Longer draft of article that appeared in the "Wiley Blackwell Companion to Semantics", Gutzmann, Matthewson, Meier, Rullmann and Zimmermann (eds.), Wiley Blackwell, 2020.

Counterfactuals:
If kangaroos had no tails…..

Ana Arregui

https://onlinelibrary.wiley.com/doi/book/10.1002/9781118788516
Print ISBN: 9781118788318| Online ISBN: 9781118788516| DOI: 10.1002/9781118788516

1. Introduction

David Lewis begins his famous book *Counterfactuals* (Lewis 1973) with the following sentence:

(1) If kangaroos had not tails, they would topple over.

He continues with the claim that the sentence means something like this:

(2) "In any possible state of affairs in which kangaroos have no tails, which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over." (Lewis 1973: 1)

The example in (1) is considered a prototypical example of a counterfactual conditional and analyses of counterfactuals building on the account sketched in (2) have had a profound influence on the linguistic literature on counterfactuals. This chapter provides a survey of some of the issues in this domain by exploring key examples that have shaped current debates. The term 'counterfactual' will be used to talk in a narrow sense about counterfactual conditionals, though in the linguistic literature the term is sometimes used more broadly to include other constructions that invoke non-actual circumstances (such as *wish*-constructions). Occasionally, I will use the term '*would*-conditional' instead of 'counterfactual' to avoid confusion. The term counterfactual is also sometimes used to talk about non-actual possible worlds (*counterfactual worlds*), in particular when it is known that they are non-actual.

To quote another opening sentence, this time from Goodman (1947), *"The analysis of counterfactual conditionals is no fussy little grammatical exercise."* Counterfactuals sit at the cross-roads of a wide set of theoretical interests. Current linguistic debates on counterfactuals address problems in compositionality and the syntax-semantics interface, theories of context change, cross-linguistic variation in the construction of modal meanings, presupposition projection, scalar implicatures, the semantic ontology relevant for natural language semantics, and the interaction between temporal meanings and modal meanings, amongst others. This is a literature that has crucial interaction with ongoing debates regarding counterfactuals in the philosophy, where many of the discussions originated, as well as important connections with cognitive sciences. As Kratzer (2012) notes: "*Investigating the laws of counterfactual reasoning is an empirical enterprise. It is no longer just a logical endeavor".* Studying counterfactuals we come to understand the devices that languages place at our disposal to access remote possibilities and follow through to their conclusions.

This chapter provides an introduction to some of the key data discussed in the literature on counterfactuals. The aim has not been to provide a thorough theoretical overview, but rather to focus on a wide range of examples, circumventing most technical details. As we will see, the literature on counterfactuals is (perhaps unavoidably) fragmented, with numerous proposals that tackle a narrow set of problems and sidestep all other issues. The literature on counterfactuals is also rather vast, and this survey has had (unavoidably) to set many topics aside. These include *might* counterfactuals (barely mentioned), an overview discussion of dynamic approaches to counterfactuals (nonexistant) and the interaction between counterfactuals and topics that concern conditionals and modality more broadly, such as donkey anaphora, gradability, the syntax-semantics interface, etc. (not

mentioned at all). There are excellent overviews of counterfactuals (and conditionals more broadly) that have provided inspiration for this chapter and would doubtlessly be of interest to readers wishing to follow up on this topic, including Bennett (2003), Edgington (1995), von Fintel (2011, 2012), etc. Readers looking for an overview of syntactic matters are referred to Bhat and Pancheva (2006).

The most important philosophical influences on current linguistic research on counterfactuals have arguably been the proposals by Robert Stalnaker (e.g. 1968, 1975, 1984) and David Lewis (e.g. 1973, 1979). Section 4 in this chapter provides a brief introduction to the empirical generalizations and issues addressed by their work, as well as some of the discussion that predated and set the stage for their proposals, and discussions that have followed challenging some of their views. Sections 2 and 3 provide a broad frame for the debate: Section 2 examines possible criteria for the classification of counterfactuals and Section 3 investigates some of their logical properties. Section 5 provides an introduction to the interpretation of counterfactuals in the framework of premise semantics. Section 6 investigates some current debates surrounding the interpretation of temporal morphology in counterfactuals, object of much current interest. Section 7 presents some concluding remarks.

2. Counterfactuals as a type of conditional

Let us begin by pointing out that the classification of counterfactuals is problematic. The example in (1) is usually considered to be a prototypical example of a counterfactual, but there isn't actually agreement as to the criteria to classify it as such nor on the examples that should be grouped together in this category. Indeed, the literature offers remarkable variability. As a way of getting started, we will discuss three possible criteria to identify counterfactuals: (a) a distinct semantics, (b) a distinct morphology, (c) a distinct status for the antecedent (i.e. false). None of them will prove conclusive, but the discussion will help set the stage for what is to come.

2.1 Semantics

We might consider distinguishing counterfactuals as a type on the basis of their semantics. From this perspective, counterfactuals would be a kind of conditional with a specific interpretation, different from the interpretation of other conditionals. This would be in opposition to a unified view that aimed to account for the interpretation of counterfactuals on the basis of a general semantics for conditionals, enriched perhaps with a 'counterfactual twist' independent of the core interpretation. There is, however, no consensus in this matter. The philosophical literature, where much of the discussion has originated, is divided on this issue, and the linguistic literature too.

At the level of intuitions about interpretation, there can be clear differences between examples we would in principle wish to classify as counterfactuals and other types of conditionals. For example, we perceive a contrast between a possible utterance of counterfactual *If kangaroos had not tails, they would topple over* (which could be construed as resulting from careful deliberation about the geometry of kangaroos and gravity) vs. indicative *If kangaroos have no tails, they will topple over* (which would probably be construed as resulting from surprising ignorance). Interpretative difference between counterfactuals and indicatives are brought out forcefully in the famous examples in (3), which were offered by Adams 1970, and related variants such as (4) (cited by von Fintel 2012):

(3) a. If Oswald didn't kill Kennedy, somebody else did.
b. If Oswald had not killed Kennedy, somebody else would have.
(Adams 1970: 90)

(4) a. If Shakespeare didn't write Hamlet, somebody else did.
b. If Shakespeare hadn't written Hamlet, somebody else would have.

(von Fintel 2012: 466)

Knowing that US president Kennedy was killed in 1963 and that Lee Harvey Oswald was arrested, our intuition is to agree that one would be justified in uttering (3a). But we would probably not agree in the case of (3b) (unless, for example, we suspect there was a conspiracy to kill Kennedy involving another gunman). Similarly, knowing that the play Hamlet is part of the patrimony of world literature, we may agree with (4a) while disagreeing with (4b). Differences such as these are not limited to the past. Edgington (1995) has argued that comparable differences can also be found with conditionals that make hypotheses the future. Consider the following example from Edgington: we are sure that one of two prisoners, Smith or Jones, will try to escape tonight. We know that Smith is very unadventurous whereas Jones is very daring and persistent, leading us to expect that it is Jones who will try to escape. As Edgington notes, in this scenario we would not accept (5a), but would be willing to accept (5b):

(5) a If Jones were not to try to escape tonight, Smith would.
b. If Jones doesn't try to escape tonight, Smith will. (Edgington 1995: 239)

Contrasts like these have been used to argue for a specialized semantics for counterfactuals distinct from that of indicatives. Indeed, Lewis (1973) explicitly appealed to (3) to argue against a unified account of conditionals. Commenting on the examples, Lewis claimed *"Therefore there really are two different sorts of conditional; not a single conditional that can appear as indicative or counterfactual depending on the speaker's opinion about the truth of the antecedent."* (Lewis 1973: 3). However, not everyone has been moved to support the view that two different semantic analyses are required, with authors varying in terms of the reasoning supporting a unified analysis and their views on the classification itself. One kind of reason that might be put forward in favour of a unified analysis of counterfactuals and indicatives is the observation that there are examples in which the difference appears to be (merely?) a temporal difference rather than a difference in core semantics. Bennett (2003) (who actually disfavours this view) discusses the following pair: imagine that I (rightly) tell you (6a), you do not swim in the sea, and the following day I tell you (6b):

(6) a. If you swim in the sea today, your cold will get worse.
b. If you had swum in the sea today, your cold would have gotten worse.

Intuitively, it may seem reasonable to think that the two conditionals 'stand and fall together', with the difference being merely one of temporal perspective. But this view is difficult to maintain. Bennett (2003, Chapter 15) points to a number of problems, including the following variant of a coin-tossing scenario: a coin is tossed randomly and comes up heads. You tell me: "*If you had bet on heads, you would have won*". This may well be true even if at no time in the past, before the tossing of the coin, could somebody reasonably have accepted "*If you bet on heads, you will win*" (for a discussion of the role of tense and temporal operators in counterfactuals, see Section 6). Another kind of reason that has been put forward supporting a unified analysis is the fact that counterfactuals and indicatives share many logical properties, licensing common patterns of inference (Stalnaker 1968, see Section 3 for discussion). In fact, according to Stalnaker's early proposal, we follow a common strategy in evaluating conditionals in general, one that abstracts away from our beliefs about the truth or falsity of the antecedent. Adapting 'Ramsey's test', Stalnaker formulates the strategy as follows:

(7) First, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally, consider whether or not the consequent is then true.
(Stalnaker 1968: 102)

Stalnaker (1968) appealed to this common strategy in the domain of beliefs to support a unified account of truth-conditions for conditionals in the framework of possible-worlds semantics. Adopting a unified semantics, Stalnaker considered that the difference between examples like (3a) and (3b) was pragmatic.[1] The unified conditional semantics underlying both cases appealed to quantification over possible worlds, but whereas in the case of (3a) quantification was restricted to worlds compatible with the presupposed information in the conversational common ground (a pragmatic default), in the case of (3b) the domain of quantification could include possibilities known to be incompatible with the common ground. Indeed, according to Stalnaker, the distinctive morphological markings observed in examples like (3a) fulfil a pragmatic function by indicating a potentially wider domain of quantification for the conditional:

(8) I take it that the subjunctive mood in English and some other languages is a conventional device for indicating that presuppositions are being suspended, which means in the case of subjunctive *conditional* statements, that the selection function is one that may reach outside the context set.
(Stalnaker 1975: 276)

The distinctive morphology observed in (1) and (3a) thus becomes necessary when the antecedent of a conditional is incompatible with the information presupposed in the conversational common ground. But this is not a result of distinctive semantics for these conditionals, it is rather a pragmatic matter of conventions regarding how to identify domains of quantification (see also Section 2.3). Von Fintel (1998) argues for an elaboration of Stalnaker's views on the indicative/subjunctive distinction according to which the indicative is semantically vacuous, putting no specific constraints on the domain of quantification of the counterfactual. The domain of quantification is considered, by default, to be included within the context set. The subjunctive is the pragmatically marked case. The role of subjunctive is to signal *"that there is a domain of quantification which contains at least some worlds outside the context set*" (von Fintel 1998: 41). This allows non-actual worlds to be considered if reasons to do so arise (i.e. if the antecedent is false or if there are other reasons to widen the domain of quantification). A Stalnaker-inspired approach to the indicative/subjunctive distinction is defended also in Schlenker (2004), who argues for an analysis of *if*-clauses in terms of definite descriptions. According to Schlenker, in indicative conditionals the *if*-clause must denote a world in the context set, carrying a presupposition that indicates closeness similar to that of the demonstrative *this*: *"From the present perspective, this is just to say that indicative mood expresses a presupposition similar to that of the word "this", but in the domain of worlds rather than of individuals. The notion 'close to the context of utterance' is rendered, following Stalnaker, as: 'within the Context Set'."* (Schlenker 2004). In subjunctive conditionals, *if*-clauses are allowed to denote a world outside the context set in a manner comparable to the presupposition associated with *that* that the referent isn't close to the speaker.

2.2 Morphology

Considering examples like (1), it may appear that morpho-syntax provides adequate criteria for the classification of counterfactuals. What appears special about counterfactuals is the presence of the modal *would* as well as apparently odd behavior for tense and aspect morphology in the antecedent clause: whereas the antecedent surfaces with past verbal morphology, it can be interpreted as making a hypothesis about the present. However, morphology has not provided an ultimately satisfying solution to the classification problem.

There is a tradition in the philosophical literature to use the term 'subjunctive conditionals' to refer to conditionals headed by *would*, and some authors even use the term 'subjunctive'

[1] But see Stalnaker (1984, Chapter 6) for alternative discussion and Bennett (2003: 358-361) for comparison of views.

interchangeably with 'counterfactual'. In this tradition, the verbal morphology we see in the antecedents of counterfactuals like (1) is characterized as subjunctive mood associated with the modal, while tense is ignored. Examples with third person singular morphology have traditionally been brought up in support of this view, since the antecedents of *would*-conditionals sometimes bear remnants of subjunctive morphology which can actually be distinguished from indicative, e.g. "*If he were here, he would be delighted*". Further intuitive justification for the subjunctive view may come from the common conception that subjunctive mood is, in some sense, 'irrealis', invoking circumstances that are known to be non-actual.

There are problems with this view. One concern is that even if we granted that conditionals headed by *would* were subjunctive, not all *would*-conditionals fall under the scope that we would intuitively wish to assign to the label 'counterfactual'. This point was made early on by Lewis, who discussed empirical problems for extending his proposal for counterfactuals to what he called 'subjunctive conditionals' in general:

(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 can see the point of Lewis's argument by putting the example in context:

(10) A: Our troops will be recalled next year.
B: That's a good thing. #If our troops entered Laos next year, there would be trouble.
B': That's a good thing. If our troops had entered Laos next year, there would have been brouble.
(Arregui 2005: 10)

As the unacceptability of B indicates, *would* by itself does not guarantee counterfactuality, the other morphology in the conditional matters (see Section 6).

There is converging agreement that the verbal morphology we see in so-called 'subjunctive' conditionals in English, including counterfactuals, is temporal morphology and not subjunctive mood morphology (in spite of disagreement regarding its compositional role in the interpretation, see Section 6). As *The Cambridge Grammar* notes, it would be very implausible to assume that the morphology we find in the antecedents of counterfactuals is subjunctive/irrealis (as opposed to past tense), since the only case in which we actually see a difference with past indicative is in the *was/were* alternation. *"It is much more plausible to say that irrealis 'were' is an unstable remnant of an earlier system – a system which has otherwise been replaced by one in which the preterite has expanded its use in such a way that it now serves to express modal remoteness as well as past time."* (The Cambridge Grammar of the English Language 2002: 88). Indirect evidence against the view that English counterfactuals have subjunctive morphology is provided by cross-linguistic considerations. As has been noted by Iatridou (2000), languages that do have independently identifiable and morphologically robust subjunctive paradigms often disallow subjunctive in counterfactuals. Iatridou illustrates this with French, where the antecedent of a counterfactual making a present hypothesis has past indicative morphology, and not present subjunctive, in spite of the fact that present subjunctive occurs in many other constructions:

(11) Si Marie avait /*ait un parapluie rouge, ….
If Marie have-PastInd / have-PresSubj a red umbrella,
'If Marie had a red umbrella, …..'
(Iatridou 2000: 265)

Examples like this show that languages with fully developed subjunctive paradigms do not require it in counterfactuals, in the case of French choosing a past indicative form to make a hypothesis about the present. This weakens the rationale to characterize English *would*-conditionals as subjunctives instead of simply past tense with a 'non standard' interpretation. Further support for the view that what we see in English *would*-conditionals is actually tense and aspect morphology with an (apparently) non-standard interpretation is provided indirectly by languages such as Spanish. Spanish also has a clearly identified subjunctive paradigm, with clear present and past tense alternations. Contrary to what we see in French, subjunctive morphology is obligatory in counterfactuals, but the morphology also displays non-standard tense behaviour. We can see this in the examples below. The sentence in (12) illustrates standard tense behavior in subjunctive embedded under the attitude verb 'want/hope'. Subjunctive is obligatory in this context and the present vs. past alternation is reflected in the temporal anchoring of the hoped-for proposition. The sentence in (13) illustrates the antecedent of a counterfactual. Subjunctive is again obligatory, but present subjunctive is clearly ungrammatical. In this context, past subjunctive morphology results in an interpretation that allows the antecedent proposition to be anchored on the speech time, reproducing the temporal puzzle observed in (1):

(12) Espero que Kiyomi esté / estuviera en su casa.
Hope that Kiyomi be-PresSubj / be-PastSubj in her house
'I hope Kiyomi is/was at home'

(13) Si Kiyomi *esté/ estuviera (ahora) en su casa, …..
If Kiyomi *be-PresSub / be-PastSub (now) in her home,…..
'If Kiyomi was (were) (now) in her home, …..'
(Arregui 2005)

The examples show that the problem of accounting for non-standard interpretations of temporal morphology remains even in the case of languages that transparently require subjunctives in counterfactuals. Appealing to mood does not really allow us to bypass the interpretative problems posed by past tense.

Languages actually differ rather widely in terms of the morphological make-up of sentences we would intuitively wish to characterize as counterfactuals, displaying a variety of strategies. It is not possible to characterize counterfactuals in terms of a specific cross-linguistically stable surface morpho-syntactic property. Some languages, as we have seen, use combinations of tense and aspect morphology which, in conjunction with special complementizers or modals, can appear to receive non-standard interpretations in counterfactuals. English, French and Spanish are examples of this kind. Additional examples from Zulu and Korean are given below as illustration:

(14) [ukuba be- ngi- thimul- ile] be-ngi-zo-dinga ithishi — Zulu
if past.impf- 1sg- sneeze- pfv impf-1sg-fut-need 5tissue
'If I had sneezed, I would have needed a tissue.'
(Bronwyn & Halpert quoting Halpert and Karawani 2012)

(15) Kim-i tap-ul alkoiss-ess-(ess)-tamyen, ne-eykey — Korean
Kim-Nom answer-Acc know-Past-Past-if you-to
malhaycwu-ess-ul kesi-ta.
tell-Past Fut-Decl
'If Kim had known the answer, he would have told you.'
(Han 2006: 173)

Other languages have dedicated markers for counterfactuality (see Nevins 2002). Ippolito and Su (2014), for example, investigate two conditional markers in Chinese: *yaoshi* (if) and *yaobushi* (if-not).

While there are no tense or aspect distinctions between the two types of conditionals, they do differ in meaning: *yaoshi*-conditionals are 'ambiguous' between so-called indicative and counterfactual interpretations but *yaobushi*-conditionals, with negation *bu* present in the complementizer, are obligatorily counterfactual.

(16) yaoshi ta zuotian you renhe wuqi, haiguan jiu hui kouliu ta.
YAOSHI he yesterday have any weapon Customs then will detain he '
'If he had any weapons yesterday, Customs detained him.'
' 'If he had had any weapons yesterday, Customs would have detained him.'

(17) yaobushi ta qunian qu jianada, ta jiu hui geng ni jiehuan.
YAOBUSHI he last.year go Canada he then will with you get.married
'If he had not gone to Canada last year, he would have married you.'
(Ippolito and Su 2014)

There has been much recent interest in the typological study of counterfactuals, greatly inspired by Iatridou (2000) (see Section 6.1).

2.3 Counterfactuality

Considering examples like (1), it may be tempting to characterize counterfactuals as conditionals that carry the presupposition that the antecedent clause is false. However, as the famous example from Anderson (1951) in (18) illustrates, *would*-conditionals are not necessarily interpreted as having false antecedents.

(18) 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)

In examples like this, counterfactuals are used in a kind of 'detective reasoning' to actually provide support for the truth of the antecedent. Edgington (1995: 240) makes the same point with the example in (19) (where the counterfactual is used to draw an inference 'to the best explanation'):

(19) A: A bus is coming.
B: How do you know? (for we can't see the oncoming traffic.)
A: People in line are picking up their bags and inching forward – *and that's what they would be doing if a bus were coming.*

Examples like these show that counterfactuality of the antecedent is defeasible and argue against the view that it is presupposed. An additional argument against counterfactuality as a presupposition is provided by Stalnaker (1975). He puts forward an example of a modus tollens argument (see Section 3), noting that if the antecedent were presupposed to be false, the argument would be rendered pointless, counter to our intuitions:

(20) 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)

Stalnaker argues that in examples like this the evaluation of the antecedent requires considering worlds incompatible with the common ground, motivating the use of 'subjunctive' morphology (see Section 2.1).

The observation that counterfactuality in counterfactuals is defeasible has led to the common view that it is actually an implicature. This is compatible with the classic proposals for counterfactuals put forward by Lewis and Stalnaker (see Section 4), where the semantics does not require that the antecedent be false (i.e. a counterfactual with a true antecedent can be true). An early discussion of how the counterfactuality implicature is drawn can be found in Iatridou (2000), where it is linked to the choice of morphology that carries an exclusion feature (see section 6.1). An alternative is found in Leahy (2011), who builds on Stalnaker (1975) and argues that the counterfactuality implicature is derived in terms of anti-presuppositions generated by the interaction between indicatives and counterfactuals (see e.g. Percus 2006, Chemla 2008 for discussion of antipresuppositions).

Given an implicature account of counterfactuality, examples in which counterfactuality proves impossible/difficult to cancel are theoretically of great importance. Various types of examples of this nature have been addressed in the literature. One class of examples which has received particular attention is the case of 'mismatched' counterfactuals in which antecedents with past perfect morphology make hypotheses about future events (Ogihara 2000, Ippolito 2003, ff.). Ogihara (2002) addressed examples like (22), with focused future adverbials:

(22) a. If John had given flowers to Mary TOMORROW$_F$, she would have been pleased.
b. If we had gone out for a walk TOMORROW$_F$, we would have had a good time.
(Ogihara 2000)

Ogihara noted that examples like these appeared to be "*truly counterfactual*". Ogihara's discussion centers on the role of focus, attributing the 'truly counterfactual' interpretation to the fact that focus induces a contrast with an actual eventuality somehow incompatible with the antecedent. In the case of (22a), for example, we may imagine a scenario as follows: Mary's birthday is tomorrow, but her boyfriend John mistakenly thought it was yesterday and gave her flowers then, resulting in Mary's bitter disappointment. An utterance of (22a) contrasts the event described by the antecedent clause with the actual event of John giving Mary flowers yesterday. The truly counterfactual nature of the interpretation is thus claimed to be derived from focus-induced contrast, allowing for a standard, similarity-based, not 'truly counterfactual' interpretation of the conditional (see Ippolito 2003, 2013 for a critical overview, also Ogihara 2013 for comments). Contra Ogihara (2000), Ippolito considers that focus is not a necessary component of mismatched counterfactuals. Ippolito (2003, 2013) investigates the contrast between examples like (23a) and (23b):

(23) Students who registered for the Advanced Italian class had to choose when to take their final test, either last Monday or tomorrow (but not both). Charlie took his final exam last Monday and didn't pass. He would have benefited a lot from extra study and practice, and if he could have taken the exam a few days later, he would have passed.

a. *Future past perfect subjunctive counterfactual*
If Charlie had taken his Advanced Italian test next Monday, he would have passed.
b. *Past simple past subjunctive counterfactual*
#If Charlie took his Advanced Italian test next Monday, he would pass.

Ippolito considers that future past perfect subjunctive counterfactuals must be uttered "*when the speaker regards the hypothetical event described by the antecedent as unrealizable (impossible)"* (Ippolito 2003: 147). They differ from non-future past perfect subjunctive counterfactuals in giving rise to a strong sense

of counterfactuality that is hard/impossible to cancel. Indeed, the attempt to construct Anderson-type examples like (24) with future past perfect subjunctive counterfactuals appears problematic, supporting the view that counterfactuality has a distinct status in this type of examples:

(24) #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 tomorrow.
(Ippolito 2003: 147)

What differentiates the past perfect cases from the simple past is a difference in the satisfaction conditions for presuppositions. According to Ippolito (e.g. Ippolito 2013), whereas simple past subjunctive counterfactuals (like 23b) require that the presuppositions associated with the antecedent be compatible with what is historically possible at the moment of utterance, past perfect subjunctive counterfactuals (like 23a) impose the weaker constraint that the presuppositions be compatible with what was historically possible at some past time: *Since the set of possibilities shrinks over time, being compatible with the set of worlds historically accessible at the utterance time entails being compatible with the set of worlds historically accessible at any time before the utterance time ."* (Ippolito 2013: 90) (see Section 6.2 for more details regarding Ippolito's analysis). Given the asymmetry in the strength of presuppositions, choosing a future past perfect subjunctive counterfactual over a future simple past subjunctive counterfactual gives rise to the anti-presupposition that the antecedent is false. Ippolito considers that anti-presuppositions cannot be cancelled, accounting for the 'strong counterfactuality' intuition associated with examples like (23a) and the observations in (24).[2]

'Mistmatched' counterfactuals are no the only types of examples in which counterfactuality proves hard /impossible to cancel. Iatridou and Embick (1994), for instance, discuss the case of counterfactuals with inverted antecedents, claiming that inversion makes it impossible to cancel counterfactuality. The contrast with non-inverted antecedents is illustrated below:

(25) a. If he had broken his leg in his childhood, which, in fact, he did, he would have exactly this type of scar.
b. #Had he broken his leg in his childhood, which, in fact, he did, he would have exactly this type of scar.
(Iatridou and Embick 1994: 201)

Iatridou and Embick provide an account of the contrast in (25) in terms of information structure: inversion in (25b) signals that at the time of utterance, it is known that the antecedent proposition is false (this is 'old information'). Biezma (2012) offers an alternative in terms of Schwarzchild-style (Schwarzschild 1999) 'givenness', noting that counterfactuality can be cancelled even in inverted conditionals if certain discourse conditions are met. An example is provided below in which A's ignorance would be incompatible with counterfactuality:

(26) A: I wonder if Maria is at the meeting?
B: I just saw John coming out of the conference room smiling.
A: Well, then she is probably not there
B: Why do you say that?
A: Had Maria been there, John would not have been that happy.

[2] Note that whereas both Leahy and Ippolito appeal to antipresuppositions to account for counterfactuality, they consider comparisons amongst different groups of conditionals and develop different accounts of how the antipresuppositions work. For additional discussion of counterfactuality anti-presuppositions, see also Arregui and Biezma (forthcoming).

Another kind of syntactically-marked example in which counterfactuality can be hard to cancel is the case of dialectal variants of counterfactuals with modal morphology in the antecedent clause:

(27) 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, Carnie and Siddiqi 2014)

Attempts to cancel counterfactuality in this kind of examples in accordance with the Anderson recipe fail, as illustrated by the contrast below (for speakers whose dialects include forms such as (28)):

(28) a. If Jones had taken arsenic, he would have shown exactly those symptoms that he does in fact show (so, he probably took arsenic).
b. #If Jones would've/had've taken arsenic, he would have shown exactly those symptoms that he does in fact show (so, he probably took arsenic).
(Biezma, Carnie and Siddiqi 2014)

Biezma et al. account for the contrast in (28) while maintaining an implicature-based account of counterfactuality by proposing principles of pragmatic economy brought into play when speakers choose morpho-syntactically more complex structures. Other examples in which dialectal variation plays a role are found in Dancygier and Sweetser (2005), cited here from Ippolito (2008) (see also Ippolito 2013: 119). The conditionals below appear as part of some American and British dialects:

(29) a. If I hadd-a known you were coming, I would-ha stayed home.
b. If I hadn't a-been ill, I'd a-got him away all right, ….
(Ippolito 2008: 269)

Dancygier and Sweetser report that in examples such as these it is obligatorily understood that the antecedent is false.

Variability in the cancellability of counterfactuality is observed more broadly at the cross-linguistic level. We have already noted the example of Chinese (Ippolito and Su 2014), in which some conditional markers result in an interpretation that is obligatorily counterfactual. Ogihara (2014) also discusses examples in which counterfactuality is obligatory, with data from Japanese. Ogihara argues that the *ta* (past) morpheme in Japanese can receive an irrealis/subjunctive interpretation (in addition to its interpretation as a past tense marker). When interpreted as 'subjunctive', it gives rise to non-cancellable counterfactuality in conditionals. This is illustrated in (30), where attempts to cancel counterfactuality lead to infelicity:

(30) (Mosi) Saburoo-ga ima nihon-ni i-ta ra,
If Saburo-NOM now Japan-at be-PAST RA,
yuumeizin dat-ta(-daroo)-ne.
celebrity be-PAST-probably-
#Zituwa Saburoo-wa ima nihon-ni iru-noda.
Actually Saburo-TOP now Japan-at be-PRES
'If Saburo were in Japan now, (he would) be a celebrity. Actually, Saburo is now in Japan."
(Ogihara 2014: 12)

The example in (31) further illustrates this point:

(31) #(Mosi) kono kanzya-ga gan dat-ta ra,
If this patient-NOM cancer be-PAST RA,
ima (zissai) aru-itami-ga (ima) at-ta daroo.

now (in reality) exist-pain-NOM now exist-PAST probably.
Dakara kare-wa gan-da.
Therefore he-TOP has-cancer.PRES.
[intended] 'If this patient had cancer, he would suffer from the pain that he is now suffering. Thus he has cancer.'
(Ogihara 2014: 14)

Example (31) shows that attempts to reproduce Anderson-style examples with conditionals with the *–ta* marker are claimed to lead to anomaly.

2.4 Conclusions

As the discussion in this section has shown, the category 'counterfactuals' is rather unstable in the literature. While there is agreement regarding some core examples, quite different criteria can be provided and authors may group together diverse examples under this label. Ultimately, variation in the classification responds to differences in theoretical views and readers must remain vigilant.

3. On the logic of counterfactuals

This section presents an introduction to discussions of inference patterns validated by counterfactuals. Intuitions about the inference patterns form part of the empirical domain that a theory of counterfactuals should account for, and inference patterns have long been studied in the philosophical literature in counterfactuals, providing important insights.

The interpretation of conditionals is often discussed in relation to the inference pattern of *Modus Ponens*. This is the pattern that guarantees that if *p* together with a conditional premise *if p, q* are both true, then *q* will also be true. This inference pattern targets some of our basic intuitions about the interpretation of conditional claims and has been adopted in a variety of conditional logics (in the case of propositional logic, for example, Modus Ponens is the inference rule that allows the inference of *q* from *p* together with the material conditional $p \supset q$). In combination with negation, it gives rise to the 'contrapositive' inference that if a conditional *if p, q* is true, and *not q* is true, then *not p* is also true (the rule of *Modus Tollens*). The influential semantic proposals by Stalnaker and Lewis for counterfactual conditionals validate Modus Ponens, predicting the truth of the consequent from the truth of the counterfactual and its antecedent. Intuitions about the pattern, however, can be rather mixed, which is unsurprising given that an utterance of a counterfactual is usually taken to convey the falsehood of its antecedent (see Section 2.3). As Lewis notes, an utterance of a counterfactual with a true antecedent is perceived as misleading, so that: *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). So, if we believe someone who utters the counterfactual *If Oswald had not killed Kennedy, Kennedy would not have been killed*, and then come to learn that it is true that Oswald did not kill Kennedy, it is much more likely that we will conclude that the original counterfactual was actually false rather than conclude that Kennedy was not killed. However, such a response does not present an argument against modus ponens, which makes a claim about the truth of the consequent when both the counterfactual and the antecedent are true. Lewis suggests that we bypass potential confounds regarding the falsehood of the antecedent by considering examples with two speakers who disagree about the truth of the antecedent, but agree about the conditional claim. Someone says: *If Caspar had come, it would have been a good party.* And someone else replies: *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: 27). Examples like this can lend support for the view that counterfactuals validate modus ponens.

A different kind of challenge to modus ponens in counterfactuals could be seen to arise from the 'counterfactual' version of examples provided by McGee (1985) for indicative conditionals.

McGee (1985) famously argued that modus ponens was not valid for indicative conditionals with nested antecedents, concluding thus that modus ponens was not actually valid for indicatives (since it had exceptions). McGee rather more tentatively extended the argument to counterfactuals (aware of the judgment difficulties when the antecedent is true). McGee's examples have the following pattern: we start off with a conditional with a conjunctive antecedent (*if (p & q), r*), then we note that, given the inference rule of *Exportation*, it implies a conditional with a nested antecedent (*if p,* (*then if q, r*)), finally we realize that we are not actually willing to infer *if q, r* from the truth of *p*, leading to the conclusion that modus ponens is not valid. Consider (32), put forth as an illustration of the law of Exportation: according to McGee, we are happy to conclude (32b) given the truth of (32a).

(32) a. If Reagan hadn't won the election and a Republican had won, it would have been Anderson.
b. If Reagan hadn't won the election, then if a Republican had won, it would have been Anderson. (McGee 1985: 467)

Imagine what would actually have been the case if Reagan had not won the elections. In that case, most likely, Carter would have come first and Reagan second. In such circumstances, in which Reagan did not win and the antecedent of (32b) is true, we would probably judge the counterfactual *If a Republican had won the elections, it would have been Anderson* as false. In these circumstances, if a Republican had won the elections, it would have been Reagan! Thus modus ponens fails. Responses to McGee's challenge to modus ponens have been mixed (see Piller 1996 for an overview of some of the arguments). Noting that philosophers remain divided, Gillies (2004) claims that in McGee's counterexamples to modus ponens, *'one feels like the victim of some kind of trick'* (Gillies 2004: 593). We can try to informally extend Gillie's diagnosis of the problem to the case of counterfactuals by noting a tension between using the antecedent (in 32a) as an assumption to reach a certain conclusion and the fact that we actually ignore the assumption in the evaluation of the consequent. This problem is overcome in many current linguistic accounts of counterfactual that would challenge the view that examples like (32b) have the structure of embedded conditionals, arguing instead for a view according to which there are multiple restrictors that together affect the interpretation of a modal operator (see e.g. Kratzer 1991 for a restrictor approach to *if*-clauses). See also Huitink (2012), who presents experimental research on McGee-type examples.

Discussions of the inference patterns validated by counterfactuals often begin by measuring our intuitions against what would be expected if counterfactuals had the rather plausible semantics of a 'strict' conditional. A strict conditional *if p, q* is characterized as true iff q is true at every accessible p-world, given some (relevant) accessibility relation (i.e. if the material conditional is necessarily true). A strict conditional analysis predicts, for example, that the sentence *If kangaroos had not tails, they would topple over*, will be true iff in all accessible possible worlds in which kangaroos have no tails, they topple over. The influential proposals by Stalnaker and Lewis veered away from a strict analysis after noting that it validates inference patterns that appear to be incorrect for counterfactuals. One example is the inference known as 'Strengthening of the Antecedent': if the strict conditional *if p, q* is true, then the strict conditional *if p & r, q* will also be true (an example of monotonic reasoning). Our intuitions about counterfactuals, however, seem to violate this pattern, giving rise to the claim that the semantics of strict conditionals is not actually correct for counterfactuals. An example illustrating this claim is provided below:

(33) a. If Sophie had gone to the parade, she would have seen Pedro.
b. But if Sophie had gone to the parade and been stuck behind a tall person, she would not have seen Pedro.
(Gillies 2007)

Sequences of examples like these have come to be known as *Sobel-sequences* (or *Lewis-Sobel sequences*)

after Lewis (1973) offered the following combinations, noting that J.Howard Sobel had brought the matter to his attention:

(34) If the USA threw its weapons into the sea tomorrow, there would be war; but if the USA and the other nuclear powers all threw their weapons into the sea tomorrow, there would be peace; but if they did so without sufficient precautions against polluting the world's fisheries there would be war; but if, after doing so, they immediately offered generous reparation for the pollution there would be peace; … (Lewis 73: 10)

In addition to the apparent discrepancies observed with antecedent-strengthening, our intuitions about counterfactuals also appear to violate other inference patterns validated by strict conditionals: transitivity ('hypothetical syllogism') and contraposition. Strict conditionals validate transitivity: if the strict conditional *if p, q* is true, and the strict conditional *if q, r* is true, then the strict conditional *if p, r* will also be true. However, our intuitions about counterfactuals do not seem to reflect this fact. In Stalnaker's example below, (35c) is not taken to follow from (35a) and (35b). As Stalnaker (1968: 106) puts it, *'It seems reasonable to affirm these premises and deny the conclusion'.*

(35) a. If J Edgar Hoover were today a communist, then he would be a traitor.
b. If J Edgar Hoover had been born in Russia, then he would today be a communist.
c. If J Edgar Hoover had been born in Russia, he would be a traitor.

Lewis (1973: 33) offers another example illustrating failure of transitivity:

(36) a. If Otto had gone to the party, then Anna would have gone.
b. If Anna had gone, then Waldo would have gone.
c. If Otto had gone, then Waldo would have gone.

Lewis helps us understand the example with the following story: "*The fact is that Otto is Waldo's successful rival for Anna's affections. Waldo still tags around after Anna, but never runs the risk of meeting Otto. Otto was locked up at the time of the party, so that his going to it is a far-fetched supposition; but Anna almost did go. Then the premises are true and the conclusion false.*"

Strict conditionals also validate contraposition: the truth of the strict conditional *if p, q* guarantees the truth of the contraposed strict conditional *if not q, not p*. However, we find again that our intuitions about counterfactuals do not seem to reflect this fact. In von Fintel's example below (von Fintel 2001, attributed to Angelika Kratzer), (37b) is not considered to follow from (37a).

(37) a. (Even) if Goethe hadn't died in 1832, he would still be dead now.
b. If Goethe were alive now, he would have died in 1832

Another example from Lewis (1973: 35) illustrates failure of this inference pattern:

(38) a. If Boris had gone to the party, Olga would still have gone.
b. If Olga had not gone, Boris would still not have gone.

Lewis fills in the relevant background as follows: *"Suppose that Boris wanted to go, but stayed away solely in order to avoid Olga, so the conclusion is false; but Olga would have gone all the more willingly if Boris had been there, so the premise is true."*

Observations about the inference patterns validated by counterfactuals have been taken to argue against a strict conditional analysis for the construction. However, inference patterns themselves do not argue in favor of a dedicated semantics for counterfactuals, since our intuitions appear to pattern alike for both counterfactuals and indicatives. This is illustrated below for the

strengthening of the antecedent (39), transitivity (40) and contraposition (41): in none of these examples do we consider that the last sentences follows on the basis of the earlier ones.

(39) a. If Jones wins the elections, then Smith will retire
b. If Smith dies before the elections and Jones wins, then Smith will retire.
(Adams 1975: 17)

(40) a. If Smith dies before the elections, Jones will win.
b. If Jones wins, then Smith will retire.
c. If Smith dies before the elections, he will retire.
(Adams 1975: 16)

(41) a. If the US halts the bombing, then North Vietnam will not agree to negotiate.
b. If North Vietnam agrees to negotiate, then the US will not have halted the bombings.
(Stalnaker 1968: 107)

So, in spite of differences we have observed between counterfactuals and indicatives (Section 2.1), they do appear to share some common logic. The impact of this observation has been varied. As Edgington (1995: 273) points out: "*Coincidence in logic does not guarantee coincidence in interpretation.*". Some philosophers, however, have considered that the logical similarities make a convincing point in favor of a unified analysis. The semantic proposal in Stalnaker (1968), for example, provides a unified truth-conditional account within a possible worlds framework (see Section 4.2, 4.3). The common behaviour with respect to inference patterns is accounted for in terms of conditions on an accessibility function. But it is also possible to capture observations regarding the inference patterns outside truth-conditional accounts. Probabilistic semantics designed to capture conditions for acceptance and assertability instead of conditions for truth were originally formulated for indicative conditionals. They also provide an account for the failure of strengthening of the antecedent, transitivity and contraposition (see e.g. Adams 1965, 1966, 1975). Some authors have argued for a unified view for conditionals by extending probabilistic accounts to the case of counterfactuals (e.g. Adams 1975, Edgington 1995). However, the similarities in logic have not convinced everyone that a unified account is needed. Gibbard (1981: 211), for example, considered that the apparent similarities between counterfactuals and indicatives '*hides a profound semantic difference*', with a probabilistic account correct for indicatives but a possible-worlds account correct for counterfactuals (see comments in Stalnaker 1984). A stroll through the literature (in particular philosophical) makes the tension clearly visible: either place the logical similarities between counterfactuals and indicatives at the core of a unified semantic analysis and add a (pragmatic?) 'twist' to account for the differences; or give independent semantic accounts for counterfactuals and indicatives, trying to provide plausible explanations for the similarity in logic. (It is worth noting that, in practice, the burden has been asymmetrical, since many authors have specialized on one type of conditional without committing themselves with respect to the other).

The discrepancies observed between the predictions made by a strict conditional analysis of counterfactuals vs. what appear to be our intuitions regarding inference patterns has historically lead authors to move away from a strict analysis, with Stalnaker and Lewis proposing influential 'variably strict' accounts (Section 4.2, 4.3). There is, however, renewed interest in strict proposals (Section 4.4). This is largely due to the observation in von Fintel (2001) that variably-strict proposals make some incorrect predictions, in particular with reversed Sobel-sequences (see Section 4.4).

(42) ??If all the nuclear powers threw their weapons into the sea tomorrow, there would be peace; but if the USA threw its weapons into the sea tomorrow, there would be war.

(43) a. If Sophie had gone to the parade and been stuck behind a tall person, she would not have seen Pedro.

b. ??But if she had gone to the parade, she would have seen Pedro.

Examples like (42) and (43) seem to throw doubt on the interpretation of intuitions regarding inference patterns, and have revived interest in strict proposals. As von Fintel (2001) points out, a strict analysis of counterfactuals has the advantage of predicting that the antecedent is downward monotonic, a welcome prediction in light of the observation that negative polarity items (NPIs) are licensed in that environment (see also a.o. Heim 1984, Kadmon and Landman 1993, Schlenker 2004):

(44) If you had left any later, you would have missed the plane.
(von Fintel 2001: 14)

There is much agreement in the linguistics literature with the claim that the logical property of 'downward monotonicity' is required for the licensing of NPIs (following Ladusaw (1979)). A variably strict analysis, according to which the antecedents of counterfactuals are not monotonic, leaves examples like (44) unexplained.

Some authors have raised objections to variably strict analysis on the basis of observations regarding disjunction in the antecedent of counterfactuals (see Nute 1984 for an overview). Descriptively, counterfactuals appear to validate a pattern of simplification of disjunctive antecedents, allowing us to infer from the counterfactual *if p or t, q* that both the counterfactual *if p, q* and the counterfactual *if t, q* are true. This pattern is not validated by the variably strict analyses proposed by Lewis and Stalnaker under the assumption that *or* receives a standard truth-conditional interpretation. The potential problem was raised already in Fine's review of Lewis's *Counterfactuals* (1975: 453), where he noted that the counterfactual in (45a) does seem to imply (45b):

(45) a. If Thorpe or Wilson were to win the next General Election, Britain would prosper.
b. If Thorpe were to win the next General Election, Britain would prosper.

Fine considered this a 'minor objection', and went on to sketch potential solutions while keeping within Lewis's framework. Other authors (Nute 1975, Ellis, Jackson & Pargetter 1977) have considered the matter to raise more serious problems for the variably strict analyses of Lewis and Stalnaker. Nute (1975), for example, claimed that we find counterfactuals like (46a) false since (46c) is false, though, as Nute (1984) noted, we judge (47) true (in spite of the fact that if the US devoted more than half of its budget to education, it could not devote more than half of it to defense). Ellis, Jackson & Pargetter (1977) raised a similar point with (47), claiming that in order for the counterfactual in (47a) to be true, both (47b) and (47c) would have to be true.

(46) a. If we had had good weather this summer or the sun had grown cold, we would have had a bumper crop.
(cited from Alonso-Ovalle 2009, a variation on an example in Nute 1975.)
b. If we had had good weather this summer, we would have had a bumper crop.
c. If the sun had grown colder this summer, we would have had a bumper crop.

(47) If the U.S. devoted more than half of its budget to defense or to education, it would devote more than half of its budget to defense.

(48) a. If New Zealand had either not sent a rugby team to South Africa or had withdrawn from the Montreal games, Tanzania would have competed.
b. If New Zealand had not sent a rugby team to South Africa, Tanzania would have competed.
c. If New Zealand had withdrawn from the Montreal games, Tanzania would have competed.
(Ellis, Jackson & Pargetter 1977: 355)

In replying to Ellis, Jackson & Pargetter, Lewis (1977) noted that the issues raised by disjunction go beyond what happens in the antecedents of counterfactuals, and speculated that one possible approach would be to question the logical form of the sentence, proposing instead to treat 'either … or…..' as setting up two different antecedents ('*The seeming disjunctive antecedent is an illusion of surface structure'* Lewis 1977: 360). This view is in line with independently motivated views in the current semantic literature according to which *or* introduces alternatives into semantic computation (e.g. Zimmermann 2000, Aloni 2003, Simons 2005, Alonso Ovalle 2008). The problem of disjunction in the antecedent clauses of counterfactuals has recently been addressed by Alonso-Ovalle (2004, 2009) and van Rooij (2006), both of whom argue against abandoning a Lewis-Stalnaker approach to counterfactuals. Alonso-Ovalle (2009) provides a solution to the puzzle of disjunctive antecedents in terms of a Hamblin-style alternatives analysis of disjunction while van Rooij (2006) argues in favor of a proposal that extends a dynamic approach to counterfactual donkey sentences to the case of disjunctive antecedents.

4. Variably strict semantics

The variably strict semantics proposed by Stalnaker and Lewis have arguably been the most influential in the linguistics tradition. There are differences between the proposals, some noted already with respect to the scope of the accounts, and others to be detailed below. But they are similar enough to warrant the label of a Stalnaker-Lewis (or Lewis-Stalnaker) semantics for counterfactuals.

Stalnaker and Lewis argued for a truth-conditional interpretation for counterfactuals in terms of quantification over possible worlds. Their proposals were spelled out as a logic for a conditional connective but their insights have been incorporated into compositional accounts of natural language by current work in linguistics. Roughly, the Stalnaker/Lewis proposals claim that a (counterfactual) conditional *if p, q* is true iff *q* is true in the worlds most similar to the actual world in which *p* is true (I will talk informally about the most similar *p*-worlds). This account captures the intuition, already noted in (2), that in evaluating a counterfactual we consider the status of the consequent in circumstances in which the antecedent is true that differ minimally from what is actually going on.

The strategy for this section will be as follows. We will begin in Section 4.1, rather indirectly, with a discussion of theories of counterfactuals before the formulation of Stalnaker's and Lewis's proposals. Our focus will be on Goodman (1947). This will prove helpful in terms of understanding some of the motivation for the Stalnaker-Lewis proposals, and it will also set the stage for the discussion of Kratzer's and Veltman's premise semantics in Section 5. In Section 4.2, we will address the common core of the Stalnaker-Lewis proposal with a discussion of similarity; and in Section 4.3, we will focus on the differences between the accounts. The last section, Section 4.4, will follow up with recent critical assessments of the variably strict view. The focus will be on von Fintel's dynamic strict account (von Fintel 2001).

4.1 Background: Premise Semantics

It can be helpful to understand the Stalnaker-Lewis proposal against the background of earlier accounts. In this section we will focus on the proposal set out in Goodman (1947),[3] which has greatly influenced the linguistic literature via its impact on the premise semantic frameworks developed by Kratzer and Veltman (see Section 5). Goodman (1947) proposed an account according to which a counterfactual is true iff it is possible to infer the consequent from the antecedent conjoined with other statements that describe relevant conditions in the evaluation world. Here is the idea:

[3] See also Chisholm (1946)

(48) When we say "If the match had been scratched, it would have lighted" we mean that conditions are such – i.e. the match is well made, it is dry enough, oxygen enough is present, etc. – that "That match lights" can be inferred from "That match is scratched" (…) We do not assert that the counterfactual is true if the circumstances obtain; rather, in asserting the counterfactual we commit ourselves to the actual truth of the statements describing the relevant conditions.
(Goodman 1947: 116).

These 'statements' are characterized as *co-tenable* with the antecedent, and, together with the antecedent, lead to the consequent on the basis of law-like connections (usually invoking natural laws). In the example above the connection would be established by a law along the lines of *Every match that is scratched, well made, dry enough, etc., lights.* Goodman highlighted that constraints need to be imposed on the set of cotenable statements (and current work on premise semantics emphasizes the importance of this point). For example, it is important to exclude statements that, though (logically) compatible with the antecedent of the counterfactual, would be false if the antecedent were true. To see this, consider the match example again. We understand the counterfactual *"If the match had been scratched, it would have lighted"* to be true in the circumstances described: if the antecedent were true, the consequent would be true (given the laws, if the match had been scratched (+ enough oxygen + dry +….), it would have lighted). But we don't understand the counterfactual "*If the match had been scratched, it would not have been dry"* to be true. Yet it is predicted to be true if the set of cotenable statements includes the statement that the the match did not light: given the laws, if the match had been struck (+enough oxygen + not lighted + ….), it would not have been dry! As the example illustrates, it is important that the set of 'cotenable statements' statements not include statements, such as the statement that the match did not light, which would be false if the antecedent were true. In addition to the problem of accounting for general constraints on the set of cotenable statements, Goodman notes it is also important to factor in the role of context in determining the set of cotenable statements, adding that it is not always easy to recover the set of conditions associated with the antecedent in a particular context. He offers the following example:

(49) If New York City were in Georgia, then New York City would be in the South.
(Goodman 1947: 121):

The counterfactual in (49) will be true if it means that if New York City were in Georgia and the boundaries of Georgia remained unchanged, then New York City would be in the South. The counterfactual will be false if it means that if New York City were in Georgia and the boundaries of New York City remained unchanged, then New York City would be in the South. Contextual considerations and subtle cues can be important in recovering the meaning of a counterfactual.

In spite of careful fine-tuning regarding the set of cotenable statements (see e.g. Goodman 1947: 117-120), Goodman noted that his proposal for the semantics of counterfactuals faced two important problems. On the one hand, the actual characterization of the set of relevant co-tenable conditions did not seem independent of the interpretation of the counterfactual itself, leading to a problem of circularity (to determine whether a counterfactual is true it is necessary to determine whether there is a suitable set of cotenable statements, but this set cannot be identified without first determining whether the counterfactual itself is true). On the other hand, there didn't appear to be a principled way of distinguishing between the (law-like) true generalizations that support counterfactuals and permit an inference from antecedent to consequent, from other (accidental) true generalizations that do not actually support counterfactuals. We'll examine two examples. First, one from Chisholm (1946). It begins like this: *"Suppose that one afternoon two men, quite independently of each other (as we should ordinarily say), were to sit on the same park bench, that they were alone there, and that, as it happened, each of them was Irish."* As Chisholm points out, the generalization *'for all x, if x is on park bench y at time t, x is Irish'* is true. But from this true generalization we do not infer the truth of the counterfactual *"If Ivan were on park bench y at time t, Ivan would be Irish"* (the assumption being that Ivan

is not Irish). *"The difficulty is that our universal conditional described what are accidents or coincidences."* (Chisholm 1946: 301) Here is a second example making the same point, one set up by Goodman (1947: 122-127): Imagine that the (accidental) generalization *"Everything in my pocket on VE day[4] was silver"* is true. We would not wish to conclude that, for a given penny P, the counterfactual *"If P had been in my pocket on VE day, it would have been silver"* would true. On the other hand, consider the true (non-accidental, law-like) generalization *"All dimes are silver"*. From the truth of this generalization, we would be willing to conclude, for a given penny P, that the counterfactual *"If P had been a dime, it would have been silver"* is true. Goodman was able to offer some criteria to distinguish true lawlike generalizations from true accidental generalizations, but, as he himself noted, he was not able to offer a fully worked out account. (As we will see in Section 5, Kratzer's premise semantics offers a principled way of distinguishing the roles played by accidental and non-accidental generalizations in the semantics of counterfactuals.)

4. 2 Towards a variably-strict account: putting similarity in the picture

Counterfactuals make claims about worlds that are distinct from the actual world but that are relevantly similar, and it is in this way that facts about the actual world affect the truth-value of counterfactuals. In Goodman's proposal, the task of 'fixing' the actual world facts that matter for a counterfactual was left to the set of cotenable statements and laws invoked by the counterfactual. But, as we have seen, this was problematic. The accounts proposed by Stalnaker and Lewis completely bypass the problem of spelling out the cotenable premises completely, proposing instead a contextually provided similarity measure that, together with the antecedent, directly identifies the worlds to be quantified over. Roughly, quantification takes place over the worlds most similar to the actual (evaluation) world in which the antecedent is true. The counterfactual itself will be true iff all such antecedent worlds are also worlds in which the consequent is true (see Section 4.3). Stalnaker and Lewis argued in favor of this analysis, dubbed 'variably strict' by Lewis (1973), by noting that it made correct predictions regarding inference patterns such as antecedent strengthening, contraposition and transitivity. To see that the variably strict analysis fails to validate antecedent strengthening consider Lewis's original kangaroo example and other Sobel sequences (repeated below):

(50) a. If kangaroos had no tails, they would topple over.
b. If kangaroos had no tails but used crutches, they would topple over.

(51) a. If Sophie had gone to the parade, she would have seen Pedro.
b. But if Sophie had gone to the parade and been stuck behind a tall person, she would not have seen Pedro.

Let us consider (50). The reported intuition is that (50a) is true while (50b) is false. A variably-strict proposal can account for this because the worlds most similar to the actual world in which kangaroos have no tails need not be worlds in which they have crutches. The domains of quantification of the two conditionals can be completely independent and it is possible for (50a) to be true while (50b) is false. The same account can be given for (51) and similar examples. In a similar manner, similarity based proposal has been argued to account for failure of contraposition and failure of transitivity.

The relation of comparative similarity between possible worlds that is at the heart of the Stalnaker and Lewis proposals is vague. Lewis explicitly welcomes this vagueness, noting that counterfactuals themselves are *"notoriously vague"*. *"The truth conditions for counterfactuals are fixed only within rough limits; like the relative importance of respects of comparison that underlie the comparative similarity of worlds,*

[4] VE day was a public holiday held in May 8 to commemorate the end of World War II in Europe.

they are a highly volatile matter, varying with every shift of context and interest." (Lewis 1973: 92). Stalnaker comments on the truth conditions as follows: *"These are vague conditions which are largely dependent on pragmatic considerations for their application."* (Stalnaker 1968: 104). Interesting early discussion of the context-dependent nature of the evaluation of counterfactuals had already been provided by Chisholm (1946: 303):[5]

(52) Let us refer to "If Apollo were a man, he would be mortal" as *a* and to "If Apollo were a man, at least one man would be immortal" as *b*. Knowing Apollo to be immortal and all men to be mortal, should we assert *a* or *b*? The answer depends upon whether we are supposing our beliefs about Apollo, or our beliefs about men, to be contrary to fact. (…). Ordinarily, the context of inquiry determines which supposition is being made.

As illustration of context-dependence, Lewis 1973 discusses Quine's famous examples (Quine 1960: 222) about Caesar in the war between the US and Korea:

(53) a. If Caesar had been in command (in Korea), he would have used the atom bomb.
b. If Caesar had been in command (in Korea), he would have used catapults.

The vagueness of the relation of comparative similarity between worlds is resolved by context. In one context, the resolution of similarity may place greater weight on similarity with respect to Caesar's character and the knowledge of weapons available to commanders in Korea, in another context similarity may be resolved in a manner that attributes greater importance to the knowledge and weapons that were available to Caesar during his lifetime. Depending on the context, one or other of the conditionals in (53) may be true.

In his review of Lewis (1973), Kit Fine (Fine 1975) raised a concern for similarity-based proposals.[6] According to Fine an analysis based on overall similarity with the actual world made incorrect predictions when faced with examples like (54):

(54) If Nixon had pressed the button, there would have been a nuclear holocaust.

Fine was concerned that an overall similarity account like Lewis's predicted that a counterfactual like (54) would be false, whereas in fact it is true (or can easily be imagined to be true). Suppose that there will never be a nuclear holocaust. To evaluate the truth of (54), we consider what happens in the most similar worlds in which Nixon pushed the button. Given the huge divergence from what actually happens that would arise if in such worlds there was a nuclear holocaust, the most similar worlds will be those in which something exceptional happens, a 'small miracle', that prevents it (e.g. the signal fails to actually launch the missiles). *"For, arguably, a world with a single miracle but no holocaust is closer to reality than one with a holocaust and no miracle"* (Fine 1975: 452). But we do not even have to go as far as miracles: *"An undetermined break in an electrical connection will do or even a determined break, so long as the consequences are insignificant in comparison with a holocaust"* (Fine 1975: 452). Lewis characterized Fine's objection as the 'future similarity objection' and discussed it in Lewis (1979), providing greater details about the similarity relation that is taken to underlie the resolution of counterfactuals.[7] The relevant notion of similarity may differ from naïve intuitions about similarity, being subject instead to diverse constraints that balance divergence from actual facts and violations of actual laws ("*Do not assume that just any respect of similarity you can think of must enter into the balance of overall similarity with positive weight."* Lewis 1979: 465). In evaluating similarity in counterfactuals, it is most important to avoid big and widespread violation of the laws, it is of second importance to maximize a match with what actually

[5] See also Goodman's example (49).
[6] Fine was one of many authors to raise concerns, see also e.g. Bennett (1974) for another review of Lewis (1973), and Bennett (2003) for discussion.
[7] Some of this was already discussed in Lewis (1973)

happens, it is of third importance to avoid small violations of the law and it is of least importance to approximate (but not match) what actually happens. The upshot is that potential counterexamples like Fine's are not problematic: the overall similarity metric is such that the most similar worlds in which the antecedent is true will be worlds that are just like the actual world up to the moment in which Nixon pushes the button (the 'branching off time' or 'fork') and will follow the laws afterwards, allowing all the divergence from actual facts predicted by the laws once the button has been pushed. Worlds in which the electricity was cut off beforehand will never perfectly match the actual world, they will only match it approximately: differences will linger, e.g. the state of the cables will never perfectly match the actual state of the cables, there will be complaints against the electricity company which will not match actual complaints, whoever cut off the electricity will feel superior for having saved the world in a manner that will not match his/her actual emotional state, etc. Since approximate match to what is actually happening is not a high priority, and is not favoured over great divergence regarding matters of particular fact, worlds which intuitively are more similar to the actual world will not actually be evaluated as more similar by the similarity metric.[8]

The proposal to weigh the similarity metric along the lines indicated by Lewis thus generates a temporal asymmetry in the way similarity is evaluated in relation to the actual world, and temporal concerns have played an important role in the discussion of similarity in counterfactuals. Up to the branching-off time/ 'fork' matters of particular fact are of great importance, but after the branching off time, respect for the laws (the drive to avoid the multiple miracles presumably needed for 'convergence' with the actual world) actually trumps facts.[9] Similarity with respect to facts after the branching-off time does count, but for little. This view of similarity has been challenged in different ways by a variety of examples (as discussed already by Lewis and Stalnaker). Pavel Tichy, for instance, raised a challenge with the following scenario: *"(…) consider a man – call him Jones- who is possessed of the following dispositions as regards wearing his hat. Bad weather invariably induces him to wear his hat. Fine weather, on the other hand, affects him neither way: on fine days he puts his hat on or leaves it on the peg, completely at random. Suppose, moreover, that actually the weather is bad, so Jones is wearing his hat."* (Tichy 1976: 271). Against this background, Tichy asks us to consider (55):

(55) If the weather were fine, Jones would be wearing his hat.

Our intuition is that (55) is false in the scenario described. Tichy's point was that the similarity-based accounts provided by Stalnaker and Lewis predict that it will be true. Intuitively, in Tichy's example, some of the facts that intuitively appear to come after the branching-off time should not count at all towards similarity (the fact that Jones actually wore his hat) (see also Jackson 1977, Bowie 1979 for related examples and discussion). But it is unclear that balancing off distinct aspects of similarity would lead to the desired results. Here is Stalnaker's characterization of the problem: *"In choosing between possible worlds in which it rains and by chance the man wears his hat and possible worlds in which it rains and by chance the man does not, there is no question of trade-offs between different respects of similarity. One can choose a world in which the man wears his hat, as he does in the actual world, without giving up any respects of similarity at all."* (Stalnaker 1984: 128). While in Tichy's example it becomes problematic to give weight to what happens after the fork, other examples appear to indicate that we really do need to pay attention. In his discussion of time in counterfactuals, Slote (1978) brings up the following example (attributed to Sidney Morgenbesser): *"Imagine a completely undetermined random coin. Your friend offers you good odds that it will not come up heads; you decline the bet, he flips, and the coin comes up heads. He then says: "you see, if you had bet (heads), you would have won."* (Slote 1978: 27). As Slote notes, our intuition is that the counterfactual is perfectly natural in this context. Under a similarity analysis, this would require a similarity metric that gives weight to what has actually happened with the coin after the fork. In this case, post-fork facts do seem to matter. Examples with similar flavour have also been brought up by Pollock, who noted

[8] See Lewis (1979) for a discussion of a broader range of scenarios.

[9] These matters have received wide attention in the philosophical literature, see Bennett 2003 chapter 19 for an overview.

that counterfactuals may be true simply because the consequent is true, without there being any 'connection' between the antecedent and consequent: *"For example, we might say of a witch doctor, 'It would not rain if he did not do a rain dance, but it would not rain if he did either'."* (Pollock 1976: 26). Bennett terms this type of example 'non-interference' conditionals, and offers the following version:

(56) *A village sits near the bottom of a dam that has a lake backed up behind it. An agronomist thinks the village's aridity problem would be solved by cutting channels in the top of the dam so that useful amounts of water would flow to where they would do most good to the crops. Critics worry that cutting the channels will weaken the dam enough for it to collapse and destroy the village. The agronomist reassures them: 'No, if the channels were cut, the village would be safe.'* (Bennett 2003: 238)[10]

The challenge, from the point of view of Lewis's weighted similarity, is to provide insights into the differences between the cases in which features of what actually happens after the branching-off time matter for the evaluation of similarity, and cases in which they appear not to. Our intuitions regarding the examples point to a difference between the 'connection' between antecedent and consequent: in Tichy's example there is a connection between Jones wearing of his hat and the weather (Jones was actually wearing his hat because the weather was bad), but in the Slote/Morgenbesser example there isn't (the outcome of the tossing of the coin wasn't dependent on betting) (see also Edgington 2003 for discussion of hindsight). Some authors have given up on the idea of characterizing this distinction without appealing to causality in the semantics of counterfactuals, e.g. Bennett 2003 proposes a similarity metric that pays attention to 'causal chains' (Bennett 2003: 235).

Yet another kind of example that presents a temporal challenge to Lewis's weighted similarity metric is the case of 'backtracking' counterfactuals. In discussing the challenge, Lewis 1979 offers the following version (based on examples by Downing 1959):

(57) 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)

Backtracking counterfactuals are predicted to be false by Lewis's weighted similarity, yet some of them (*If Jim had asked Jack for help today, there would have been no quarrel yesterday*) can sound rather natural. According to Lewis, the examples sound natural only under a non-standard resolution of similarity, often highlighted by special syntactic structures (*there would have to have been*). There has been much debate regarding this type of examples in philosophy and linguistics, in terms of what they tell us about the similarity relation, what their actual interpretation is, the role of the special syntax, etc. (see e.g. Davis 1979, Slote 1978, Bennett 1984, 2003, Woods 1997, Arregui 2005, Schulz 2007, Frank 1997). It has been noted in the literature that whereas special syntax is often of great importance for a smooth resolution of backtracking (58), in some examples backtracking can proceed quite smoothly with simpler structures (59). Arregui (2005) and Schulz (2007) have argued that backtracking without special syntax is possible when the relation between the antecedent and consequent is not contingent (analytically or logically necessary).

(58) a. If the die had fallen six uppermost, it would (have to) have been thrown differently.

[10] As noted by Pollock (1976) and others, these types of examples are often rendered most naturally in English with 'even if' or 'still'. Though the illustrated examples are with non-counterfactuals, the same point can be made with more straightforwardly counterfactual conditionals: The plan for cutting channels is abandoned due to lack of funds. The villagers are relieved as they had feared for the village. But the agronomist points out: *'No, if the channels had been cut, the village would (still) have been safe.'*

(Bennett 1984)

b. If the plane had arrived at 2:00, it would have to have departed at 1:00. (Davis 1979)

(59) a. If Stevenson were President in February 1953, he would have been elected in November 1952. (Bennett 1984)

b. If he were a bachelor, he wouldn't have married. (Arregui 2005)

b. If Clarissa were 30 years now, she would have been born in 1966. (Frank 1997)

Debates amongst linguists regarding the interpretation of counterfactuals echo many of the issues that were first presented in the philosophical literature as problematic for the characterization of the similarity metric relevant to a Stalnaker/Lewis analysis. Kratzer's bi-dimensional proposals for the interpretation of modals on the basis of a modal base and ordering source (Kratzer 1979, 1981, etc.) provided an early account that integrates the interpretation of counterfactuals within a broader system that is able to shed light on modal interpretations more generally. The similarity intuition is addressed by means of an ordering imposed on the set of possible worlds on the basis of propositions true in the actual world. The challenges posed by posed by Tichy's puzzle are at the core of premise-semantics for counterfactuals as developed by Kratzer and Veltman (see Section 5). Temporal asymmetry reminiscent of Lewis's weighted similarity can be found in numerous proposals that identify the modal domain of quantification on the basis of a relation of historical accessibility, hardwiring into the semantics the fact that quantification takes place over worlds that match the history of the actual world up to some (past) moment (e.g. Ippolito 2003, 2006, 2013, Frank 1997, Condoravdi 2002).[11] The Stalnaker-Lewis characterization of similarity has been subject of much recent interest (e.g. Ippolito 2003, 2013, Arregui 2005, 2009, Schulz 2007). And, as noted, backtracking has also been of interest to linguists e.g. Frank (1997), Arregui (2005), Schulz (2007).

4.3 Differences between Stalnaker's and Lewis's proposals

The similarity-based accounts set forth by Stalnaker and Lewis share enough features to allow us to talk about a Lewis-Stalnaker / Stalnaker-Lewis semantics for counterfactuals. But there are important differences that have been subject to debate, in particular in the philosophical literature. In this section we will pin down the semantic proposals made by Stalnaker and Lewis more precisely, and focus on the differences.

According to Stalnaker (1968), the truth conditions for counterfactuals (and conditionals more generally) are determined as in (60) (where > stands in for the conditional connective, A is the antecedent proposition, B the consequent proposition, and α the evaluation world):[12]

(60) a. A > B is true in α if B is true in f(A, α)

b. A > B is false in α if B is false in f(A, α).

(Stalnaker 1968: 103)

In the proposal in (60), f is a selection function that takes a proposition and a possible world as arguments and delivers a possible world as its value. Given an antecedent A, the selection function f selects a possible world in which A is true. The conditional claims that B is also true in the selected A-world (thus, the conditional will be true in the actual world if it turns out that B is indeed true in the selected A-world). Intuitively, the selected antecedent world will be the world in which the

[11] See also Kaufmann 2005 for historical accessibility in the analysis of indicative conditionals.

[12] See also Stalnaker 1984, chapter 7

antecedent is true that differs minimally from the actual world (the evaluation world). *"This implies, first, that there are no differences between the actual world and the selected world except those that are required, implicitly or explicitly, by the antecedent. Further, it means that among the alternative ways of making the required changes, one must choose one that does the least violence to the correct description and explanation of the actual world."* (Stalnaker 1968: 104) Stalnaker imposes a series of formal constraints that must be satisfied by a selection function in order to be suitable in explicating the interpretation of conditionals. The constraints ensure, for example, that given an antecedent A, the selected world will be a world in which A is true; they ensure also that if A is true in the evaluation world, then the selected world will be the evaluation world; as well as ensuring that the function establish a total ordering amongst possible worlds. The formal constraints do not ensure that that there is a unique selection function for a conditional, and the problem of selecting a selection function from the set of possible ones remains as a pragmatic problem for counterfactuals.

Lewis's proposal for the semantics of counterfactuals was given in relation to a system of accessible worlds that was established in relation to the evaluation world. The system, that established a series of 'spheres of accessibility' around the evaluation world, was meant to capture information about the relations of overall comparative similarity between worlds. Each possible world i was assigned a set of spheres $i, consisting of a set of sets of possible worlds. The sets within $i were 'nested' (i.e. for any S and T members of $i, either S is a subset of T or T is a subset of S). Each sphere contains worlds that resemble the evaluation world i at least to a certain degree. The smaller the sphere, the more similar to i its members will be. The set $i is 'centered' in the evaluation world, containing the singleton set {i} as a member. These, as well as other formal constraints imposed by Lewis, ensure that a system of spheres carried information about comparative similarity between worlds in relation to an evaluation world. Having characterized a system of spheres, Lewis proposed the truth-conditions for counterfactuals as follows (where $\Box\!\!\rightarrow$ is the counterfactual conditional connective, $\supset$ is the material implication conditional connective, and $ is an assignment of spheres to possible worlds):

(61) $\phi \Box\!\!\rightarrow \psi$ is true at the world i (according to a system of spheres $) if and only if either
(1) no ϕ-world belongs to any sphere S in $i, or
(2) some sphere S in $i does contain at least one ϕ-world, and $\phi \supset \psi$ holds at every world in S.
(Lewis 1973: 16)

(61-1) spells out the vacuous case in which ϕ is either true in no possible world or only in worlds outside the sphere of accessibility (Lewis considers that in this case ϕ is not 'entertainable' as a counterfactual supposition at i). (61-2) spells out the main case in which ϕ is an entertainable supposition, true in world(s) at some sphere of accessibility surrounding i. In this case, the counterfactual is true if and only if the consequent is true in every antecedent-world in that sphere. If neither (61-1) nor (61-2) are satisfied, the counterfactual will be false at the evaluation world i.

The proposals made by Stalnaker and Lewis both appeal to a similarity ordering between worlds to identify the worlds that are 'relevant' to the evaluation of counterfactuals, i.e. the worlds in which the consequent is claimed to be true. As we have seen, the proposals differ with respect to the assumptions about that ordering. In Stalnaker's account, the truth-conditions are stated in terms of a function that takes the antecedent and evaluation world and delivers a unique world in which the antecedent is true as output (a unique relevant world). In Lewis's account (in the non-vacuous case), the truth-conditions are stated in terms of possibly multiple worlds and there may be multiple worlds in which the antecedent is true that are relevant. Stalnaker's proposal thus carries a *uniqueness assumption*, which Lewis explicitly argued against. Part of Lewis's arguments against the uniqueness assumption came from examples originally set forth by Quine (Quine 1982: 23), who was pessimistic about our chances regarding natural language counterfactuals: *"It may be wondered, indeed, whether any*

really coherent theory of the contrafactual conditional of ordinary usage is possible at all, (…)" and worried in particular about examples such as these:

(62) a. If Bizet and Verdi had been compatriots, Bizet would have been Italian.
b. If Bizet and Verdi had been compatriots, Verdi would have been French.

The issue, according to Lewis, is about ties in similarity: "Comparative similarity permits ties, and Stalnaker's function does not." (Lewis 1973: 80). Out of the blue, we are slightly perplexed about (62). We may feel unable to commit to ether. But Stalnaker's original proposal appears to predict, erroneously, that one of the examples in (62) must be true and the other false. Lewis's own proposal, allowing multiple antecedent worlds in one sphere, allows for ties in similarity, predicting instead that both (62a) and (62b) could be false. In later work, Stalnaker's position has been to defend the uniqueness assumption in the semantics, but to supplement it with a theory of vagueness according to which our intuitions regarding examples like (62) arise because in practice the selection function remains undefined. "*But given such a theory, we can reconcile the uniqueness assumption, as an assumption of the abstract semantics for conditionals, with the fact that it is unrealistic to assume that our conceptual resources are capable of well ordering the possible worlds*." (Stalnaker 1984: 134) (Stalnaker's own choice is to recommend a supervaluation theory of vagueness to account for the resolution of the similarity selection function.) A theory that allows vagueness to affect the resolution of similarity while adopting the uniqueness assumption in the semantics could provide an account of our blurry intuitions regarding the conditionals in (62), while keeping the original conditional logic intact. The debate about (62) thus turns on whether both conditionals are false (the original Lewis prediction), or whether their truth-value is indeterminate (Stalnaker with-vague-resolution prediction). The adoption of the uniqueness assumption in the semantics leads to a corollary difference with Lewis's account, since it allows Stalnaker's proposal to validate the law of *Conditional Excluded Middle* or CEM ("*The principal virtue and the principal vice of Stalnaker's theory"* in Lewis's words (Lewis 1973: 79)). CEM states that, given *A* and *B*, either the counterfactual *if A, would B* is true, or the counterfactual *if A, would not B* is true. This can be seen as a good outcome given that, as predicted by the law, our intuitions appear to support the claim that we negate a counterfactual by negating the consequent. Citing Stalnaker 1968: *'This explains the fact, noted by both Goodman and Chisholm in their early papers on counterfactuals, that the normal way to contradict a counterfactual is to contradict the consequent, keeping the same antecedent. To deny "If Kennedy were alive today, we wouldn't be in this Vietnam mess," we say "If Kennedy were alive today, we would so be in this Vietnam mess."* The validity of CEM also predicts that, in spite of our blurred intuitions regarding (62), we would not actually be able to negate both conditionals nor assert both conditionals. The fact that our intuitions appear to go in this direction was already observed by Lewis, who noted that it appears we cannot truthfully say:

(63) It is not the case that if Bizet and Verdi were compatriots, Bizet would be Italian; and it is not the case that if Bizet and Verdi were compatriots, Bizet would not be Italian; nevertheless, if Bizet and Verdi were compatriots, Bizet either would nor would not be Italian.
(Lewis 1973: 80)

Validating CEM, Stalnaker's theory predicts our intuitions in (63), whereas under Lewis's account, (63) could actually be true.[13] Debates regarding CEM continue (for recent discussion see e.g. Bennet 2003, Cross 2009, Williams 2010, Swanson 2012). Some of the debates in the linguistics literature have centered on the interpretation of quantified conditionals such as *No student will pass if she goofs off:* some authors have considered that CEM provides insights into their interpretation (e.g. von Fintel and Iatridou 2002, Higginbotham 2003, Klinedienst 2011) whereas others have favored other accounts (e.g. Leslie 2009) [see Huitink 2010 for an overview of issues]. Within the linguistic

[13] See Bennett 2003 183-189 for discussion of CEM

tradition, authors have weakened the link between the uniqueness assumption and CEM. For example: Schlenker 2004 offers an account of conditionals in which the antecedent clause functions as a referential plural definite description, appealing to a choice function analysis which does not validate CEM; von Fintel 1997 discusses conditional interpretations that allow for multiple antecedent-worlds but carry a homogeneity presupposition which validates CEM without uniqueness.

One of the costs counted by Lewis against the uniqueness assumption and CEM is that it rendered impossible an account of *might*-conditionals as duals of *would*-conditionals. Within Lewis's own proposal, a conditional of the form *if* ϕ*, might* ψ can be defined in terms of corresponding counterfactuals and negation (where $\Diamond\!\!\rightarrow$ is the logical connective corresponding to *might,* $\Box\!\!\rightarrow$ the counterfactual connective, and $\neg$ negation):

(64) $\phi \Diamond\!\!\rightarrow \psi =_{def} \neg (\phi \Box\!\!\rightarrow \neg \psi)$.
[Lewis 1973: 21]

A *might* counterfactual is thus characterized as the negation of the corresponding *would*-counterfactual with a negated consequent (so, *if* ϕ *might have* ψ will be true if and only if it is <u>false</u> that if ϕ it would have been the case that <u>not</u> ψ). But within Stalnaker's CEM-validating framework, the definition in (64) actually collapses the distinction between *might* and *would* counterfactuals, since they will mutually entail each other (if the *might*-conditional is true, then the consequent will be true in the unique relevant antecedent world, making the *would*-conditional true as well and if the *would*-conditional is true, then it will be false that in the unique relevant antecedent world the consequent will be false, making the *might*-conditional true as well). Lewis 1973 counted the collapse of the distinction between *might* and *would* –conditionals, given a dual analysis, as an argument against CEM and the uniqueness assumption. Stalnaker's response was to challenge Lewis's dual analysis, offering examples that conjoin the negation of a *would*-counterfactual with a *might*-counterfactual and noting that such examples sound odd ('paradoxical'), though nothing in Lewis's dual analysis predicts they should. Consider the following examples:

(65) a. Would President Carter have appointed a woman to the Supreme Court last year if a vacancy had occurred?
b. #No, certainly not, although he might have appointed a woman.
[Stalnaker 1984: 144]

(66) #It is not the case that John would have come to the party if he had been invited. But he might have.
[von Fintel 2012: 9]

Instead of a dual analysis, Stalnaker proposes that *might*-counterfactuals have a logical form according to which an epistemic possibility *might* operator scopes over a counterfactual. Under this proposal, the sentence in (67a) could be paraphrased as (67b) (paraphrase from von Fintel 2012):

(67) a. If John had been invited, he might have come to the party.
b. It might be the case that if John had been invited, he would have come to the party.
[von Fintel 2012: 9]

Such a wide-scoping epistemic *might* analysis was challenged already in Lewis 1973, who discussed the example *If I had looked in my pocket, I might have found a penny.* According to Lewis, the truth of the counterfactual depends exclusively on the facts: if I did not look, and there was no penny, the conditional is 'plainly' false. And this is so even if it was possible, given my epistemic state, that if I had looked in my pocket, I would have found a penny. Stalnaker's response was to defend the availability of a true epistemic reading but to allow also for other readings based on idealized states of

knowledge in which the conditional is false. The debate regarding the interpretation of *might*-conditionals continues, including discussions about their logical form, semantics, pragmatics and relations to *would*-conditionals, e.g. Bennett 2003, DeRose 1999, Hawthorne 2005, Eagle 2007, Gillies 2007, Hajek 2009, Williams 2010, Moss 2012. Early discussion of *might*- and *would*-conditionals set within a broader theory of modality in natural language can be found in Kratzer 1979, 1981.

Another important difference between the proposals put forward by Stalnaker and Lewis concerns the so-called *Limit Assumption*. This is the assumption that given any similarity ordering amongst possible worlds there will be a set of the most similar antecedent-worlds to the actual world. Stalnaker's proposal makes the limit assumption, which follows from the uniqueness assumption. But it is possible to make the limit assumption while rejecting the uniqueness assumption, and discussions of the limit assumption have been carried out independently. Lewis (1973) presented an empirical argument against the limit assumption:

(68) Suppose we entertain the counterfactual supposition that this point

---

there appears a line more than an inch long. (Actually, it is just under an inch.) There are worlds with a line 2" long; worlds presumably closer to ours with a line 11/2" long; worlds presumably still closer to ours with a line 11/4" long; worlds presumably still closer…. But how long is the line in the *closes* worlds with a line more than one inch long? If it is 1+x" for any x, however small, why are there no other worlds still closer to ours in which it is 1+1/2x", a length still closer to its actual length? (…) Just as there is no shortest possible length above 1", so there is no closet world to ours among the worlds with lines more than one inch long (….).
[Lewis 1973: 20]

Examples which challenge the limit assumption appear to show that in some cases there isn't a 'minimal change' to be made in order to make the antecedent true. Stalnaker's response was to argue that a similarity measure appropriate for the interpretation of conditionals should take into account what is *relevant*, making it possible to ignore differences that are not relevant in the context in which the conditional claim is made:

(69) Even if, in terms of some general notion of overall similarity (word) *i* is clearly more similar to the actual world than (world) *j*, if the ways in which it is more similar are irrelevant, then *j* may be as good a candidate for selection as *i*. In the example, it may be that what matters is that the line is more than one inch long, but still short enough to fit on the page. In this case, all lengths over one inch, but less than four or five inches, will be equally good.
[Stalnaker 1984: 141]

Moreover, argued Stalnaker, abandoning the limit assumption led to unintuitive conclusions: *"The point is not just that there is no particular length the line would have had. More than this, there is not even any length that it might have had."* [Stalnaker 1984: 142] (see also discussion and references in Bennett 2003: 175-180, as well as recent discussion in Swanson 2012 examining the limit assumption in both counterfactuals and counterpart theory). While the limit assumption is tied to debates in the philosophical literature, linguists often consider that making the limit assumption is a relatively harmless simplification that can render the semantics of counterfactuals more intuitive, and boldly take that step. Proposals for counterfactuals in linguistics often strike a middle ground between Stalnaker's and Lewis's proposals, setting aside the limit assumption and allowing for ties in similarity in many instances for the case of ease of exposition.

4.4 Back to a strict (but now dynamic) analysis

As we have seen, the variably-strict analyses put forward by Stalnaker and Lewis targeted intuitions that seemed problematic for a strict implication analysis of counterfactuals. However, following von Fintel 2001, there has been much interest in reviewing the view that the strict analysis is to be dismissed. As von Fintel observes, there are data that are in principle problematic for the variably strict view, that would receive a more straightforward explanation under a strict analysis that also incorporated context change. Von Fintel notes the contrast between examples like (70), the type of Lewis-Sobel sequences that had been taken to argue in favor against a strict implication analysis, and examples like (71) (see earlier discussion in Section 3):

(70) If the USA threw its weapons into the sea tomorrow, there would be war; but if all the nuclear powers threw their weapons into the sea tomorrow, there would be peace.

(71) ??If all the nuclear powers threw their weapons into the sea tomorrow, there would be peace; but if the USA threw its weapons into the sea tomorrow, there would be war.
(von Fintel 2001: 11)

The sequence in (71), contrary to what we find in (70), appears unacceptable. The utterance of the first counterfactual in (71) appears to highlight possibilities that end up being relevant in the interpretation of the second counterfactual. But variably-strict proposals do not offer an immediate explanation for this contrast, since in such accounts the two conditionals can be true and consistent with each other in the same context, leaving the effect of order unexplained.

A similar order-effect appears to be found in other data that were taken to support a variably-strict account. We have already seen that syllogistic reasoning appears to collapse with counterfactuals, arguing against a strict implication proposal:

(72) a. If Hoover had been a Communist, he would have been a traitor.
b. If Hoover had been born in Russia, he would have been a Communist.
c. -not-> If Hoover had been born in Russia, he would have been a traitor.

Von Fintel notes that our intuitions are affected by the order of the premises:

(73) a. If Hoover had been born in Russia, he would have been a Communist.
b. ??If Hoover had been a Communist, he would have been a traitor.
c. -not-> If Hoover had been born in Russia, he would have been a traitor.
[von Fintel 2001: 13]

In this case we are not obviously prepared to agree with the premise that if Hoover had been a Communist, he would have been a traitor. The more natural way to read the antecedent in (9b) is as including situations in which Hoover was a Communist and born in Russia, which would render the counterfactual false.

Von Fintel characterizes the problems posed by (71) and (73) in terms of a dynamic framework. Building on Warmbrod (1981), von Fintel proposes a strict analysis of counterfactuals that incorporates a proposal for context change. The evaluation of counterfactuals is established with respect to a 'modal horizon' that identifies the accessible worlds. As part of the dynamic process, the modal horizon is taken to change throughout a discourse, 'expanding' to include more possibilities in a way that ensures that the domain of quantification of counterfactuals is not empty: "*The procedure will be this: If a conditional is accepted as an assertion, the context will first be changed to expand the modal horizon if the antecedent wasn't already considered a relevant possibility. Then, the conditional will be interpreted in the new context.*" (von Fintel 2001: 19). The semantics for counterfactuals is given in terms of strict implication, it is the modal horizon function that is responsible for the apparent non-monotonicity of

the construction. This shift in the analysis allows von Fintel to explain why order matters in examples like (71) and (73). Context-shift is expected to go 'one-way': once a possibility has become relevant, it will stay relevant. Indeed, as von Fintel notes, this is something familiar from discussions of context-shifts with respect to other types of examples (such as knowledge attributions, see Lewis 1996). The dynamic strict proposal thus attributes the behavior of counterfactuals to the general dynamics of context changes. It is not impossible for the modal horizon to shrink in the course of a discourse, but that typically has to be signaled explicitly (e.g. 'but that would never happen'):

(74) If the USA threw its weapons into the sea tomorrow, there would be war. Well, if all the nuclear powers threw their weapons into the sea tomorrow, there would be peace. But of course, that would never happen. So, as things stand, if the USA threw its weapons into the sea tomorrow, there would be war.
[von Fintel 2001: 23]

On occasions, however, resetting can be done more indirectly:

(75) a. A: If John had been at the party, it would have been much more fun.
b. B: Well, If John had been at the party and had gotten into a fight with Perry, that wouldn't have been any fun at all.
c. A: Yes, but Perry wasn't there. So, if John had been at the party, he wouldn't have gotten into a fight with Perry.
[von Fintel 2003: 23]

In addition to accounting for the order-effects in Sobel sequences, von Fintel's account incorporated a limited notion of entailment that he terms 'Strawson entailment' (see also von Fintel 1999) under which inference patterns such as Strengthening of the antecedent, contraposition and hypothetical syllogism are valid. The proposed analysis characterizes the antecedents of counterfactuals as (Strawson) downward monotonic, correctly predicting that NPIs will be acceptable in this environment (see Section 3):

(76) If you had left any later, you would have missed the plane.
[von Fintel 2001: 14]

Gillies 2007 offers an elaboration of von Fintel 2001's dynamic strict analysis, extending the proposal to sequences with *might*-counterfactuals. As the examples below illustrate, Sobel-style sequences appear possible in such cases:

(77) a. If Sophie had gone to the parade, she would have seen Pedro dance; but of course,
b. if Sophie had gone to the parade, she might have been stuck behind someone tall and then wouldn't have seen Pedro dance.

(78) a. If Hans had come to the party he would have had fun; but of course,
b. if Hand had come to the party, he might have run into Anna and they would have had a huge fight, and that would not have been fun at all.
[Gillies 2007: 342]

As before, reversing the order has serious effects. As the examples below illustrate, possibilities made salient by consequents also affect counterfactual accessibility:

(79) ??If Sophie had gone to the parade, she might have been stuck behind someone tall and then wouldn't have seen Pedro dance; but of course, if Sophie had gone to the parade, she would have seen Pedro dance.

(80) ?? If Hans had come to the party, he might have run into Anna and they would have had a huge fight, and that would not have been fun at all; but of course, if Hans had come to the party he would have had fun.
[Gillies 2007: 343]

An analysis of the non-monotonicity of counterfactuals in terms of pragmatic constraints on context shift (as opposed to a non-monotonic semantics) had been entertained in abstract terms, and set aside, by both Stalnaker and Lewis. Indeed, Lewis considered a proposal of this kind "defeatist": "*It consigns to the wastebasket of contextually resolved vagueness something much more amenable to systematic analysis than most of the rest of the mess in that wastebasket."* (Lewis 1973: 13). The observation that order matters in Sobel sequences, however, has revived interest in alternatives to Stalnaker and Lewis's variably strict proposals. What could the variably strict analysis say with respect to these examples? Moss 2012 has argued that it is possible to account for order effects in Sobel sequences within a variably-strict proposal for counterfactuals without appealing to ad-hoc manipulations of the similarity relation. According to Moss, the fact that order matters in Sobel sequences should be understood in terms of general constraints on assertability according to which it is epistemically irresponsible for a speaker to make claims that are incompatible with salient possibilities raised by some utterance. Reverse Sobel-sequences are considered to be bad because the speaker of the second counterfactual cannot rule out possibilities incompatible with his/her utterance. To see that this is a general constraint, consider the following situation: "Suppose we are enjoying a perfectly normal day at the zoo, looking at an animal in the zebra cage that seems to have natural black and white stripes. It has not recently crossed our minds that the zoo may be running a really low-budget operation, where they paint mules to look like zebras." (Moss 2012). Moss claims that in this context, the conversational exchange in (81) seems reasonable (though you are oddly pedantic), whereas the exchange in (82) does not work:

(81) a. That animal was born with stripes.
b. But cleverly disguised mules are not born with stripes.

(82) a. Cleverly disguised mules are not born with stripes.
b. #But that animal was born with stripes.

According to Moss, once the possibility that the animal is a cleverly disguised mule has been made salient, it is epistemically irresponsible to claim that it was born with stripes (and thus infelicitous). It is this general principle that is responsible for the unacceptability of reverse Sobel-sequences (see Moss 2012 for technical details of the proposal applied to counterfactuals).

It seems fair to say that investigations in this domain are ongoing, with variably-strict proposals facing the challenge posed by ordering effects. At stake is a bigger picture of the mechanisms of context change and the delimitation between semantics and pragmatics.

5. Premise semantics for counterfactuals

The similarity-based proposals of Stalnaker and Lewis moved away from Goodman's attempt to actually spell out the actual world features that were considered relevant in the interpretation of counterfactuals. Instead of appealing to cotenable statements/propositions, Stalnaker and Lewis shifted the analysis towards a contextually-given notion of global similarity, bypassing problems noted by Goodman. But the alternative approach of so-called 'premise semantics' for counterfactuals has kept the focus on premise sets in the interpretation of counterfactuals, continuing within a tradition inspired by Goodman. As we will see, premise semantics have allowed researchers to probe into the more fine-grained details of how we reason in interpreting counterfactuals, details often obscured by the Stalnaker-Lewis global similarity metric.

The most influential proposals of this type have been the premise semantics developed by Kratzer and Veltman. They bear some similarity to Goodman's ideas in interpreting counterfactuals in relation to a set of propositions associated with a possible world (a 'premise set'). Different choices of premise sets can result in different interpretations. In the case of Kratzer 1979, for example, the interpretation of counterfactuals depends on a premise set constituted by propositions true in the actual (evaluation) world. Premise semantics are thus able to capture the world-dependent aspect of the interpretation of counterfactuals through the identification of the premise set (reminiscent of Goodman's 'cotenable' propositions). Lewis 1981 showed that premise semantics and an ordering semantics based on similarity were essentially equivalent, able to deliver the same results for the truth-conditions of counterfactuals. Premise semantics differs, however, from similarity-based proposals such as Stalnaker's and Lewis's in providing explicit details regarding the choices made to identify the premise set and the principles that govern the relations between propositions in the premise set. While contextually given similarity relations remain a (more or less) opaque black box, premise semantics for counterfactuals makes the underlying machinery and our reasoning more visible. Discussing her inspiration, Kratzer (2012) noted: *"(…) I wanted to show that there could be truth-conditional theories of modals and conditionals based on principles for reasoning from possibly inconsistent premises. I believed that such theories could yield analyses of counterfactuals that, as far as their logical properties were concerned, were as good as the similarity based theory of Lewis."*

Counterfactuals establish a relation between the propositions in the antecedent and consequent that is mediated by what is true, but it is clear that when evaluating a counterfactual *if p, q* it cannot be the case that we simply consider whether *q* would follow if we were to just add *p* to the set of true propositions. When *p* is false, this would lead to an inconsistent set. We need, in a sense, to 'revise' our original set. Kratzer proposes that we add *p* to sets of (true) propositions that are consistent with *p*. We then check: if we were to continue adding consistent true propositions to the result, would we end up with a set in which *q* is true (that is, would *q* be entailed)? Iff so, the counterfactual will be true. Kratzer presents the idea as follows:

(83) A *would*-counterfactual is true in a world *w* if and only if every way of adding propositions that are true in *w* to the antecedent while preserving consistency reaches a point where the resulting set of propositions logically implies the consequent. (Kratzer 2012)

This idea was made precise in Kratzer 1981 in the following manner:

(84) A *would*-counterfactual with antecedent *p* and consequent *q* is true in a world *w* if and only if for every set in $\mathbf{F}_{w,\,p}$ there is a superset in $\mathbf{F}_{w,\,p}$ which logically implies *q*. (Kratzer 1989, 2012)

The definition of $\mathbf{F}_{w,\,p}$ is given in terms of the definition of $\mathbf{F}_{w}$, which Kratzer (2012) calls the 'base' set. This is the set of propositions that satisfies constraints of truth, persistence, cognitive viability and non-redundancy. Given $\mathbf{F}_{w}$, Kratzer then defines a set of premise sets $\mathbf{F}_{w,\,p}$ (which she calls the 'crucial set') as the set of all subsets of the set $\mathbf{F}_{w} \cup \{p\}$ that satisfy certain constraints (see Kratzer).

Many of the results achieved by premise semantics result from constraints imposed on the base and premise sets. In their evolving work on counterfactuals, both Veltman and Kratzer have made proposals regarding how to best formulate those constraints. Kratzer 1981, for example, provided evidence of how the choice of propositions in a premise set affects the outcome of counterfactual reasoning. We'll consider one of her examples as illustration:

(85) Hans and Babette spend the evening together. They go to a restaurant called "Dutchman's Delight", sit down, order, eat and talk. Suppose now, counterfactually, that Babette had gone to a bistro called "Frenchman's Horror". Where would Hans have gone? (I have to add that Hans rather likes that bistro.)
(Kratzer 1981: 206)

We naturally conclude that if Babette had gone to "Frenchman's Horror", Hans would have gone there too (after all, they are spending the evening together). But there is a puzzle here that we can intuitively think of in this way: if we include in our premise set the proposition that Babette went to "Dutchman's Delight" and the proposition that Hans went to "Dutchman's Delight" *separately*, then when we evaluate our counterfactual *If Babette had gone to "Frenchman's Horror", Hans would have gone there too,* it will come out false. Because amongst the true propositions that we will consider together with the antecedent will be the proposition that Hans went to "Dutchman's Delight" and it will not be possible to reach the conclusion that Hans went to "Frenchman's Horror". In her 1981 discussion, Kratzer pointed out that it would be a mistake to add the two propositions as 'independent facts' to the premise set. Given our example, we intuitively think of them as connected, 'one fact'. Removing one from the premise set automatically removes the other. To use Kratzer's terms, if these two facts are "lumped" together, they stand or fall together. Kratzer investigated "lumping" in greater detail in Kratzer 1989, arguing that premise sets should be closed under lumping: if we put a proposition in a premise set, we should also put in all the propositions that it lumps.[14] Lumping will thus have consequences for the interpretation of counterfactuals, affecting what facts are taken into account (and, in this way, which aspects of 'similarity' matter). The lumping relation is vague, and its vagueness will infect the semantics of counterfactuals. As Kratzer shows, we may have sharp intuitions for some cases but murky intuitions for others: "*Consider the following example: My neighbor's house burnt down. His kitchen burnt down as part of it. The proposition that his house burnt down, then, lumps the proposition that his kitchen burnt down in the actual world. My neighbor's barn was destroyed by the same fire. Was the barn part of the house?"* [Kratzer 1989: 610] In cases like this we may not have clear intuitions as to whether the proposition that his house burnt down lumps the proposition that his barn burnt down in the actual world. But this simply predicts, correctly, that our intuitions regarding counterfactuals will sometimes be vague. But making the semantics of counterfactuals sensitive to lumping does allow Kratzer to deliver crisp results where Stalnaker/Lewis similarity might mistakenly lead us to expect vagueness.

(86) Last year a zebra escaped from the Hamburg zoo. The escape was made possible by a forgetful keeper who forgot to close the door of a compound containing zebras, giraffes, and gazelles. A zebra felt like escaping, and it took off. The other animals preferred to stay in captivity.
[Kratzer 1989: 625]

As Kratzer notes, given this scenario we would <u>not</u> automatically conclude that the counterfactual *If a different animal had escaped, it would have been a zebra* is true. If we were to address this in terms of a Stalnaker/Lewis view on similarity, we would have to conclude that in evaluating the counterfactual, similarity with the actual escaped animal (let's call him John, following Kratzer) does not matter. But why doesn't it? There is nothing in the similarity theory that predicts our intuition. *"Note that it is not that the similarity theory says anything false about examples of this kind. It just doesn't say enough. It stays vague where our intuitions are relatively sharp."* (Kratzer 1989: 262) A premise semantics with a premise set constrained by lumping delivers better results. Even though the proposition that a zebra escaped is true in the actual world, it cannot be consistently added to the antecedent of the counterfactual, because in the actual world it lumps the proposition that John escaped (the actual escapee!), which is incompatible with the antecedent of the counterfactual. Kratzer 1989 analyses a wide range of examples from the perspective of premise semantics, discussing the role of the lumping relation. Many of the puzzles addressed by Kratzer relate to observations already made by Goodman 1947, including the example of King Ludwig of Bavaria, Kratzer's version of Goodman's match example

[14] I will not present the formal details of Kratzer's proposal here, the reader is referred to Kratzer 1989, 2012. Kratzer provides an explicit definition of the lumping relation between propositions in a situations-semantic framework

discussed earlier (Kratzer's puzzles goes as follows: King Ludwig of Bavaria likes to spend his weekends at Leoni castle. Whenever the flag is up and the lights are on, the King is in the castle. Right now the lights are on, the flag is down, and the King is away. If the flag were up, would the King be in the castle or would the lights be off?). In the 1989 paper, Kratzer shows that it is possible to resolve problems noted by Goodman regarding differences between accidental and non-accidental generalizations in terms of their lumping properties, predicting differences in how they affect the premise set for counterfactuals. As poor lumpers, non-accidental generalizations can more easily be added to antecedents without generating inconsistencies. The result is that they play an important role in the semantics of counterfactuals, parallel, in a sense, to the primacy of laws in the similarity-metric proposed by Lewis. In the 2012 paper, Kratzer significantly expands on her 1989 proposal, modifying some of the constraints that affect the base set and premise set to include, for example, a non-redundancy constraint on the base set and the idea that non-accidental generalizations carry information about confirming situations (readers are referred to Kratzer 2012 for the details regarding differences with the 1989 account). Some of Kratzer (2012)'s discussion addresses points brought up by Veltman (2005), whom we turn to next.

Veltman 2005 provides an extension of earlier premise semantics to the case of counterfactuals. The proposal is set within a dynamic framework that aims to account for the update effects of counterfactual hypothesis.[15] The empirical discussion focuses on Tichy's challenge to the Stalnaker/Lewis similarity relation, echoing at the same time some of the concerns raised by Goodman 1947. The well-known puzzle is to account for when we do pay attention to particular facts when evaluating counterfactuals and when we don't. Tichy provided an example of when we don't (55). Veltman's version below provides an example of when we do (see also Slote's Morgenbesser example in Section 4.2):

(87) Suppose Jones always flips a coin before he opens the curtains to see what the weather is like. Heads means he is going to wear his hat in case the weather is fine, whereas tails means he is not going to wear his hat in that case. Like above [Tichy's original example], bad weather invariably makes him wear his hat. Now suppose that today heads came up when he flipped his coin, and that it is raining. So, again, Jones is wearing his hat.
(Veltman 2005: 164)

In this context, contrary to our judgments in Tichy's original scenario, we would say that the counterfactual *If the weather had been fine, Jones would have been wearing his hat* is true. Veltman's point is that it would be difficult to appeal to similarity in a manner that explains why the facts about Jones wearing his hat affect the evaluation of similarity in this example but do not affect it in Tichy's original case. The intuition, claims Veltman, is that the cases differ because in Tichy's original example, contrary to (87), we have to give up the *reason* why Jones was wearing his hat and Stalnaker/Lewis similarity does not capture that. "(…) *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. But we want to keep in as many independent facts as we can."* (Veltman 2005: 164). In Veltman's proposal, an important distinction is made between facts that are independent from other facts, and facts that are 'brought along' by general laws (where general laws are whatever the agent considers as such, and may include natural laws, conventional laws, etc.). In setting up the semantics of counterfactuals, Veltman proposes to carry out a revision on premise sets that identifies the basic facts compatible with the antecedent of a counterfactual and then allows the laws to bring along the dependent facts. This ensures that the semantics for counterfactuals will give priority to the laws of the actual/evaluation world (as we have seen in other proposals). In Veltman's account, a premise set for the evaluation of a counterfactual in w is not the set of all propositions true in w but the *basis* for w, characterized as a

[15] I will not be able to go into the details of the dynamic proposal here, the reader is referred to Veltman 2005 that offers a fully explicit analysis, as well as the earlier Veltman 1985.

minimal set of propositions such that, together with the general laws, they completely determine all the facts in w (Veltman notes that a world may have more than one basis, but I set this aside here). To evaluate a counterfactual we first retract from the premise set (basis) the facts that are incompatible with the antecedent and we then let the laws bring along the dependent facts. The resulting worlds will be the worlds quantified over by the counterfactual. By differentiating between independent and dependent facts, Veltman's proposal can account for our varying intuitions in Tichy's example (55) and the version in (87). In (55), the fact that Jones wore his hat will not make it into the premise set since it is not an independent fact (it depends on the fact regarding the weather). Nor will it be brought along by any of the laws. As a result, the worlds quantified over by the counterfactual will not all be worlds in which Jones wore his hat and the counterfactual will be false. In the case of (87), however, the basis will include the independent fact regarding the outcome of the tossing of the coin. It won't include the (dependent) fact regarding Jones wearing his hat. But that does not matter, because this fact will be brought along by the laws regarding the coin coming up heads and Jones wearing his hat. So all the worlds quantified over will be worlds in which Jones wears his hat, and the counterfactual will be true.

As becomes evident from the above discussion, both the premise semantics set up by Kratzer and Veltman address the issue that facts 'stand and fall' together, but do so differently (and indeed differ with respect to some examples, see discussion of the King Ludwig example in Veltman 2005 and Kratzer 2012). While Veltman appeals to the generalizations that the agent considers to be general laws, Kratzer introduces 'lumping'. The new lumping relation permits Kratzer to provide a quantificational characterization of dependencies between facts independently of laws, allowing for some dependencies to also be law-like while others are grounded in the accidental distribution of facts in the actual world.

One of the central aims of premise semantics for counterfactuals is to account for dependencies between facts as they play out in the interpretation of counterfactuals. In recent work, Schulz (2011) proposes to include causal dependencies between facts directly into premise semantics by making premise sets sensitive to a causal notion of consequence formalized along the lines of Pearl 2000's causal models developed in computer science (see also Schulz 2007). *"One could argue that the sensibility of conditionals to causal dependencies is only an epiphenomenon (see Lewis 1973). Contrary to Lewis, I will claim that the truth conditions of conditional sentences build on the contextually salient causal dependencies."* [Schuzl 2011: 242]. Schulz builds on Veltman 2005, modifying the notion of basis for a world so as to make it sensitive to a 'dynamics' (roughly, a causal model). According to Schulz, a (counterfactual) conditional of the form *if alpha, beta* will be true in a world w given a dynamics D if alpha together with the basis for w causally entails beta given D (this is the basic interpretation rule for conditionals presented in Schulz 2011: 247, see also Definition 8, p. 249). Schulz 2007 argues that adding a causal component to the analysis provides a better handle on variants of the Oswald-Kennedy examples, as well as providing insights into causality puzzles in the literature. The reader is referred to Schulz 2007 for a comparison between Veltman's original premise semantics and a Pearl-causal-models-based account. (Readers are also referred to Pearl 2000, who offers a proposal for the interpretation of counterfactuals in terms of causal models and a comparison with Lewis-style proposals, to Kaufmann 2013, who develops a semantics for indicative conditionals on the basis of Pearl-style causal models, and to Santorio 2014 who seeks to develop a bridge between premise semantics and causal models via 'filtering' semantics).

6. Tense and aspect in counterfactuals

Most accounts of counterfactuals that address issues of tense and its interpretation have been inspired by the Stalnaker-Lewis view of similarity (as opposed to focusing on the reasoning strategies addressed by premise semantics, Section 5). A recurring idea in this literature has been that the presence of past tense in counterfactuals has consequences regarding how the most similar worlds to the actual world are identified. The similarity view has thus provided a natural home for this debate.

Tense made its way into the discussion of counterfactuals very early on. Part of the interest had to do with the perceived temporal asymmetry in the interpretation of counterfactuals that seemed to favor 'past facts' (see Section 4.2), with early proposals that investigated the interaction between counterfactuals and tense logic (e.g. Thomason and Gupta 1980, Thomason 1984). Such discussions were grounded on observations about the interpretation of counterfactuals that were established independently of considerations of linguistic form, leading to semantic proposals justified independently of tense morphology. An early proposal to take tense morphology in counterfactuals more seriously is found in Dudman (1983, 1984), who noted that tense morphology in English counterfactuals does not appear to have its usual interpretation, receiving instead a 'shifted' reading. This shift allows '-ed' verbal morphology in antecedents that make future or present hypotheses and 'had –ed' verbal morphology in antecedents that make future, present or past hypothesis:

(88) *-ed.*
a. If Grannie missed the last bus tomorrow, she would walk home. [future hypothesis]
b. If Her Majesty was here now, she would be revolted. [present hypothesis]

(89) *had –ed*
a. If Grannie had missed the last bus on Friday /next Friday, she would have walked home (she is actually dead). [future hypothesis]
b. If Her Majesty had been here now, she would have been revolted. [present hypothesis]
c. If Grannie had missed the last bus on Friday (last Friday), she would have walked home (luckily, she caught it). [past hypothesis]
[Dudman 1984: 150]

Dudman argued that the shift in tense interpretation was actually crucial in getting the classification of counterfactuals right, and proposed an account that moved away from the traditional 'indicative' vs. 'subjunctive'/counterfactual terminology and appealed instead to standard vs. 'shifted' interpretations of tense (see Edgington 1995 for discussion of Dudman's classification of conditionals).

While the observation that some languages mark counterfactuals with temporal morphology has been familiar for some time, linguists have found the explanation rather elusive. In *Modality*, Palmer commented, pessimistically, on attempts to link past tense and 'unreality': "*The relation between past and unreality has often been noted, but the explanations seem to be largely circular*" (Palmer 1986: 211). Recent proposals investigating the interpretation of tense morphology in counterfactuals have been more optimistic, with, broadly, two kinds of views: some authors claim that past tense is in indeed inherently able to play two roles, one temporal and one modal, with remoteness in the modal domain relevant for the interpretation of counterfactuals; while others have maintained a temporal semantics for past tense, accounting for its interpretation in counterfactuals in terms of how it interacts compositionally with modals. We will explore these views in the following sections. Section 6.1 presents a discussion of the 'modal remoteness' view, addressing work by Iatridou (2000), Anand and Hacquard (2009) and Schulz (2014). The section also provides a brief mention of typological work inspired by Iatridou (2000), represented here by new proposals by Bjorkman and Halpert (forthcoming). Section 6.2 turns to the view that past morphology in counterfactual is truly interpreted as a temporal past tense. The section focuses on proposals by Ippolito (2003, etc.) and Arregui (2005, etc.), including also reference to Gronn and von Stechow (2011), Gronn (2013), and Romero (2014). There is a growing body of work focused on the relation between tense and modality in counterfactuals, and the discussion presented here is not meant to be taken as an exhaustive description of a lively domain.

6.2 Modal 'past'

Iatridou (2000) presented a highly influential proposal for the interpretation of past tense in counterfactuals. The paper argued that tense morphology in counterfactuals fails to receive a temporal interpretation, and was thus '*fake*' (linking the interpretation of tense in counterfactuals to other 'counterfactual' constructions such as wish-constructions). Iatridou's view was supported with a broad range of data from a variety of languages and included aspect as well as tense. The main empirical focus of the paper was on Modern Greek, which is similar to English in also having past perfect morphology in the antecedent clause of counterfactuals. In Modern Greek counterfactuals, the matrix clause bears a future particle and a verb bearing past morphology:

(91) An iχ$_1$e pari to siropi θa iχ$_1$e θ$_1$ni kala
if had taken the syrup FUT had become better
'If he had taken the syrup, he would have gotten better' (Iatridou 2000:233)

Past marking plays a crucial role in the interpretation. With non-past perfect morphology, the conditional is interpreted as 'still open', and could, for example, be used as instructions to a caretaker:

(92) An pari afto to siropi θa γ$_1$ni kala
if take.non-past.Perf this syrup FUT become.non-past.perfective well
'If he takes this syrup, he will get better' (Iatridou 2000:234)

With past (imperfective) morphology, conditionals are interpreted as 'less vivid' (to use Iatridou's term) but still epistemically open:

(93) An eperne afto to siropi θa γ$_1$inotan kala
if take.past.imperf this syrup FUT become.past.imperfective well
'If he took this syrup, he would get better' (Iatridou 2000:234)

The 'less vivid' examples bring along a sense that the actual world is less likely to be a world in which the antecedent is true than a world in which the antecedent is false. Iatridou relates the interpretation of examples like (91) and (93), noting that there appears to be a layer of past morphology that does not receive a past tense interpretation: "*I will refer to such occurrences of past morphology that do not receive a temporal past interpretation as 'fake past' or 'fake tense'*." (Iatridou 2000: 235). How to account for the *fake past* phenomenon? Iatridou advocates for view according to which "*the past tense morpheme always has the same meaning, but the domain it operates on varies according to the environment*" (Iatridou 2000: 245). The core of the interpretation of past tense morphology is provided by an *exclusion feature* that locates a topic outside a domain associated with the speaker, where the topic is allowed to range over times or worlds. In its temporal interpretation, the exclusion feature establishes that the topic time excludes the time interval that as far as is known is the time of the speaker (i.e. the utterance time). This is the interpretation we normally think of as 'past tense'. In its modal interpretation, the exclusion feature establishes that the topic worlds exclude the worlds that as far as is known could be the actual world of the speaker (i.e. the worlds that could be the actual world). This is the interpretation that leads to the intuition of modal remoteness. In Iatridou's characterization, past perfect (pluperfect) examples like (91) have two layers of 'past' and while one is used temporally, the other one leads to counterfactuality. The 'counterfactuality intuition' arises as an implicature due to the fact that the speaker is signaling that s/he has chosen to predicate the antecedent proposition over worlds that are members of a set that excludes the worlds that are epistemic contenders for the actual world.

In addition to discussing fake tense in examples like (93), Iatridou also argues for *fake aspect,* noting that imperfective morphology in the antecedent and consequent of (93) appear to lack typical 'event in progress' features. In a sense, the clauses in (93) appear to be interpreted 'perfectively': "[4]

does not mean that if the patient will be in the process of taking the syrup, he will be in the process of getting better; instead, it means that after he takes the syrup, he will get better." (Iatridou 2000: 236, see also other data). In spite of the fact that the aspectual morphology appears in a sense 'bleached' of it's standard interpretation, imperfective morphology is actually obligatory in Modern Greek counterfactuals. Iatridou proposes that this responds to a generalization that requires imperfective aspectual morphology when the temporal coordinates of an eventuality are not set with respect to the utterance time. Building on some of the issues noted by Iatridou, Anand and Hacquard (2009) propose an account of the imperfective in Romance counterfactuals. The starting point is the observation that in French counterfactuals both the antecedent and consequent have obligatory imperfective morphology (conditional morphology is taken to spell out a future operator scoping over an imperfective operator, see also Iatridou 2000):

(94) Si Jean arrivait demain, il rencontrerait Jane.
If Jean arrive-impf tomorrow, he met-cond Jane.
If Jean arrived tomorrow, he would meet Jane.
(Anand & Hacquard 2009: 46)

Building on Hacquard's event-related modality (e.g. Hacquard 2006, 2009, 2011), they argue that the modal semantic contribution of imperfective becomes vacuous when imperfective and future operators are stacked in counterfactuals (echoing Iastridou's 'fake aspect'). The ongoing event interpretation associated with the modal component is therefore lost. The presuppositions associated with the imperfective operator, however, remain and, in the interpretation of counterfactuals, they end up putting temporal constraints on a 'fork' event (in the sense of Bennett 2003).

In recent work, Bijorkman and Halpert (forthcoming) have presented a broader typology of tense and aspect marking in counterfactuals, noting cross-linguistic variation in the identification of '*fake'* tense and aspect (see also e.g. Nevins 2002, Bjorkman 2011, Bjorkman and Halpert 2012, Halpert and Karawani 2011). Bjorkman and Halpert argue that languages that mark counterfactuality with temporal morphology do so with either past or imperfective, but not both (though languages may *appear* to require both). Exploring a broad range of data, Bjorkman and Halpert investigate languages that allow perfectives in counterfactuals, including languages in which it is claimed to be interpreted (Zulu) and languages in which it is not (Palestinian Arabic). They also provide examples from Russian, which is classified as a language that marks counterfactuality with past, allowing both perfective (95a) and imperfective (95b) with the expected 'non-fake' aspectual interpretations:[16]

(95) a. Esli by Džon umer, my poxoromi-l-i by ego na gor-e
if subj John die.Pfv.Pst we bury.Pfv-Pst-Pl subj he.acc on mountain-Loc
If John died, we would bury him on the mountains
b. Esli by Džon umira-l, s nim by-l by doktor
if subj John die.Impf-Pst with he.Instr be-Pst subj doctor
If John were dying, the doctor would be with him

Bjorkman and Halpert conclude that their typological studies lend support for views according to which across languages a single syntactic position is associated with the composition of counterfactual semantics, even though no single tense or aspect is required across all of them.

Schulz (2014) presents a recent influential proposal to account for 'fake tense'. Schulz sides with Iatridou in following the view that, in a sense, tense in counterfactuals ends up as 'modal', working out an explicit proposal for the compositional role played by tense morphology plays in the interpretation (see also Schulz 2007). Schulz differs from proposals such as Iatridou's that take past morphology to correspond to an under-specified operator. The proposal instead is that the uninterpretable features associated with past morphology can be checked by both past tense

[16] Examples cited from a presentation by Sabine Iatridou, 2009.

operators and (certain) modal operators, allowing for both temporal and 'modal remoteness' interpretations. Schulz proposes that this variability can be made sense of by the fact that there are structural similarities between the different types of operators and their domains. In the case of modals, the 'past' feature establishes that the worlds quantified over 'precedes' the epistemic center for the utterance. Given Schulz's proposal, this means that they are not optimal with respect to an epistemic ordering and thus are not expected to be the actual world. Thus, by manipulating how features are checked, Schulz provides an account of the contrast between examples like (96) and (97):

(96) If Peter left in time, he will be in Frankfurt this evening.
(97) If Peter left in time, he would be in Frankfurt this evening.
(Schulz 2014: 135)

In (96) tense morphology is interpreted temporally as past and quantification takes place over epistemically optimal worlds that satisfy beliefs and expectations, whereas in (97) tense morphology is interpreted 'modally' as indicating quantification over worlds that are not epistemically optimal. With echoes of Iatridou, Schulz notes: "*In other words, the sentence presupposes that the speaker is honest, and that he/she does not expect the actual world to be among the closest antecedent worlds."* (Schulz 2014: 136)

6.3 'Real' tense

Authors who have argued that tense morphology in counterfactuals receives a standard temporal interpretation have often linked it to the view that the past history of the world is particularly important in determining the quantificational domain of the modal. But this is not to be understood in a simplistic manner. We have already noted many examples (some discussed already by Lewis 1973) that illustrate that it is actually problematic to propose that the domain of quantification is made up of all law-like worlds that match the actual world up to some past branching-off time (remember, for example, discussion of facts that 'stand and fall' together in Section 5, as well as the puzzles posed by Tichy's examples (wearing a hat), Slote's (tossing a coin), backtracking (Jim asking for help), etc.). Authors who argue for a 'real past' interpretation for past morphology in counterfactuals have proceeded cautiously, developing theories that spell out a role for past interpretation while at the same time incorporating the refinements needed to address the more complex cases. Much of the current work on the interpretation of tense in counterfactuals owes an important debt to Condoravdi's account of the temporal interpretation of modals (Condoravdi 2002). While Condoravdi's focus was not on conditional constructions, the insights shed by Condoravdi's discussion of the relation between temporal operators and metaphysical modals, as well as observations regarding constraints on domains of quantification, have proven of great importance.

In an evolving series of works, Ippolito has put forward a compositional account of the semantic role of tense and aspect in counterfactuals (e.g. Ippolito 2003, 2006, 2008, 2013a, b).[17] Ippolito has linked the semantic impact of tense and aspect in counterfactuals to an impact in pragmatics, establishing a relation between the choice of temporal morphology and the projection of presuppositions and felicity conditions in counterfactuals. Ippolito (2006), for example, notes the following contrast:

(98) Suppose that John had been training to run the Boston marathon for several months when he died.
a. Trainer: # If John ran the Boston marathon next spring, he would win.
b. Trainer: If John had run the Boston marathon next spring, he would have won.

[17] I cannot do justice to the full complexity of Ippolito's work here, and will only mention some highlights. Readers are referred to Ippolito for details of the analysis and a broader set of data, as well as discussion of alternative accounts. Particularly relevant is Ippolito (2013), a book-length treatment.

[Ippolito 2006: 633-634]

In the context in (98), (98b) is felicitous, but (98a) is not. Ippolito attributes this difference to a contrast in the behavior of the presuppositions associated with the antecedent. Running the Boston marathon presupposes being alive. Ippolito argues that simple past subjunctive counterfactuals require that presuppositions be compatible with the history of the world at the utterance time, while past perfect subjunctive counterfactuals only require that they be compatible with the history of the world at an earlier time. Other examples illustrating this contrast are provided below: with the anaphoric presupposition trigger *too* in (99) and with *quit* in (100) [see Ippolito (2013), Chapter 4 for more examples and a detailed discussion of the behavior of different types of presupposition triggers].

(99) Jack hasn't seen Spielberg's most recent movie, and will never see it. He could have seen it yesterday$_1$, and…
   a. #If he saw it [tomorrow]$_F$ too$_1$, he would regret it.
   b. If he had seen it [tomorrow]$_F$ too$_1$, he would have regretted it.
   [Ippolito 2013: 122]

(100) Lucy was a heavy smoker but she quit smoking ten years ago, after she had pneumonia. A new law was passed last week requiring people who have quit smoking to take a new medical test (the law is not retroactive). This test detects long-term problems in ex-smokers but is very painful. Thinking about Lucy, I say:
   a. #Good for her. If she quit smoking tomorrow instead, she would have to take this new painful test.
   b. Good for her. If she had quit smoking tomorrow instead, she would have had to take this painful test. [Ippolito 2008: 259]

As Ippolito notes, even though simple past subjunctive counterfactuals appear problematic in the (a) examples above, it is not the case that they cannot be 'counterfactual':

(101) John is not sick now and he will not miss the final ball game. If he were, he would be devastated.
(102) John is not in love with Mary. If he were, he would ask her to marry him.
   [Ippolito 2013: 54]

On the other hand, if counterfactuals make hypotheses about events that have 'already happened in the past', they are infelicitous:

(103) I called John yesterday to wish him a happy birthday, but it was the wrong day. His birthday is tomorrow and he got really upset. I am mortified.
   #If only I called him tomorrow (instead), he would be happy.
   (cf. I am mortified. If only I had called him tomorrow (instead), he would have been happy.)
   [Ippolito 2013: 54]

Ippolito's proposal aims to provide an analysis of subjunctive counterfactuals that allows both simple past and past perfect subjunctive counterfactuals to be felicitous when the antecedent is false, but only allows past perfects to be felicitous when the presuppositions are false. The proposal has two key ingredients: a semantics that identifies the domain of quantification of counterfactuals taking into account both historical accessibility and similarity; and felicity conditions that predict that the reference time that anchors the counterfactual will impact the projection of presuppositions. Ippolito's semantic proposal derives differences in the truth-conditions of simple past and past perfect subjunctive counterfactuals on the basis of an analysis that embeds a bare conditional under

temporal operators. In Ippolito (2013), two times are claimed to have a crucial impact: the conditional's reference time (at which the presuppositions of the conditional must hold) and the conditional's historical accessibility time (see also Ippolito (2006) for a discussion in terms of perfect operators and a dynamic account of presupposition projection that builds on the dynamic account of counterfactuals in Heim 1992). According to Ippolito (2013), the reference time for simple past counterfactuals is the time of utterance, whereas past tense projected above the bare conditional anchors the historical accessibility relation in some past time. In past perfect counterfactuals, two layers of past are projected above the conditional, resulting in an interpretation in which the accessibility time for the relation of historical accessibility is shifted towards the past and the reference time is also shifted towards the past. In both cases, quantification takes place over the most similar worlds that match the history of the actual world up to some past time (i.e. the accessibility time is past), allowing both conditionals to receive 'counterfactual interpretations' (i.e. both types of conditionals can be true if the antecedent is false).[18] But the difference in reference times has crucial impact for the projection of presuppositions: in the case of simple past subjunctives, the presuppositions of the conditional must be compatible with the history of the world at the utterance time, whereas in the case of past subjunctives, the presuppositions need only be compatible with the history of the world up to some relevant past time (see Ippolito 2006 for discussion in a dynamic framework). Thus, if the presuppositions of a counterfactual are incompatible with the history of the world up to the time of utterance (as in (1), (2) and (3)), only a past perfect subjunctive can be felicitous. In this way, Ippolito provides a proposal for the interaction between the temporal anchoring of counterfactuals and the way in which presuppositions behave.

Arregui (2005, 2007, 2009) offers a different view of temporal anchoring in counterfactuals. Whereas other proposals have characterized counterfactuals with perfect aspect morphology in the antecedent clause as conditionals in the scope of two layers of past, Arregui (2005, 2007) proposed that in such examples a single past tense scopes over the conditional, but perfect aspect is interpreted within the antecedent clause. Arregui supports the claim that the layer of perfect morphology is actually interpreted as aspect within the antecedent with examples that diagnose the presence of perfect aspect, such as (104):

(104) If you had lived in this house since 1975, you would have qualified for a rent subsidy.

Following Kamp and Reyle (1995), Arregui notes that *since*-clauses are only compatible with perfect tenses (Kamp and Reyle note the contrast between *Mary had lived in Amsterdam since 1975* and **Mary lived in Amsterdam since 1975*). The acceptability of *since*-clauses in the antecedents of counterfactuals such as (2) thus provide support for the view that the perfect is interpreted within the antecedent clause. Arregui (2005, 2007) argues for a unified approach to *would*-conditionals in which the crucial difference is associated with the choice of aspect in the antecedent clause. According to this view, antecedent clauses with perfective aspect (simple eventive clauses) carry a presupposition that anchors the antecedent proposition to worlds in the context set. This accounts for the felicitous example in (105) and the contrast in (106):

(105) Suppose you will go on holidays next week, and ask me to look after your plants. I accept, but feel rather nervous. I am not very good with plants.
You: Could you look after my plants next week while I am away?
Me: Of course, but I am rather nervous. If your plants died next week, I would be very upset.

(106) (continuation) Suppose that your plants die before you leave on holidays, and you cancel your request. I would feel sorry, but also relieved.

[18] I have set aside here discussion of the temporal orientation of the antecedents and consequents themselves, see Ippolito 2013: 92ff.

You: Don't worry about looking after my plants. They died yesterday.

a. Me_1: I am sorry, but also a bit relieved. If your plants had died next week, I would have been very upset.

b. Me_2: I am rather relieved that your plants died yesterday. #If they died next week (instead), I would be very upset.

The semantics of stative antecedents and perfects does not result in the same constraint, allowing the antecedent proposition to be true in worlds outside the context set as in (106a) and (107):

(107) Suppose you keep your plants in a dark closet in the kitchen, and are worried because they are not growing. I can see what is going wrong:
You: I am worried about my plants.
Me: Oh, they simply do not have enough light. If they had enough light, they would be fine.

Arregui (2005, 2007) thus provides an analysis according to which the contrast between examples like (106a) and (106b) are not due to an 'extra' layer of past tense shifting the relation of historical accessibility or the time of conditional assessment to the past. Instead, the examples vary with respect to the epistemic sensitivity of the antecedent proposition. Perfective antecedents, such as (106a), result in proposition that can only be true in worlds in the context set. Thus, even if the modal in the conditional can in principle quantify over worlds outside the context set, a perfective antecedent will restrict the domain of quantification to worlds in the context set. To reach worlds outside the context set, a non-perfective (i.e. stative or perfect) antecedent must be used.

Whereas Arregui (2007) focused on the difference between perfective and perfect antecedents in *would*-conditionals, Arregui (2009) investigated the interpretation of past tense scoping over the modal. Building on Arregui (2005), Arregui (2009) proposed that the role of past tense is to anchor the interpretation of the conditional on particular actual world facts.[19] In this view, dubbed by Arregui a 'de re' analysis, counterfactuals make modal claims about past facts. The proposal shifts from the Stalnaker-Lewis 'global' similarity to a 'local similarity' view. Adopting a referential approach to tense, Arregui claims that tense identifies the facts that matter, with identification of past facts across worlds carried out via counterpart relations. The basic claim is that in order for a counterfactual of the from *if* α*, would* β to be actually true, the set of law-like worlds in which α is true that also contain a counterpart of the actual facts that past tense refers to must be a subset of the worlds in which β is true. Arguments in favor of the view that counterfactuals are about past facts are constructed by embedding counterfactuals in Gettier-type scenarios, making the assumption that knowledge attribution requires de re belief about facts (Kratzer 2002). Consider (108):

(108) Smith knows that if Nixon had pushed the button, there would have been a nuclear holocaust.
(Arregui 2009: 257)

Suppose that in the past the the button had been connected to an A-set of missiles, and that if those had been launched, there would have been a nuclear holocaust. Suppose also that Smith was aware that the button was connected to those missiles. But at some later point there was a change in military strategy, and the button was disconnected from the A-missiles and connected to a B-set of missiles. If those had been launched, there would also have been a nuclear holocaust, but Smith never actually found out that the wiring had been changed. In this case (as in the classic Gettier scenarios), we would not wish to say that (108) is true. This can be explained, following Kratzer's account of knowledge ascriptions, by saying that Smith is not properly acquainted with the facts the

[19] I will not enter into the technical details of the proposal nor discuss the full range of examples addressed in the paper. The reader is referred to Arregui (2009) for details.

counterfactual is about, thus providing evidence for the view that there are facts that the counterfactual is about. Further support for the claim that the role of past tense in counterfactuals is to resolve similarity by anchoring the counterfactual on past facts is provided by the observation that differences in the semantic status of tense actually affect the resolution of similarity. Tenses embedded in relative clauses have been shown to operate as 'free' temporal pronouns whereas tenses embedded under propositional attitude verbs have typically been construed as 'bound' (see e.g. Kratzer 1989). As the examples below illustrate, these differences in syntactic configuration make a difference for the interpretation of counterfactuals. Whereas we cannot judge (109a) true in the context provided, (109b) can be judged true:

(109) At the party, John met Jane and Jim met Joan. Jane and Joan had both been in the space program at NASA, though some years apart. They were both expelled.

a. #At that party, both men met a woman who would have been the first woman in space if she hadn't been expelled from NASA.
b. At that party, both men met a woman whom they believed would have been the first woman in space if she hadn't been expelled from NASA.

In (109a) there is a counterfactual in a relative clause. In this case, tense acts as a free pronoun and picks out some actual world past. But the counterfactual cannot be true if it is resolved with respect to a specific fact (neither the facts pertaining to Jane, Joan or both of them together will lead to truth). In the case of (109b), on the other hand, the counterfactual is in the complement of an attitude verb. In this case, tense is bound. The embedded proposition will be predicated of facts anchored in the belief states of John and Jim, allowing for distinct facts and permitting a true interpretation. Examples like (109) thus provide further support for the view that the resolution of similarity in counterfactuals is linked to the interpretation of tense.

One of the challenges that must be faced by theories that address the interpretation of tense in counterfactuals is to explain the mapping between syntax and semantics. In the 'real tense' proposals for counterfactuals, tense is generally taken to interact with the interpretation of the modal operator, however, tense and aspect morphology show up within the antecedent (and consequent) clause. Both Ippolito and Arregui address the syntax-semantics mismatch with theories according to which there is some form of agreement between the temporal morphology in the antecedent and the tense operators found higher in the structure. The issue has been addressed more recently by Gronn and von Stechow (2011) (who offer a typological overview, see also Gronn 2013). Gronn and von Stechow argue for a system of feature transmission that addresses tense morphology and incorporates an irrealis mood operator. The syntax-semantics interface has also been investigated by Romero (2014), who built on Dudman's original insights. According to Romero, a layer of past tense in counterfactuals is responsible for a Dudman-style 'back shift' that sets the evaluation time for an embedded conditional (other tenses are interpreted deicitically within the antecedent and consequent clauses).

7. Conclusion

It is perhaps unsurprising that counterfactuals continue to be a lively area of research. They have provided testing grounds for many of the important ideas in the semantic literature and continue to provide an arena for current and upcoming debates (see for example discussions of conditional questions (e.g. Isaacs and Rawlins 2008), gradability and degrees in modality (e.g. Lassiter 2011, forthcoming), optativity Rifkin 2000, Grosz 2010, 2012, Biezma 2011), as well as a variety of views on dynamic frameworks). At an intuitive level, counterfactuals seem to capture our imagination. As von Fintel (2011) has observed, conditionals 'whisk us away', and counterfactuals do a particularly good job.

References

Adams, Ernest. 1965. "The logic of conditionals*". Inquiry: An Interdisciplinary Journal of Philosophy*, 8:1-4, 166-197.

Adams, Ernest. 1966. "Probability and the Logic of Conditionals". In: J. Hintikka and P. Suppes (eds.), *Aspects of Inductive Logic*. Amsterdam: North-Holland, 265-316.

Adams, Ernest. 1970. "Subjunctive and Indicative Conditionals". *Foundations of Language* 6, 89-94.

Adams, Ernest. 1975. *The Logic of Conditionals*. Dordrecht: Reidel.

Aloni, Maria. 2003. 'Free Choice in Modal Contexts". In: M. Weisgerber (ed.), *Proceedings of the Conference "SuB7 – Sinn und Bedeutung"*. Arbeitspapier Nr. 114. Konstanz, 28–37.

Alonso-Ovalle, Luis. 2004. "Simplification of Disjunctive Antecedents". In: K. Moulton and M. Wolf (eds.): *Proceedings of the North East Linguistic Society*. Amherst, MA, 1–15.

Alonso-Ovalle, Luis. 2008. "Innocent Exclusion in an Alternative Semantics". *Natural Language Semantics*, 16:2, 115-128.

Alonso-Ovalle, Luis. 2009. "Counterfactuals, Correlatives, and Disjunction". *Linguistics and Philosophy* 32: 2, 207-244.

Anand, Pranav & Valentine Hacquard. 2009. "The Role of the Imperfect in Romance Counterfactuals". In: Martin Prinzhorn, Viola Schmitt and Sarah Zobel (eds.): *Proceedings of Sinn und Bedeutung 14*, Vienna, 37-50.

Anderson, Alan Ross. 1951. "A Note on Subjunctive and Counterfactual Conditionals". *Analysis* 12, 35-8.

Arregui, Ana. 2005. *On the accessibility of possible worlds: The role of tense and aspect*. Ph.D. Thesis, UMass, Amherst.

Arregui, Ana. 2005. "Layering modalities". Ms., University of Ottawa.

Arregui, Ana. 2007. "When aspect matters: the case of *would*-conditionals". *Natural Language Semantics*, 15, 221–264.

Arregui, Ana. 2009. "On similarity in counterfactuals". *Linguistics and Philosophy* 32: 3, 245-278.

Arregui, Ana & María Biezma. forthcoming. "Discourse rationality and the counterfactuality implicatures in backtracking conditionals." Proceedings of SuB.

Bennett, Jonathan. 1974. "Counterfactuals and Possible Worlds". *Canadian Journal of Philosophy* 4, 381-402.

Bennett, Jonathan. 1984. "Counterfactuals and Temporal Direction". *Philosophical Review* 93, 57-91.

Bennett, Jonathan. 2003. *A Philosophical Guide to Conditionals.* Oxford University Press, Oxford.

Bhatt, Rajesh & Roumyana Pancheva. 2006. "Conditionals". In: *The Blackwell Companion to Syntax*, Vol. I. Blackwell, 638–687.

Biezma, María. 2012. "Conditional inversion and GIVENNESS". In: *Proceedings of Semantics and Linguistic Theory* XXI, 552–571.

Biezma, María , Andrew Carnie & Daniel Siddiqi. 2014. "Counterfactuality in non-standard subjunctive conditionals". *Proceedings of NELS 43*, 25–36.

Bjorkman, Bronwyn M. 2011. *BE-ing Default: the Morphosyntax of Auxiliaries*. PhD Dissertation, MIT.

Bjorkman, Bronwyn M. & Claire Halpert. Forthcoming. "In an imperfect world: deriving the typology of counterfactual marking". In: Arregui, Rivero, Salanova (eds.), *Modality across categories*, OUP.

Bowie, G. Lee. 1979. "The Similarity Approach to Counterfactuals: Some Problems". *Noûs*, Vol. 13, No. 4, Special Issue on Counterfactuals and Laws, 477-498.

Chemla, Emmanuel. 2008. "An epistemic step for anti-presuppositions". *Journal of Semantics* 25:141–173.

Chisholm, Roderick M. 1946. "The Contrary-to-Fact Conditional". *Mind* 55. Vol. 55, No. 220, 289-307.

Condoravdi, Cleo. 2002. "Temporal interpretations of modals". In D. Beaver, S. Kaufman, & B. Clark (Eds.), *Stanford papers in semantics*. Palo Alto: CSLI Publications.

Cowper, Elizabeth & Daniel Currie Hall. 2008. "The Hungarian conditional: non-deictic counterfactuality". *Proceedings of the 2008 annual conference of the Canadian Linguistic Association*.

Cross, Charles. 2009. "Conditional Excluded Middle". *Erkenntnis*, Vol. 70, No. 2, 173-188.

Dancygier, Barbara & Eve Sweetser 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge: Cambridge University Press.

Davis, Wayne. 1979. "Indicative and subjunctive conditionals", *Philosophical Review* 88, 544-564.

DeRose, Keith. 1999. "Can It Be That It Would Have Been Even Though It Might Not Have Been?" *Philosophical Perspectives* 13, 385–413.

Downing, P.B. 1959. "Subjunctive conditionals, time order, and causation". *Proceedings of the Aristotelian Society*, New Series, Vol. 59 (1958-1959), 125-140.

Dudman, Victor. H.1983. "Tense and Time in English Verb Clusters of the Primary Pattern", *Australian Journal of Linguistics* 3, 25–44.

Dudman, Victor. 1984. "Conditional Interpretations of If-sentences". *Australian Journal of Linguistics* 4,

143-204.

Eagle, Anthony. 2007. 'Might' counterfactuals.

Edgington, Dorothy. 1995. "On Conditionals". *Mind* 104, 235-329.

Edgington, Dorothy. 2003. "Counterfactuals and the Benefit of Hindsight". In: P. Dowe and P. Noordhof (eds.), *Cause and chance: causation in an indeterministic world.* London: Routledge. 12-27.

Ellis, Brian, Franck Jackson & Robert Pargetter. 1977. "An Objection to Possible-World Semantics for Counterfactual Logics". *Journal of Philosophical Logic* 6, 35-57.

Fine, Kit. 1975. "Review of Lewis's Counterfactuals". *Mind* 84, 451-8.

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

von Fintel, Kai. 2001. "Counterfactuals in a Dynamic Context". In: M. Kenstowicz (ed.): *Ken Hale. A Life in Language.* Cambridge, MA: The MIT Press, pp. 123–153.

von Fintel, Kai. 2011. "Conditionals". In: Klaus von Heusinger, Claudia Maienborn and Paul Portnr (eds.), *Semantics: an international handbook of meaning*, vol. 2, Berlin, Boston: de Gruyter Mouton.

von fintel, Kai. 2012. "Subjunctive Conditionals". In: Gillian Russell and Delia Graff Fara (eds.), *The Routledge companion to philosophy of language*. New York: Routledge, 466-477.

von Fintel, Kai & Sabine Iatridou. 2002. "If and when if-clauses can restrict quantifiers". Ms. MIT.

Frank, Anette. 1997. *Context dependence and modal constructions*, PhD dissertation, University of Stuttgart.

Gibbard, Allan. 1981. "Two Recent Theories of Conditionals", in Harper et al. (eds.), *Ifs*, 211-47.

Gillies, Anthony. 2004. "Epistemic conditionals and conditional epistemics". *Noûs* 38, 585-616.

Gillies, Anthony. 2007. "Counterfactual scorekeeping." *Linguistics and Philosophy* 30, 329-360.

Goodman, Nelson. 1947. "The problem of counterfactual conditionals". *The Journal of Philosophy* 44, 113-118.

Grønn, Atle. 2013. "Aspect and tense in counterfactual main clauses: Fake or real?" In: Josephson & Söhrman (eds.), *Diachronic and Typological Perspectives on Verbs* [SLCS 134], Amsterdam: John Benjamins, 133-158.

Grønn, Atle & Arnim von Stechow. 2011. Tense in subjunctive conditionals across languages. Handout.

Hacquard, Valentine. 2006. *Aspects of modality*. Ph.D. Thesis, MIT.

Hacquard, Valentine. 2009. "On the interaction of aspect and modal auxiliaries". *Linguistics and Philosophy* 32, 279–315.

Hacquard, Valentine. 2010. "On the Event Relativity of Modal Auxiliaries". *Natural Language Semantics* 18: 1, 79-114.

Hájek, Alan. 2009. "Most counterfactuals are false". ms. Australian National University

Halpert, Claire & Hadil Karawani. 2012. "Aspect in counterfactuals from A(rabic) to Z(ulu)". In: *Proceedings of WCCFL* 29. Cascadilla Proceedings Project.

Harper, William L., Robert Stalnaker & Glenn Pearce (eds.). 1981. *Ifs: Conditionals, Belief, Decision, Chance, and Time*. D. Reidel Publishing Company, Dordrecht.

Han, Chung-hye. 2006. "Variation in form-meaning mapping between Korean and English counterfactuals." *Journal of East Asian Linguistics* 15, 167-193.

Hawthorne, John. 2005. "Chance and counterfactuals". *Philosophy and Phenomenological Research* 70: 2, 396-405.

Heim, Irene. 1984. "A Note on Polarity Sensitivity and Downward Entailingness." *Proceedings of NELS* 14, 98-107.

Heim, Irene. 1992. "Presupposition Projection and the Semantics of Attitude Verbs". *Journal of Semantics* 9,183–221.

Higginbotham, James. 2003. "Conditionals and compositionality". *Philosophical Perspectives* 17,181–94.

Huddleston, Rodney & Geoffery Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Huitink, Janneke. 2010. "Quantified conditionals and compositionality". *Language and Linguistics Compass* 4:1, 42-53.

Huitink, Janneke. 2012. "McGee's counterexample to modus ponens in context". In: B. Stolterfoht and S. Featherston (eds.), *Empirical approaches to linguistic theory: Studies in meaning and structure*, De Gruyter Mouton, 169-185

Iatridou, Sabine. 2000. "The grammatical ingredients of counterfactuality". *Linguistic Inquiry* 31, 231–270.

Iatridou, Sabine & David Embick. 1994. "Conditional Inversion". In: *Proceedings of NELS* 24, UMass, Amherst: GLSA, 189–203.

Ippolito, Michela. 2003. "Presuppositions and implicatures in counterfactuals". *Natural Language Semantics* 11, 145–186.

Ippolito, Michela. 2006. "Semantic Composition and Presupposition Projection in Subjunctive Conditionals." *Linguistics and Philosophy* 29, 631-672.

Ippolito, Michela. 2008. "Subjunctive conditionals". In: Grønn, Atle (ed.,) *Proceedings of Sinn und Bedeutung 12,* Oslo: ILOS, 256-270.

Ippolito, Michela. 2013. "Counterfactuals and conditional questions under discussion". In: *Proceedings of SALT* 23, 194–211

Ippolito, Michela. 2013.*Subjunctive conditionals: a linguistic analysis*. MIT Press.

Ippolito, Michela & Julia Su. 2014. "Counterfactuals, negation and polarity". In: Luka Crnicˇ and Uli Sauerland (eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim*, vol. 1, MITWPL 70, 225–243.

Jackson, Frank. 1977. "A Causal Theory of Counterfactuals". *Australasian Journal of Philosophy* 55, 3-21.

Kadmon, Nirit & Fred Landman. 1993. "Any". *Linguistics and Philosophy,* 16, 353-422

Kamp, Hans & Uwe Reyle. 1995. *From discourse to logic*, Kluwer Academic Publishers, Dordrecht.

Kaufmann, Stefan. 2005. "Conditional truth and future reference". *Journal of Semantics* 22, 231- 280.

Kaufmann, Stefan. 2013. "Causal premise semantics". Cognitive science 37: 6, 1136–1170.

Klinedienst, Nathan. 2011. "Quantified conditionals and conditional excluded middle". *Journal of Semantics 28: 1, 149-170.*

Kratzer, Angelika. 1977. "What 'must' and 'can' must and can mean", *Linguistics and Philosophy* 1, 337-355.

Kratzer, Angelika. 1979. "Conditional Necessity and Possibility". In: R. Buerle, U. Egli, and A. v. Stechow (eds.), *Semantics from Different Points of View*. Berlin: Springer-Verlag, 117–147.

Kratzer, Angelika. 1981. "The notional category of modality". In: Hans-Jurgen Eickmeyer and Hannes Rieser (eds.), *Words, worlds and contexts: new approaches in word semantics*, Berlin: Walter de Gruyter, 38–74.

Kratzer, Angelika. 1981. "Partition and revision: the semantics of counterfactuals". *Journal of Philosophical Logic* 10, 242–258.

Kratzer, Angelika. 1989. "An investigations of the lumps of thought". *Linguistics and Philosophy* 12, 607-653.

Kratzer, Angelika. 1991, "Conditionals". In: A. von Stechow and D. Wunderlich (eds.): *Semantics: An International Handbook of Contemporary Research*. Berlin: de Gruyter, 651–656.

Kratzer, Angelika. 2012. *Modals and Conditionals: New and Revised Perspectives*. Oxford University Press.

Ladusaw, William A. 1979. *Polarity Sensitivity as Inherent Scope Relations*. Ph.D. Dissertation, University of Texas, Austin.

Leahy, Brian. 2011 "Presuppositions and Antipresuppositions in Conditionals". *Proceedings of SALT* 21, 257–274.

Legate, Julie. 2003. "The morphosemantics of Warlpiri counterfactual conditionals." *Linguistic Inquiry*

34, 155-162.

Leslie, Sarah-Jane. 2009. " "If", "unless", and quantification". In: Robert J. Stainton and Christopher Viger (eds.), *Compositionality, context and semantic value: essays in honor of Ernie Lepore*. The Netherlands: Springer, 3-30.

Lewis, David. 1973. *Counterfactuals*. Cambridge, Mass.: Harvard University Press.

Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow". *Noûs* 13, 455–476.

Lewis, David. 1981. "Ordering Semantics and Premise Semantics for Counterfactuals". *Journal of Philosophical Logic* 10, 217-34.

Lewis, David. 1996. "Elusive knowledge". *Australasian Journal of Philosophy* 74: 4, 549-567.

McGee, Vann. 1985. "A Counterexample to Modus Ponens". *Journal of Philosophy* 82: 462-71.

Moss, Sarah. 2012. "On the pragmatics of counterfactuals". *Noûs* 46, 561-586.

Nevins, Andrew. 2002. "Counterfactuality without past tense." In: M. Hirotani (ed.), *The Proceedings NELS* 32. Amherst, MA: GLSA.

Nute, Donald. 1975. "Counterfactuals and the Similarity of Worlds". *Journal of Philosophy* 72, 773-8.

Nute, Donald. 1984. "Conditional Logic". In: D. Gabbay and F. Guenthner (eds.): *Handbook of Philosophical Logic,* Vol. II. Dordrecht: Reidel, 387-439.

Ogihara, Toshiyuki. 2000. "Counterfactuals, Temporal Adverbs, and Association with Focus". In: *Proceedings of SALT* 10, CLS Publications, Cornell University, Ithaca, N.Y. 115–131.

Ogihara, Toshiyuki. 2013. "Counterfactual conditionals and focus". *English Linguistics* 30: 2, 509–539.

Ogihara, Toshiyuki. 2014. "The semantics of '-ta' in Japanese future conditionals". In: Luka Crnič and Uli Sauerland (eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim*, vol. 2, MITWPL 71, 1–21.

Palmer, Frank R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

Pearl, Judea. 2000. *Causality: models, reasoning and inference*. Cambridge University Press.

Percus, Orin. 2006. "Antipresuppositions." In: A. Ueyama (ed.), Theoretical and empirical studies of reference and anaphora, Japan Society for the Promotion of Science: Report of the Grant-in-Aid for Scientific Research (B).

Piller, Christian. 1996. "Vann McGee's counterexample to Modus Ponens". *Philosophical Studies*, Vol. 82, 27-54.

Pollock, John L. 1976. *Subjunctive Reasoning*. Boston: Reidel.

Quine, Willard V. O. 1960. *Word and Object*. MIT Press.

Romero, Maribel. 2014. " 'Fake tense' in counterfactuals: a temporal remoteness approach". In: Luka Crnicˇ and Uli Sauerland (eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim*, vol. 2, MITWPL 71, 47–63

van Rooij, Robert. 2006. "Free Choice Counterfactual Donkeys". *Journal of Semantics* 23: 4, 383–402.

Santorio , Paolo. 2014. "Filtering Semantics for Counterfactuals: Bridging Causal Models and Premise Semantics", In: *Proceedings of SALT* 24.

Schlenker, Philippe. 2004. "Conditionals as Definite Descriptions (a Referential Analysis)". *Research on Language and Computation* 2: 3, 417–462.

Schulz, Katrin. 2007. *Minimal Models in Semantics and Pragmatics: Free Choice, Exhaustivity and Conditionals*. PhD dissertation, ILLC, University of Amsterdam.

Schulz, Katrin. 2011. "If you'd wiggled a, then b would've changed". *Synthese* 179: 2, 239–251.

Schulz, Katrin. 2014. "Fake Tense in Conditional Sentences. A Modal Approach". *Natural Language Semantics, 22*: 2, 117-144.

Schwarzschild, Roger. 1999. "GIVENness, AvoidF and Other Constraints on the Placement of Accent", *Natural Language Semantics* 7, 141–177.

Simons, Mandy. 2005. "Dividing Things Up: The Semantics of Or and the ModalOr Interaction". *Natural Language Semantics* 13, 271–316.

Slote, Michael A. 1978. "Time in Counterfactuals". *Philosophical Review* 87, 3-27.

Stalnaker, Robert. 1968. "A theory of conditionals". In: N. Rescher (ed.), *Studies in Logical Theory*. Basil Blackwell. Oxford, 98–112.

Stalnaker, Robert. 1975. "Indicative conditionals". *Philosophia* 5: 3, 269–286.

Stalnaker, Robert. 1978. "A Defense of Conditional Excluded Middle". In: Harper et al. (eds.), 87-104.

Stalnaker, Robert. 1984. *Inquiry*. Cambridge, Mass.: MIT Press.

Swanson, Eric. 2012. "Conditional Excluded Middle without the Limit Assumption*". Philosophy and Phenomenological Research*Vol. LXXXV No. 2, 301-321.

Thomason, Richmond. 1984. "Combinations of tense and modality". In: D. Gabby and F. Guenthner (eds.), *Handbook of Philosophical Logic*, vol. II, 135-165.

Thomason, Richmond, & Anil Gupta. 1980. "A Theory of Conditionals in the Context of Branching Time", reprinted in Harper et al. (eds.) 1981: 299-322.

Tichý, Pavel. 1976. "A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals". *Philosophical Studies* 29, 271-3.

Veltman, Frank. 1976. "Prejudices, presuppositions, and the theory of counterfactuals". In: J.

Groenendijk & M. Stokhof (eds), Amsterdam Papers in Formal Grammar. Proceedings of the 1st Amsterdam Colloquium. University of Amsterdam, 248-281.

Veltman, Frank.1985. *Logics for conditionals*. Ph.D. thesis, University of Amsterdam.

Veltman, Frank. 1996. "Defaults in update semantics". *Journal of Philosophical Logic* 25, 221-261.

Veltman, Frank. 2005. "Making counterfactual assumptions". *Journal of Semantics* 22, 159-180.

Warmbröd, Ken. 1981. "Counterfactuals and Substitution of Equivalent Antecedents". *Journal of Philosophical Logic* 10, 267-89.

Williams, J. Robert G. 2010. "Defending Conditional Excluded Middle". *Noûs* 44: 4, 650–668.

Woods, Michel. 1997. *Conditionals*. Oxford University Press.

Zimmermann, Ede. 2000. "Free choice disjunction and epistemic possibility". *Natural Language Semantics* 8, 255-290.