

Discourse Rationality and the Counterfactuality Implicature in Backtracking Conditionals¹

Ana Arregui — University of Ottawa

María Biezma — University of Konstanz

Abstract. This paper contributes to current discussions of counterfactuality implicatures in *would*-conditionals. The empirical focus is on the contrast between forward-looking and backtracking examples in Anderson-style ‘detective reasoning’ sequences. We show that differences regarding the cancellability of counterfactuality in these examples follow from general principles of discourse rationality and can be extended to provide a more general account of the cancellability of counterfactuality implicatures (e.g. in future-shifted examples).

Keywords: counterfactuality, implicatures, conditional perfection, backtracking counterfactuals.

1. Introduction

This paper investigates counterfactuality in *would*-conditionals bringing together ideas about counterfactuality implicatures, backtracking conditionals and good ‘discourse manners’. The empirical focus will be on Anderson-style (Anderson 1951) cancellation of counterfactuality, comparing ‘forward looking’ conditionals with backtrackers (in which the consequent temporally precedes the antecedent). As (1) illustrates, counterfactuality appears cancellable in ‘forward looking’ conditionals (1a), but not so in backtrackers (1b):

- (1) a. If Jones had taken arsenic yesterday, he would show lividity symptoms now. He does show lividity symptoms now. So he probably took arsenic yesterday.
- b. If Jones were showing lividity symptoms now, he would (have to) have taken arsenic yesterday. He did take arsenic yesterday. # So he is probably showing lividity symptoms now.

The paper proposes an account of the contrast between (1a) and (1b), exploring its consequences for the cancellation of counterfactuality in other types of examples. An important component of our proposal is the idea that good discourse manners are expected: if a speaker has chosen to set up an implicature, it cannot be trivially cancelled (i.e. cancelled ‘for no reason’). Conditional perfection will also be an important component of the account, as we will show that it plays an important role in the cancellation of implicatures via the detective reasoning associated with Anderson-style examples. With these ingredients in place, our account of why counterfactuality cannot be can-

¹We would like to thank the audience of SuB 20 for helpful comments and feedback. Earlier versions of this work were presented at the University of Potsdam and we would also like to thank Malte Zimmermann and our audience for their feedback, as well as the audience of TbiLLC 2013. Remaining errors are our own. This research has been partially supported by the Deutsche Forschungsgemeinschaft (DFG, Project 1836/1-1). Authors are listed alphabetically.

cancelled in examples like (1b) will also be able to account for the observed difficulties in cancelling counterfactuality in future-shifted *would*-conditionals (as discussed by e.g. Ogihara 2000; Ippolito 2013).

The paper is structured as follows: in *Section 2* we spell out our assumptions about counterfactuality, expanding on the proposal in Leahy (2011) to characterize it as an antipresupposition while at the same time noting the problem for cancelling counterfactuality in non-standard *would*-conditionals presented in Biezma et al. (2013); in *Section 3* we discuss backtracking *would*-conditionals, following Arregui (2005a, b), in *Section 4* we spell out our assumptions about conditional perfection, building on von Stechow (2001, 2009) to include a discussion of perfection in backtrackers; in *Section 5* we spell out the proposal for the cancellation of counterfactuality in forward-looking vs. backtracking *would*-conditionals, and extend our proposal to argue that nothing special needs to be said about future-shifted *would*-conditionals (contra Ippolito 2013, Martin 2015); we conclude in *Section 6*.

2. On counterfactuality

2.1. Cancelling counterfactuality

As is standardly observed, an utterance of a *would*-conditional often gives rise to the understanding that the antecedent clause proposition is false in the actual world. We will use the term ‘counterfactuality’ informally to refer to this understanding. An illustration of counterfactuality is provided by (2) (an utterance of the conditional would most likely, out of the blue, lead to the understanding that Caspar did not come to the party):

- (2) If Caspar had come to the party, it would have been fun. (Lewis 1973)

Ever since Anderson (1951), it has been accepted that counterfactuality in *would*-conditionals is an implicature. The slight variant of Anderson’s famous example provided in (1a) is repeated below:²

- (3) If Jones had taken arsenic yesterday, he would show lividity symptoms now. He does show lividity symptoms now. So he probably took arsenic yesterday.

Counterfactuality is ‘cancelled’ in (3) (as had been noted by Anderson already, the discourse in (3) actually argues in favour of the truth of the antecedent).

²Here is Anderson’s original example:

- (1) In the investigation of Jones’s death, a doctor might say, “If Jones had taken arsenic, he would have shown just exactly those symptoms which he does in fact show”. Now in this context the doctor’s statement would probably be taken as lending support to the view that Jones took arsenic - it would certainly not be held to imply that Jones did not take arsenic. (Anderson 1951: 37)

Examples like this support the view that counterfactuality is not actually part of the truth-conditional content of *would*-conditionals or a presupposition, but arises instead as a defeasible pragmatic inference.³ Anderson-type examples had already been given by Chisholm (1946), who described the ‘deliberative use’ of *would*-conditionals (that he termed ‘subjunctives’) as ‘detective reasoning’:

- (4) When we prepare for a crucial experiment, we review the situation and consider what would happen if our hypothesis were true and what would happen if it were false. The subjunctive conditional is essential to the expression of these deliberations. In defending a hypothesis, I may employ a subjunctive conditional even though I believe the antecedent to be true; I may say, “If this were so, that would be so; but, as you see, this *is so*...”. It is said that detectives talk in this manner. (Chisholm 1946: 291)

Additional support for the claim that counterfactuality is not presupposed comes from Stalnaker’s observation that at times we seem to meaningfully argue in favour of the falsehood of the antecedent, a move predicted to be trivial if counterfactuality were presupposed:

- (5) Consider the argument, ‘The murdered used an ice pick. But if the butler had done it, he wouldn’t have used an ice-pick’. So the murderer must have been someone else. The subjunctive conditional premise in this modus tollens argument cannot be counterfactual since if it were the speaker would be blatantly begging the question by presupposing, in giving his argument, that his conclusion was true. (Stalnaker 1975: 277)

But it should be said that counterfactuality is insistent. Lewis (1973) claimed that, upon finding out that the antecedent of a *would*-conditional is true, we are more likely to come to believe that the conditional was false rather than to accept that the consequent is true. So, if after hearing somebody utter (2) we were to find out that Caspar did go to the party, we would be more likely to conclude that the speaker of (2) had said something false rather than conclude that the party was indeed fun. In Lewis’s words, *The false information conveyed by using a counterfactual construction with a true antecedent eclipses the falsity or truth of the conditional itself.* (Lewis 1973: 26). The fact that it is not trivial to set aside counterfactuality is an important lesson, and we will come back to this in the following sections. Cross-speaker dialogues, as Lewis noted, can help provide suitable settings for side-stepping counterfactuality without generating weird discourses:

- (6) A: If Caspar had come, it would have been a good party.
B: That is true; for he did, and it was a good party. You didn’t see him because you spent the whole time in the kitchen, missing all the fun. (Lewis 1973)

³In following sections we will schematize the truth-conditional content of *would*-conditionals as $\alpha \Rightarrow \beta$ and the counterfactuality implicature as $\rightsquigarrow \neg\alpha$.

As (6) illustrates, two speakers may agree on a *would*-conditional without agreeing on the truth/falsity of the antecedent, again supporting the view that counterfactuality is not presupposed.

While there is broad agreement that counterfactuality in *would*-conditionals is an implicature, examples in which it is impossible/difficult to cancel (such as (1b)) present a challenge. As we will show, the explanation for the difficulties in (1b) can actually account for a family of examples.

2.2. The counterfactuality implicature

There are various views regarding how the counterfactuality implicature is generated. Stalnaker (1975) suggested that the morphology typical of *would*-conditionals served as a ‘conventional device’ to indicate that presuppositions are being suspended (more recent proposals include von Stechow 1997; Iatridou 2000; Ippolito 2013). We will build on Leahy (2011), who elaborates an account of counterfactuality as an ‘anti-presupposition’ implicature arising from competition with indicatives. According to Leahy, indicative conditionals like (7a) presuppose that the antecedents are epistemically possible for their speakers, whereas conditionals like (7b) do not carry presuppositions. At the core of Leahy’s proposal is the idea that when a speaker chooses to make an utterance that is presuppositionally weaker to a salient alternative with the same semantic content, s/he gives rise to the implicature that s/he does not believe the stronger presupposition to be felicitous (see Heim 1991 and related work in antipresuppositions). This accounts for the contrast between (7a), with counterfactuality implicature, and (7b), without:

- (7) a. If John had come, it would have been fun.
b. If John came, it was fun. (Leahy 2011)

We will add to the discussion of anti-presuppositions the contrast arising between examples with perfective (8a) vs perfect aspect (8b), investigated in Arregui (2005b, 2007):⁴

- (8) a. If your plants died tomorrow, I would be very upset.
b. If your plants had died tomorrow, I would have been very upset. (Arregui 2007)

Consider the examples against the following background: You will be away from your house tomorrow and ask me to look after your plants. I am worried, as I am very bad with plants. In this context, I could felicitously utter (8a), but (8b) would be odd. Suppose now that your plants died yesterday, and you let me know to cancel your request. In this context, I could felicitously utter (8b), but (8a) would be odd. As the second scenario illustrates, when it is known that the proposition that your plants die tomorrow is false (in this case, because it is known that they have already

⁴It is left for future work to evaluate whether aspect-based competition could provide an account of all cases.

died), perfect morphology is required in the antecedent clause. The perfective morphology in (8a) restricts the domain of quantification to worlds compatible with what is known and does not allow us to quantify over worlds in which your plants die tomorrow instead of yesterday (resulting in a hypothesis that your plants die for a second time). Lewis (1973) had already noted the indicative-like flavour of *would*-conditionals like (8a) and set them aside in his account of counterfactuals:

- (9) More importantly, there are subjunctive conditionals pertaining to the future, like ‘If our ground troops entered Laos next year, there would be trouble’ that appear to have the truth conditions of indicative conditionals, rather than of the counterfactual conditionals I shall be considering. (Lewis 1973: 4)

We will follow Arregui (2005b, 2007) in attributing the contrast between (8a) and (8b) to a view according to which perfective aspect carries epistemic consequences restricting quantification to worlds in the context set (see Arregui (2005b, 2007) for a detailed analysis, see Ippolito (2013) for critical comments). A consequence of this view is that perfect examples will lead to counterfactuality by competition with perfective examples. Counterfactuality in future-shifted examples like (8b) can thus be understood as an implicature arising from competition with examples like (8a). (The cancellation of counterfactuality in future-shifted examples has been investigated in e.g. Ippolito (2013) and Martin (2015), and we will return to this topic in Section 5.3).

2.3. Cancelling counterfactuality and discourse manners

The characterization of counterfactuality as an implicature in principle predicts that it can be cancelled. This does not, of course, predict that it will be trivial to cancel the implicature. Biezma et al. (2013) investigated another kind of *would*-conditional in which counterfactuality resisted cancellation, studying the contrast between the interpretation of ‘regular’ *would*-conditionals (10a) vs. non-standard forms (10b) below:

- (10) a. If Jones had taken arsenic, he would have shown some symptoms.
b. If Jones would have/ had have/ would’ve/ had’ve/ woulda’/ hada’/ would of / had of taken arsenic, he would have shown some symptoms. (Biezma et al. 2013)

Examples like (10b) had been discussed by Siddiqi and Carnie (2012), who developed morphosyntactic arguments showing that the extra morphology corresponds to modal heads.⁵ Biezma et al. (2013) present the novel observation that in dialects that allow both (10a) and (10b), they do not actually obtain identical interpretation, since counterfactuality cannot be cancelled in the non-standard variants:

⁵Contra the assumptions made in Ippolito (2013) and Martin (2015) that these are past markers.

- (11) a. If Jones had taken arsenic, he would have shown exactly those symptoms that he in fact shows (so, he probably took arsenic)
- b. # If Jones had've/ would've taken arsenic, he would have shown exactly those symptoms that he in fact shows (so, he probably took arsenic)

Biezma et al. (2013) provide a semantic analysis of the extra layer of modality in non-standard forms according to which the antecedent clause embeds a simple subjunctive. Adopting the proposal for simple subjunctives in Kasper (1992), they argue that non-standard forms make claims about the preconditions for the truth of the antecedent, and are informationally weaker than their standard counterparts. The proposal predicts that the utterance of a non-standard form gives rise to the implicature that the pre-conditions for the truth of the antecedent proposition are not satisfied, and thus also to the implicature that the antecedent proposition is false. Cancelling the implicature that the antecedent is false automatically cancels the implicature that the preconditions are not satisfied, but gives rise to the question of why the speaker chose the more complex form in the first place. The non-standard forms are morphologically, structurally, semantically and inferentially more complex. Their use is only (pragmatically) justified if some mileage will be obtained from this complexity. Cancelling the counterfactuality implicature in examples like (11b) would annul any advantage conferred by the more complex non-standard forms and work against 'discourse rationality' (cancelling an inference for no reason that one has chosen to trigger amounts to a 'discourse contradiction'). It would go against principles of pragmatic economy within the Gricean tradition (i.e. to use the simplest and most informative form available). In brief, this is 'bad discourse manners' and we are not willing to go along with it.

This account of why counterfactuality cannot be cancelled in non-standard *would*-conditionals like (10b) is reminiscent of proposals put forward by Abbott (2006) regarding why some presuppositions cannot be 'lifted' (cancelled). Abbott claims that it is not possible to cancel presuppositions associated with 'hard' presupposition triggers (i.e. those for which there is no reason for use other than to trigger a presupposition). We illustrate this in (12) with a cleft construction, which has been characterized as a hard trigger. As (12) shows, the associated presuppositions cannot be lifted:

- (12) I have no idea whether the problem has been solved, but # maybe it was Sue who solved it. (Abbott 2006)

There is no reason to choose the more complex form other than to generate the associated presupposition. In this kind of cases, we do not accept efforts to cancel or detach the presupposition (as we do in some examples with 'soft' triggers). Underlying this explanation is the assumption that discourse participants will behave 'reasonably' and not wilfully choose specialized more-complex forms in order to lift the associated meanings for no reason whatsoever.

In agreement with this line of thinking, we consider that counterfactuality in *would*-conditionals in general cannot be cancelled 'for no reason'. Doing so would be a case of bad discourse manners. This corresponds with Lewis's early observation about the difficulties we face when evaluating a

would-conditional with true antecedents that we noted in Section 2.1. Our account of the asymmetry between (1a) and (1b) shows that whereas we have ‘good reasons’ to cancel counterfactuality in examples like (1a), this is not the case in (1b). And, as a result, we are not able to do so.

3. Double-modality in backtrackers

3.1. On backtracking and the semantics of *would*-conditionals

In the case of backtracking *would*-conditionals, the temporal location of the antecedent follows that of the consequent (we ‘backtrack’ in time as we move from antecedent to consequent). Backtracking *would*-conditionals have been the subject of much interest, both in the philosophical and, more recently, in the linguistics literature (e.g. Slote 1978; Davis 1979; Bennett 1984, 2003; Frank 1996; Arregui 2005a, b; Schulz 2007). Backtrackers appear to pose a problem for a Lewis-style resolution of similarity in the classic semantics for counterfactuals, and were already investigated in Lewis (1979). Lewis offered the following example:

- (13) Jim and Jack quarreled yesterday, and Jack is still hopping mad. We conclude that if Jim asked Jack for help today, Jack would not help him. But wait: Jim is a prideful fellow. He never would ask for help after such a quarrel; if Jim were to ask Jack for help today, there would have to have been no quarrel yesterday. (Lewis 1979: 456)

Given the standard resolution of the similarity relation relevant for the interpretation of counterfactuals (Lewis 1973, 1979), backtracking conditionals like *If Jim were to ask Jack for help today, there would have to have been no quarrel yesterday* in (13) are expected to be false. The standard resolution of similarity identifies the domain of quantification as the worlds most similar to the actual world up to the time of the antecedent clause. In the example above, the time at which Jim asks Jack for help today. But in such worlds, there was a quarrel yesterday, so the consequent would be false. However, as Lewis pointed out, even though we often judge backtrackers as false, this is not always so. In some cases, in particular with an extra layer of auxiliaries as in (13) (*would have to have been*), backtrackers are quite easily judged as true.⁶ Lewis did not see this as a problem for his analysis of counterfactuals, as it reflected a special resolution of the similarity relation. The kind of counterfactual dependence that interested him, associated with causation, arose only under the standard resolution of similarity that predicted that backtrackers were false. In what follows we analyze the ‘special syntax’ noted by Lewis as introducing a second layer of modality.

⁶Our goal in this paper is to evaluate the backtracking version of Anderson-type examples, which are facilitated by the ‘special syntax’ noted by Lewis. We will thus limit our discussion of backtrackers to examples with special syntax. However, as has been noted in the literature (e.g. Arregui 2005a, b; Schulz 2007) some backtrackers go through quite smoothly without special syntax. We will mention them briefly in Section 5.

3.2. Double modality in backtrackers

As noted originally in Lewis (1979), backtracking in *would*-conditionals is often facilitated by a ‘special syntax’ with an extra layer of auxiliaries. Additional examples are provided below:

- (14) a. If the plane had arrived at 2:00, it would have to have departed at 1:00. (Davis 1979)
b. If the die had fallen six uppermost, it would (have to) have been thrown differently. (Bennett 1984)

In investigating the role of the special syntax, Davis (1979) argued that it introduced a second layer of modality that entailed that the consequent was true at the evaluation world (Davis characterized the consequent in (14a) as ‘*some sort of tensed modal statement*’). While the view that the consequent proposition is true at the evaluation world has proven problematic (see e.g. discussion in Bennett 2003), we will side with Davis in characterizing the consequent in examples like (14a) and (14b) as carrying an extra layer of modality with a structure that can be schematized as $\alpha \Rightarrow \Box\beta$ (see Arregui 2005a, b). The extra layer of modality in the consequent facilitates backtracking by invoking a non-reflexive modal accessibility relation. It may be marked overtly or it may remain implicit (as long as there is sufficient contextual support to allow it to be recovered). Consider Lewis’s example again:

- (15) If Jim had asked Jack for help today, there would have to have been no quarrel yesterday.

With a second layer of auxiliaries in the consequent clause, the claim made by the conditional is about what the (salient) laws would have predicted under the circumstances described by the antecedent (where by ‘laws’ we understand natural laws, but also things like the conventions governing games, regulations and accepted (defeasible) generalizations regarding what is ‘normal’ or expected). In (15), the antecedent brings us to the most similar worlds in which Jim asked Jack for help today (in those worlds there was a quarrel yesterday), while the consequent requires that all such worlds be worlds in which the salient laws (let’s say, generalizations regarding what is normal/habitual behaviour) require that there was no quarrel yesterday. The conditional may well be true without modifying the standard resolution of similarity in counterfactuals.

In summary, following Arregui (2005a, b), backtracking counterfactuals with special syntax claim that in the most similar worlds in which the antecedent is true, the worlds made accessible by (salient) law/s or generalizations are worlds in which the consequent is true (see details in Arregui 2005a, b). It is important to note two points about this proposal: (i) backtracking examples do not establish ‘causal links’ between antecedent and consequent (at least as intuitively understood); (ii) the second layer of modality in the consequent may appeal to a non-reflexive accessibility relation, which means that examples like $\alpha \Rightarrow \Box\beta$ may be true even if β is not true in the most similar worlds in which α is true. As noted, it is for this reason that the extra layer of modality facilitates

backtracking.

Examples with disagreement provide indirect support for the view that the consequents of back-trackers like (14a) and (14b) bear an extra layer of modality, even if it remains implicit.

(16) A: If Jim had asked Jack for help today, there would (have to) have been no quarrel yesterday.

B: That's not true. The other day Jim asked Jack for help after a quarrel.

As B's reply shows, accepting the truth of A's statement requires accepting the law-like generalization that Jim does not ask Jack for help after a quarrel (contrast this with a disagreement along the lines: *That's not true. There would (still) have been a quarrel yesterday!*).

The discussion of backtracking *would*-conditionals is important because it shows that when we backtrack Anderson-style detective reasoning sequences, the results are importantly different from the original. The type of 'causal reasoning' associated with the 'forward looking' examples is broken, and a second layer of modality appears. These features of the backtracking examples will play an important role in explaining why counterfactuality cannot be cancelled in these cases.

4. Conditional strengthening in forward looking vs. backtracking *would*-conditionals

In a standard semantics for conditionals, abstracting away from specific conditional flavours, a conditional of the form *if p, q* claims that all relevant *p*-situations are also *q*-situations (e.g. in the case of *would*-conditionals, under a standard Lewis-Stalnaker style analysis, that all the most similar worlds in which the antecedent is true are also worlds in which the consequent is true). If a conditional is true, *p* will provide sufficient conditions for *q*. But, as noted already by Geis and Zwicky (1971), conditionals are often understood in a stronger manner, with the antecedent considered to provide sufficient and necessary conditions for the consequent. Here is one of their examples:

(17) If you mow the lawn, I'll give you five dollars.

We naturally understand not only that I will give you five dollars if you mow the lawn, but also that I will only give you five dollars if you mow the lawn. This strengthened interpretation ('perfected') is not predicted by the standard semantics.

In his account of conditional perfection, von Stechow (2001, 2009) argued that strengthening in a conditional *if p, q* is an implicature triggered when the conditional is considered to provide an exhaustive answer to the question of how to bring about the consequent *q*. The intuition, which von Stechow noted already in Cornulier (1983), is the following: conditional perfection arises in utterance situations in which it is understood that if conditions other than the antecedent (that is, sufficient conditions other than *p*) existed for the consequent *q*, they would have been mentioned. The fact that they haven't leads to the inference that the antecedent *p* provides the 'only' conditions (sufficient and necessary) leading to the consequent *q*. von Stechow's proposal puts together

a view according to which conditionals answer questions under discussion (QUDs) in the discourse situation (which may be implicit) (Roberts 2012; Büring 2003)) together with an account of questions-answer relations that allows answers to be understood exhaustively. Thus, when the QUD is understood to be a question about the conditions for the consequent and the antecedent is understood as an exhaustive answer, the result is a perfection of the interpretation to an ‘if and only if’ meaning. Consider the example below:

- (18) QUD_(implicit): Under which conditions will you give me five dollars?
If you mow the lawn, I’ll give you five dollars.

As von Stechow points out, this proposal predicts a typology of cases in which perfection does arise and cases in which it does not. If a conditional is not understood as an answer to a QUD regarding the conditions for the consequent, the prediction is that perfection will not arise. Von Stechow illustrates this with an example of a conditional understood as a response to a QUD regarding what follows from the antecedent:

- (19) A: John is in Amherst today.
QUD_(implicit): What (of current interest) follows from John’s being in Amherst today?
B: If he is in Amherst, he’ll be home late tonight. (von Stechow 2001)

In examples like (19), in which the QUD is about the consequences, we do not expect a perfected interpretation and, as von Stechow notes, we do not obtain it. Another kind of example in which the proposal predicts absence of a perfected interpretation is in those cases in which a conditional is understood as responding to a QUD regarding the conditions for the consequent but in which the question is not understood as requiring an exhaustive answer, receiving instead a ‘mention some’ interpretation. Von Stechow suggests that ‘mention some’ interpretation for questions may contribute towards explaining contexts in which conditionals fail to receive a perfected reading. Consider the following type of example, discussed in von Stechow (2001) following an observation in Lilje (1972):

- (20) Teenager: How can I earn five dollars?
Grandmother: I’ll give you five dollars if you mow the lawn.

In cases like this, it is clear that the question does not require an exhaustive answer, and that mentioning one way of bringing about the consequent may be enough. The QUD seems to receive a ‘mention some’ interpretation and perfection is not predicted nor obtained.

We find perfected interpretations in the domains of *would*-conditionals in those cases in which we understand the conditional to address a QUD regarding what would have brought about the consequent (paraphrasing von Stechow 2001, *what are all the antecedents p such that in all p-worlds the consequent q would have been true?*).

- (21) QUD_(implicit): In what circumstances would you have given me five dollars?
If you had mown the lawn, I would have given you five dollars.

Below are further examples that can easily be understood as receiving a perfected interpretation:

- (22) a. Doctor scolding a patient for not taking medication on time: If you had taken your medication on time, you would have gotten better.
b. Sports commentator evaluating an accident on a racing course: If Jones hadn't tried to overtake Smith, that crash would not have happened.

In spite of general pragmatic differences between *would*-conditionals and indicatives such as (17), it is clear that the conditional interpretation is perfected in (22a)-(22b). The doctor's statement in (22a) is naturally understood as indicating that it was only by taking the medication on time that the patient would have gotten better. And the sports commentator in (22b) is noting that it is only by failing to try to overtake Smith that Jones would have avoided the crash. We will follow von Fintel in the view that perfection in these cases results from understanding the conditional as an exhaustive answer to an (implicit) QUD (e.g. *what would you have had to do in order to get better?*, *under what circumstances would the crash have been avoided?*). A perfected interpretation is naturally understood for our original *would*-conditional examples in (1a), repeated below :

- (23) Doctor: If Jones had taken arsenic yesterday, he would have shown lividity symptoms today.

A doctor's utterance of (23) would lead us to conclude that if Jones had not taken arsenic yesterday, he would not have shown lividity symptoms today (without the poison, he would have been symptom-free). The antecedent is thus naturally taken to provide not only sufficient but also necessary conditions for the consequent. As expected given von Fintel's proposal, if the QUD is not about what would have brought about lividity symptoms, but about the effects of arsenic, the perfected interpretation does not arise. This is illustrated in the context set up in (24):

- (24) Doctor: Arsenic works like all the other poisons. If he had taken arsenic yesterday, he would have shown lividity symptoms today, just as if he had taken cyanide or curare.

The counterfactual in (24) does not lead us to conclude that it is only if he had taken arsenic yesterday that he would have shown lividity symptoms today.

In the proposed account, the possibility of generating a perfected interpretation for a conditional depends on the conditional being understood as an exhaustive answer to an (implicit) QUD regarding the conditions for the consequent. If a conditional cannot be understood as an answer to how a consequent would have been brought about, it will not generate a perfected interpretation. This, we

claim, is an important difference between forward-looking *would*-conditionals and backtrackers. Contrary to what we have seen with forward-looking *would*-conditionals, backtrackers cannot be understood as answering a question regarding how the consequent would have been brought about. This is illustrated below with examples with the special syntax that facilitates backtracking:

- (25) a. QUD_(implicit): In what circumstances would the plane have to have departed at 1:00?
If the plane had arrived at 2:00, it would have to have departed at 1:00.
- b. QUD_(implicit): In what circumstances would the die have to have been thrown differently?
If the die had fallen six uppermost, it would have to have been thrown differently.

The discourses in (25a) and (25b) do not make coherent question-answer pairs. We cannot understand the backtrackers as providing an answer to a question about the circumstances that would have brought about the (modalized) consequent.⁷ We can see this with our other examples:

- (26) a. QUD_(implicit): In what circumstances would there have to have been no quarrel yesterday?
If Jim had asked Jack for help today, there would have to have been no quarrel yesterday.
- b. QUD_(implicit): In what circumstances would it have been the case that Jones would have to have taken arsenic yesterday?
If Jones had shown lividity symptoms today, he would have to have taken arsenic yesterday.

We again see that the discourses with backtrackers do not provide natural question-answer pairs to QUDs about the conditions for the (modalized) consequent. Since the conditionals cannot be understood as providing an exhaustive answer to a question regarding the conditions for the consequent, the prediction is that perfection will not arise. As we will see in the next section, perfection is an important ingredient in the felicitous cancellation of counterfactuality. Differences in perfection will lie at the heart of the contrast between (1a) and (1b).

5. Cancelling counterfactuality

We are now in a position to address the contrast between (1a) and (1b). We have assembled our main ingredients: (i) counterfactuality in *would*-conditionals is an implicature that cannot be cancelled trivially ('for no reason') (from Section 2); (ii) backtracking *would*-conditionals are special (from Section 3); and (iii) forward-looking conditionals and backtrackers differ in terms of perfection implicatures (from Section 4). With these in hand, we will be able to explain (1a) vs.

⁷We do not have an account of why this should be so. We could speculate that the QUD is understood as a question about 'causes', and backtrackers do not provide information about causes, but further work is needed.

(1b) and, as we will see, shed some light on counterfactuality in future-shifted *would*-conditionals more generally.

5.1. In forward-looking *would*-conditionals

Let us examine again the Anderson example:

- (27) Doctor: If Jones had taken arsenic yesterday, he would have shown lividity symptoms today. He does show lividity symptoms today. So he probably took arsenic yesterday.

What is the information provided by the *would*-conditional in (27)? In addition to the information provided by the truth-conditions of the conditional (28i), the interpretation is enriched with two implicatures: the counterfactuality implicature (28ii), and the perfection implicature that only if the antecedent is true will the consequent be true (28iii). Given this information, the utterance of the conditional leads to the inference that the consequent is false (28iv) (that is, $\rightsquigarrow \neg\beta$). But then the speaker in (27) continues by asserting the consequent (β). This must lead to a revision of the previous step and there are two options available: cancelling counterfactuality or cancelling perfection.

$$(28) \left[\begin{array}{l} \text{i. } \alpha \Rightarrow \beta \\ \text{ii. } \rightsquigarrow \neg\alpha \\ \text{iii. } \rightsquigarrow \text{ only if } \alpha \Rightarrow \beta \end{array} \right] \text{ iv. } \rightsquigarrow \neg\beta \left. + \beta \text{ hence } \left\{ \begin{array}{l} \text{cancel } \rightsquigarrow \neg\alpha \\ \text{or} \\ \text{cancel } \rightsquigarrow \text{ only if } \alpha \Rightarrow \beta \end{array} \right\} \right.$$

The assertion that Jones probably took arsenic yesterday reflects this dilemma. The speaker does not simply reject counterfactuality, aware of the possibility that perfection may be at fault (there may have been independent conditions that would have lead up to the consequent).⁸ It is important to highlight that counterfactuality in this case has not been cancelled in a trivial manner, ‘for no reason’: given perfection, counterfactuality leads to a contradiction.

In the above discussion we have made the assumption that the utterance of the conditionals triggers conditional perfection. This is a crucial component of our account of why it is possible to cancel counterfactuality. The reason why following the conditional in (27) with the consequent (β) requires revision (and eventually leads to cancelling counterfactuality) is because of the interaction between conditional perfection and counterfactuality. The prediction is that if perfection is missing, Anderson-style sequences of ‘detective reasoning’ would not be felicitous. This is borne

⁸Superficially, there seems to be an asymmetry between the two implicatures, in that we consider it probable that counterfactuality is false, as opposed to considering it probable that perfection was at fault. An investigation of preferences in this domain remains for future work.

out by the data. If, for example, as illustrated in (29), world knowledge weakens the perfection implicature, it becomes much trickier to cancel counterfactuality:

- (29) Doctor: If Jones had taken arsenic yesterday, he would have high blood pressure now. He does have high blood pressure now. # So he probably took arsenic yesterday.

It would probably be odd for a doctor to utter (29). It is well-known (and salient) that high blood pressure may result from many different conditions. We do not (obviously) understand the conditional in (29) as perfected, and cancelling counterfactuality is not straightforward. Similarly, if we overtly set up a context in which it is clear that perfection will not be an option, counterfactuality is again difficult to cancel:

- (30) Doctor: All poisons work the same. If he had taken arsenic yesterday, he would have shown lividity symptoms today, just as if he had taken cyanide or curare. He is actually showing lividity symptoms today. # So he probably took arsenic.

Contrary to what we saw in (27), the discourse in (30) is not obviously coherent. In the absence of perfection, the truth of the consequent is not informative with respect to the status of the antecedent. It does not lead to a revision of the counterfactuality implicature (since, in the absence of perfection, it does not generate contradictions).

5.2. In backtracking *would*-conditionals

We turn now to the backtracking Anderson example, repeated below:

- (31) If Jones had shown lividity symptoms today, he would have to have taken arsenic yesterday. He did take arsenic yesterday, # so he probably shows lividity symptoms today.

As noted before, perfection is absent in backtracking examples. This means that the conditional will not give rise to the strengthening inference (\rightsquigarrow only if $\alpha \Rightarrow \beta$). There is, however, a kind of particular ‘strength’ associated with backtrackers. As we noted earlier, in the case of backtracking examples like (31), backtracking is facilitated by the modalized consequent. The truth conditional import of backtrackers is not just $\alpha \Rightarrow \beta$, but, in more detail, $\alpha \Rightarrow \Box\beta$ (32i). The assertion of this conditional appeals to the lawlike dependency between α and β . Given the relevant laws, α brings about β . This means that there is also in (31) a necessary condition associated with the *would*-conditional: given the laws, it is only if β is true that α will be true (32iii). But, contrary to what we saw in (27), this is not a case of pragmatic strengthening, but an entailment that follows

from the lawlike dependency between α and β . In the absence of the perfection implicature, the utterance of β following the conditional in (27) does not give rise to any inconsistency and thus revision is not justified.

$$(32) \left[\begin{array}{l} \text{i. } \alpha \Rightarrow \Box\beta \\ \text{ii. } \rightsquigarrow \neg\alpha \\ \text{iii. Only if}_{LAW} \beta \rightarrow \alpha \end{array} \right] + \beta$$

There is thus no ‘reason’ to cancel the counterfactuality implicature. Given the modalized consequent and absence of perfection, the utterance of β carries no consequences for the information provided by the *would*-conditional. In terms of the information provided at that point, there is no reason to cancel counterfactuality. It would be bad discourse-manners to do so. We are not allowed.⁹

5.3. In future-shifted conditionals

Future-shifted *would*-conditionals have been subject of much interest in the literature (e.g. Ogihara 2000; Ippolito 2003, 2013; Arregui 2005b, 2007), with perfect aspect examples like (8b) receiving particular attention. Below is an illustrative future-shifted past perfect *would*-conditional:

- (33) If Charlie had taken his Advanced Italian test tomorrow, he would have passed. (Ippolito 2003)

Suppose that Charlie took his Advanced Italian test yesterday and failed. He could well have benefitted from extra time to study. In such a context, we could utter (33) to claim that if he had taken the exam tomorrow instead of yesterday, he would have passed. In proposing an analysis examples like (33), Ippolito follows Ogihara (2000) in claiming that counterfactuality in these cases is not defeasible. Ippolito illustrates the point with examples like (34):

- (34) #If Charlie had gone to Boston by train tomorrow, Lucy would have found in his pocket the ticket that she in fact found. So, he must be going to Boston by train tomorrow.

⁹At this point we would like to remind our readers of our earlier observation that backtracking is sometimes allowed without special syntax. Following Arregui (2005a, b); Schulz (2007), these cases can be characterized as invoking an ‘analytic’ or necessary relation between antecedent and consequent. For example (inspired by Frank 1996):

- (1) If she were 30 years old today, she would have been born in 1985.

Counterfactuality is hard to cancel in these examples too. A continuation with: *She was born in 1985 so she is probably 30 years old today* is odd. This can be explained along the same lines as the examples above, since backtrackers do not give rise to perfection, so the assertion of the consequent does not generate a contradiction.

There is, as Ippolito points out, the possibility of a confound when faced with examples like (34). In this case we are not only attempting to cancel counterfactuality in a future-shifted conditional, we are also cancelling counterfactuality in a backtracking conditional. Ippolito considers that this confound is not actually problematic, since there are many examples of backtracking conditionals which are acceptable. The difficulty really is with cancellation in future-shifted past perfect examples: *My claim is that what makes ((34)) infelicitous is the attempt to cancel its counterfactuality. Why are past perfect nonfuture subjunctive conditionals and past perfect future counterfactuals so different?* (Ippolito 2013: 26) However, having noted that many backtrackers are acceptable, Ippolito does not actually consider the possibility that what is problematic about examples like (34) is the attempt to cancel counterfactuality in a backtracker, as opposed to a problem in cancelling counterfactuality in a future-shifted past perfect conditional. We have shown that, independently of what happens with future-shifted conditionals, counterfactuality in backtrackers cannot be cancelled. This alone would account for examples like (34). Do we still need to address counterfactuality in future-shifted past perfect conditionals as something special?

We would like to argue that cancelling counterfactuality in future-shifted past perfect conditionals is not in itself a problem, and that the difficulties that have been observed arise as secondary issues associated with other things (like backtracking). It will not be possible to make this point with Anderson-style detective reasoning sequences, since it is necessary to consider examples shifted towards the future in which the consequent isn't true at the speech time (for that would lead to backtracking). We need other ways of showing that future-shifted *would*-conditionals do not entail counterfactuality. A proposal for an argument of this kind is provided below, following the modus-tollens model of Stalnaker's discussion in (5):

- (35) a. It is very unlikely that Susan will go to Washington next Tuesday. If she had gone next Tuesday, she would have met with Obama on Wednesday. But she has an appointment with him for next Saturday. So she will probably will not go next Tuesday.
- b. I really don't think they will have a baby in 2017. If they had had a baby in 2017, they would have had to pay more taxes in 2017. They are very fiscally minded. I think they will wait another year to save up. But then, you never know!

The small discourses in the examples above actually present arguments in favor of the falsity of the future-shifted *would*-conditional (while at the same time leaving open a small door to the possibility that the antecedent may actually be true). The argumentation against the antecedent would not be expected if it was entailed/undefeasible that the antecedent was false (this was the original point made by Stalnaker's argument in (5)).

We may also turn to Lewis's cross-speaker dialogues in the search of conducive environments in which to cancel counterfactuality in future-shifted *would*-conditionals:

- (36) A: It is such a pity Susan will not come for Christmas this year!
B: Uh?!

A: Well, if she had come for Christmas this year, her mother would have put together an enormous feast.

B: Well, you are right about that!! In fact, she has already started the shopping! Didn't you hear? Susan phoned yesterday to say she will come after all!

If, following Lewis's discussion in (6), we consider that examples like this show that that there is cross-speaker agreement on the *would*-conditional, the exchange shows that such agreement does not depend on agreement regarding the falsehood of the antecedent. This would be unexpected if the counterfactuality implicature in future-shifted conditionals was not cancellable.

6. Conclusion

This paper integrates a view of counterfactuality implicatures with discourse-level considerations. We have shown that discourse structure, in particular pertaining to the relation between conditional assertions and QUDs, affects the possibility of cancelling counterfactuality. We have also argued in support of good discourse manners, showing that the cancellation of counterfactuality implicatures cannot be done 'for no reason'. Along the way we have developed a general view of counterfactuality implicatures that sheds light on some well-known examples in the literature without requiring specific constraints.

References

- Abbott, B. (2006). Where have some presuppositions gone? In B. J. Birner and G. Ward (Eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, Volume 80 of *Studies in Language Companion Series (SLCS)*, pp. 1–20. Philadelphia, PA: John Benjamins.
- Anderson, A. (1951). A note on subjunctive and counterfactual conditionals. *Analysis* 12, 35–38.
- Arregui, A. (2005a). Layering modalities. Ms. University of Ottawa.
- Arregui, A. (2005b). *On the Accessibility of Possible Worlds*. Ph. D. thesis, University of Massachusetts Amherst, Amherst. GLSA publications.
- Arregui, A. (2007). When Aspect Matters: the Case of *Would*-conditionals. *Natural Language Semantics* 15, 221–264.
- Bennett, J. (1984). Counterfactuals and temporal direction. *Philosophical Review* 93, 57–91.
- Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Clarendon Press.
- Biezma, M., D. Siddiqi, and A. Carnie (2013). Counterfactuality in non-standard subjunctive conditionals. In H.-L. Huang, E. Poole, and A. Rysling (Eds.), *Proceedings of North Eastern Linguistics Society 43rd*, Volume 1, pp. 25–36.
- Büring, D. (2003). On D-Trees, Beans, and B-Accents. *Linguistics & Philosophy* 26(5), 511–545.
- Chisholm, R. M. (1946). The contrary-to-fact conditional. *Mind* 55(220), 289–307.
- Cornulier, B. (1983). If and the presumption of exhaustivity. *Journal of Pragmatics* 7, 247–249.
- Davis, W. (1979). Indicative and subjunctive conditionals. *Philosophical Review* 88, 544–564.

- von Fintel, K. (1997). The presupposition of subjunctive conditionals. In U. Sauerland and O. Percus (Eds.), *The interpretive tract, MITWPL 25*, pp. 29–44. Cambridge, MA: MIT Press.
- von Fintel, K. (2001). Conditional strengthening: a case study in implicature. Manuscript, MIT.
- von Fintel, K. (2009). Conditional strengthening, again. Invited talk at XPRAG 2009; Lyon, France; April 24, 2009.
- Frank, A. (1996). *Context Dependence in Modal Constructions*. PhD thesis, Institut für maschinelle Sprachverarbeitung, Universität Stuttgart, Stuttgart.
- Geis, M. L. and A. M. Zwicky (1971). On invited inferences. *Linguistic Inquiry* 2(4), 561–566.
- Heim, I. (1991). Artikel und Definitheit. In A. von Stechow and D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, pp. 487–535. Berlin: De Gruyter.
- Iatridou, S. (2000). The Grammatical Ingredients of Counterfactuality. *Linguistic Inquiry* 31, 231–280.
- Ippolito, M. (2003). Presuppositions and implicatures in counterfactuals. *Natural Language Semantics* 11(2), 145–186.
- Ippolito, M. (2013). *Subjunctive Conditionals*. MIT Press.
- Kasper, W. (1992). Presuppositions, composition, and simple subjunctives. *Journal of Semantics* 9, 307–331.
- Leahy, B. (2011). Presuppositions and antipresuppositions in conditionals. In D. L. Neil Ashton, Anca Chereches (Ed.), *Proceedings of the 21st Semantics and Linguistics Theory*, pp. 257–274. elanguage.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.
- Lewis, D. (1979). Counterfactual dependence and time’s arrow. *Noûs* 13(4), 455–476.
- Lilje, G. W. (1972). Uninvited inferences. *Linguistic Inquiry* 3(4), 540–542.
- Martin, F. (2015). The imperfective in subjunctive conditionals: Fake or real aspect? In T. Brochhagen, F. Roelofsen, and N. Theiler (Eds.), *Proceedings of the 20th Amsterdam Colloquium*, pp. 266–275.
- Ogihara, T. (2000). Counterfactuals, temporal adverbs, and association with focus. In *Semantics and Linguistic Theory Semantics and Linguistic theory (SALT) 19*, Cornell University, Ithaca, N.Y., pp. 115–131. CLS Publications.
- Roberts, C. (2012, December). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6), 1–69.
- Schulz, K. (2007). *Minimal Models in Semantics and Pragmatics: Free Choice, Exhasutivity and Conditionals*. Ph. D. thesis, ILLC, University of Amsterdam.
- Slote, M. A. (1978). Time in counterfactuals. *Philosophical Review* 87, 3–27.
- Stalnaker, R. (1975). Indicative Conditionals. *Philosophia* 5(3), 269–86.