SUBJUNCTIVE CONDITIONALS AND POLARITY REVERSALS*

Lauri Karttunen
The University of Texas at Austin

1. 'In his two recent papers on polarity reversals (1970a and 1970b), C. L. Baker has discussed the problem that affirmative polarity items (already, some, would rather, etc.) may also under certain conditions appear in a negative context. Similarly, words of negative polarity (yet, any, bother, etc.) sometimes occur in an affirmative sentence. As Baker has shown, a great number of such exceptions to the general polarity rule can be accounted for in terms of the semantic notions "entailment" and "presupposition". Roughly, Baker's principle is that items with "wrong" polarity are tolerated in contexts where a corresponding sentence with reversed polarity is entailed or presupposed by the offending sentence. Counterfactual conditionals provide striking examples of this phenomenon, as shown below [examples (10a) and (10b) in Baker 1970b].

(1) (a) If no one had asked my sister to the dance yet, I would be trying to find someone to take her.
(b) If someone hadn't already asked my sister to the dance, I would be trying to find someone to take her.

Since the antecedent clauses are negative in both (1a) and (1b), they should not contain any affirmative polarity items. (1b) is a violation of this general polarity rule. In order to explain its well-formedness, we have to make use of the fact that (1b) is a counterfactual conditional. As such, it
presupposes the negation of its antecedent clause. Thus (1b) - as well as (1a) - presupposes (2).

(2) Someone has already asked my sister to the dance.

In (2), already appears in an affirmative context. Since (1b) presupposes (2), which conforms to the general polarity rule, Baker's principle correctly predicts that (1b) should also be acceptable.

The two conditionals in (1) appear virtually synonymous. However, Baker discovered that, if (1a) and (1b) are embedded in a certain larger context, a surprising difference manifests itself. Consider the following examples [(10f) and (10g) in Baker 1970b].

(3) (a) [Granted that I don't have a sister.] However, if I did have a sister, and if no one had asked my sister to the dance yet, I would be trying to find someone to take her.

(b) [Granted that I don't have a sister.] *However, if I did have a sister, and if someone hadn't already asked my sister to the dance, I would be trying to find someone to take her.

Why is the latter example ungrammatical? For some reason, already is a true polarity violation in (3b), contrary to what Baker's principle would seem to predict in the light of the previous example. I will show below that the explanation which Baker himself gives to this phenomenon cannot be right. Nevertheless, I claim that the ungrammaticality of (3b) is not a counterexample to Baker's theory. On the contrary, it seems to me that the examples in (3) provide further evidence in favor of Baker's principle. At the same time, they reveal something about the nature of embedded subjunctive conditionals which often has gone unnoticed.
2. Let us first see why Baker's own account of the ungrammaticality of (3b) is mistaken. He relates this phenomenon to a difference which, he believes, exists between (1a) and (1b). Baker assumes that there are actually two kinds of counterfactual conditionals: (1b) presupposes that its antecedent is false, whereas (1a) presupposes that the antecedent is not true. Consequently, (1b) presupposes the truth of (2), which is the negation of its antecedent, while (1a) only presupposes the "non-factuality" of (4).

(4) No one has asked my sister to the dance yet.

Furthermore, he observes that there are two ways in which (4) may fail to be factual. One possibility is that (2) is true, in which case (4) must be false. The other possibility is that (5) is false.

(5) I have a sister.

Baker concludes that one of the ways in which the presupposition of (1a) can be met is that (5) is false. Since (1b) presupposes (2), which in turn presupposes (5), the crucial difference between (1a) and (1b) is that only the latter presupposes that I have a sister.

This part of Baker's argument is open to two objections. First, no justification is given for the view that there are two different types of counterfactual conditionals: those that presuppose that the antecedent is false vs. those that presuppose its "non-factuality". This distinction, as well as the assignment of (1a) and (1b) to their respective categories appears entirely ad hoc. Secondly, what ultimately follows from this distinction, namely the claim that (1a) does not presuppose (5), goes against one's intuitive judgement. I find it hard to believe that there is any difference between (1a) and (1b) in this respect.²

Even if Baker were right in assuming that (1b) is the only one of the two sentences in question that presupposes
(5), the rest of his argument would still fail. According to Baker, (3b) is unacceptable for the following reason. The truth of (5) is presupposed by (1b), which constitutes a part of (3b); however, in (3b) (5) is "explicitly marked as false". Therefore, (3b) is semantically anomalous. Since (1a) does not presuppose (5), the same contradiction does not arise in (3a), which is grammatical. What Baker does not take into account is that, in (3), both (1a) and (1b) are embedded as consequent clauses of a conditional in which the antecedent, "if I did have a sister", consists of (5). For this reason, the supposed difference between (1a) and (1b) is irrelevant. Whatever is explicitly given as the antecedent clause of a conditional - counterfactual or not - may always be presupposed by the consequent. For instance, consider counterfactual conditionals where the consequent clause has a factive predicate and thus presupposes its complement. The example in (6) is perfectly grammatical, although the underlying complement of the factive verb regret is "explicitly marked as false", as far as the actual world is concerned. 3

(6) [Granted that Frances is not Swedish.] If Frances were a Swede, she would regret it.

If Baker were right in his argument, (6) would have to be ungrammatical. The difference in grammaticality between (3a) and (3b) cannot be explained by any difference between the presuppositions of (1a) and (1b), even if such a discrepancy existed.

3. What, then, is the correct explanation? Let us take a more careful look at the examples in (3). On the surface, they are of the form (7).

(7) If $S_1$, and if $S_2$, then $S_3$. 
As it stands, (7) does not match any well-formed formula in standard propositional calculus. At first, one might be tempted to think that (7) is ambiguous. There are two equally plausible ways to make it to a logical well-formed sentence, namely (8a) and (8b).

(8) (a) If $S_1$ then (if $S_2$ then $S_3$).
(b) If ($S_1$ and $S_2$) then $S_3$.

However, a moment's reflection shows that it should not matter which representation we choose, since the two formulas in question are logically equivalent: $p \supset (q \supset r) \equiv (p \land q) \supset r$. One can easily convince oneself of this equivalence by considering examples like (9a) and (9b). There does not seem to be any difference in meaning.

(9) (a) If there had been a Communist revolution in Austria, then, if Johnson had been President at that time, the US would have occupied Vienna.
(b) If there had been a Communist revolution in Austria and Johnson had been President at that time, the US would have occupied Vienna.

Let us, therefore, adopt the form (8a) as basic. The original examples in (3) can now be rewritten as in (10).

(10) (a) If I had a sister, then, if no one had asked my sister to the dance yet, I would be trying to find someone to take her.
(b) *If I had a sister, then, if someone hadn't already asked my sister to the dance, I would be trying to find someone to take her.

At first, it appears that the examples in (10) both contain a counterfactual conditional that is embedded in another
counterfactual conditional. But what do we actually mean by this term? According to the usual definition, a counterfactual conditional presupposes the negation of its antecedent. There is no doubt that (10a) and (10b) both presuppose (11), which is the negation of the first antecedent.

(11) I do not have a sister.

But do they really presuppose the negation of the embedded antecedent? It might just be that, while the subjunctive form is a necessary characteristic of a counterfactual conditional, in this case it is not a sufficient one. This suspicion is confirmed by such less complicated examples as (9a). In (9a) the embedded conditional is identical to (12).

(12) If Johnson had been President at that time, the US would have occupied Vienna.

Although it is certainly true that, in isolation, (12) presupposes the negation of this antecedent, one can show that no such presupposition accompanies this conditional when it occurs embedded in (9a). Consider the adverb at that time in (12). In the context of (9a) it refers to the time at which there was a Communist revolution in Austria. However, (9a) as a whole presupposes that, in our actual world, there never was such a revolution. The antecedent of the embedded conditional in (9a), "Johnson has been President at that time", can be neither true nor false, as far as the actual world is concerned, since the expression at that time does not specify any point at all in its past history. Consequently, the embedded conditional in (9a) is not counterfactual in the usual sense.

Of course, we can argue that it is mistaken to try to interpret the embedded conditional in (9a) in terms of the actual world. The first antecedent clause of (9a), "There has been a Communist revolution in Austria", invites us to consider a possible world whose history differs from ours
in just that respect. However, it seems that the embedded conditional cannot be considered as counterfactual in that possible world either. For all we know, at the time the revolution took place, Johnson may or may not have been President. It is certainly not the case that the embedded conditional presupposes that Johnson was not President at the time of the hypothetical revolution.

Under either one of the suggested interpretations, when it occurs embedded in (9a), (12) turns out to be something else than counterfactual. Its subjunctive form is simply due to the fact that it constitutes the consequent clause of a "real" counterfactual conditional. The subjunctive mood is a matter of required agreement. Note that we cannot say (9a').

(9a') *If there had been a Communist revolution in Austria, then, if Johnson has been President at that time, the US has occupied Vienna.

It is clear from the preceding discussion that the two antecedent clauses in (9a) have a different status. Either way we want to look at it, the second antecedent clause presupposes the first. Consequently, and contrary to what the logical equivalence \( p \supset (q \supset r) \equiv q \supset (p \supset r) \) would lead us to expect, (9a) and (13) are not equivalent.⁴

(13) If Johnson had been President at that time, then, if there had been a Communist revolution in Austria, the US would have occupied Vienna.

Assuming that the phrase at that time refers to time of the revolution, (13) is ungrammatical. In natural language, the equivalence given in (14) seems to hold only if the two antecedent clauses, \( S_1 \) and \( S_2 \), are independent from each
other; that is, the truth value of one does not depend on
the truth or falsity of the other.

\[(14) \text{ If } S_1 \text{ then } (\text{ if } S_2 \text{ then } S_3) \equiv \text{ If } S_2 \text{ then } (\text{ if } S_1 \text{ then } S_3)\]

For instance, the equivalence holds in (15).

\[(15) \begin{align*}
\text{(a) If Chicago were in Texas, then, if Austin} \\
\text{were in Illinois, Chicago would be south} \\
\text{of Austin.}
\end{align*}\]

\[(\text{b) If Austin were in Illinois, then, if Chicago} \\
\text{were in Texas, Chicago would be south of} \\
\text{Austin.}\]

It is easy to see that neither one of the two antecedent
clauses in (15) presupposes the other. It also makes no
difference which of the two clauses we choose first in
defining a possible state of affairs. In the actual world,
Austin is in Texas and there is no reason to believe that
Austin would be located elsewhere even if Chicago were in
Texas too. We know that the two antecedent clauses are
both false in the actual world and either one would also
be false in a possible world where the other is true,
assuming that everything else we know about geography
remains constant. Therefore, nothing prevents us from
regarding the embedded subjunctive conditionals as counter-
factual in (15). On the other hand, in (9a) the embedded
conditional is not counterfactual in spite of its subjunctive
form.

4. Let us now go back to the original examples. In addition
to saying that (3a) and (3b) are equivalent to (10a) and
(10b), respectively, we also have to find out what the
status of the embedded conditional is.
(10) (a) If I had a sister, then, if no one had asked my sister to the dance yet, I would be trying to find someone to take her.

(b) *If I had a sister, then, if someone hadn't already asked my sister to the dance, I would be trying to find someone to take her.

If the examples in (10) were like those in (15), the embedded conditional could be regarded as counterfactual. However, (10a) and (10b) both are similar to (9a). The relation between the independent conditional (1a) and its embedded duplicate in (10a), for example, is the same that we observed in the case of (12) and (9a). In isolation, (1a) presupposes (2).

(1a) If no one had asked my sister to the dance yet, I would be trying to find someone to take her.

(2) Someone has already asked my sister to the dance.

As far as the actual world is concerned, the antecedent of the embedded conditional in (10a), "No one has asked my sister to the dance yet", can be neither true nor false, since (10a) as a whole presupposes that, in the actual world, the speaker does not have a sister. What about the hypothetical state of affairs which (10a) invites us to consider? In that state of affairs, the speaker has a sister, thus (2) might be true or it might be false. Neither of the alternatives is presented as more likely than the other in (10a). Consequently, no matter how we look at it, the embedded conditional in (10a) is not counterfactual in the proper sense of the term. Note also that, since the first antecedent is presupposed by the second one, the equivalence in (14) fails here just as it does in case of (9a). As shown by (10a'), the inversion
of the antecedents in (10a) yields an ungrammatical sentence, not an equivalent one.

(10a') *If no one had asked my sister to the dance yet, then, if I had a sister, I would be trying to find someone to take her.

Therefore, the technique that was used to justify the claim that the embedded conditional in (15) can be regarded as counterfactual yields a negative result here.

By the same reasoning, although (1b) in isolation undoubtedly presupposes (2), the negation of its antecedent clause, no such presupposition accompanies the embedded conditional in (10b).

(1b) If someone hadn't already asked my sister to the dance, I would be trying to find someone to take her.

In (10b), the embedded conditional is not counterfactual. This observation solves immediately the problem why already, an affirmative polarity item, is acceptable in (1b) but is not acceptable in (10b) under seemingly identical conditions. By Baker's principle, items with wrong polarity are tolerated if the corresponding sentence with reversed polarity is presupposed. Since (1b) presupposes (2), already is acceptable in spite of the negative environment. But (10b) does not presuppose (2), therefore, it must be ungrammatical. What first appeared to be a counterexample is actually further evidence in support of Baker's principle.

A similar analysis can be given to the second set of examples that Baker discusses in this connection [examples (11a) and (11b) in Baker 1970b]. As Baker points out, (16a) and (16b) are interchangeable.
(16) (a) If, as of January 1, John hadn't yet stopped beating his wife, he could have been sent to the county jail.

(b) If, as of January 1, John hadn't already stopped beating his wife, he could have been sent to the county jail.

(16b) with already in a negative environment violates the general polarity rule. However, it presupposes (17), in which already appears in an affirmative environment.

(17) As of January 1, John had already stopped beating his wife.

Therefore, the well-formedness of (16b) can be explained by Baker's principle.

The problem for Baker lies in explaining what is wrong with (18).

(18) *If John had been beating his wife, then, if, as of January 1, he hadn't already stopped beating her, he could have been sent to the county jail.

The explanation that Baker himself gives for this phenomenon is mistaken for reasons that are analogous to those we already discussed in connection with the previous examples. The correct explanation is quite simple. Although (16b) in isolation presupposes (17), the embedded subjunctive conditional in (18) is not counterfactual. For all we know, in some possible world which differs from the actual one in the respect that John has been beating his wife, he may or may not have stopped beating her on January 1. Since (18) does not presuppose (17), Baker's principle cannot apply. (18) is a simple violation of the general polarity rule.
5. There are some interesting consequences that follow from the observation that an embedded subjunctive conditional may not be counterfactual in the usual sense. It is no longer clear whether we can explain the well-formedness of (19) by Baker's principle. Nevertheless, (19) is considered acceptable by anyone who is aware of the fact that Johnson retired as President in 1969.

(19) In 1970, if there had been a Communist revolution in Austria, then, if Johnson had not already retired as President, the US would have occupied Vienna.

First of all, note that the two antecedent clauses in (19) are independent from each other. For this reason, it is possible for the equivalence given in (14) to hold here, although it does not hold in the case of (9a), (10a), and (18). In this respect, (19) is similar to (15). The inversion of the antecedent clauses in (19) yields an equivalent sentence, (19').

(19') In 1970, if Johnson had not already retired as President, then, if there had been a Communist revolution in Austria, the US would have occupied Vienna.

Because of the equivalence of (19) and (19'), one can argue that it is presupposed by (19) that the negations of the two antecedent clauses, "There has not been a Communist revolution in Austria" and "Johnson has already retired as President", are both true in the actual world. In this case, the acceptability of already in (19) might be explainable by Baker's principle.

Another way of looking at the problem is the following. In (19), we are invited to consider a possible state of affairs that might have taken place in 1970. This state of affairs differs from any actual one - hence the
subjunctive mood of the first antecedent - in the respect that there has been a Communist revolution in Austria. Assuming that all the other things that we know about the world history remain constant, since Johnson retired as President in 1969, it must be true in our hypothetical state of affairs that Johnson has already retired as President. In this case, the occurrence of the affirmative polarity item already in (19) would be justified, not because of any logical relation between the sentences (19) and (20), but because our knowledge of history leads us to regard (20) as necessarily true in the possible state of affairs that (19) invites us to consider.

(20) Johnson has already retired as President.

Note that the two ways to explain (19) which are outlined above are quite different. In the first case, I have attempted to show that (19) in fact presupposes that (20) is true in the actual world. This argument relies on the use of the equivalence (14), If $S_1$ than (If $S_2$ then $S_3$) = If $S_2$ then (If $S_1$ then $S_3$), which in turn depends on the knowledge that $S_1$ and $S_2$ are both false in the actual world, hence neither one presupposes the other. This makes it possible to use Baker's principle in explaining the acceptability of already. In the second case, it is argued that (19) is acceptable because it invites us to consider a state of affairs in which the truth of (20) is entailed by independent premises. Here we are appealing to something more general than Baker's requirement that the truth of (20) be entailed or presupposed by the offending sentence, that is, by (19).

Given these two approaches, the question arises whether there are any data that would force us to choose between these alternative ways of looking at the problem. In many cases either one of the two explanations will work. For example, the unacceptability of (21) can be explained
either by pointing out that here the inversion of the antecedents will not yield an equivalent sentence or by observing that, in the state of affairs we are invited to consider here, (20) is necessarily false given the fact that Johnson retired in 1969.

(21) *In 1965, if there had been a Communist revolution in Austria, then, if Johnson had not already retired as President, the US would have occupied Vienna.

However, it seems that only the latter approach is capable of accounting for the grammaticality of examples like (22).

(22) I hate all women. If I had a sister, then, if I didn't already hate my sister as a woman, I would certainly start disliking her as a close relative.

In the case of this misogynic statement, the first antecedent, "I have a sister", is presupposed by the antecedent of the embedded conditional. As we pointed out earlier, it seems that in all such cases the inversion of the antecedents yields an ungrammatical sentence, not an equivalent one. Thus (22') is clearly ungrammatical.

(22') I hate all women. *If I didn't already hate my sister as a woman, then, if I had a sister, I would certainly start disliking her as a close relative.

Furthermore, as far as the actual world is concerned (23) can be neither true nor false, since it is presupposed by (22) that the speaker does not actually have a sister.

(23) I already hate my sister as a woman.
For these reasons, the grammaticality of (22) cannot be accounted for within the framework of the first proposal. Since the equivalence (14), If \( S_1 \) then (If \( S_2 \) then \( S_3 \)) \( \equiv \) If \( S_2 \) then (If \( S_1 \) then \( S_3 \)), clearly fails in case of (22) and (22'), one would predict that (22) is anomalous just as the other cases, (10b), (18), and (21), where the same equivalence has failed. Nevertheless, (22) is acceptable.

On the other hand, we can account for (22) in the framework of the second proposal. Given the additional premise "I hate all women", it logically follows that (23) must be true in the possible state of affairs that (22) invites us to consider. It is easy to see that, if the speaker hates all women, it is necessarily true that he would hate his sister, if he had one. But here we are appealing to a more general principle than what Baker has proposed, since the truth of (22) is not entailed or presupposed by the offending sentence as such.

Why is it then, that the previous examples, (10b) and (18), were found ungrammatical? If no logical relation between (10b) and (2) is required, why is (10b) ungrammatical?

(10b) *If I had a sister, then, if someone hadn't already asked my sister to the dance, I would be trying to find someone to take her.

(2) Someone has already asked my sister to the dance.

My guess is that (10b) is not so much ungrammatical as unreasonable. Consider what sort of independent assumption it would take to justify already in this case. The speaker would have to assume that, in all possible courses of events where he has a sister, it is the case
that someone has already asked her to the dance at the time of his uttering (10b). If there are speakers who are able to entertain such beliefs, I would predict that, for them, (10b) is an acceptable statement.

Similarly, (18) would require one to believe that (17) be a necessary truth in the hypothetical situation in which John has been beating his wife.

(18) *If John had been beating his wife, then, if, as of January 1, he hadn't already stopped beating her, he could have been sent to the county jail.

(17) As of January 1, John had already stopped beating his wife.

It seems to me that (22) is different from (10b) and (18) only for the following reason. In (22), the independent assumption which is required to establish (23) as a necessary truth in the suggested hypothetical state of affairs is a very simple one and is, furthermore, explicitly given by the context. On the other hand, the corresponding independent premises required by (10b) and (18) are very peculiar and hard to figure out. If these two sentences are ungrammatical at all, they are ungrammatical because they force one to make unreasonable assumptions.

6. I believe that this discussion has established the following two points. First of all, what Baker says about the examples (10b) and (18) cannot be right. Nevertheless, they are not counterexamples to his theory but actually provide more evidence for it. Secondly, examples like (22) show that, as 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 of that principle ultimately turns out to be, it will have to cover a larger class of cases.
BIBLIOGRAPHY


FOOTNOTES

*Work on this paper was supported by the Linguistics Research Center under a grant for Theoretical Study Effort of High Quality Translation (Grant F30602-70-C-0129) from the Rome Air Development Center. I want to expres my thanks to C. L. Baker, Orvokki Heinämäki, Frances Karttunen, and Stanley Peters for discussing these matters with me. Especially I want to thank C. L. Baker for his helpful comments on an earlier version of this paper.

1 Here I will only discuss the word already in negative contexts, since it is this class of sentences where Baker's principle has proved most successful in predicting grammaticality. I will not be concerned with examples where the principle in its present form is already known to fail. For a discussion of such difficulties, see Baker 1970b and Horn 1970.

2 What may be confusing the issue is that conditionals like (1a) are also colloquially used, as it were, to abbreviate longer expressions of the type (3a). A person who does not have a sister might well utter (1a) to someone in a conversation. However, I would claim that, in such cases, a prefix like "If I had a sister and ..." or "If I were you and ..." is, if not explicitly there, at least tacitly understood. I take it to be a fact that, unless (1a) is construed as this sort of elliptical utterance, it presupposes (5). Note that (1b) cannot be used elliptically in the same way, presumably because the corresponding "complete" statement, (3b) is ungrammatical.
For more examples of the same sort, see Karttunen 1969 and Morgan 1969.

Similarly, in spite of the logical equivalence

\[(p \land q) \supset r \equiv (q \land p) \supset r,\]

in natural language (i) appears to fail whenever \(S_2\) presupposes \(S_1\).

(i) If \((S_1 \text{ and } S_2)\) then \(S_3 \equiv \text{If } (S_2 \text{ and } S_1) \text{ then } S_3.\)

For example, the inversion of the conjoined antecedents in (ii) results in an ungrammatical sentence, as shown in (iii).

(ii) If I had a sister and no one had asked my sister to the dance, I would be trying to find someone to take her.

(iii) *If no one had asked my sister to the dance and I had a sister, I would be trying to find someone to take her.

For this reason, it wouldn't have mattered if he had chosen to discuss (9b) instead of (9a).

Robin Lakoff (1969) discusses many examples that also seem to involve the speaker's beliefs and expectations which cannot be explained in terms of any logical relations between the sentences in question.