Had I'd've known better: "Extra" auxiliaries in English conditionals

---

A Thesis
Presented to
The Division of Philosophy, Religion, Psychology, and Linguistics
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

Aidan J. Malanoski

May 2020

Approved for the Division
(Linguistics)

Matt Pearson

Acknowledgments

Thank you to Matt and Sameer, for teaching me to think like a linguist.
Thank you to Roslyn, for treating me like a peer.
Thank you to Maya, for listening.
Thank you to Teo, for putting up with me.
Thank you to Mom, for supporting me.
Thank you to Eli, my love love, for always being there for me.
And finally, thank you to Lillie Case, Liv Johnson, Danny Riso, Timothy Shinners, Elena Turner, and Max Wink, for sharing your intuitions.

Table of Contents

Abstract

This thesis investigates English conditional constructions that seem to have "extra" auxiliaries. These conditionals fall into two types. The first type, **modal antecedent conditionals**, is characterized by the sequence of a modal and a perfect auxiliary in the conditional adjunct. These conditionals are distinctive in that they generate a counterfactual inference that cannot be cancelled, and that for some speakers, *would* seems to obligatorily receive a volitional interpretation in this environment. The second type of conditional with "extra" auxiliaries is **copy conditionals**, inverted conditionals in which the auxiliary appears both in C and in T. I argue that in copy conditionals, there is morphological reanalysis of the subject and T-head as a single head #D+T#, allowing the auxiliary to appear both in C, and as part of the Vocabulary Item inserted in #D+T#.

Dedication

To my family, to my mom, and to Eli

Chapter 1

Introduction to conditionals

This thesis concerns conditional constructions, found in some varieties of English, in which there seem to be "extra" auxiliaries. There are several variations, as demonstrated in the following examples:

(1) a. If he$_i$'d done the readings, John$_i$ wouldn't have failed the class.
    b. If he$_i$'d've done the readings, John$_i$ wouldn't have failed the class.

(2) a. Had he$_i$ done the readings, John$_i$ wouldn't have failed the class.
    b. Had he$_i$'d done the readings, John$_i$ wouldn't have failed the class.

The constructions in (1b) and (2b) have "extra" auxiliaries in the sense that they have more auxiliary verbs in the adjunct clause than do their counterparts in (1a) and (2a), despite being truth-conditionally equivalent. Additionally, all of the sentences implicate that John did not do the readings. However, this implicature is cancellable in all but (1b), as shown in (3b) below. The conditionals in (1a) and (1b) also contrast in that in the former, *he'd* unambiguously corresponds to *he had*, whereas in the latter, *he'd've* may correspond to *he had have* or *he would have*.

(3) a. If he'd done the readings, he wouldn't have failed the class. Since he didn't fail the class, he must have done the readings.
    b. If he'd've done the readings, he wouldn't have failed the class. #Since he didn't fail the class, he must have done the readings.

On the other hand, there is no difference in cancellability between (2a) and (2b). Additionally, whereas (1b) is also acceptable with *would* and (sometimes) *had* instead of nonsyllabic *'d* in the antecedent, (2b) is ungrammatical unless the auxiliary surfaces in its non-syllabic form, as shown in (4). Also, the contracted auxiliary in (2b) seems unambiguously to be *had*, not *would*.

(4) *Had he had done the readings, he wouldn't have failed the class.

I propose that despite their surface similarities, (1b) and (2b) represent distinct phenomena. The former construction, which I term a **modal antecedent conditional (MAC)**, is characterized by the sequence of a modal auxiliary – *would* or modal *had* – and the perfect auxiliary *have* in the conditional adjunct. Semantically,

modal antecedent conditionals differ from conditionals without an "extra" auxiliary in that they do not allow the counterfactual implicature to be cancelled. Furthermore, for some speakers, MACs with *would* in the antecedent seem to force a volitional reading of that modal.

For conditionals like (2b), on the other hand, I argue that the two instances of the auxiliary are copies of the same element. For this reason, I call such conditionals **copy conditionals**. Multiple copies can surface here because Morphology fuses the subject and the auxiliary, rendering the auxiliary invisible to linearization, which otherwise would force deletion of the lower copy.

This chapter lays out a syntactic and semantic framework for discussing conditionals. Section 1.1 gives an overview of the assumptions I make about the syntax. Section 1.2 presents the strengths and weaknesses of different approaches to the semantics of conditionals. Section 1.3 gives the compositional semantics that will be used in this thesis.

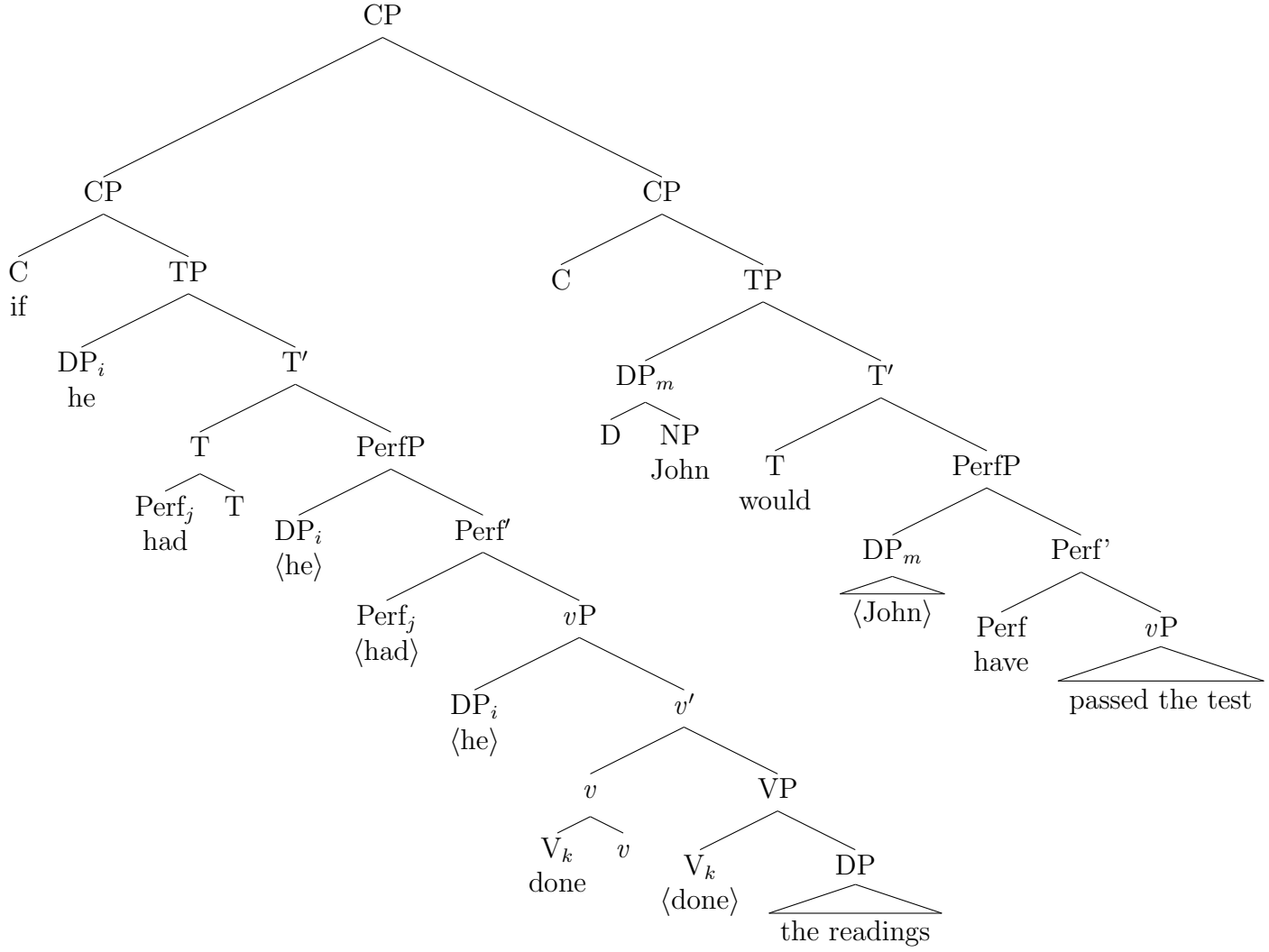
Before beginning our discussion, let us specify terminology. A prototypical English conditional has the form *if p, then q*. The proposition p is called the **antecedent**, and the proposition q is the **consequent**. The CP *if p* is known as the **conditional adjunct**. The full sentence *if p, then q* is a **conditional construction**, or more simply a **conditional**. For a proposition p, the term ***p*-world** refers to a possible world in which p is true. The **evaluation world** is the world with respect to which a statement is evaluated. Unless otherwise noted, I will assume that this is the actual world. Likewise, the **evaluation time** is the time with respect to which a statement is evaluated. As with the evaluation world, unless otherwise noted, I will assume that this is the actual time, *i.e.*, the present.

1.1 Syntax

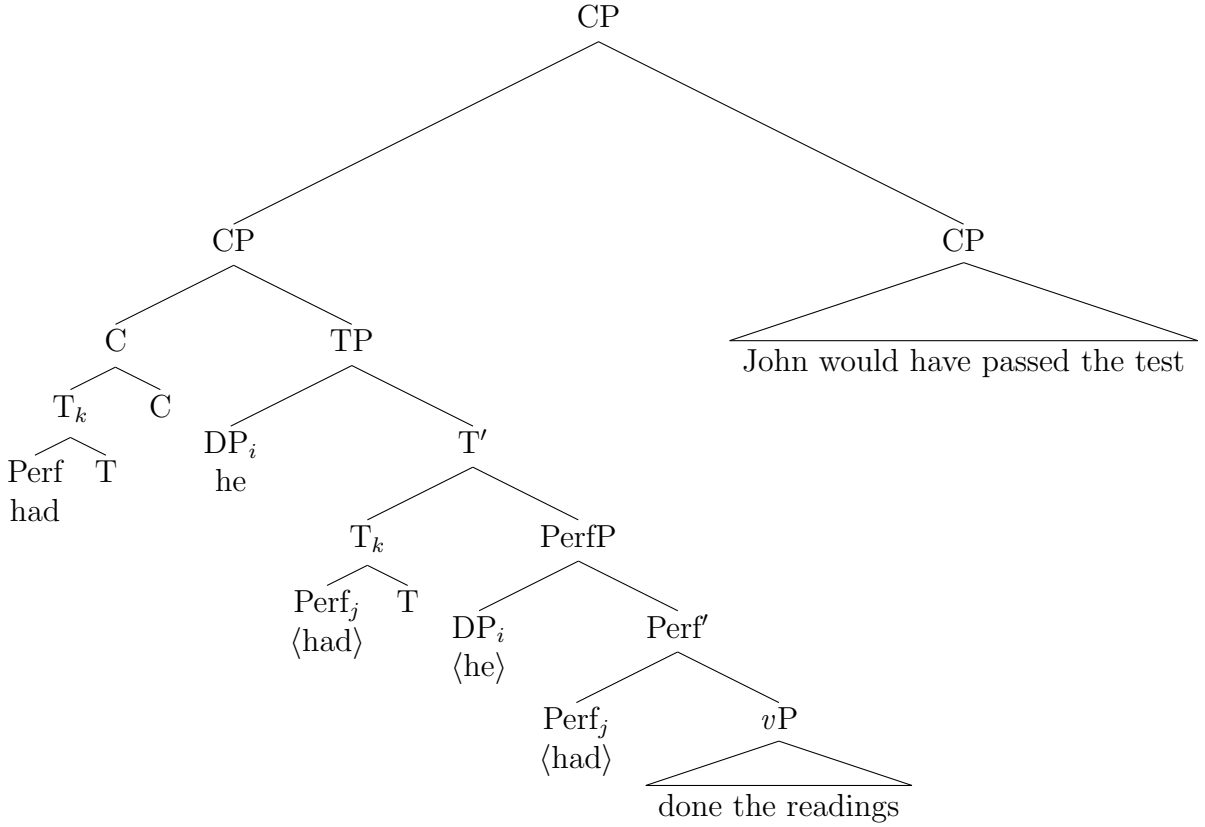
Regarding the syntax, I follow standard Minimalist assumptions about the structure of a sentence, including that a sentence is a projection of T(ense), and that there is functional head v that takes VP as its complement. The verb raises from V to v in the course of the derivation. Additionally, in transitive and unergative sentences, the subject base-merges (enters the tree) in Spec,vP, while in unaccusative and passive sentences, the subject base-merges inside VP. Finally, I assume that movement leaves behind a copy of the moved element.

Regarding conditionals, I assume that *if* is a complementizer, and that the clause introduced by *if* is an adjunct. Additionally, English allows the conditional adjunct to be marked by subject-auxiliary inversion, as seen in (5b). I assume that inversion here is T-to-C movement, that is, head adjunction of T to C. Finally, I assume that the perfect auxiliary *have* base merges below TP and moves to T in the course of the derivation. Thus, (1a), repeated below, has the structure in (5a), and (2a) has the structure in (5b). Indices are used not as a theoretical construct, but as a notational convenience for indicating movement.

(5) a. If he had done the readings, John would have passed the test.



b. Had he done the readings, John would have passed the test.



1.2 Conditional semantics

This section outlines several approaches to the semantics of conditional constructions. I first present three approaches – material implication, strict implication, and the Stalnaker/Lewis analysis – to how the truth of a conditional is determined. I then discuss an important distinction between types of conditionals.

1.2.1 Approaches to conditionals

A straightforward approach to the semantics of conditionals would be to treat them as **material implication**, so that *if p, then q* is true iff $p \rightarrow q$. In other words, *if p, then q* is false only when p is true and q is false. This analysis has a number of limitations, however. First, it predicts that any sentence with a false antecedent is true, which is not the case (von Fintel, 2011):

(6) If circles were squares, then rectangles would be circles.

Most would judge (6) to be false, but since the antecedent is false, this sentence is true under material implication. A material implication analysis also fails to accurately describe the behavior of negated conditionals, as in the following example from von Fintel & Heim (2011, 51):

(7) a. It's not true that if there is a major earthquake in Cambridge tomorrow, my house will collapse.

b. There will be a major earthquake in Cambridge tomorrow, and my house will fail to collapse.

(7a) is the natural language version of $\neg(p \rightarrow q)$ ("it is not the case that p implies q"), where p is the proposition *there is a major earthquake in Cambridge tomorrow* and q is the proposition *my house will collapse*. However, $\neg(p \rightarrow q)$ is truth-conditionally equivalent to $p \wedge \neg q$ ("p is true and q is false"), which is expressed by (7b). Thus, material implication falsely predicts that (7a) and (7b) are equivalent.

Analyzing conditionals as an instance of **strict implication** addresses some of the shortcomings of the material implication analysis. Under a strict implication analysis, *if p, then q* is true if and only if the possible worlds in which p is true are a subset of the worlds in which q is true – in other words, if and only if q is true in all worlds where p is true (von Fintel, 2011). However, von Fintel & Heim (2011) note that this raises the problem of contingency: not all p-worlds will be relevant to the evaluation of a conditional. Consider the following:

(8) If you were born in the United States, then you are American.

One can imagine a possible world in which the United States does not have birthright citizenship; in such a world, p may be true while q is false. Thus, strict implication predicts (8) to be false. Yet the sentence is true (at least at the time of writing), so it seems that not all possible worlds are relevant to the analysis of a conditional.

Strict implication makes other false predictions. For example, under strict implication, if a conditional *if p, then q* is true, then *if* $p \wedge r$*, then q* will also be true,

since the worlds in which p and r are true are a subset of the worlds in which p is true (von Fintel, 2011). This prediction, called *strengthening the antecedent*, does not hold, however, as (9b) below does not follow from (9a):

(9) a. If you met Molly, you would like her.
b. If you met Molly and she was mean to you, you would like her.

Strict implication also predicts that transitivity ((*if p, then q*) $\wedge$ (*if q, then r*) $\Rightarrow$ *if p, then r*) holds, since if the p-worlds are a subset of the q-worlds, and the q-worlds are a subset of the r-worlds, then the p-worlds are a subset of the r-worlds. For example, even if both (10a) and (10b) below are true, it does not follow that (10c) is true as well.

(10) a. If Ronald gets sick, he will lose weight.
b. If Ronald loses weight, he will be healthier.
c. If Ronald gets sick, he will be healthier.

Finally, strict implication predicts that contraposition (*if p, then q* $\Rightarrow$ *if* $\neg q$*, then* $\neg p$) holds, since if the p-worlds are a subset of the q-worlds, then the $\neg q$-worlds are a subset of the $\neg p$-worlds (von Fintel, 2011). To illustrate, it is easy to imagine a situation in which (10a) above is true, but (11) is not, so contraposition does not hold:

(11) If Ronald does not lose weight, he will not get sick.

The problem with strict implication is that in evaluating a conditional, we only consider possible worlds that are in some way contingent on the evaluation world, a fact which a strict implication analysis fails to account for. Robert Stalnaker and David Lewis independently addressed the problem of contingency by proposing that in evaluating a conditional *if p, then q*, speakers consider only a subset of the worlds in which p is true (von Fintel, 2011). In particular, under such a **Stalnaker/Lewis** semantics, the set of possible worlds is ordered by each world's similarity to the evaluation world. A conditional is then true if q is true in all the worlds where p is true that otherwise do not differ from the evaluation world. This solves the problematic inferences that arise under strict implication. First, in evaluating (8), we consider not all the worlds where you were born in the United States, but rather just those worlds where you were born in the United States but that otherwise do not differ from the actual world. These are worlds where the United States has birthright citizenship, just as in the actual world, so we judge (8) to be true.

Additionally, we no longer expect strengthening the antecedent to hold, because the $(p \wedge r)$-worlds that are most similar to the evaluation world are not necessarily a subset of the p-worlds that are most similar to the evaluation world. Transitivity and contraposition also do not hold under a Stalnaker/Lewis analysis. Consider again the examples from above:

(12) a. If you met Molly, you would like her.
b. If you met Molly and she was mean to you, you would like her.

(13) a. If Ronald gets sick, he will lose weight.

b. If Ronald loses weight, he will be healthier.

c. If Ronald gets sick, he will be healthier.

d. If Ronald does not lose weight, he will not get sick.

The possible worlds in which you meet Molly and which are maximally similar to the actual world are presumably not worlds where she is mean to you, or else I would not say that you would like her. So, the worlds described by the antecedent of (12b) are not a subset of those described by the antecedent of (12a), and we no longer expect the former to follow from the latter. Thus, strengthening the antecedent does not hold.

Likewise, the p-worlds of (13b) are the possible worlds in which Ronald loses weight which are otherwise maximally similar to the actual world. Presumably, Ronald is not sick in the actual world, or else (13a) would be infelicitous. The p-worlds of (13b) are thus worlds where Ronald loses weight but is not sick. By contrast, the q-worlds of (13a) are worlds where Ronald loses weight because he is sick. So, the q-worlds of (13a) are not the same as the p-worlds of (13b), and we no longer expect transitivity to hold.

Finally, contraposition does not hold: (13d) does not follow from (13a). This is because (13a) asserts that the possible worlds in which Ronald gets sick which are maximally similar to the actual world are worlds where he will lose weight. Because of the similarity restriction, this is only a subset of the possible worlds where Ronald gets sick, so although the most similar worlds where Ronald gets sick are worlds where he loses weight (q), in less similar worlds where Ronald gets sick, he will not lose weight ($\neg q$). Since some worlds where Ronald does not lose weight ($\neg q$-worlds) are still worlds where he is sick (p-worlds), the truth of (13d) does not follow from (13a).

Because it solves the problem of contingency, the Stalnaker/Lewis approach is commonly assumed in studies of conditionals, and it is the approach I will assume here. In 1.3 below, I will give an implementation of these ideas in compositional semantics. But first, let us consider the varieties of conditionals.

1.2.2 Types of conditionals

Several types of conditional are distinguished in the semantic literature; of these, the opposition between **open** and **remote conditionals** is most relevant to the present study.[1] In English, these two types of conditionals can be distinguished morphosyntactically. Remote conditionals are typically marked by past tense in the antecedent, and a modal such as *would, could, should*, or *might* in the consequent, as in (14).

[1]Open and remote conditionals are more commonly referred to as *indicative* and *subjunctive* conditionals, respectively. This terminology is based on Latin morphology, and does not reflect the morphosyntactic expression of these conditionals in many other languages (Huddleston & Pullum, 2002). It is thus misleading to those not versed in the literature on conditionals. Consequently, I follow Huddleston & Pullum (2002) in referring to these as open and remote conditionals.

(14) If Saul left at noon, he would/could/should/might make it to the airport in time for his flight.

Some authors have treated *would* as the past tense form of *will* (Iatridou, 2000; Ippolito, 2013). I follow this analysis, and furthermore assume that at least in remote conditionals, *could, should* and *might* act as past tense forms of *can, shall* and *may*. We can then generalize our morphosyntactic description of remote conditionals as follows: remote conditionals require a modal in the consequent, and past tense in both the antecedent and the consequent. Note that this correctly allows for modals in the antecedent, as seen in (15)-(16) (although *should* and *might* are, to me, highly formal, and largely restricted to politeness contexts). Indeed, modal antecedent conditionals, under study here, are a type of remote conditional.

(15) If Debra would/could go the party tomorrow, Sara would be really happy.

(16) If you should/might fetch me a coffee, I would be most pleased.

Finally, note that the past tense morphology in a remote conditional does not serve to locate the eventualities described by the antecedent in the past, as indicated by the compatibility of past tense with future-oriented adverbials in (17a) (Iatridou, 2000). To form a remote conditional about the past, the perfect auxiliary *have* is required in the consequent, and in some varieties of English, past perfect morphology is required in the antecedent, as demonstrated in (17b) (Ippolito, 2013). In other varieties of English, past perfect morphology is not required in the antecedent, as shown in (17c).

(17) a. If Alonso left tomorrow, he would miss the ball game.
b. If Beck had found out about it yesterday, he would have told me.
c. % If Beck found out about it yesterday, he would have told me.

Under the morphosyntactic definitions I am using here, an open conditional is any conditional that is not a remote conditional. Consequently, open conditionals allow a greater variety of structures than do remote conditionals, as seen in (18a).

(18) a. If Matt is a linguist, then he knows a lot of languages.
b. If Matt is a linguist, then he will/can/may/must know a lot of languages.
c. If Matt is a linguist, then he could/should/might/??would know a lot of languages.

In addition to their morphosyntactic differences, open and remote conditionals also differ in meaning. Informally, remote conditionals suggest that the eventuality expressed by the antecedent is false or unlikely; this is the so-called **counterfactual** inference. Open conditionals, on the other hand, seem to leave the truth of the antecedent open (von Fintel, 2011). Thus in the remote conditional (19a), I seem to view my winning the lottery as unlikely, whereas in (19b), I am much more hopeful.

(19) a. If I won the lottery, I would pay off my student loans.
b. If I win the lottery, I will pay off my student loans.

That said, the counterfactual inference is an implicature, and thus cancellable. This can be seen in "detective reasoning" contexts, as in (20) (Anderson, 1951). Additionally, open conditionals with false antecedents like (21) are also "counterfactual" in the sense that they are actually untrue (Ippolito, 2013). In short, the falsity of the antecedent is insufficient to distinguish open and remote conditionals.

(20) If John had got his car fixed, he would have driven himself to the party. Since he came by himself, his car must be working again.

(21) If you are famous, then I am the Queen of Sheba.

1.3 A compositional semantics for conditionals

This section provides an implementation of the Stalnaker/Lewis analysis of conditionals in Fregean compositional semantics (see Heim & Kratzer 1998). Specifically, it presents the semantics given by Ippolito (2013), which builds on Angelika Kratzer's influential approach to conditionals (Kratzer, 1981, 1986). Under Kratzer's approach, a conditional consists of a modal operator quantifying over possible worlds. The antecedent restricts the domain of quantification of the modal, and the consequent provides the nuclear scope. Furthermore, conditionals are interpreted relative to a **conversational background**, which identifies the possible worlds relevant to the interpretation of the conditional, and an **ordering source**, which imposes an ordering on a set of worlds. As we will see, these parameters provide a way of ranking similarity to the evaluation world, as in a Stalnaker/Lewis analysis.

Ippolito (2013) implements the conversational background and the ordering source as a **historical accessibility function** and a **similarity function**, respectively. The historical accessibility function $[\![\text{HIST}]\!]$ takes as input a proposition p and a world w' and returns true if p is true in w', and w' has the same history as the evaluation world up to the evaluation time t. This is formalized in (22) (Ippolito, 2013, 57).[2] The similarity function $[\![\text{SIM}]\!]$ takes as input a proposition p and a world w' and returns true if p is true in w' and w' is as or more similar to the evaluation world w than any other world w'' (Ippolito, 2013, 58). This is formalized in (23).

(22) $[\![\text{HIST}]\!]^{c,g,t,w} = \lambda p_{<s,t>}.\lambda w'.$ w' has the same history as w up to t and $p(w') = 1$

(23) $[\![\text{SIM}]\!]^{c,g,t,w} = \lambda p_{<s,t>}.\lambda w'.p(w') = 1 \wedge \neg\exists w''[p(w'') = 1 \wedge w'' <_w w']$, where $w' <_w w''$ reads "w' is more similar to w than w''"

In plainer language, HIST says that the possible worlds relevant to the interpretation of a conditional are those worlds in which the antecedent p is true and which share the evaluation world's history up to time t. These worlds are thus contingent on the evaluation world, since they must share history with the evaluation world. SIM requires that we only consider possible worlds that are as similar to the evaluation world as possible. Specifically, if these possible worlds differ from the evaluation

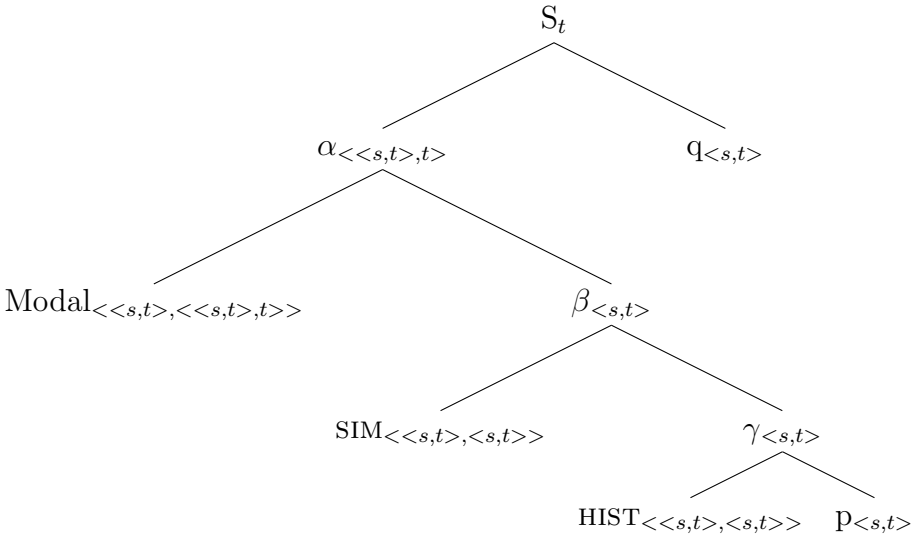
[2] $[\![\text{HIST}]\!]^{c,g,t,w}$ reads "the denotation of HIST with respect to *c, g, t,* and *w*." The superscript c is the context of utterance; g is the assignment function, which maps deictic elements such as pronouns to their referents; t is the evaluation time; and w is the evaluation world.

world, they differ only in that p and its presuppositions are true. The proposition p and its presuppositions may or may not be true in the evaluation world. Together, these parameters restrict the domain of quantification to the set of p-worlds maximally similar to the actual world, as desired in a Stalnaker/Lewis semantics.

As mentioned above, Kratzer proposes that a conditional consists of a modal operator quantifying over possible worlds. Under this analysis, even in conditionals with no overt modal operator, there is a covert operator. Thus in (24b), there is a covert epistemic modal, so that that (24a) and (24b) are roughly equivalent (von Fintel, 2011). The basic structure of a conditional is given in (25). I use the term **bare conditional** to refer to structures such as (25) in which there is no temporal operator scoping over the rest of the structure, as will be the case in the next section. Generally, in a bare conditional, the root of the tree immediately dominates q and the mother of the modal operator.

(24) a. If Rafael is not here, he must be at home.
b. If Rafael is not here, he is at home.

(25)



It is worth noting that under this semantics, the modal does not merge with q. This is at odds with the syntax, where the modal occurs in the CP corresponding to the proposition q, as in (24a), raising the question of how the semantic structure is derived from the syntactic one. Such a question is beyond the scope of this paper. That said, this general structure for conditionals is fairly well-established in the literature.

1.3.1 Simple past remote conditionals

Having established a general semantics for conditionals, we can now turn to the question of the role of past tense. In a remote conditional with simple past tense morphology, which I will refer to as a **simple past remote conditional (SPRC)**, Ippolito argues that the tense provides the **accessibility time**, the time with respect to which $[\![\text{HIST}]\!]$ is evaluated. Thus, the set of historically accessible worlds is now determined at some time t' that precedes the **reference time**, the evaluation time of the entire conditional structure. Additionally, the historical accessibility function will

take an argument of type i (an interval of time) in addition to its other arguments, so that it will be of type $< i, << s, t >, << s, t >>>>$.

Now, Ippolito (2013) assumes a **branching futures** semantics, the basic idea of which is as follows. At any given moment in time, there are many possibilities for what the next moment will be, and each of these possibilities is a different possible world. For example, say that I am at the grocery store, deciding whether to buy milk. This choice gives rise to a possible world in which I buy milk at that time, and a possible world in which I do not, only one of which is the actual world. Consequently, if a possible world w' diverges from the actual world before the present, then that world is known to be false: it is counterfactual. However, the future is unknown, so if a possible world diverges from the actual world at or after the present, that possible world is a **possible future** of the actual world.

Given a world w and times t, t' such that t' precedes t, the possible worlds accessible from w at t' include all the possible worlds accessible from w at t, since t is a continuation of t'. However, there are also possible continuations of w accessible at t' that have been been foreclosed at t. Thus, there are more possible worlds available at t' than at t – the number of accessible possible worlds shrinks over time. To make this concrete, suppose that I have bought my milk, and am deciding whether to walk home or take the bus. There are two possible worlds currently accessible from the actual world at the present time: the world where I bought milk and walked home, and the world where I bought milk and took the bus. However, at the point where I was deciding to buy milk, there were four possible worlds available:

- buy milk, walk
- buy milk, take bus
- do not buy milk, walk
- do not buy milk, take bus

Thus, the possible worlds currently accessible from the actual world are a subset of the possible worlds available in the past.

In a simple past remote conditional, historical accessibility is determined at some time t' prior to the reference time t, so the domain of quantification of a remote conditional consists of worlds that share the evaluation world w's history up to t'. Since t' precedes t, this includes all the possible worlds accessible at t as well as counterfactual worlds that have been foreclosed at t. This contrasts with an open conditional, where the accessibility time is the present time t, and so we consider only worlds that share w's history up to t. The domain of quantification of an open conditional is thus a subset of the domain of quantification of a remote conditional, and so a remote conditional is less informative than an open conditional. By the maxim of quantity (Grice, 1975), the choice of a remote conditional over an open conditional thus gives rise to the implicature that the relevant worlds are those that diverge from w before t – that is, counterfactual worlds. Thus, the counterfactual inference emerges as a conversational implicature, and is therefore cancellable.

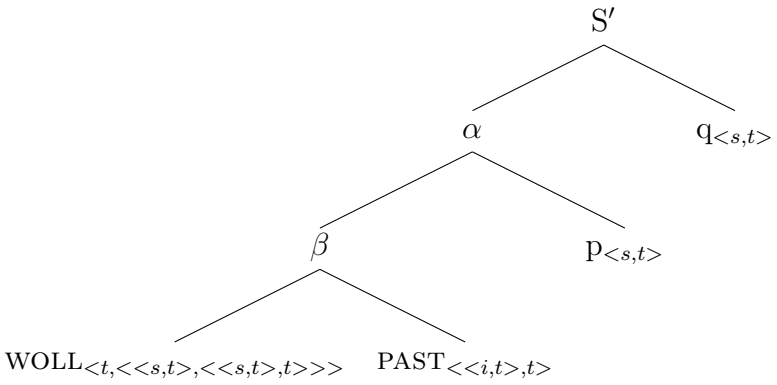
We mentioned above the antecedent of a conditional restricts the domain of quantification of some modal operator. Ippolito only considers remote conditionals with *would* as their modal, so for her, this modal is WOLL, a label used in the semantic literature for the tenseless form of *will* and *would*. (26) gives her definition for WOLL (Ippolito, 2013, 60).[3] ⟦WOLL⟧ takes as input a time t', a proposition p, and a proposition q, and returns true if all the p-worlds that are historically accessible from the evaluation world at time t' and that are maximally similar to the evaluation world are worlds where q is true. This is a function of type $< i, << s, t >, << s, t >, t >>>$

(26) $[\![\text{WOLL}]\!]^{c,g,t,w} = \lambda t' \in D_i.\lambda p_{<s,t>}.\lambda q_{<s,t>}.\forall w'[w' \in \text{SIM}_w(\text{HIST}_{w,t'}(p)) \rightarrow w' \in q]$

Ippolito then gives the definition of past tense in (27) (Ippolito, 2013, 60). ⟦PAST⟧ takes as input a predicate of times and returns true if there exists a time prior to the evaluation time where that predicate is true. Since ⟦PAST⟧ is of type $\langle\langle i, t\rangle, t\rangle$, however, it cannot serve as an argument to ⟦WOLL⟧, which requires an argument of type i. Consequently, PAST undergoes quantifier raising, moving to the top of the tree and leaving behind a trace of type i. This trace can then serve as the input to ⟦WOLL⟧. (28) gives the semantic structure before quantifier raising has applied; (29) gives the structure after.[4]

(27) $[\![\text{PAST}]\!]^{c,g,t,w} = \lambda P_{<i,t>}.\exists t' < t : P(t') = 1$, where $t' < t$ reads "t' precedes t"

(28)



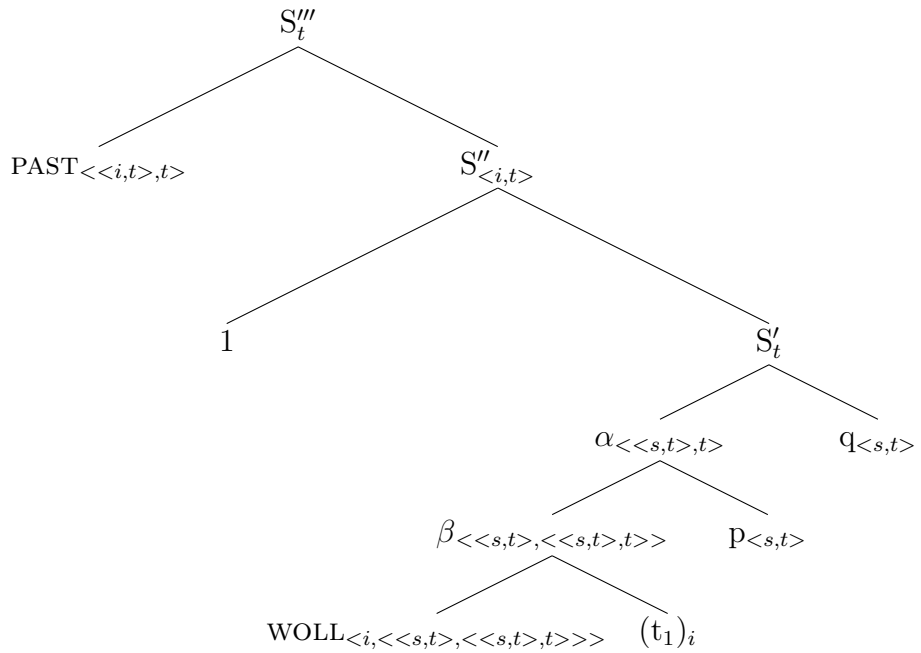
[3] To save space in her trees, Ippolito incorporates HIST and SIM into the definition of WOLL, rather than giving them their own nodes. I will follow this convention.

[4] The index 1 is introduced in the course of quantifier raising. The node S′ composes with the index 1 to form a node of type <i,t> by *predicate abstraction*, defined as follows (Heim & Kratzer, 1998, 186):

i. Let α be a branching node with daughters β and γ, where β is a numerical index i. Then for any assignment function a, $[\![\alpha]\!]^a = \lambda x.[\![\gamma]\!]^{a^{x/i}}$.

An assignment function a maps indices onto referents. The assignment function $a^{x/i}$ is the assignment function a modified so as to map the index i onto x.

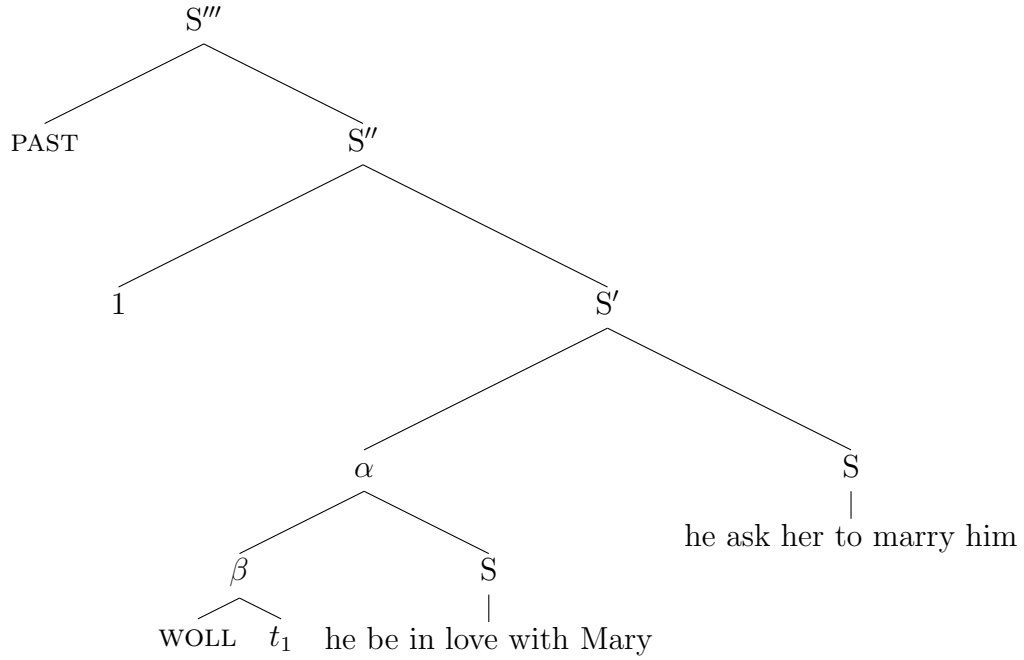
(29)



To take a concrete example, consider the sentence in (30), whose denotation is given in (32). According to this, (30) is true just in case all the worlds [i] where John is in love with Mary that are [ii] historically accessible from the actual world at some past time and [iii] maximally similar to actual world are worlds where John asks Mary to marry him.

(30) If John were in love with Mary, he would ask her to marry him.

(31)



(32) $[\![\text{PAST}\ [1[[[\text{WOLL}\ t_1][\text{he be in love with Mary}]]\ [\text{he ask her to marry him}]]]]\!]^{c,g,t,w}$ = 1 iff $\exists t' < t[\forall w'[w' \in \text{SIM}_w(\text{HIST}_{w,t'}(\lambda w''.\text{John is in love with Mary in } w'')) \rightarrow w' \in \{w''' : \text{John asks Mary to marry him in } w'''\}]]$

To summarize, in a remote conditional, the accessibility time (the time at which the set of historically accessible worlds is identified) precedes the reference time (the time with respect to which the entire conditional is evaluated), so the domain of quantification of the conditional includes possible worlds not available at the reference time: counterfactual worlds. The maxim of quantity leads to the implicature that the worlds relevant to evaluating the conditional are the counterfactual worlds, thus generating the counterfactual implicature.

However, this is not all there is to the meaning of a remote conditional. Consider the following example (Ippolito, 2013, 55):

(33) John is dead. #If he were in love with Mary, he would ask her to marry him.

Ippolito argues that the reason (33) is anomalous is because the predicate *be in love* presupposes that "the subject of the predicate exists as the time of predication" (Ippolito, 2013, 75), an idea she takes from Musan (1997). Based on examples like (33), Ippolito argues that the presuppositions of a conditional need to be compatible with the set of worlds historically accessible not at the accessibility time of $[\![\textsc{woll}]\!]$, but at the reference time. In other words, a remote conditional presupposes that there exists a possible world, historically accessible from the actual world at the reference time, such that the presuppositions of the bare conditional are true in that world.

In a simple past remote conditional, Ippolito argues that the reference time is the utterance time, so that in (33), the existence presupposition must be satisfied at the utterance time. Since at the utterance time, John is dead, and thus does not "exist" in the relevant sense, (33) is infelicitous. Incorporating this into the definition of WOLL gives (34). Note that t here is the reference time.

(34) $[\![\textsc{woll}]\!]^{c,g,t,w} = \lambda t' \in D_i.\lambda p_{<s,t>} : \textsc{hist}_{w,t} \cap ps(p) \neq \varnothing.\lambda q_{<s,t>} : (\textsc{hist}_{w,t} \cap p) \cap ps(q) \neq \varnothing.\forall w'[w' \in \textsc{sim}_w(\textsc{hist}_{w,t'}(p)) \rightarrow w' \in q]$

The restriction on p in (34) requires that there be worlds historically accessible from the evaluation world at time t where the presuppositions to p (abbreviated $ps(p)$) are true. The restriction on q requires that there be p-worlds historically accessible from the evaluation world at time t where the presuppositions to q are true. These restrictions reflect the fact that the presuppositions of a conditional *if p, then q* are the presuppositions of p and the presuppositions of q not entailed by p.

Returning to the example in (30), repeated below, the truth conditions for the conditional are the same. However, according to the modified definition of WOLL, (30) is defined only if its presuppositions are satisfied at the utterance time. So, John must be alive at the utterance time, explaining the infelicity of (33).

(35) If he were in love with Mary, he would ask her to marry him.

In short, the past tense in a remote conditional shifts the accessibility time to the past. Consequently, the domain of quantification of a remote conditional includes both counterfactual worlds and possible futures of the actual world. This contrasts with an open conditional, where the domain of quantification only includes possible futures of the actual world. Because the domain of quantification of an open conditional is a subset of that of a remote conditional, the choice of a remote conditional

over an open conditional gives rise to the implicature that only the counterfactual worlds are relevant to interpreting the conditional, rather than the possible futures of the actual world. Additionally, to be felicitous, the presuppositions of a conditional must be compatible with the actual world at the reference time, the time with respect to which WOLL is evaluated.

1.3.2 Past perfect remote conditionals

This section concerns remote conditionals with past perfect morphology (**past perfect remote conditionals (PPRCs)**), which differ from simple past remote conditionals in important ways. To begin the discussion, Ippolito notes three facts about remote conditionals with past perfect morphology. First, such conditionals can describe both past and nonpast eventualities (36). Second, it is not possible to cancel the counterfactual inference in a nonpast remote conditional with past perfect morphology – they are **strongly counterfactual** ((37), from Ippolito 2013, 92). Third, they are felicitous in different contexts than the corresponding conditionals with simple past morphology (38).

(36) If you had arrived yesterday/tomorrow, I would have been there to greet you.

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

(38) Simone is competing for a promotion at work. Unfortunately, she missed an important business meeting yesterday due to a traffic jam caused by inclement weather. Tomorrow's weather is expected to be perfect.

a. #If the meeting were tomorrow (instead), Simone would make it.

b. If the meeting had been tomorrow (instead), Simone would have made it.

Generally, when an eventuality is described with past perfect morphosyntax, this indicates that the eventuality precedes some salient past time t'. In (39), for example, the eventuality of John leaving precedes Mary's arrival, which is itself a past event. Because past perfect morphosyntax describes eventualities that are anterior to some past time, Ippolito views past perfect remote conditionals as introducing another layer of past tense. Furthermore, because past perfects are infelicitous without a salient reference time, as seen in (40), Ippolito proposes that this second layer of past tense is deictic, unlike PAST above, which is quantificational. She gives this second layer of past tense the denotation in (41) (Ippolito, 2013, 81). According to this, the assignment function g maps $[\![\mathbf{past}_k]\!]$ to its referent $g(k)$, a point in time that precedes the evaluation time.

(39) John had (already) left by the time Mary arrived.

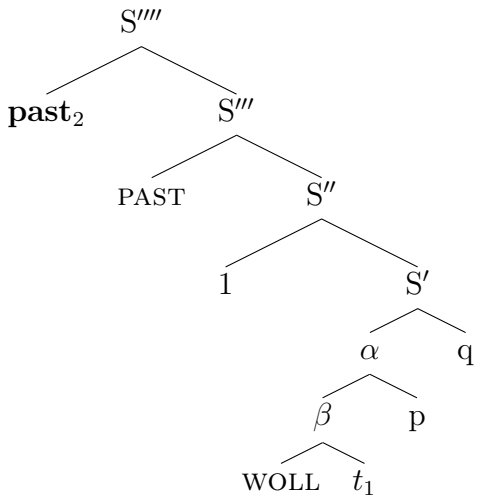
(40) # John had left.

(41) $[\![\mathbf{past}_k]\!]^{c,g,t,w}$ defined only if $g(k) < t$; if defined, $[\![\mathbf{past}_k]\!]^{c,g,t,w} = g(k)$

Ippolito proposes (42) as a possible structure for past perfect remote conditionals. The quantificational past $[\![\text{PAST}]\!]$ shifts the accessibility time to the past, just as it

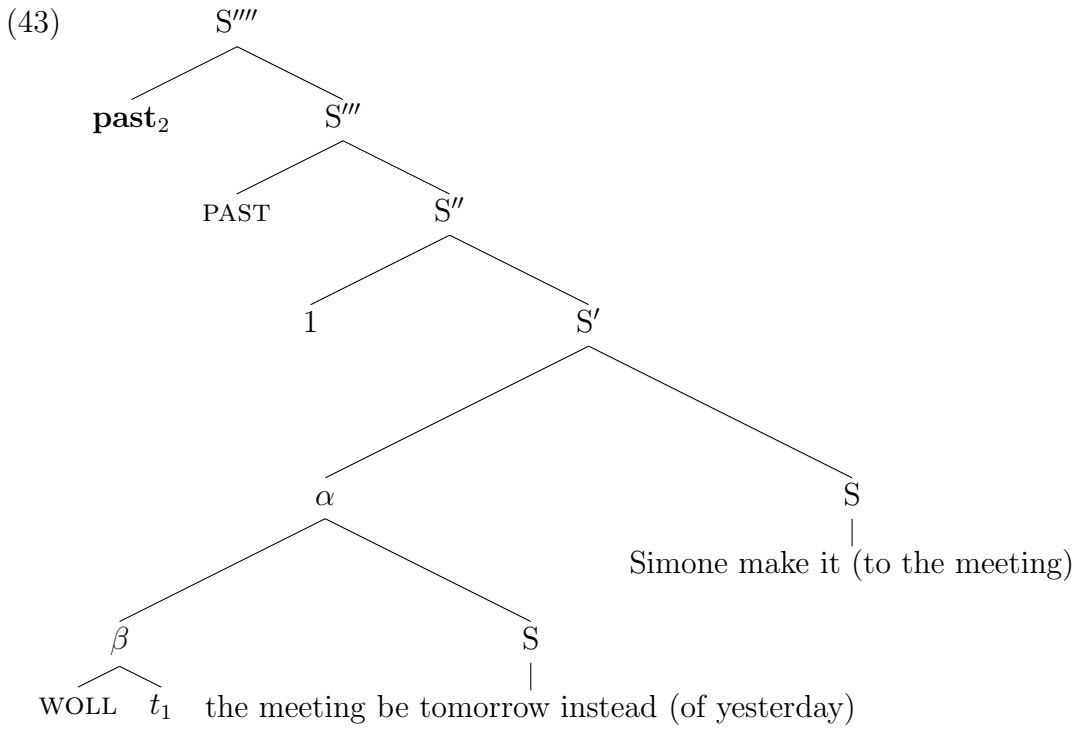
does in a simple past remote conditional. The deictic past ⟦**past**⟧ then shifts the reference time – the time at which the presuppositions of the conditional must be satisfied – to the past. The accessibility time precedes the reference time. Thus, a past perfect remote conditional differs from a simple past remote conditional in that the presuppositions of the conditional must be satisfied in the past rather than the present.

(42)



As an example, let us return to (38b), whose structure and truth conditions are given in (43) and (44), respectively. According to (44), (38b) is true just in case all the worlds [i] where the meeting happens tomorrow that are [ii] historically accessible from the actual world at some time before $g(2)$ (which is a past time) and [iii] maximally similar to actual world are worlds where Simone makes it to the meeting. This conditional is only defined if the presuppositions of the bare conditional are true at $g(2)$.

(43)



(44) $[\![\textbf{past}_2\ [\textsc{past}\ [1\ [\ [\ [\textsc{woll}\ t_1][\text{the meeting be tomorrow instead (of yesterday)}]]\ [\text{Simone make it to the meeting}]]]]]\!]^{c,g,t,w} = 1$ iff $\exists t' < g(2)[\forall w'[w' \in \textsc{sim}_w(\textsc{hist}_{w,t'}(\lambda w''.\text{the meeting is tomorrow instead of yesterday in } w'')) \rightarrow w' \in \{w''' : \text{Simone makes it to the meeting in } w'''\}]]$, defined iff $\textsc{hist}_{w,g(2)} \cap ps(p) \neq \varnothing$ and $(\textsc{hist}_{w,g(2)} \cap p) \cap ps(q) \neq \varnothing$

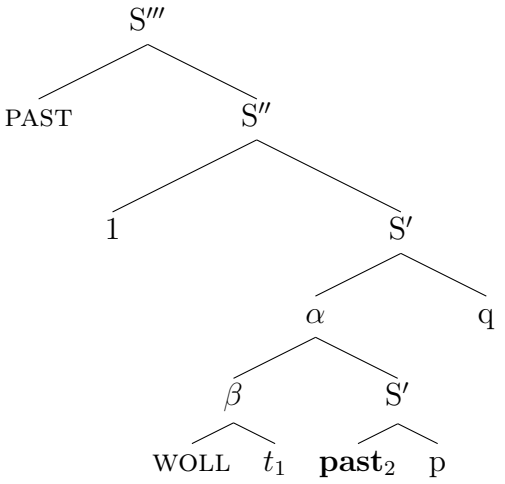
Now, in (38), for the meeting to happen tomorrow, it must not already have happened in the past. In other words, (38a) and (38b) presuppose that the meeting has not yet happened at the reference time. Ippolito calls this the *possibility presupposition* (Ippolito, 2013, 75). Because (38a) is a simple past remote conditional, the reference time is the present. In the scenario in (38), the meeting has already happened by the present time, so the possibility presupposition is not satisfied. Consequently, (38a) is infelicitous. In contrast, the reference time of (38b) is some contextually salient past time – presumably, the time before the meeting happened. Because the meeting has not yet happened, the possibility presupposition is satisfied, and (38b) is felicitous.

This explains why SPRCs and PPRCs are felicitous in different contexts. However, I have not yet explained why past perfect remote conditionals can describe nonpast eventualities and why they are strongly counterfactual. Regarding the former observation, note that in (42), **past** scopes over the rest of the conditional. It is not interpreted within the antecedent or consequent clauses, and so does not affect the temporal reference of those clauses. Thus, the antecedent and consequent are themselves nonpast propositions.

That said, when describing past eventualities, past perfect remote conditionals are actually ambiguous between two possible readings. There is one reading, given in (45), in which $[\![\textsc{past}]\!]$ scopes over the bare conditional and $[\![\textbf{past}]\!]$ is interpreted

within the antecedent. Because there is only one layer of past scoping over the bare conditional, this is semantically a simple past remote conditional about the past. Because this reading is available, the counterfactual inference of a PPRC describing a past eventuality is cancellable, as in (46), repeated from (20) above.

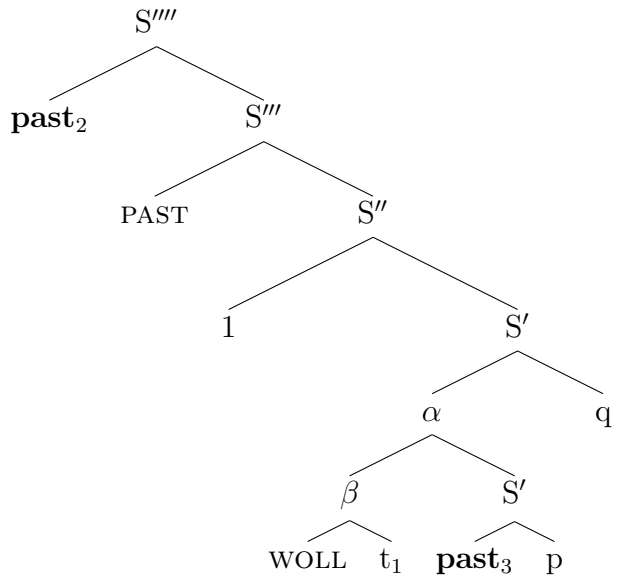
(45)



(46) If John had got his car fixed, he would have driven himself to the party. Since he came by himself, his car must be working again.

In the other possible reading of a past perfect remote conditional about the past, there is one deictic **past** scoping over the bare conditional, and another deictic **past** interpreted within the antecedent. A conditional with this structure would be strongly counterfactual, in direct contrast with the structure in (45). Although in mainstream English, these two structures can only be disambiguated by context, we will see in chapter 2 that modal antecedent conditionals may provide an unambiguous way of expressing the structure in (47). Because of this ambiguity, I will hereon use the term **past perfect remote conditional** to refer to structures with two layers of past scoping over the bare conditional, and **simple past remote conditional** to refer to structures with only one layer of past scoping over the bare conditional. In other words, I use the terms to describe a conditional's semantic structure rather than its overt morphosyntax. Consequently, (45) represents a simple past remote conditional, and (47) represents a past perfect remote conditional, even though these two structures could correspond to the same pronounced form.

(47)



Having explained why PPRCs and SPRCs are felicitous in different contexts, and why PPRCs can describe nonpast future eventualities, we are left with the puzzle of why past perfect remote conditionals are strongly counterfactual, *i.e.*, why they do not allow the counterfactual implicature to be cancelled. Ippolito proposes that this is because the counterfactual implicature in a PPRC is an **antipresupposition** rather than a conversational implicature. An antipresupposition arises as follows. Suppose that we have two expressions A and B, that A and B are truth-conditionally equivalent, and that A has presuppositions but B does not. Heim (1991) proposed a conversational principle **Maximize Presupposition** which requires that given two truth-conditionally equivalent options, a speaker choose the option with stronger presuppositions – in this case, A. If the speaker were to choose B over A, because of the Cooperative Principle (Grice, 1975), this would lead to the implicature that either the speaker does not believe the presuppositions $ps(A)$ of A to be true or that the speaker does not believe themself to be an authority on the truth of $ps(A)$. This implicature is an antipresupposition. To derive the conclusion that the speaker believes $ps(A)$ to be false, Ippolito assumes that the addressee makes two assumptions: that the speaker believes $ps(A)$ to be either true or false (the **competence assumption**), and that the speaker believes themself to be an authority on $ps(A)$ (the **authority assumption**) (Ippolito, 2013, 91). The authority assumption rules out the implicature that the speaker does not believe themself to be an authority on the truth of $ps(A)$. Then, given the competence assumption, it follows from the fact that the speaker does not believe $ps(A)$ to be true that the speaker believes $ps(A)$ not to be true. Importantly, antipresuppositions are not cancellable.

According to Ippolito, SPRCs and PPRCs are truth-conditionally equivalent. Additionally, recall that both types of conditional presuppose that there exists a possible world [i] historically accessible from the actual world at the reference time such that [ii] the presuppositions of the bare conditional are true in that world. In less technical terms, remote conditionals presuppose that the presuppositions of the bare conditional are compatible with the actual world at the reference time. In a SPRC,

the reference time is the present, whereas in a PPRC, the reference time is the past. Because in a branching futures semantics, the set of accessible possible worlds shrinks over time, the requirement that the presuppositions of the bare conditional be compatible with the actual world at the present is stricter than the requirement that the presuppositions of the bare conditional be compatible with the actual world at some past time. This is because at a later time, there are fewer accessible possible worlds overall, and thus fewer accessible possible worlds where the presuppositions of the bare conditional are true. Thus, SPRCs have stronger presuppositions than PPRCs, so the choice of a PPRC combined with the authority and competence assumptions generates the antipresupposition that the presuppositions of the bare conditional are not satisfied at the present. Recall that the presuppositions of the bare conditional include the presuppositions $ps(p)$ of the antecedent p. If $ps(p)$ is false, then p is false. Moreover, because this is an antipresupposition, the implicature that $ps(p)$ and thus p are false is not cancellable, giving us the strong counterfactuality of past perfect subjunctive conditionals.

To take a concrete example, let us return to the conditionals from (38), repeated below in modified form. (48a) presupposes that at the present time t, the meeting has not yet happened. On the other hand, (48b) presupposes that the meeting has not yet happened at some past time t'. If the meeting has not yet happened at t, then it follows that the meeting has not yet happened at t'. Consequently, the presupposition that the meeting has not yet happened at t is stronger than the presupposition that the meeting has not yet happened at t'. The choice of a PPRC over a SPRC hence generates the antipresupposition that the meeting has already happened at t ($\neg ps(p)$), and thus that it will not take place tomorrow ($\neg p$).

(48) a. If the meeting were tomorrow, Simone would make it.

b. If the meeting had been tomorrow, Simone would have made it.

To summarize, past perfect remote conditionals have two layers of past scoping over the bare conditional. One of the layers of past shifts to the past the accessibility time, when the set of historically accessible worlds is determined. The other layer shifts the reference time, the time at which the presuppositions of the bare conditional must be satisfied. Because the reference time is in the past, PPRCs can be used to describe eventualities that have been precluded in the actual world by the present time. Additionally, since the two layers of past scope over the bare conditional, and neither is (necessarily) interpreted in the antecedent or consequent, PPRCs can describe nonpast eventualities as well as past ones. Finally, because PPRCs are truth-conditionally equivalent to simple past remote conditionals but have weaker presuppositions, the choice of a PPRC over a SPRC generates the antipresupposition that the presuppositions of the bare conditional are false. Because antipresuppositions are not cancellable, this causes PPRCs to have an uncancellable counterfactual implicature.

1.4 Conclusion

This chapter provided an overview of the syntactic and semantic frameworks used in this thesis. Section 1.1 laid out the assumptions made about the syntax. Section 1.2 presented different approaches to the semantics of conditionals, culminating in a Stalnaker/Lewis semantics, in which a conditional *if p, then q* is true just in case q is true in all the p-worlds that differ minimally from the actual world. Also discussed in this section was the difference between open and remote conditionals. Finally, section 1.3 gave a compositional implementation of a Stalnaker/Lewis semantics. Two types of remote conditionals were distinguished: simple past remote conditionals (SPRCs) and past perfect remote conditionals (PPRCs). In SPRCs, the counterfactual inference is a conversational implicature and thus cancellable, whereas in PPRCs, the counterfactual inference is an antipresupposition and thus not cancellable.

Chapter 2

Modal antecedent conditionals

This chapter concerns remote conditionals with a modal and a perfect auxiliary in the antecedent. I have termed such conditionals **modal antecedent conditionals (MACs)**. As mentioned in chapter 1, these conditionals differ from conditionals with *had* alone in that they are **strongly counterfactual**: they generate a counterfactual implicature that cannot be cancelled. This is demonstrated in (49) below, repeated from example (3) in chapter 1.

(49) My professor thinks my friend Harry didn't do any of the readings she assigned him last semester. I don't know if Harry did or didn't, but I reason:

a. If he'd done the readings, he wouldn't have failed the class. Since he didn't fail the class, he must have done the readings.

b. If he'd've done the readings, he wouldn't have failed the class. #Since he didn't fail the class, he must have done the readings.

I argue that in modal antecedent conditionals, there are two layers of past scoping over the bare conditional, just as in a past perfect remote conditional. As explained in section 1.3.2, such a structure renders these conditionals strongly counterfactual. Furthermore, I argue that modal antecedent conditionals with *would* embed a modal WOLL in the antecedent clause that some speakers interpret volitionally. On the other hand, modal antecedent conditionals with *had* may embed a past tense operator in the antecedent, with the consequence that they can only describe past eventualities. Finally, modal antecedent conditionals with *had* face a syntactic OCP restriction that prevents the projection headed by perfect *have* from immediately following a projection headed by modal *had*.

The chapter begins with an overview of the data collection process. Section 2.2 concerns the historical development of the modal antecedent conditional construction, while section 2.3 concerns the syntax. Finally, section 2.4 discusses the semantic characteristics of MACs.

2.1 Methodology

To investigate these constructions, I collected acceptability judgments from eight college-age native speakers of English. One of these speakers was from England; the rest were American. The sentence set evolved over the course of the investigation, with the most substantive revision coming between the first and second versions. This was because several sentences in the first version that were intended to collect information on the cancellation of the counterfactual inference were incorrectly worded, with the consequence that they did not actually test for the cancellation of that inference. Additionally, there was a widening of the empirical scope between the first and second versions. All modifications after the second version were stylistic. Appendix A contains the first version of the questionnaire; appendix B gives the final version. Orthographic choices (*e.g.*, *would've* rather than *would have*) reflect my intended pronunciation in the elicitation.

All elicitations were presented orally. Certain elicitations were accompanied by a background scenario, shown in italics in the appendices. In such cases, I told participants that there was set-up, read the set-up, indicated that the set-up was over, and then read the sentence or pair of sentences on which I would collect judgments. I coded speakers' responses as acceptance, rejection, or intermediate judgment. In cases where speakers rejected a sentence or gave an intermediate judgment, I typically asked them to elaborate on why.

The questionnaire was divided into three parts, with the first part primarily consisting of structures I expected most speakers to accept, the third part consisting of structures I did not expect all speakers to accept, and the second part intermediate between these extremes. For the first three participants, the order of elicitation was kept constant. For the remaining participants, the order was scrambled within each part, although the three parts were still presented sequentially. Additionally, the first two elicitations listed in Part 1 were always presented first.

In reporting judgments, I will give the number of participants who accepted a construction outright, the number of participants who rejected a construction, and the number of participants who gave intermediate judgments, coded as *y,* *n*, and *m*, respectively.

2.2 Historical development

Little work has investigated the origins of modal antecedent conditionals. Dancygier & Sweetser (2005) propose that the evolution of *would have* and *had have* in conditional antecedents followed the trajectory in (50) (I use the pronoun *I* for illustration). They conjecture that these forms arose by analogy with the auxiliaries in the consequent. These auxiliaries are often reduced, so that, for example, *would have* is pronounced as *woulda* [ˈwʊɾə]. According to Dancygier & Sweetser (2005), *I'd'a* developed to match the auxiliaries in the consequent, as shown in (51a). The [ə] in *I'd'a* was then expanded to [əv], again by analogy with the consequent (51b), and then the most advanced speakers expanded the contracted auxiliaries in the antecedent to

give *would've* or *had've* (51c).

(50) I'd [a͡ɪd] → I'd-a [ˈa͡ɪdə] → I'd've [ˈa͡ɪdəv] → I had've/would've

(51) a. If I'd'a left earlier, I'd'a made it in time.
b. If I'd've left earlier, I'd've made it in time.
c. If I would've/had've left earlier, I would've made it in time.

However, Molencki (2000), notes that *had have* has been occurring in conditional antecedents since at least the 15th century, as in (52), from William Caxton's 1481 translation of the *Roman de Renart* (Molencki, 2000, 325). I have modernized the orthography to improve the crystallinity of the example.

(52) Had Tybert the cat have been there, he should also somewhat have suffered.

Molencki notes that the earliest examples of this form are contemporaneous with the loss of a morphologically distinct past subjunctive in English, implying that this form was filling the gap left by the past subjunctive, and was not motivated solely by phonological parellelism between the antecedent and consequent. That said, it is possible that modern modal antecedent conditionals arose independently from these early examples, perhaps following the development proposed by Dancygier & Sweetser. While more work is needed to understand the history of this construction, such work is beyond the scope of the present study.

2.3 Syntax

Kayne (1997) proposes that the perfect auxiliary *have* found in MACs and other constructions is not in fact the auxiliary *have*, but rather a complementizer *of*. He argues this on the basis of similarities between this supposed *of* and infinitival *to*, which he also considers to be a complementizer. However, his arguments are compatible with an analysis under which *of* is a reduced form of the auxiliary *have*, and Kayne himself acknowledges that in some varieties of English, what he identifies as *of* may indeed be *have*. Rather than introduce a new complementizer that licenses a perfect participle, I will stick to the analysis that the *'ve* in modal antecedent conditionals and other conditionals is in fact the perfect auxiliary *have*.

Siddiqi & Carnie (2012) argue that *had* in conditionals with *had have* is not a perfect auxiliary, but rather a modal. For evidence, they cite [i] the complementary distribution of *had* in this construction with other modals (53); [ii] the variation between *would have* and *had have* in modal antecedent conditionals; [iii] the fact that *had* is morphologically a past tense form (a property it shares with *would, could, should* and *might*; see section 1.2.2); [iv] the fact that *had* licenses the infinitival form of *have* rather than the perfect participle *had* (54); and [v] the fact that *had* licenses the same reduced forms of perfect *have* that the other modals do. (56) gives the structure for the conditional adjunct of the modal antecedent conditional in (49b). The adjunct is repeated in (56).

(53) a. * If I would have have been there, I could have helped.
(Siddiqi & Carnie, 2012, 19)

b. * If I would had have been there, I could have helped.

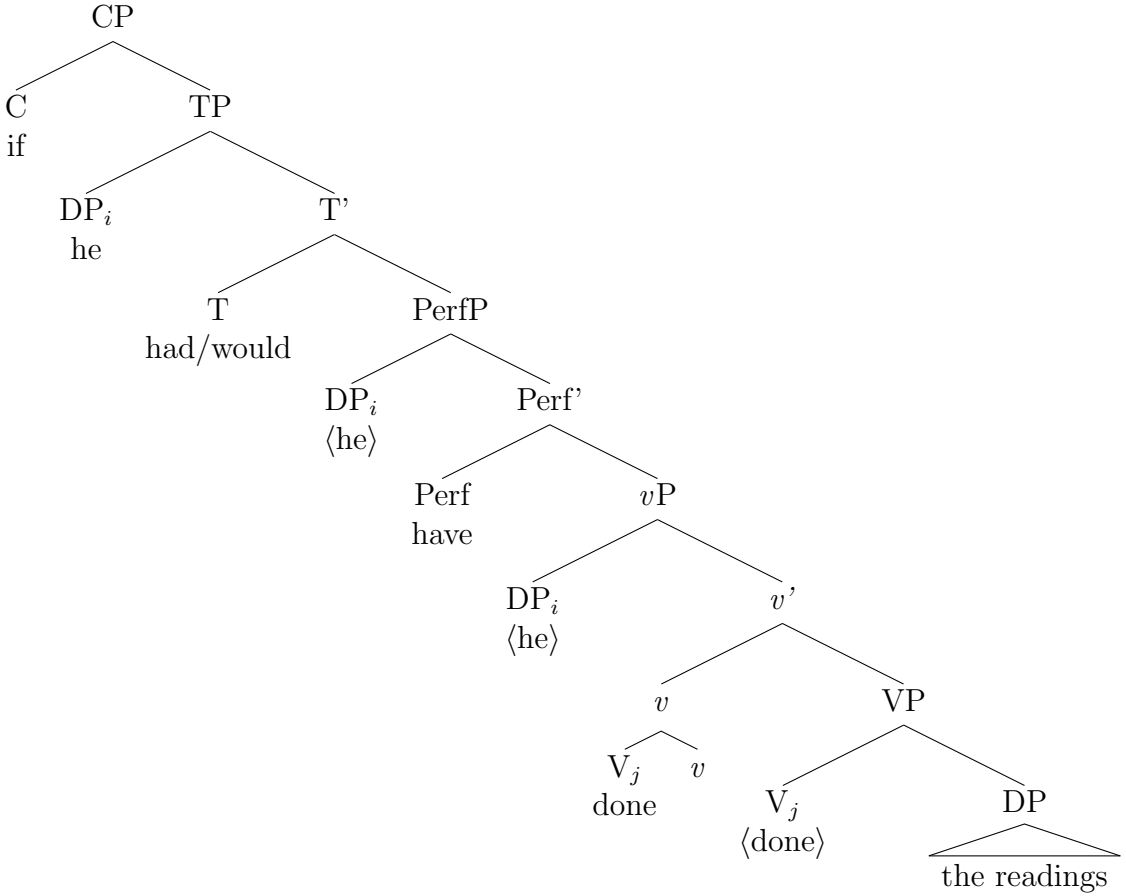
(54) * If I had had been there, I could have helped.

(55) a. If I had have been there, I could have helped.

b. If I had've been there, I could have helped.

c. If I had'a been there, I could have helped.

(56) If he'd've done the readings...



As seen above, the structure of the antecedent of a modal antecedent conditional is almost identical the structure of the corresponding conditional with past perfect morphology (*If he'd done the readings*; *cf.* (5a) in chapter 1). The key difference is that whereas in a conditional with past perfect morphology, there is a single auxiliary *had* which base-merges in Perf and moves to T, in a modal antecedent conditional, there are two auxiliaries. One of these auxiliaries, *have*, base-merges in Perf and does not move. The other auxiliary, *had* or *would*, base-merges in T.[1]

2.3.1 OCP effects in modal antecedent conditionals

There is an interesting asymmetry in acceptability between modal antecedent conditionals with *had* and those with *would*. For example, all but one participant accepted the MAC with *would* in (57). In contrast, a minority of participants accepted the MAC with *had* in (58). However, acceptability for MACs with *had* increased dramatically when *had* is negated, as in (59).

(57) If they would've called me, I would've come immediately. ($y = 7, m = 0, n = 1$)[2]

(58) If he had've knocked one more time, I would've answered the door. ($y = 3, m = 0, n = 5$)

(59) If they hadn't've pressed "reply all," they would've avoided a lot of embarrassment. ($y = 7, m = 1, n = 0$)

I propose that this restriction on the co-occurrence of *had* and *have* is an instance of the **obligatory contour principle (OCP)**. OCP was originally conceived as a phonological constraint restricting the co-occurrence of segments with shared features (Leben, 1973).[3] However, some authors have argued that one or more similar constraints operate in the syntax (*e.g.* Richards 2010; Hiraiwa 2014). Following such authors, I propose that most speakers reject (58) because there is a featural overlap between *had* and *have*, and this triggers an OCP effect when *had* and *have* head adjacent projections, with the definition of **adjacency** given in (60).[4] Under this definition, the TP headed by *had* is adjacent to the PerfP headed by *have* because *had* merges with PerfP. Likewise, PerfP is adjacent to TP because it merges with *had*, the head of TP. The adjacency relation is symmetric, so in general, if XP is adjacent to YP, YP is adjacent to XP.

[1] It is possible that *had* or *would* base-merge in a modal projection and then move to T. What is relevant for the present analysis is that these auxiliaries do not base-merge in Perf.

[2] As mentioned in section 2.1, y indicates the number of speakers who accepted the sentence, m the number of speakers who gave intermediate judgments, and n the number of speakers who rejected the sentence. Here, seven speakers accepted the sentence, and one speaker rejected it.

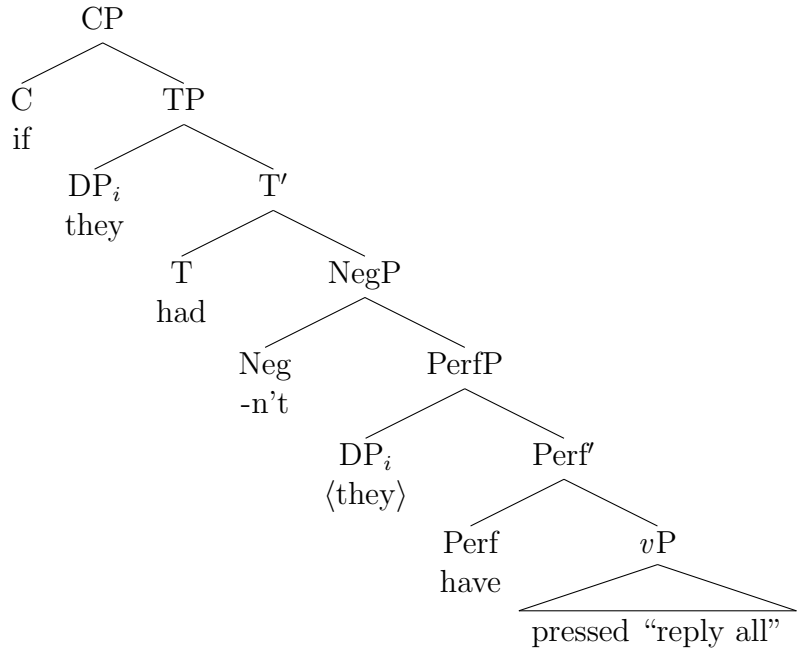
[3] In its original conception, it was a restriction on sequences of the same tone. It is now understood to apply to features other than tones.

[4] It is beyond the scope of this study to determine what the features of *had* and *have* are. However, because diachronically *had* developed from *have*, it seems reasonable to propose that synchronically there is featural overlap between the two auxiliaries.

(60) A projection XP with head X is **adjacent** to a projection YP with head Y if X merges with YP or Y merges with XP.

This explains why (59) is acceptable. There is presumably a NegP dominating the PerfP headed by *have*, so *had* takes NegP rather than PerfP as a complement. Consequently, TP is adjacent to NegP but not PerfP. By the symmetry of the adjacency relation, PerfP is not adjacent to TP either. Because TP and PerfP are no longer adjacent, OCP does not apply, and (59) is grammatical. The structure of (59) is given in (61).[5]

(61)



Now, it is in principle possible that the restriction on the co-occurrence of *had* and *have* is phonological rather than syntactic, *i.e.*, that *had* and *have* are phonetically too similar, and thus cannot be string-adjacent. Then (59) is acceptable because the negative affix *-n't* breaks the near-identity relationship between *had* and *have* (or alternatively, because it makes it so that *had* and *have* are no longer adjacent). However, there are two pieces of evidence suggesting that the restriction is syntactic, not phonological. First, *had* and *have* can co-occur in the opposite order, as shown in (62). From a phonological perspective, the acceptability of (62) is unexpected, since changing the order of *had* and *have* does not affect the near-identity relationship between the elements. Syntactically, however, *had* in (62) is a verb, rather than a modal auxiliary. As evidence that this is true, this instance of *had* cannot reduce like the auxiliary uses of *had* can. For example, *had* in (62) cannot be reduced to [əd]. Presumably, the featural content of verbal *have* differs significantly enough from that of the perfect auxiliary *have* that syntactic OCP is not triggered.

(62) John and Mary have had their hair cut three times this year.

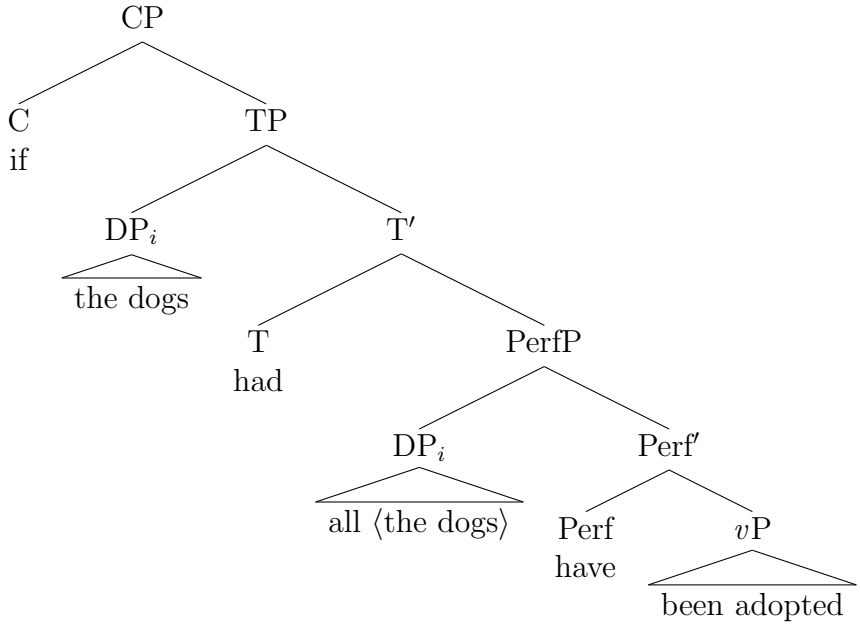
The second piece of evidence against a phonological analysis of the restriction on modal antecedent conditionals with *had* is exemplified in (63), where a stranded

[5] I leave open the question of why and how the negative suffix *-n't* is realized on *had*.

quantifier appears between *had* and *have*. Speakers largely rejected (63). If the relevant form of OCP were phonological, we would expect speakers to accept (63), because *had* and *have* are no longer string-adjacent. Syntactically, however, the rejection of (63) is expected. By assumption, the stranded quantifier *all* occupies Spec,PerfP. Thus, TP and PerfP are still syntactically adjacent, triggering OCP (64).

(63) If the dogs had all have been adopted, the SPCA would be closed. ($y = 1, m = 3, n = 4$).

(64)



To review, I propose that the relative unacceptability of modal antecedent conditionals with *had* is due to an OCP effect preventing a TP headed by modal *had* from being adjacent to a PerfP headed by perfect *have*. I argue that the relevant constraint is syntactic rather than phonological because the perfect auxiliary *have* can appear adjacent to the verb *had*, and because pronouncing a stranded quantifier between *had* and *have* does not make an MAC with *had* acceptable.

2.4 Semantics

This section discusses the semantics of modal antecedent conditionals. I first discuss MACs with *had*, then MACs with *would*.

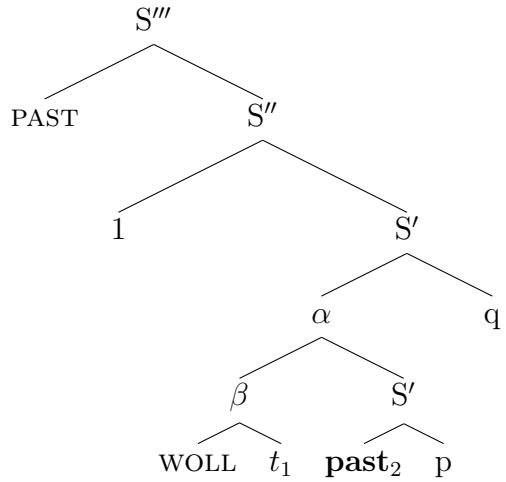
2.4.1 Modal antecedent conditionals with *had*

Recall from section 1.3.2 that when describing past eventualities, remote conditionals with past perfect morphology are semantically ambiguous. On the one hand, such a conditional could be a simple past remote conditional, with one past-tense operator scoping over the bare conditional.[6] This structure is shown in (65) (repeated from

[6]Recall that the terms **simple past remote conditional (SPRC)** and **past perfect remote conditional (PPRC)** refer to semantic rather than morphosyntactic structure. A SPRC has one

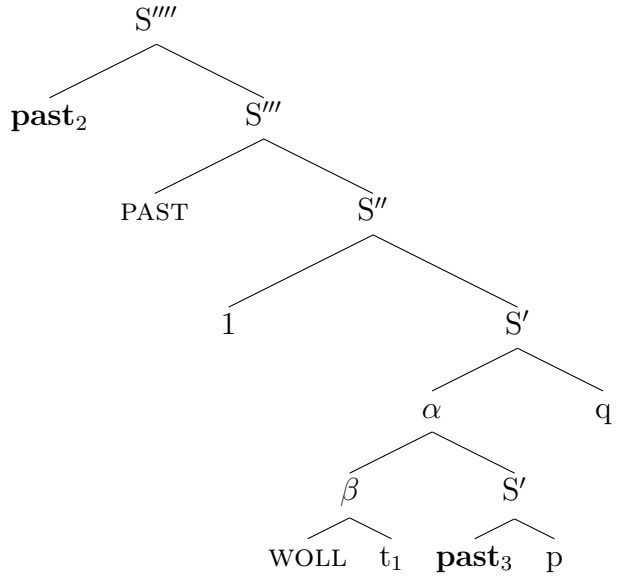
(45) in chapter 1).

(65)



The other option for a remote conditional with past perfect morphology describing a past eventuality is that it is a past perfect remote conditional, giving it the structure in (66) (repeated from (47)).

(66)



Ippolito (2013) suggests that modal antecedent conditionals with *had* are an unambiguous expression of the structure in (66). In other words, modal antecedent conditionals with *had* are past perfect remote conditionals about past eventualities. This hypothesis makes two empirical predictions. First, for the reasons set out in section 1.3.2, a modal antecedent conditional with *had* should be strongly counterfactual. Thus, we expect more speakers to reject (68), where the counterfactual inference is cancelled, than (67) (repeated from (58) above), where it is not. This is indeed the case: one fewer speaker accepted (68) than (67). Of course, the effect is very slight, so I would not conclude on the basis of this data alone that MACs are strongly counterfactual. However, because prior research (Dancygier & Sweetser, 2005; Biezma et al.,

layer of past scoping over the bare conditional, while a PPRC has two layers of past scoping over the bare conditional.

2014) has found MACs with *had* to be strongly counterfactual, and because my data do not contradict that finding, I feel confident in concluding that modal antecedent conditionals with *had* are indeed strongly counterfactual.

(67) If he had've knocked one more time, I would've answered the door. ($y = 3, m = 0, n = 5$)

(68) If she had've been there, Rob would have been really peeved. Since Rob is looking grumpy, Dani must have been there. ($y = 2, m = 0, n = 6$).

The second prediction made by Ippolito's hypothesis is that MACs with *had* can only be used to describe past eventualities. Consequently, we expect more speakers to reject (69) than (67), since (69) describes a future situation and should hence be pragmatically anomalous. My data do not strongly support this prediction. While fewer speakers accepted (69) than (67), fewer speakers rejected it either. Instead, there were more intermediate judgments. Based on this evidence, I do not think it reasonable to conclude that (69) is any less acceptable than (67). Thus, we cannot conclude that MACs with *had* necessarily describe past eventualities.

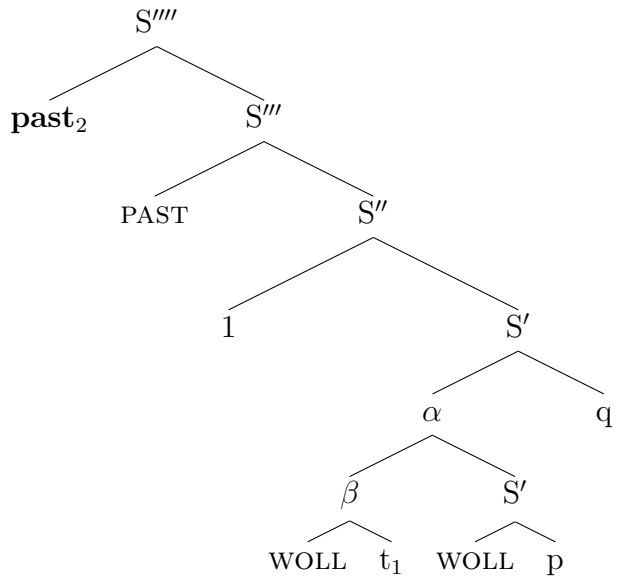
(69) If he had've come tomorrow instead, I would've been there to greet him. ($y = 1, m = 3, n = 4$)

In summary, we predicted that modal antecedent conditionals with *had* are PPRCs about past eventualities. Consequently, we predict that they are strongly counterfactual, and that they are infelicitous when describing nonpast eventualities. While my data and other research are in line with the former prediction, I did not find strong support for the second. Thus, we can conclude that MACs with *had* are PPRCs, but not that they only describe past eventualities.

2.4.2 Modal antecedent conditionals with *would*

Ippolito (2013) does not comment on modal antecedent conditionals with *would* as the modal. These conditionals differ from past perfect remote conditionals in that there is a modal *would* and a perfect auxiliary *have* in the antecedent, rather than just the inflected perfect auxiliary *had*. Given their surface structure, it seems logical to propose that the *would* that appears in the antecedent is interpreted in the antecedent clause, so that an MAC with *would* might have the structure in (70).

(70)



According to this structure, MACs with *would* are a type of PPRC, just as MACs with *had* are. Consequently, we predict that they are strongly counterfactual. Prior research (Biezma et al., 2014) has confirmed this, as has my own. To illustrate, compare (71) (repeated from (57) above) with (72). The latter received far more intermediate judgments than did (71), but did not receive any more rejections. This pattern is explicable if (72) is syntactically well-formed but pragmatically marked by the cancellation of the counterfactual inference.

(71) If they would've called me, I would've come immediately. ($y = 7, m = 0, n = 1$)

(72) If Jones would've taken arsenic, he would've shown exactly those symptoms he did show. So, he must have taken arsenic. ($y = 4, m = 3, n = 1$)

So, just like MACs with *had*, MACs with *would* do not allow the cancellation of the counterfactual inference. However, it is not yet clear what the interpretive role of the second WOLL is.

Now, WOLL has a number of readings. First, it has an epistemic reading, where it conveys predictability (73) (Collins, 2007). Indeed, the WOLL found in conditionals is epistemic, as it indicates that the consequent q is predictable from the fact that p is true. The modal WOLL also has a temporal reading, where it refers to future temporalities (74). Collins (2007) considers future WOLL a special case of epistemic WOLL, since it involves predictions about the future, and thus falls under the epistemic umbrella of predictability.[7]

(73) Mrs. Lee will be in her office right now.

(74) Roark will leave tomorrow.

WOLL also has a "volitional" reading, in which it denotes intention or willingness (75) (Collins, 2007). Finally, WOLL has an aspectual reading, where it denotes

[7]Collins discusses *will* only. However, I believe the points he makes about *will* extend to *would* as well.

"characteristic or habitual behavior" (76) (Huddleston & Pullum, 2002, 194). Collins (2007) considers this to be an instance of volitional WOLL when the behavior is intentional as in (76a), and an instance of epistemic WOLL when it is not (76b), since a "characteristic" behavior is predictable.

(75) a. I won't let Bert in no matter how much he knocks.
b. Will you take me to soccer practice?
c. Michelle wouldn't wash the dishes (even if I asked).

(76) a. When he was in high school, Dylan would take three naps a day.
b. Ice will melt at room temperature.

Incidentally, WOLL is often ambiguous between different readings. For example, it can be difficult to distinguish its intentional use from its futurate use, since if someone intends to do something, then we predict that that someone will do that thing. Thus, (75a) can be read as an assertion of my intention not to let Bert in, but it can equally be read as a statement of the fact that at no future time am I going to let Bert in. Additionally, the volitional reading of WOLL tends to be most available in questions and when negated (Axel-Tober & Gergel, 2015).

Let us consider which of these readings of WOLL we might find in the antecedent of a modal antecedent conditional. If WOLL is interpreted epistemically, then we expect that the truth of the antecedent p is predictable. In (77), for example, we expect the "fact" that I would have left earlier to be predictable. However, uttering a remote conditional indicates that p may or may not be true in the actual world, and implicates that it is not. Consequently, it is not predictable from what the speaker knows of the actual world that p is true. Thus, the conditions for uttering a remote conditional are not generally compatible with the conditions for epistemic WOLL, suggesting that the WOLL that occurs in the antecedent of an MAC is not epistemic.

(77) If I would've left earlier, I would've made it to work on time.

To make this concrete, consider (77). If the occurrence of *would* in the antecedent is interpreted epistemically, then it should be predictable from what I, the speaker, know of the actual world that I left earlier. However, in uttering (77), I imply that I did not leave earlier. Thus, it is not predictable that I left earlier, so the epistemic reading of *would* is not possible.

If futurate WOLL is indeed a subcase of epistemic WOLL as Collins (2007) claims, then it should be ruled out in the antecedent of an MAC. In any case, the futurate reading is not forced, or else (77) would be anomalous, due to the co-occurrence of futurate *would* with the past-oriented adverbial *earlier*. Likewise, the epistemic cases of "aspectual" WOLL should be ruled out as well.

The two remaining options for the interpretation of WOLL is that it receives a volitional interpretation, or that it is semantically vacuous. As it turns out, both options are instantiated in different speakers. To test whether WOLL receives a volitional interpretation, I investigated whether it was compatible with an expletive subject (78). Volitional WOLL should not be compatible with an expletive subject, because it expresses the subject's volition, and thus requires a subject capable of volition.

(78) If it wouldn't have snowed, she would be in Cancun by now. (y=5, m=1, n=2)

Two speakers rejected (78), and one speaker gave an intermediate judgment. This contrasts with (71), which seven speakers accepted and only one rejected. Moreover, the rejection of (78) does not reduce to a rejection of MACs with *would* in general, since all three of the speakers who did not accept (78) did accept MACs with *would* when they had animate subjects. Furthermore, one speaker who rejected the sentence suggested the antecedent be changed to *If it hadn't snowed*, which crucially does not allow a volitional reading. Even more telling is the speaker who commented that they "feel like *would've* is for people." It seems that for these speakers, WOLL requires an animate (or perhaps human) subject in this construction, suggesting that the modal retains its volitional reading.

That said, these data do not necessitate that the speakers interpret *would* volitionally here – it could be, for example, that their grammars require that *would* take an animate subject irrespective of its semantic interpretation. It is possible that *would* in MACs is in the process of losing its volitional meaning. Indeed, the volitional reading of WOLL is only dominant in questions and when negated (Axel-Tober & Gergel, 2015), suggesting that volitional WOLL is a negative polarity item. Certain negative polarity items are acceptable in questions and negated sentences, but are degraded in conditionals. For me, *either* and *anymore* are two such examples, as demonstrated in (79) and (80) respectively.[8] It is possible that for some speakers, volitional WOLL is among the class of negative polarity items that can appear in conditionals, and for other speakers, it is in the class that cannot.

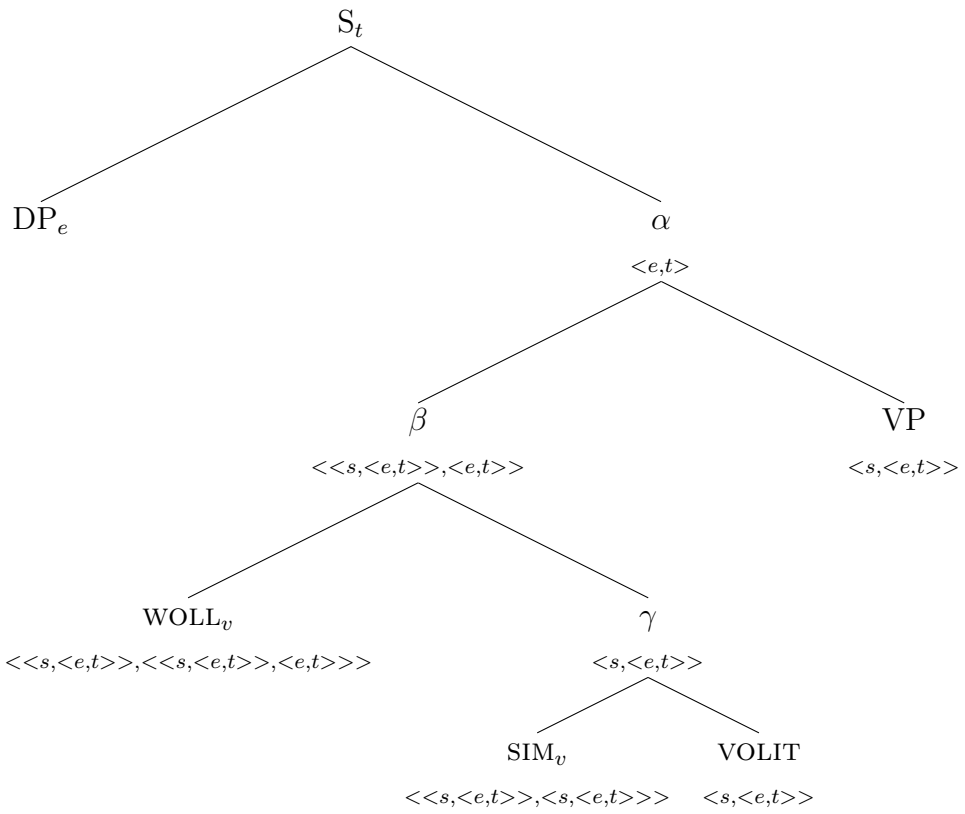
(79) a. I won't go to the party. John won't go either.
b. I don't know if I will go to the party. I wonder if John will go either.
c. I am probably going to the party. If John is going too/*either, then I'm definitely going.

(80) a. She isn't going to school anymore.
b. Is she going to school anymore?
c. ? If she was going to school anymore, I would know about it.

We have seen that for at least some speakers, the WOLL interpreted in the antecedent of a modal antecedent conditional seems to necessarily receive a volitional interpretation. The question remains of how to model the semantics of volitional WOLL. Unlike the WOLL found in conditionals, volitional WOLL (hereon WOLL_v) must take an individual (of type e) as an argument, because it describes the volition of the subject. It follows that it must be interpreted in place, between the subject and the predicate, because if it composed with the entire antecedent, then it would not receive an argument of type e, but rather one of type t or $<s,t>$. This is demonstrated in (81) below.

[8] For me, *either* is more acceptable in embedded questions than matrix questions, which is why I demonstrate its use in an embedded question in (79b), rather than a matrix question.

(81)



Like conditional ⟦WOLL⟧, volitional ⟦WOLL_v⟧ also needs to compose with an accessibility function. In this case the necessary form of accessibility is not historical but rather volitional accessibility, identifying the worlds that are compatible with what the subject is willing to do in the actual world. The denotation for such a volitional accessibility function VOLIT is given in (82).

(82) $[\![\text{VOLIT}]\!]^{c,g,t,w} = \lambda w'_s.\ \lambda x_e : x$ is capable of volition. w' is compatible with what x is willing to do in w.

According to (82), ⟦VOLIT⟧ takes as input a world w' and an individual x that is capable of volition, and returns true if w' is compatible with what the individual x is willing to do in the evaluation world w.

Additionally, in evaluating WOLL_v, we presumably quantify over worlds that are maximally similar to the actual world, so we will still need the similarity function ⟦SIM⟧. However, here ⟦SIM⟧ will need to take an argument of type <s,<e,t>> rather than <s,t> to be able to compose with ⟦VOLIT⟧. Likewise, it will need to return a function of that same type to be able to compose with WOLL_v, so ⟦SIM⟧ will be of type <<s,<e,t>>,<s,<e,t>>> and will have the denotation in (83).[9] I will hereon refer to this instance of SIM as SIM_v. With these ingredients in place, WOLL_v should have the denotation in (84).

(83) $[\![\text{SIM}_v]\!]^{c,g,t,w} = \lambda P_{<s,<e,t>>}.\ \lambda w'_s.\ \lambda x_e.\ P(w')(x) = 1 \wedge \neg\exists w''[P(w'')(x) = 1 \wedge w'' <_w w']$

[9] This suggests that SIM may be an abstract function that takes an intensional function as its input and returns an intensional function of the same type as its output.

(84) $[\![\text{WOLL}_v]\!]^{c,g,t,w} = \lambda P_{<s,<e,t>>}.\ \lambda Q_{<s,<e,t>>}.\ \lambda x_e.\ \forall w'[P(w')(x) \rightarrow Q(w')(x)]$

According to (83), $[\![\text{SIM}_v]\!]$ takes as its input (the intension of) a property P, a world w', and an individual x, and returns true if P is true of x in w', and there is no other world w'' where P is true of x and that is more similar to the evaluation world w than w' is. $[\![\text{SIM}_v]\!]$ takes as input two properties P and Q and an individual x, and returns true if in all worlds w' where P is true of x, Q is also true of x.

To make this clearer, let us consider again the sentence in (85), modified to have a name a name instead of a pronoun. The antecedent of that sentence, shown in (86), has the truth conditions in (87).

(85) If John would've called me, I would've come immediately.

(86) John WOLL_v call me

(87) $[\![\text{WOLL}_v(\text{SIM}_v(\text{VOLIT}))(\text{call me})(John)]\!]^{c,g,t,w} = 1$ iff $\forall w'[\text{SIM}(\text{VOLIT})(w')(John) \rightarrow$ call-me$'(w')(John)]$, defined only if John is capable of volition

According to (87), (86) is true if and only if all the possible worlds compatible with what John is willing to do in the actual world are worlds where you call me. The sentence is only defined if John is capable of volition.

2.5 Conclusion

This chapter began with an overview of my data collection methodology, followed by a brief discussion of the historical development of modal antecedent conditionals. In section 2.3, I presented evidence that the *had* found in modal antecedent conditionals is a modal, rather than a perfect auxiliary, so that MACs have the structure [modal+*have*+perfect participle] in their antecedent. Furthermore, after observing that MACs with *had* are typically not accepted in the affirmative, but are accepted when negated, I proposed that a syntactic OCP effect prevents adjacency between the modal *had* and perfect *have* due to featural overlap between the two auxiliaries.

Section 2.4 dealt with the semantic characteristics of modal antecedent conditionals. I proposed that MACs are a type of past perfect remote conditional, with two layers of past tense scoping over the bare conditional. This analysis explains why the counterfactual inference is not cancellable in MACs. I also hypothesized that modal antecedent conditionals with *had* as the modal embed a past tense operator in the antecedent, so that they necessarily describe past eventualities. However, the data were equivocal on this point. As for MACs with *would* as the modal, I proposed that *would* is either interpreted volitionally in the antecedent of the conditional, or it is semantically vacuous. The data supports both hypotheses, as there were speakers who followed both patterns. I ended the chapter with a model of the semantics of volitional WOLL.

Chapter 3

Copy conditionals

This chapter concerns **copy conditionals**: inverted conditionals with two copies of the auxiliary *had*, one of which surfaces as non-syllabic *'d* [d].[1] This phenomenon is demonstrated in (88), repeated from the first chapter.

(88) Had he'd done the readings, he wouldn't have failed the class.

Unlike in modal antecedent conditionals, I have found that the addition of the second auxiliary does not affect the cancellability of the counterfactual implicature. Nor does the addition of a second *had* affect the meaning in any other clear way. Additionally, for almost all speakers, copy conditionals are only acceptable if the lower auxiliary is in its non-syllabic form. In consideration of these facts, I propose that the lower auxiliary is a copy of the auxiliary *had* that has undergone T-to-C movement, and that pronunciation of multiple copies is allowed due to morphological Fusion between a subject pronoun and the copy of the auxiliary in T.

This chapter proceeds as follows. First, I give an overview of the Copy Theory of movement. I then give an overview of Distributed Morphology, followed by an analysis of auxiliary contraction. Finally, I consider the consequences of the proposal that copy conditionals arise as a result of morphological Fusion.

3.1 The Copy Theory of movement

The **Copy Theory of movement** came to prominence in the early 1990s as an alternative to the **Trace Theory of movement**, then the dominant approach to movement in generative syntax. Under the Trace Theory, when an item moves, it leaves behind a silent category co-indexed with the moved element; this silent category is a **trace**. This contrasts with the Copy Theory, which proposes that when something moves, it leaves behind an identical copy indistinct from the moved item. As an example, the sentence in (89a) has the structure in (89b) under the Trace Theory,

[1] I did not explicitly test whether the contracted auxiliary in conditionals like (88) can be *would*. However, when conducting elicitations, many participants expanded sentences like (88) to *had he had*. They never expanded such sentences to *had he would*. Likewise, my intuition on the construction is that the contracted auxiliary is *had*, not *would*. This evidence suggests that the contracted auxiliary in a copy conditional can only be *had*.

while it has the structure in (89c) under the Copy Theory (where strikeout indicates unpronounced material).[2]

(89) a. What did you read?
b. What$_i$ did you read t$_i$?
c. What$_i$ did you read ~~what$_i$~~?

A question arises under the Copy Theory: if a moved element and its copy are identical, including in their phonological features, then why are not all copies pronounced? Nunes (2004, 2011) proposes that the pronunciation of multiple copies is generally ruled out because it generates contradictory linear orderings.

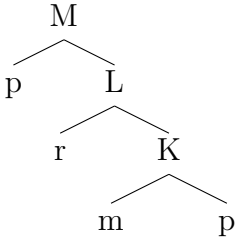
According to Kayne (1994), linear order is determined by asymmetric c-command, so that if a constituent A asymmetrically c-commands B, then A precedes B in the linear order of the sentence. According to Chomsky (1995), **linearization**, the process of imposing linear order, applies post-syntactically. To illustrate the problem this poses for the Copy Theory, consider (90).

(90) *What$_i$ did you read what$_i$?

The higher copy of *what* asymmetrically c-commands the lower copy, so the higher copy should precede the lower copy in the linear order of the sentence. However, the copies are non-distinct, so this amounts to requiring that *what* precede itself, which is not possible. Likewise, *what* should both precede and be preceded by *did*, *you*, and *read*, since these words are c-commanded by the higher copy of *what* and c-command the lower copy. Because of these ordering contradictions, (90) is ungrammatical.

Normally, the computational system avoids this problem by deleting all copies except one, as was shown in (89c) above (Nunes, 2004, 2011). However, there are other strategies for resolving the problem of linearizing identical copies, one of which is morphological Fusion. Suppose we have the structure in (91a), where p, r, and m are terminals. Note that there are two copies of p. As shown in (91b), the morphological component can fuse the two terminal nodes m and p into a single atomic terminal $\#m{+}p\#$ whose internal structure is invisible to the computational system (Chomsky, 1995; Nunes, 2004).[3] So, the lower copy of p is not itself visible to the linearization algorithm, avoiding the problems associated with linearizing identical copies.[4]

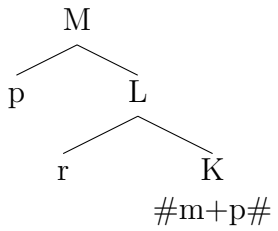
(91) a.



[2]For more information on why the Trace Theory was abandoned in favor of the Copy Theory, I refer the reader to Nunes (2011).

[3]Following Nunes (2004, 2011), I indicate Fusion by enclosing the fused heads in hashes. The choice to put a plus sign between the fused heads is my own.

[4]Chomsky (1995) only notes the possibility of Morphology converting two sister terminals into a "'phonological word' not subject internally to [linearization]". It is Nunes (2004) who proposes that Morphology accomplishes this through the process of Fusion, outlined in section 3.2.

b.


In principle, then, languages can allow identical copies to co-occur if one of those copies undergoes Fusion with another element. As one piece of evidence that languages make use of this option, Nunes (2004, 2011) notes that in some languages, when a wh-phrase from an embedded clause moves to Spec,CP of the matrix clause, it is also pronounced in Spec,CP of every embedded clause it passes through. In many such languages, only simplex wh-phrases, which are simultaneously minimal and maximal, can surface multiple times. For example, in varieties of German that allow this phenomenon, a monomorphemic wh-phrase like *wen* 'who' can be pronounced multiple times (92a), but not a complex wh-phrase like *wessen Buch* 'whose book' (92b). To account for this, Nunes (2004) proposes that intermediate copies of wh-phrases are able to be realized in these languages because they undergo Fusion with the C-head of the embedded clause they are realized in. This explains why only simplex wh-phrases are pronounced multiple times: Fusion applies to heads, and so can apply to a simplex wh-phrase (which is simultaneously a phrase and a head), but not a complex wh-phrase (which is not a head). As further evidence in support of this analysis, in Frisian, intermediate copies of wh-phrases are pronounced as a single phonological word with the complementizer, which is expected if they have undergone Fusion (93).

Non-standard German

(92) a. Wen glaubt Hans wen Jakob gesehen hat?
whom thinks Hans whom Jakob seen has
'Who does Hans think Jakob saw?' (Nunes 2004, 38, citing McDaniel 1986)

b. *[Wessen Buch] glaubst du [*wessen Buch] Hans liest?
whose book think you whose book Hans reads
'Whose book do you think Hans is reading?' (Nunes 2004, 38, citing McDaniel 1986)

Frisian

(93) Wêr tinke jo wêr't Jan wennet?
where think you where-that Jan lives
'Where do you think that Jan lives?' (Nunes 2004, 38, citing Hiemstra 1986)

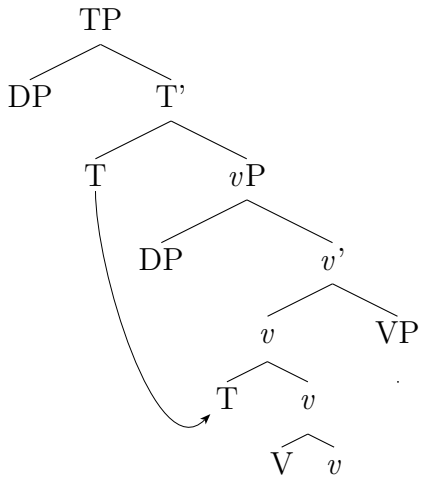
To summarize, having multiple copies of the same element in a phrase marker leads to ordering contradictions. While these ordering contradictions are typically blocked by deletion of all but one copy, they can also be resolved by morphological Fusion. When the latter occurs, multiple copies appear.

3.2 Distributed morphology

I have proposed that a morphological operation, Fusion, can allow multiple copies of the same constituent to appear in a sentence. However, I have not yet spelled out my assumptions about the morphology. I am assuming the framework of **Distributed Morphology (DM)** (Halle & Marantz, 1993). Under DM, syntax does not manipulate words, but rather bundles of features referred to as **heads** or **morphemes**. Syntax is responsible for organizing morphemes into hierarchical structures, and Morphology is responsible for **Spell-Out**, the process of mapping these structures onto their **Phonological Form (PF)**. One part of this process is inserting appropriate Vocabulary Items (phonological expressions) for each morpheme; this process is referred to as **Vocabulary Insertion**. Another part of Spell-Out is linearization, as described in the previous section.

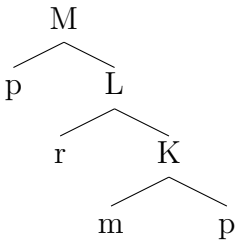
Morphology has a number of operations for manipulating structural relationships. One such operation is **Morphological Merger (M-merger)**, also known as **Lowering**. This operation occurs before Vocabulary Insertion and adjoins a head to the head of its sister (Halle & Marantz, 1993; Embick & Noyer, 2001). As an example, the realization of tense features on the verb in English involves Lowering of T to v, as shown in (94). According to Embick & Noyer (2001), M-merger does not leave behind a copy. To avoid confusion, I will mark Lowering with arrows, as shown in (94), while I will continue to mark raising (*i.e.*, syntactic movement) with indices.

(94)

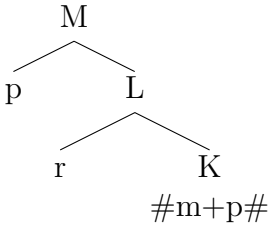


Another operation that occurs before Vocabulary Insertion is **Fusion**, illustrated in (95) (repeated from (91) above) (Halle & Marantz, 1993). In Fusion, two sister heads are combined into a single head whose internal structure then becomes invisible to Morphology (Halle & Marantz, 1993). Both head-raising and Lowering feed Fusion, since both cause two heads to form a single X° constituent.

(95) a. Before Fusion



b. After Fusion



3.3 Auxiliary contraction and copy conditionals

This section presents an analysis of auxiliary contraction, and discusses its relevance for copy conditionals. Discussions of the English auxiliary system generally distinguish between three possible realizations of an auxiliary: the **full form**, which has not undergone phonological reduction; the **weak form**, which has undergone phonological reduction; and the **clitic form**, which is non-syllabic and consists only of a consonant.[5]

At first glance, the weak form and clitic form both seem to be reduced forms of the full auxiliary. Indeed, they both may be represented in the orthography as contraction, as shown in (96). However, there are important differences between clitic and weak forms. First, the weak forms of the auxiliaries are syllabic, whereas the clitic forms are non-syllabic. Second, the weak forms can be derived from the full forms by phonological processes that apply elsewhere in the language, whereas the clitic forms cannot (Close, 2004). Third, the clitic forms can only occur when the subject is a simplex DP (*i.e.,* a pronoun).[6] Fourth, the clitic forms can trigger changes in the pronunciation of the subject pronoun not triggered by the weak form. For example, the subject pronoun *I* can be realized as [ɑ] instead of [a͡ɪ] when it occurs with the clitic form of *will*, as shown in (96a-i). Finally, only a proper subset of the auxiliaries that have distinct weak forms have clitic forms. For example, *could* has a weak form [kʰəd] but no clitic form. The table in (97) gives possible realizations of the strong, weak, and clitic forms of those auxiliaries that have clitic forms.

(96) a. I'll go.

[5]The term *clitic form*, which I take from Close (2004), refers only to the fact that this form of the auxiliary forms a phonological word with the subject. It does not necessarily refer to the syntactic or morphological behavior of the auxiliary.

[6]There is an exception to this generalization. The clitic forms of *is* and *has* can occur with non-pronominal subjects, perhaps due to analogy with the possessive and plural markers.

i. With clitic form: [a͡ɪl go͡ʊ], [ɑl go͡ʊ]

ii. With weak form: [a͡ɪ əl go͡ʊ]

b. Sky'll go.

i. With clitic form: *[ska͡ɪl go͡ʊ]

ii. With weak form: [ska͡ɪ əl go͡ʊ]

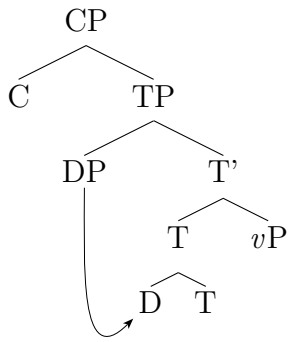
(97)

Auxiliary	Full form	Weak form	Clitic form
am	[æm]	[əm]	[m]
are	[ɑɹ]	[ə˞]	[ɹ]
is	[ɪz]	[ɪz]	[z]
have	[hæv]	[əv]	[v]
has	[hæz]	[əz]	[z]
had	[hæd]	[əd]	[d]
will	[wɪl]	[əl]	[l]
would	[wʊd]	[əd]	[d]

Because it is the clitic form rather than the weak form that I am interested in, I will use the term **auxiliary contraction** to refer only to the realization of the auxiliary in its non-syllabic form.

As mentioned above, the clitic forms of the auxiliaries require that the subject be a pronoun. To account for this fact, I propose that for auxiliary contraction to occur, there must be M-merger between the subject DP and the T-head, so that D and T form a single T°. This is schematized in (98).

(98)

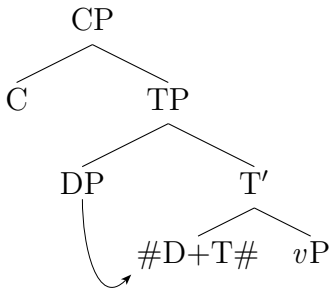


This analysis explains why auxiliary contraction can only occur with pronoun subjects. Recall that M-merger adjoins a head to the head of its sister. Pronouns are simplex DPs, simultaneously minimal (a head) and maximal (a phrase), so they can undergo M-merger and adjoin to T. On the other hand, a complex DP is not a head, and thus cannot undergo M-merger. If we assume that T can only be spelled out as a clitic form if this M-merger has occurred, then it follows that the clitic form is only licensed when the subject is a simplex DP, *i.e.*, a pronoun.

While this analysis explains why auxiliary contraction requires the subject to be a pronoun, it does not explain why copy conditionals are acceptable, since the internal structure of a complex head created under M-merger is still visible to morphophonological processes, including linearization. Thus, we still have the problem of linearizing two copies of the same auxiliary. To derive the correct results, M-merger must feed a Fusion operation between the pronoun and the auxiliary. This Fusion creates a head

#D+T# whose internal structure is invisible to linearization. Because its internal structure in invisible, an ordering statement in which T precedes #D+T# will not be a contradiction. (99) shows what (98) would look like after Fusion has taken place.

(99)

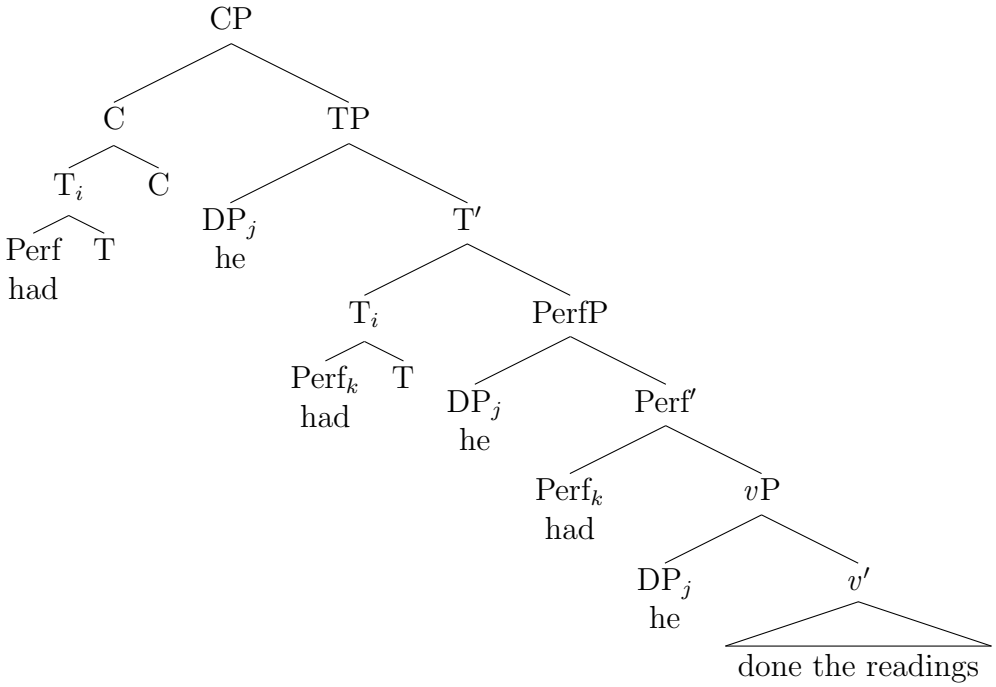


To make this more concrete, let us return to (88), repeated below in (100a). (101) gives the structure of the conditional adjunct before Morphology has applied.[7] If no operations applied in the Morphology before linearization, then linearization would produce an ordering paradox in which T precedes itself. To resolve this ordering paradox, the lower copy of T would not be pronounced, giving (100b).

(100) a. Had he'd done the readings, he wouldn't have failed the class.

b. Had he done the readings...

(101)

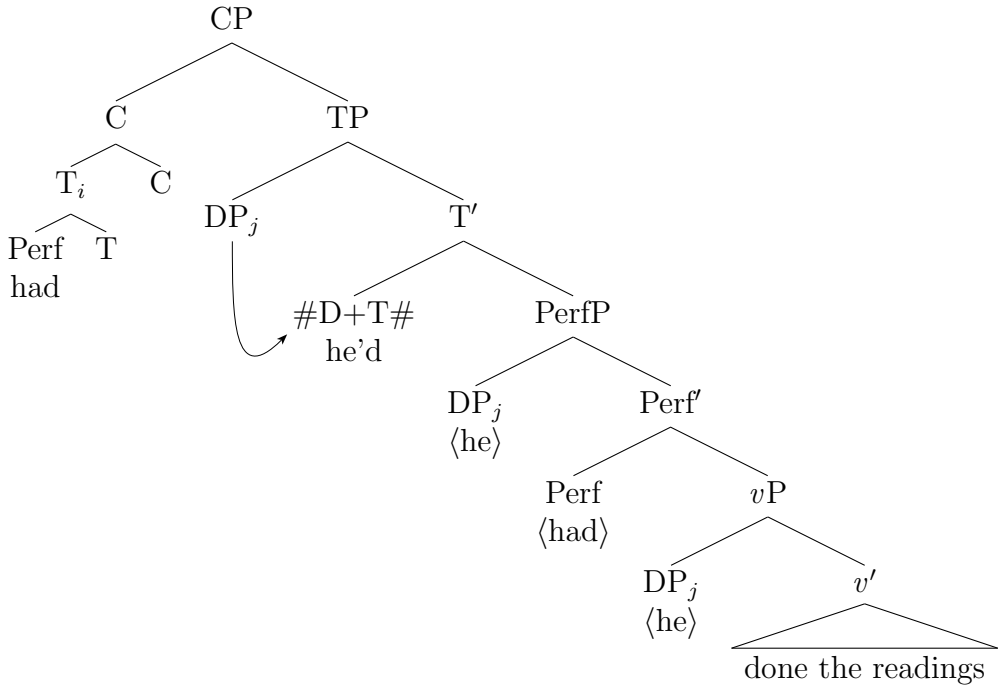


Of course, I have claimed that operations do apply in the Morphology before linearization. Specifically, M-merger occurs between the subject (which is a simplex DP) and T. This M-merger feeds Fusion between the subject and T, creating a head #D+T#. Linearization then produces an ordering in which T precedes #D+T#. However, because the internal structure of #D+T# is inaccessible, this does not

[7] I have included Vocabulary Items in the tree for expository purposes only; at this point in the derivation, the heads do not yet have any phonological content.

produce an ordering paradox. Consequently, both T and #D+T# will be spelled out, the former as *had*, and the latter as *he'd*.[8]

(102)



To summarize, I have proposed that auxiliary contraction requires M-merger between a pronominal subject and T. This M-merger feeds a Fusion operation, creating a head #D+T#. Because ordering T relative to #D+T# does not cause an ordering contradiction, the auxiliary is pronounced both in the copy of T that has moved to C, and as part of the subject/auxiliary complex inserted in #D+T#. The fact that #D+T# is spelled out as the combination of the pronoun and clitic form of the auxiliary (*e.g., he'd*) explains why the majority of speakers require that the lower auxiliary surface in its clitic form.

3.4 Questions and consequences

This section considers the consequences of the proposal that morphological Fusion allows multiple copies of the auxiliary to appear in copy conditionals, and the questions that this proposal raises.

3.4.1 Mechanisms of auxiliary contraction

I have proposed that in copy conditionals, the Fused head #D+T# is spelled out as the combination of the subject pronoun and the auxiliary. In (102) above, for example, it was spelled out as *he'd*. So, the clitic form of the auxiliary can appear as a result of Fusion. However, the clitic form can also appear in circumstances where

[8] A relevant question is why the subject pronoun is only pronounced as part of #D+T#, when this analysis predicts that it should be pronounced more than once. This will be addressed below.

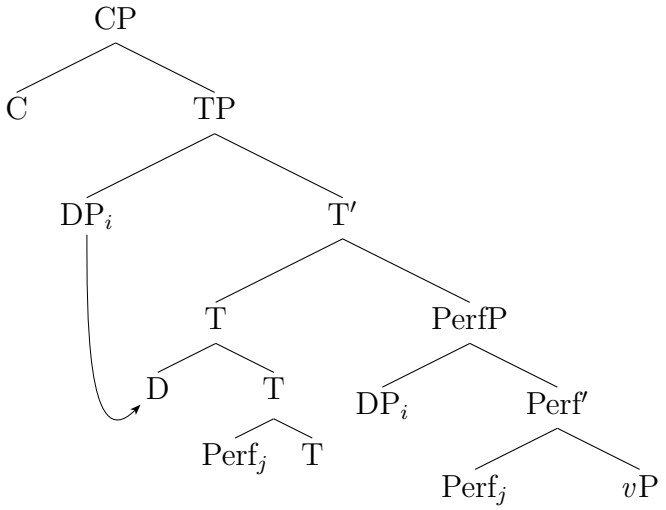
it is not obvious that Fusion has occurred. In (103a), for example, Fusion cannot have happened before linearization, or else Perf would have been spelled out, giving (103b).[9]

(103) a. He'd left already.
b. He'd had left already.

Embick & Noyer (2001) hypothesize that linearization happens at Vocabulary Insertion. If this is true, then the fact that Fusion has not occurred before linearization in (103a) entails that Fusion did not occur before Vocabulary Insertion, and thus that *he'd* in (103a) consists of two separate Vocabulary Items, *he* and *'d*, corresponding to D and T respectively. In other words, there are two distinct processes of auxiliary contraction: one in which the subject and auxiliary are inserted as a single Vocabulary Item, and one in which the subject and auxiliary are inserted as two separate Vocabulary Items. This conclusion is undesirable on Minimalist assumptions, since it introduces a redundancy into the linguistic system.

If linearization precedes Vocabulary Insertion, however, then we can provide a unified account of auxiliary contraction. In copy conditionals and sentences like (103b), where the auxiliary is pronounced multiple times, Fusion happens before linearization. In cases like (103a), on the other hand, Fusion happens after linearization. To see how this produces the desired results, consider again (103a). Its structure when linearization applies is given in (104).

(104)



When linearization applies, it will give an ordering in which the copy of Perf in T precedes the copy of Perf in its base-merge position. Because the copies are non-distinct, this creates an impossible ordering. To resolve this, Vocabulary Insertion is blocked for the lower copy of Perf. After linearization has happened, D and T can Fuse into #D+T#. Then when Vocabulary Insertion happens, *he'd* will be inserted at #D+T#. On the other hand, no Vocabulary Item will be inserted at Perf, because Vocabulary Insertion was already blocked for Perf at linearization. Thus, if linearization precedes Vocabulary Insertion, then we can provide a unified account for

[9]Interestingly, some speakers accept sentences like (103b), as discussed below.

auxiliary contraction, in which pronouns with contracted auxiliaries are single Vocabulary Items. Moreover, the proposal that auxiliary contraction involves insertion of a single Vocabulary Item could explain why the clitic forms can trigger stem changes in the pronoun: the "stem changes" actually reflect a difference between Vocabulary Items, rather than the outcome of a synchronic phonological process.[10] Finally, I believe that the assumption that linearization precedes Vocabulary Insertion is reasonable, since Vocabulary Insertion can be blocked for certain heads as a result of linearization.

3.4.2 Multiple copies in other environments

To reiterate an earlier point, if Fusion precedes linearization, then multiple copies of the auxiliary should surface: one as part of #D+T#, and one elsewhere. The conditions for multiple copies to occur reference neither conditionals nor inversion. Consequently, we predict that multiple copies of the auxiliary should be able to occur outside of conditionals and without inversion. My data support this prediction. Multiple copies of the auxiliary are as acceptable in inversion triggered by the fronting of *not only* (106) as they are in conditional inversion (105), and only slightly worse in yes-no questions (107) and in conditional adjuncts without inversion (108).

(105) Had I'd set aside more time for my assignment, I would've done better in the end. $(y = 6, m = 2, n = 0)$

(106) Not only had he'd met Kelly before, but they actually used to date. $(y = 6, m = 2, n = 0)$

(107) Had he'd been to Europe before? $(y = 4, m = 2, n = 2)$

(108) If I'd had done the readings, I would've passed the test. $(y = 5, m = 2, n = 1)$

It is not clear to me why (107) is less acceptable than other cases with multiple copies of *had*. One participant commented that it would be better if there were another clause, so perhaps the ungrammaticality is not because of the multiple copies of the auxiliary, but rather because not enough context was provided for the past perfect construction, which is generally infelicitous out of context.

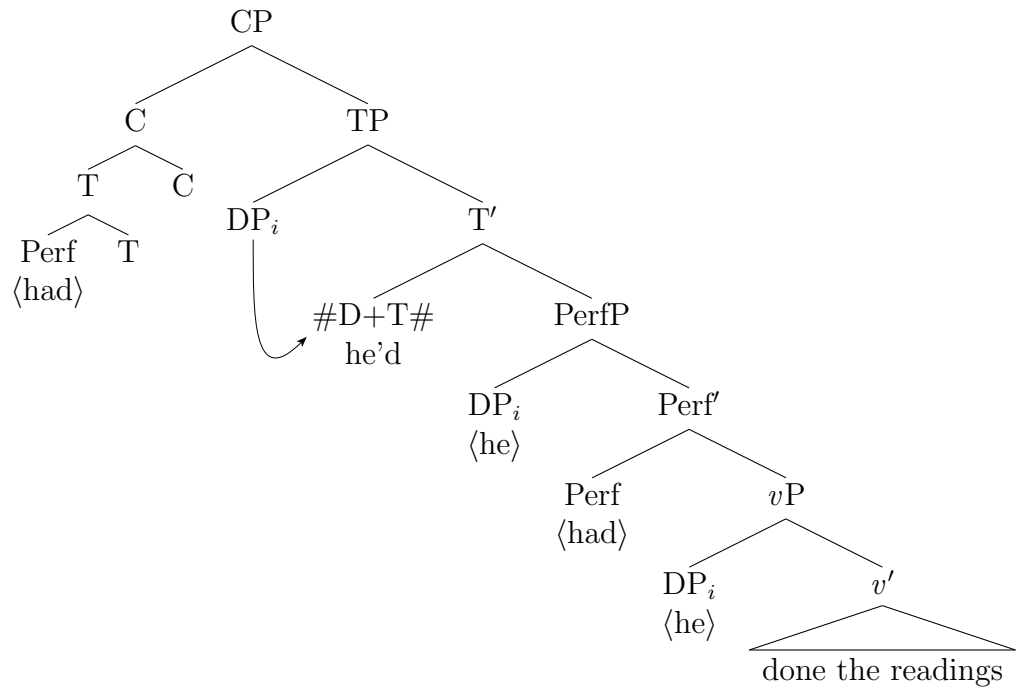
3.4.3 Multiple copies of subject pronouns

An important question is why we see multiple copies of *had* in copy conditionals, but only one copy of the subject. After all, because the internal structure of #D+T# is inaccessible to linearization, it would not create an ordering contradiction for the pronoun to be pronounced both as part of #D+T# and elsewhere (*e.g.* in Spec,PerfP). In this section, I will present some possible explanations for why the subject is only pronounced as part of #D+T#.

[10] This is only an advantage inasmuch as the stem changes triggered by auxiliary contraction are idiosyncratic. If the stem changes are regular, then they may will be the result of a phonological process rather than a lexical difference.

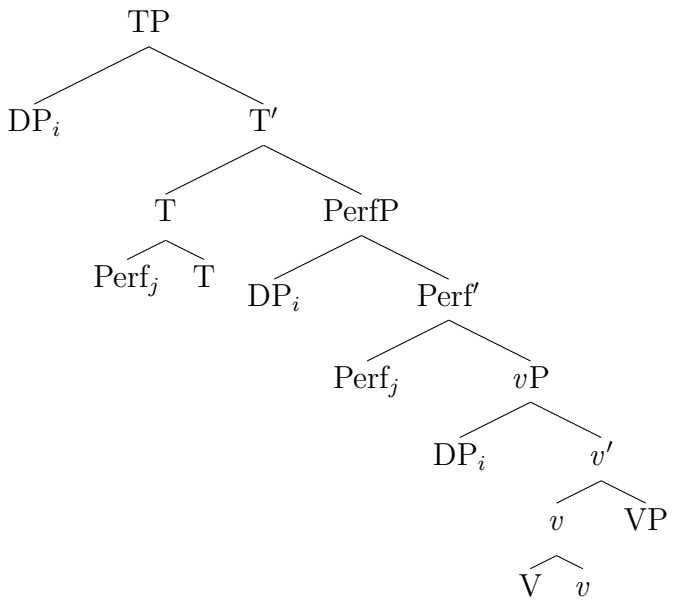
One possibility appeals to economy considerations. Let us return to (102), repeated here as (109). We see that the subject pronoun is spelled out as part of the Vocabulary Item *he'd* inserted in #D+T#, but is not spelled out as *he* in Spec,PerfP. Now, the Vocabulary Item *he'd* presumably has all the features of *he*, as *he'd* is unambiguously a third-person singular nominative pronoun (at least in part). Consequently, there is no meaning loss if *he* is omitted. Economy considerations will then block Vocabulary Insertion of *he* in Spec,PerfP, since this operation is redundant and thus unnecessary. On the other hand, insertion of *had* in C will not be blocked, because this Vocabulary Item serves the functional purpose of marking the clause as a conditional adjunct.

(109)



A second possibility appeals to phases. According to Chomsky (2001), Spell-Out does not happen at once, but rather proceeds in parallel with syntax. Each time a **phase** (CP, vP, and maybe other categories) is produced, the complement of the phase head is sent to Morphology. Thus, when the CP in (109) is completed, (110) is sent to Morphology. M-merger between Spec,TP and T will occur as usual. If Fusion precedes linearization, then we predict that Vocabulary Items will be inserted in both #D+T# and Spec,PerfP, as before. However, if Fusion follows linearization, then Vocabulary Insertion will be blocked for Spec,PerfP, so the pronoun will only be pronounced as part of the Vocabulary Item *he'd* inserted in #D+T#. This happens before C is spelled out, so by the time C has entered Morphology, D and T will have already Fused. Consequently, the auxiliary will appear both as *had* in the copy of Perf embedded in C, and as part of *he'd*.

(110)

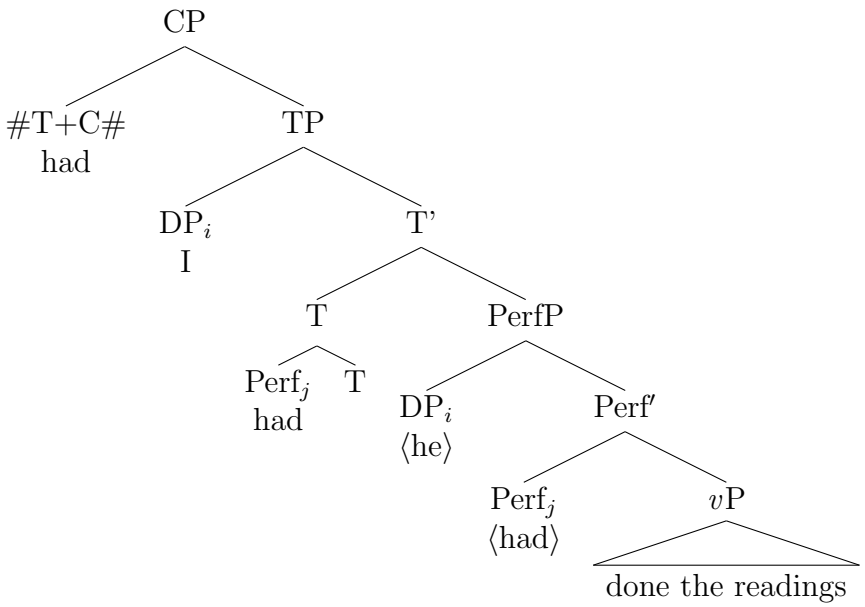


This analysis has a number of issues. For example, if linearization precedes Fusion, then (108) (*If I'd had done the readings*...) should be unacceptable with the auxiliary pronounced twice. This is at odds with the fact that five speakers accepted the sentence. However, my goal here is not to defend a particular account of why the subject pronoun is only pronounced once, but rather to demonstrate the possibility of finding a principled account of this fact.

3.4.4 Other possible Fusions

Up to this point, I have explored the idea that multiple copies of the auxiliary *had* are permitted in copy conditionals because of Fusion between the subject and the auxiliary in T. However, head-to-head movement also creates an environment for Fusion, so *a priori* we might expect multiple copies of the auxiliary to be permitted due to Fusion between the raised auxiliary and C. This would give the structure in (111).

(111)



If Fusion happens along these lines, then the lower instance of *had* should be able to surface in any form, not just the clitic form. While most speakers only accepted sentences with multiple copies of *had* if one of those copies was in its clitic form, one speaker accepted such sentences regardless of the form of the auxiliaries (112). Possibly, for this speaker, the higher copy of *had* can fuse with C.

(112) Had they had finished their homework earlier, they wouldn't have missed the game. ($y = 1, m = 3, n = 4$)

3.4.5 Other open questions

There remain a number of important questions that are beyond the scope of this thesis. Up to this point, I have only considered instances in which multiple copies of *had* are pronounced, raising the question of whether multiple copies of other auxiliaries surface. Unfortunately, I did not collect data on this question.

Two more questions are why Fusion occurs, and to what degree it is optional. As discussed in section 3.4.1, if we adopt the hypothesis that linearization happens concurrently with Vocabulary Insertion, then Fusion must be optional. If not, then we would expect multiple copies of the auxiliary to surface whenever auxiliary contraction occurs, which is not the case. On the other hand, if linearization precedes Vocabulary Insertion, and Fusion either precedes or follows linearization, then Fusion could be mandatory. In short, the answers to this question are still unclear.

3.5 Conclusion

I have proposed that auxiliary contraction involves M-merger between a simplex DP subject and T. This M-merger feeds a Fusion operation, creating a head #D+T# whose internal structure is inaccessible to the linguistic system, and which is spelled out as the combination of the pronoun and the clitic form of the auxiliary. When

auxiliary contraction occurs in an inverted conditional, there is no ordering contradiction between the copy of T in C and the copy of T that became part of #D+T#, since the latter is invisible. Consequently, the auxiliary seems to appear twice – as *had* in C, and as part of *I'd*, *he'd*, *etc.*, in #D+T# – creating a copy conditional. This explains why for most speakers, copy conditionals are only acceptable when the lower copy of the auxiliary appears in its clitic form.

Conclusion

In this thesis, I have investigated two conditional constructions in which there seem to be "extra" auxiliaries in the conditional adjunct. The first of these constructions, the modal antecedent conditionals, are characterized by a modal auxiliary followed by the perfect auxiliary *have* in the antecedent. Distinctively, one of the possible modals here is a modal *had*, although for many speakers the use of modal *had* is restricted by OCP. Modal antecedent conditionals are semantically distinct from past perfect conditionals in that they do not allow the counterfactual inference to be cancelled. Furthermore, for some speakers, modal antecedent conditionals with *would* in the antecedent seem to require that the modal be interpreted volitionally.

The second construction is the copy conditional, so called because it involves the appearance of multiple copies of the auxiliary *had*. Normally linearization prevents multiple copies of the same element from being pronounced. Here, however, morphological Fusion renders one of the copies invisible to linearization, allowing multiple copies of the auxiliaries to surface. This analysis predicts that multiple copies of *had* should be able to appear whenever one of those copies appears in clitic form, which preliminary data suggest to be the case.

Appendix A

Initial questionnaire

A.1 Part 1

1. If James wins the lottery, he will pay off his student loans.
2. If Mary went to London, she would visit Big Ben.
3. If I had been clearer during the interview, I would have gotten the job.
4. *We're at a party, talking about our friend John whose car broke down recently, and we're wondering if it's been fixed yet. We see John walk in with his friend Andy, and I say*: If John had got his car fixed, he would've driven himself to the party. Since he caught a ride with Andy, his car must still be in the shop.
5. Had Robert worn his coat, he wouldn't have caught cold.
6. If I'd've checked the weather, I wouldn't have got caught in the rain.
7. *We threw our friend Pat a surprise party, and you're worried she knew about it beforehand. I don't think so, so I say:* Had Pat known we were throwing her a surprise party, she wouldn't have been surprised, which she definitely was. So she couldn't have known.
8. Had I'd set aside more time for my assignment, I would've done better in the end.
9. *Simone missed an important business meeting yesterday.* If the meeting had been tomorrow instead, she would have been able to make it.
10. If she'd've come over earlier, we would've had time to play Monopoly.

A.2 Part 2

1. *My mom thinks I'm hiding her phone from as a prank, but I think she just misplaced it or something. I say to her:* Had I HAD seen it, I would've told her where it was.

2. *I tried to deliver a package to a professor, but they weren't in their office just then. A sign on the door says they have office hours in a few hours.* If I'd've come a few hours later instead, the professor would've been there.

3. *I'm trying to figure out if my boss got the email I sent him.* If he'd've seen my email, he would have sent a reply. Since he didn't send a reply, he must not have seen it.

4. Had she'd heard the news, she wouldn't look so cheery.

5. If he hadda knocked one more time, I would've answered the door.

6. If they would've called me, I would've come running.

7. Had he'd read the book, he wouldn't have liked the movie.

8. If they hadn't've pressed "reply all," they would've avoided a lot of embarrassment.

9. *Jennifer was supposed to go on a trip to Mexico today, but her flight was cancelled because of inclement weather.* If it wouldn't have snowed, she would be in Cancun by now.

10. Had they'd been more polite, I wouldn't have given them a time-out.

A.3 Part 3

1. *Rob dropped by to see me yesterday, but I was out of town until late tonight, so I wasn't there to greet him.* If he hadda come tomorrow instead, I would've been there to greet him.

2. *My mom tried to call me yesterday to chat, but I was too busy then. Tomorrow, I have much less going on.* If she would've called me tomorrow, I would've had time to talk to her.

3. *We're wondering if our friend Linda got the promotion she was competing for.* Had she'd got the promotion, she would've replaced her old car. So judging by her new Toyota, she must have got the promotion.

Appendix B

Final questionnaire

B.1 Part 1

1. If James wins the lottery, he will pay off his student loans.

2. If Mary went to London, she would visit Big Ben.

3. If I had spoken more clearly during the interview, I would have gotten the job.

4. *We're at a party, talking about our friend John whose car broke down recently, and we're wondering if it's been fixed yet. We see John walk in alone:* If John had got his car fixed, he would've driven himself to the party. Since he came by himself, his car must be working again.

5. Had Robert worn his coat, he wouldn't have caught a cold.

6. If I'd've checked the weather, I wouldn't have been stuck in the rain.

7. *We threw our friend Pat a surprise party, and I think she knew about it beforehand. I say:* Had Pat known we were throwing her a surprise party, she wouldn't have been surprised. Since she didn't act surprised, she must have known about the party.

8. Had I'd set aside more time for my assignment, I would've done better in the end.

9. *Simone missed an important business meeting yesterday.* If the meeting had been tomorrow instead, she would have been able to make it.

10. If she'd've come over earlier, we would've had time to play Monopoly.

11. Had they had finished their homework earlier, they wouldn't have missed the game.

B.2 Part 2

1. *My mom thinks I'm hiding her phone from her as a prank, but I think she just misplaced it or something:* Had I HAD seen it, I would've told her where it was.

2. *I tried to deliver a package to a professor, but they weren't in their office just then. A sign on the door says they have office hours in a few hours.* If I'd've come a few hours later instead, the professor would've been there.

3. If he'd've got the job, he would have brought champagne. So, since he brought champagne, he must have got the job.

4. Had she'd heard the news, she wouldn't look so cheery.

5. If he had've knocked one more time, I would've answered the door.

6. If they would've called me, I would've come immediately.

7. If the dogs had all have been adopted, the SPCA would be closed.

8. Had he'd read the book, he wouldn't have liked the movie.

9. *Max is feeling lonely. Gina thinks her friend Kelly is a perfect fit. But when Gina goes to introduce them, it turns out they already knew each other. She tells me:* Not only had he'd met Kelly before, but they actually used to date!

10. If they hadn't've pressed "reply all," they would've avoided a lot of embarrassment.

11. *Jennifer was supposed to go on a trip to Mexico today, but her flight was cancelled because of inclement weather.* If it wouldn't have snowed, she would be in Cancun by now.

12. Had they'd been more polite, I wouldn't have given them a time-out.

13. *Dani and Rob are rival co-workers who had a meeting scheduled for today. I wonder if Dani managed to make it to the meeting.* If she had've been there, Rob would have been really peeved. Since Rob is looking grumpy, Dani must have been there.

B.3 Part 3

1. *Rob dropped by to see me yesterday, but I was out of town until late tonight, so I wasn't there to greet him.* If he had've come tomorrow instead, I would've been there to greet him.

2. If I'd had done the readings, I would've passed the test.

3. *My mom tried to call me yesterday to chat, but I was too busy then. Tomorrow, I have much less going on.* If she would've called me tomorrow, I would've had time to talk to her.

4. If the teachers all had've quit, there wouldn't be anyone left in the schools.

5. Had they'd've left earlier, they would've avoided the traffic.

6. Had I hadn't left the oven on, my house wouldn't have burned down.

7. *We're wondering if our friend Linda got the promotion she was competing for.* Had she'd got the promotion, she would've replaced her old car. So judging by her new Toyota, she must have got the promotion.

8. Had we had've chosen our classes earlier, we would've had more options.

9. *I'm a detective, trying to determine what happened to Mr. Jones, who was found dead.* If Jones would've taken arsenic, he would've shown exactly those symptoms he did show. So, he must have taken arsenic.

10. If the politicians would all have voted against the bill, it would've prevented an unnecessary war.

11. *Your friend Marc went to France last week, and I'm wondering if it was his first time or not:* Had he'd been to Europe before?

References

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

Axel-Tober, K., & Gergel, R. (2015). *The Oxford Handbook of Modality and Mood*, chap. Modality and mood in formal syntactic approaches. Oxford University Press.

Bhatt, R., & Pancheva, R. (2006). *The Blackwell Companion to Syntax*, vol. 1, chap. Conditionals, (pp. 638–687). Blackwell Publishing.

Biezma, M. (2012). Conditional inversion and GIVENNESS. In N. Ashton, A. Chereches, & D. Lutz (Eds.), *Proceedings of SALT 21*, (pp. 552–571). Linguistic Society of America.

Biezma, M., Carnie, A., & Siddiqi, D. (2014). Counterfactuality in non-standard subjunctive conditionals. In H.-L. Huang, E. Pool, & A. Rysling (Eds.), *NELS 43: Proceedings of the Forty-Third Annual Meeting of the North East Linguistic Society*, vol. 1, (pp. 25–36). City University of New York, Graduate Linguistics Student Association.

Bruening, B. (2017). *The Blackwell Companion to Syntax*, chap. Subject-auxiliary inversion, (pp. 4250–4277). Hoboken: John Wiley & Sons.

Chomsky, N. (1995). *Government and Binding Theory and the Minimalist Program: Principles and Parameters in Syntactic Theory*, chap. Bare phrase structure, (pp. 385–439). Oxford: Blackwell.

Chomsky, N. (2001). *Ken Hale: A Life in Language*, chap. Derivation by phase, (pp. 1–52). Cambridge, Massachusetts: MIT Press.

Close, J. (2004). *English Auxiliaries: A Syntactic Study of Contraction and Variation*. Ph.D. thesis, University of York, York.

Collins, P. C. (2007). *Functional Perspectives on Grammar and Discourse: In honour of Angela Downing*, chap. Modality across World Englishes. Benjamins.

Dancygier, B., & Sweetser, E. (2005). *Mental Spaces in Grammar: Conditional Constructions*. No. 108 in Cambridge Studies in Linguistics. Cambridge, England: Cambridge University Press.

Embick, D., & Noyer, R. (2001). Movement operations after syntax. *Linguistic Inquiry*, *32*(4), 555–595.

Grice, P. (1975). *The Logic of Grammar*, chap. Logic and conversation, (pp. 64–75). Encino, CA: Dickinson.

Halle, M., & Marantz, A. (1993). *The View from Building 20: Essays in Honor of Sylvain Bromberger*, chap. Distributed Morphology and the pieces of inflection, (pp. 111–176). MIT Press.

Harley, H., & Noyer, R. (1999). Distributed Morphology. *Glot International*, *4*(4), 3–9.

Heim, I. (1991). *Semantik: Ein internationales handbuch der zeitgenösischen forschung*, chap. Artikel und definitheit, (pp. 487–535). Berlin: de Gruyter.

Heim, I., & Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell Publishers.

Hiemstra, I. (1986). Some aspects of *wh*-questions in frisian. In *North-Western European Language Evolution (NOWELE)*, vol. 8, (pp. 97–110).

Hiraiwa, K. (2014). *Identity Relations in Grammar*, chap. Constraining doubling, (pp. 225–254). de Gruyter.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge, England: Cambridge University Press.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, *31*(2), 231–270.

Iatridou, S., & Embick, D. (1994). Conditional inversion. In M. Gonzàlez (Ed.), *Proceedings of the North Eastern Linguistics Society 24*. University of Massachusetts, Amherst: Graduate Linguistics Student Association.

Ippolito, M. (2008). Subjunctive conditionals. In A. Grønn (Ed.), *Proceeds of SuB12*, (pp. 256–270). Oslo, Norway: ILOS.

Ippolito, M. (2013). *Subjunctive Conditionals: A Linguistic Analysis*. Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.

Kayne, R. S. (1994). *The Antisymmetry of Syntax*. No. 25 in Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.

Kayne, R. S. (1997). The english complementizer of. *The Journal of Comparative Germanic Linguistics*, *1*(1), 43–54.

Kratzer, A. (1981). *Words, worlds, and contexts*, chap. The notional category of modality, (pp. 38–73). de Gruyter.

Kratzer, A. (1986). Conditionals. In A. M. Farley, P. T. Farley, & K.-E. McCullough (Eds.), *CLS 22: Papers from the Parasession on Pragmatics and Grammatical Theory*, (pp. 1–15). Chicago: Chicago Linguistic Society.

Leben, W. R. (1973). *Suprasegmental Phonology*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

McDaniel, D. (1986). *Conditions on wh-chains*. Ph.D. thesis, City University of New York.

Molencki, R. (2000). *Pathways of Change: Grammaticalization in English*, chap. Parallelism vs. asymmetry: The case of English counterfactual conditionals, (pp. 311–328). Amsterdam: John Benjamins Publishing Company.

Musan, R. (1997). Tense, predicates and lifetime effects. *Natural Language Semantics*, *5*, 271–301.

Nunes, J. (2004). *Linearization of Chains and Sideward Movement*. No. 43 in Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.

Nunes, J. (2011). *The Oxford Handbook of Linguistic Minimalism*, chap. The copy theory. Oxford University Press.

Ogihara, T. (2007). Tense and aspect in truth-conditional semantics. *Lingua*, *117*, 392–418.

Richards, N. (2010). *Uttering Trees*. Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.

Siddiqi, D., & Carnie, A. (2012). The english modal had. *Snippets*, *25*, 18–20.

von Fintel, K. (2011). *Semantics: An International Handbook of Meaning*, vol. 2, chap. Conditionals, (pp. 1515–1538). Berlin: de Gruyter Mouton.

von Fintel, K. (2012). *The Routledge Companion to Philosophy of Language*, chap. Subjunctive conditionals, (pp. 466–477). New York: Routledge.

von Fintel, K., & Heim, I. (2011). Intensional semantics. Unpublished manuscript.

Zwicky, A. M., & Pullum, G. K. (1983). Cliticization vs. inflection: English n't. *Language*, *59*(3), 502–513.