕ-worlds w' in c in which the conditional is false – which means that $S(w',\mathcal{A}(c,\neg\phi'))$ is not comprised of worlds which are uniformly ψ-worlds – the same could be the case of the alternative conditional *if* $\neg\phi'$, *would* ψ at worlds like w.

5.5 Indicative conditionals

If polarity reversal required that the prejacent of the negation be entailed by the context, as it does on Baker's and Schwarz & Bhatt's proposals, then polarity reversal would be completely excluded from indicative conditionals, which must be used in contexts compatible with their antecedent. No context can simultaneously satisfy both conditions, and hence it would be predicted that polarity reversal cannot be observed in the antecedents of indicative conditionals. However, polarity reversal can be observed with indicative conditionals provided the context is of a certain kind. The indicative conditionals with polarity reversal in (46-a) and (47-a) are acceptable and give rise to a kind of uniformity implication, spelled out in (46-b) and (47-b), respectively.

(46) a. If he was not already in his office when they got there, they left without him.
b. If he was not in his office when they got there, they left without him, regardless of when he came to the office.

(47) a. If he was not still in town during their meeting, they did not see him.
b. If he was not in town during their meeting, they did not see him, regardless of when he left town.

The reason polarity reversal is not *generally* acceptable in indicative conditionals is structurally equivalent to the reason polarity reversal is not acceptable with subjunctive conditionals when their antecedent is consistent with the context. Before showing this, we need to fix the semantics of the indicative conditional. Following Stalnaker's (1975) uniform treatment of subjunctive and indicative conditionals, I will assume that an indicative conditional *if* ϕ, ψ uttered in context c has the semantics in (48).

(48) $[\![\textit{if}\ \phi, \psi]\!]^S_c = \{w \in c \mid S(w, c \cap [\![\phi]\!]_c) \subseteq [\![\psi]\!]_c\}$

As with the semantics proposed for subjunctive conditionals in (33), the proposition expressed by an indicative conditional is a subset of the context's context set and the proposition that is given as an argument to the selection function encodes the presuppositions of the antecedent, in addition to its truth-conditional content.

The constraint in (43) amounts to (49) for an indicative conditional. (49) constrains the selection functions involved in (48) and plays a role in determining the plain and alternative content of conditionals like (46-a) and (47-a).

(49) For any $w \in c$, $S(w, c \cap [\![\phi]\!]_c) \subseteq c$.

To see why polarity reversal is generally disallowed in indicative conditionals, we need to show that the relation of informational ordering between the plain content and the content of the alternatives is broken under polarity reversal. Adapting our notation from the previous section, let us now use $\mathcal{A}(c,\chi)$ to abbreviate the proposition $c \cap [\![\chi]\!]_c$. For those $w \in c$ in which χ is true, $S(w, \mathcal{A}(c,\chi)) = \{w\}$, given the centering condition in (26-b). Suppose the antecedent exhibits polarity reversal, as in (46-a) and (47-a). Consider a context c which does not settle when the transition happens and with some $w \in c$ such that $w \in [\![\neg\phi]\!]_c$, $w \in [\![\psi]\!]_c$, $w \notin [\![\neg\phi']\!]_c$ for some $\phi' \in Alt(\phi)$. Then $S(w, \mathcal{A}(c,\neg\phi)) = \{w\}$ but $w \notin S(w, \mathcal{A}(c,\neg\phi'))$. As before, this breaks the subset relation between the values of S for the plain value of the antecedent and one of its alternatives, which in turn disrupts the informational ordering between the plain semantic value for the conditional and its alternative semantic values. To the extent that the conditional is informative relative to c, say due to there being a $\neg\phi$-$\neg\psi$-world in c, such a world may well be in $S(w, \mathcal{A}(c,\neg\phi'))$, which, by (49), is a subset of c.

Now suppose we have a context c favoring selection functions S satisfying the condition in (50).

(50) For any $w \in c$ and any $\phi' \in Alt(\phi)$, $S(w, \mathcal{A}(c,\neg\phi')) \subseteq S(w, \mathcal{A}(c,\neg\phi))$

This ensures that any antecedent worlds ($\neg\phi$-worlds) in the context behave uniformly across the alternatives. This means that the conditional supports limited strengthening of the antecedent inferences: when *if* ϕ, ψ is true, the conditionals with strengthened antecedents *if* ϕ *and* ϕ', ψ and *if* ϕ *and* $\neg\phi', \psi$ are also true, for any $\phi' \in Alt(\phi)$. For instance, when (46-a) is true, so is (51), which is an instantiation of the uniformity implication of (46-a).

(51) If he was not in his office when they got there and only came later, they left without him.

In such contexts, use of (46-a) or (47-a) would constitute a scalar assertion. For the countercase we constructed above, even though $w \notin S(w, \mathcal{A}(c,\neg\phi'))$, any world in

$S(w, \mathcal{A}(c, \neg\phi'))$ would be a ψ world, just like w is.

Conditionals like (46-a) and (47-a) convey that there is a bound on how late or how early the transition can occur and still have the consequent be true. Those features of a context that support a use of (46-a) as a scalar assertion, including the kinds of selection functions relevant for the interpretation of (46-a), would also verify the conditional in (51).

5.6 Polarity reversal and types of conditionals

We may now wonder whether non-counterfactual subjunctive conditionals would not be fine in the same contexts as indicatives are. A conditional like (46-a) can still be informative relative to contexts that satisfy the condition in (50), as long as the context contains some $\neg\phi$-$\neg\psi$-worlds. On the other hand, non-counterfactual subjunctive conditionals exhibiting polarity reversal would be uninformative relative to any context in which they would constitute a scalar assertion. Take, for example, the subjunctive conditionals in (44-b) and (45-b) and suppose there is a context that encodes the following regularity: no lights throughout the negative phase of John's not being in the office, lights throughout the positive phase of John's being in the office. In contrast to the case of indicative conditionals, in any context in which the subjunctive conditional exhibiting polarity reversal would constitute a scalar assertion, there can be no John-not-in-the-office and lights-on worlds (i.e., worlds where the antecedent holds but the consequent does not) given the assumption that the context encodes the regularity. Similarly, the selection function would deliver for any John-not-in-the-office world only worlds in which the consequent is true. So the conditional is not informative relative to such a context. But in the kinds of reasoning non-counterfactual subjunctive conditionals are employed for, the conditional premise has to be informative.

Thus there is an interesting asymmetry between indicative and non-counterfactual subjunctive conditionals, just as there is an asymmetry between counterfactuals and all other conditionals with respect to polarity reversal. Counterfactuals allow polarity reversal as a matter of course because the use of the conditional constitutes a scalar assertion in any context in which it is felicitous, and this is so because contextual entailments must be given up once the information that gives rise to them is revised. Indicative conditionals allow polarity reversal only relative to contexts with additional information that would guarantee that the conditional constitutes a scalar assertion. By contrast any context in which non-counterfactual subjunctive conditionals would constitute scalar assertions are contexts that defeat the purpose of their being used.

References

Anderson, Alan Ross. 1951. A note on subjunctive and counterfactual conditionals. *Analysis* 12(2). 35–38.

Baker, C. Lee. 1970a. Double Negatives. *Linguistic Inquiry* 1. 169–186.

Baker, C. Lee. 1970b. Problems of polarity in counterfactuals. In Jerrold Sadock & Anthony Vanek (eds.), *Studies Presented to Robert B. Lees by his Students*, 1–15. Edmonton: PIL Monograph Series 1, Linguistic Research Inc.

Chierchia, Gennaro. 2006. Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry* 37(4). 535–590.

Edgington, Dorothy. 1995. On conditionals. *Mind* 104(414). 235–329.

Edgington, Dorothy. 2008. Counterfactuals. *Proceedings of the Aristotelian Society* 108(1, part 1). 1–21.

von Fintel, Kai. 1998. The presupposition of subjunctive conditionals. In Uli Sauerland & Orin Percus (eds.), *The interpretive tract (MIT Working Papers in Linguistics 25)*, 29–44. Cambridge, MA: MITWPL.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221.

Horn, Laurence. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.

Karttunen, Lauri. 1971. Subjunctive conditionals and polarity reversals. *Papers in Linguistics* 4. 279–296.

Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1. 181–194.

Kratzer, Angelika. 1981. Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic* 10(2). 201–216.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12(5). 607–653.

Krifka, Manfred. 1991. Some remarks on polarity items. In Dietmar Zaefferer (ed.), *Semantic Universals and Universal Semantics*, 150–189. New York: Foris Publications.

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25. 209–257.

Ladusaw, William A. 1980. *Polarity Sensitivity as Inherent Scope Relations*. New York: Garland Press.

Lewis, David. 1973. *Counterfactuals*. Oxford: Blackwell.

Lewis, David. 1979. Counterfactual dependence and time's arrow. *Noûs* 13(4). 455–476.

Lewis, David. 1981. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic* 10(2). 217–234.

Löbner, Sebastian. 1989. German *Schon - Erst - Noch*: An integrated analysis. *Linguistics and Philosophy* 12(2). 167–212.

van der Sandt, Rob. 1991. Denial. *Papers from the Annual Meeting of the Chicago Linguistic Society: Parasession on Negation* 27(2). 331–344.

Schwarz, Bernhard & Rajesh Bhatt. 2006. Light negation and polarity. In Raffaella Zanuttini, Héctor Campos, Elena Herburger & Paul H. Portner (eds.), *Crosslinguistic Research in Syntax and Semantics: Negation, Tense, and Clausal Architecture*, 175–197. Washington, D.C: Georgetown University Press.

Stalnaker, Robert. 1968. A theory of conditionals. *Studies in Logical Theory* 2. 98–112.

Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophical Logic* 2(4). 447–457.

Stalnaker, Robert. 1975. Indicative conditionals. *Philosophia* 5(3). 269–286.

Szabolcsi, Anna. 2004. Positive Polarity — Negative Polarity. *Natural Language and Linguistic Theory* 22(2). 409–452.

Tichý, Pavel. 1976. A counterexample to the Stalnaker-Lewis analysis of counterfactuals. *Philosophical Studies* 29(4). 271–273.

Veltman, Frank. 2005. Making counterfactual assumptions. *Journal of Semantics* 22(2). 159–180.

Zeijlstra, Hedde. 2012. Not a light negation. *Semantics and Linguistic Theory* 22. 545–564.