Attitudes, aboutness, and fake restricted readings*

Kai von Fintel
MIT

Robert Pasternak
Leibniz-Center for General Linguistics (ZAS)

Abstract A sentence like *If Laura becomes a zombie, she wants you to shoot her* seems to assert that Laura has a certain kind of *restricted attitude*: her current desires, when considering only possible states of affairs in which she becomes a zombie, are such that you shoot her. We discuss several intuitively plausible accounts that attribute these readings to restrictions imposed on the world-quantification of the attitude or some modal operator in its scope. We then provide evidence against these proposals and argue that such "restricted" readings are not a result of any quantificational restriction of this sort. Instead we adopt the view that attitudes are about (possibly non-actual) situations, and illusory restricted readings are a reflection of certain relationships between an attitude-holder and an "about-situation". This new theory accounts for all of the data that were problematic for restriction-based accounts and invites reconsideration of conditionals more generally, on the grounds that other cases of apparent *if*-clause restriction might similarly be reducible to "fake" restricted readings.

1 Introduction

Consider (1), based on an example from Pasternak (2018):

(1) If Laura becomes a zombie, she wants you to shoot her.

On its surface, (1) appears to be a conditional with the attitude verb *want* in its consequent. However, the interpretation that it gives rise to is not a claim about what Laura's desires *would be* were she to become a zombie. Presumably, zombie-Laura would not want to be shot: she would either want *not* to be shot, or lack the cognitive disposition to care one way or the other. Rather, (1) seems to be a statement about what Laura's *current* desires are in light of the possibility of her potentially becoming a zombie. That is, (1) claims that Laura has a certain kind of *restricted attitude*: when restricting her attention only to those possible states of affairs in which she becomes a zombie, Laura wants to be shot.

* **Comments welcome and much appreciated.** Acknowledgments to be added later. The authors are listed alphabetically and share joint lead authorship of this paper. Pasternak's research is funded by DFG Grant #387623969 (*DP-Border*, PIs: Artemis Alexiadou and Uli Sauerland).

Constructions of the sort exemplified by (1), which we will refer to as *attitude conditionals*, can be usefully contrasted with examples like (2), which we will refer to as *would-attitude conditionals*.

(2) a. If Laura becomes a zombie, she will want you to shoot her.
b. If Laura became a zombie, she would want you to shoot her.

The would-attitude conditionals in (2), unlike their attitude conditional counterpart in (1), are at least by default read as being about the desires of zombie-Laura, and thus presumably false. That being said, it seems that these latter examples do not fully prohibit a restricted or restricted-like reading relating Laura's actual, present desires, rather than her hypothetical zombie-desires. This is an interesting fact in and of itself, and one that we will eventually gesture toward an explanation for later in this paper. But of greater interest for our purposes is that this marginal interpretation of (2) is not only the most prominent reading of (1), but seemingly the only reading.

Perhaps unsurprisingly, restricted readings are not confined to *want* and can arise with a host of embedding verbs, as exemplified by the conditionals in (3):

(3) a. If it rains tomorrow, then I **expect/permit/forbid/need** you to bring an umbrella.
b. If it rained yesterday, then I **hope/think/know/suspect/pray** that you brought an umbrella.

Restricted readings of attitudes are also not confined to conditionals. For instance, Blumberg & Holguín (2019) provide examples in which disjunctions can give rise to similar sorts of readings, as can be seen in their (4):

(4) Blumberg & Holguín 2019: p. 380:
Either a lot of people are on the deck outside, or I regret that I didn't bring more friends.

As Blumberg & Holguín note, the interpretation of (4) is more or less the same as that of the attitude conditional in (5), suggesting that such *attitude disjunctions* can also give rise to restricted readings.

(5) If there aren't a lot of people on the deck outside, then I regret that I didn't bring more friends.

So how does one account for these restricted readings? When focusing our attention on attitude conditionals, some obvious candidates come to mind. Attitudes and modals are commonly held to denote quantifiers over possible worlds, and since the work of Lewis (1975) and Kratzer (1981, 1991a,b, 2012) *if*-clauses are generally taken to serve as restrictors of quantificational operators, including world-quantifiers. It thus may be the case that restricted readings in attitude conditionals

involve the antecedent restricting the quantification of either the attitude or some modal in the scope of that attitude. That is, in an attitude conditional like (1), the world-quantification of *want* (or an embedded modal) is restricted so that it ranges only over possible worlds in which Laura becomes a zombie. This could potentially be generalized to other types of restricted attitude readings such as attitude disjunctions; the important point is that attitude restriction involves world-quantificational restriction of either the attitude or some embedded modal. We will refer to such accounts of restricted attitude readings as *restriction-based* accounts.

After some preliminary discussion in Section 2 of our basic assumptions about the semantics of attitudes and conditionals, in Section 3 we will outline the three broad types of restriction-based account that we take to be most plausible, focusing primarily on attitude conditionals: **embedded conditional**, **attitude restriction** (Pasternak 2018), and **double restriction** (Blumberg & Holguín 2019). These three types of analysis share in common the aforementioned stance that restricted attitude readings come from restricting world-quantifiers, but differ in what world-quantifiers the *if*-clause is taken to restrict. However, in Section 4 we will provide a variety of evidence suggesting that none of these restriction-based accounts make the right predictions. In short, we show that restricted readings of attitude conditionals cannot plausibly be derived from restricting the quantification of the attitude or an embedded modal, and must come from elsewhere.

In Section 5 we will offer an analysis of restricted readings of attitudes in which there is no actual grammatical or pragmatic operation restricting the quantification of the attitude or any modals in its scope. In fact, on our analysis, attitude conditionals are simply as they appear: they are plain conditionals with attitude verbs in their consequents. The apparent restricted reading thus might be more appropriately referred to as a *fake restricted reading*. As far as how these fake restricted readings arise, we will follow Barwise & Perry (1983) and Kratzer (2002) in positing that attitudes are about *situations*, where situations are partial possible worlds. The difference between attitude and would-attitude conditionals then comes down to the fact that attitude conditionals make claims about real-world attitudes toward hypothetical situations (e.g., Laura's actual desires about possible zombie-situations), while would-attitude conditionals make claims about hypothetical attitudes toward hypothetical situations. We will show how an analysis of (fake) restricted readings built around this use of "about-situations" makes the right predictions for all of the cases that were problematic for restriction-based accounts. After providing our account of attitude conditionals, would-attitude conditionals, and attitude disjunctions, we will briefly discuss the possibility of generalizing our analysis of attitude conditionals to all conditionals: should the idea of *if*-clauses as quantificational restrictors be replaced with fake restriction across the board, for all of the varied operators that *if*-clauses supposedly restrict? We will conclude that while this is an appealing

possibility for multiple reasons, certain semantic differences between attitudes and other operators entail that finding concrete evidence in favor of fake restriction over real quantificational restriction can be trickier for the latter than for the former. As a result, the possible extension to other forms of conditionals is left as an intriguing possibility, rather than a fully fleshed-out analysis. Finally, in Section 6 we will offer some concluding remarks, as well as some potential lines for future inquiry.

2 Basic assumptions about attitudes and conditionals

First we put forward some basic assumptions about the lexical and compositional semantics of attitudes and conditionals. It is worth noting that many of these assumptions are orthogonal to the restriction-based analyses discussed in Sections 3 and 4, as well as to our own analysis in Section 5. We will try to be explicit about which assumptions are and are not crucial.

2.1 Conditionals

We start with conditionals. As noted in the introduction, for most of this paper we will adopt the standard view that *if*-clauses serve to restrict quantificational operators, whether they be adverbs of quantification (e.g., *always*, *sometimes*), modals, or perhaps some other operators. Take, for example, the deontic conditional in (6):

(6) If Laura becomes a zombie, you must shoot her.

Deontic *must* is generally taken to quantify over those worlds that are *circumstantially accessible*. For example, on Kratzerian approaches (Kratzer 1981, 1991a,b, 2012), it universally quantifies over the deontically ideal worlds in this domain: the antecedent-less *You must shoot Laura* is true iff among circumstantially accessible worlds, all of the deontically best ones are worlds in which you shoot Laura. The *if*-clause then restricts the domain of circumstantially accessible worlds to those in which the antecedent is true. Thus, (6) is true iff among those circumstantially accessible worlds *in which Laura becomes a zombie*, all of the best ones are worlds in which you shoot her.

Note that while we are framing this discussion in terms of a Kratzerian "best-worlds" analysis of modals—and will continue to do so for the duration of the paper—this conception of *if*-clauses as restrictors of quantificational operators is not confined to a Kratzer-style semantics for deontic *must*. For example, Goble (1996) and Lassiter (2011) argue that deontic modality is best captured by means of the decision-theoretic notion of *expected utility*. But decision-theoretic analyses are equally compatible with the conception of *if*-clauses as restrictors. If *must q* is evaluated by determining the expected utility of *q* relative to the whole domain of

possible worlds (or perhaps just those deemed circumstantially accessible), then *if p, must q* can be evaluated by determining the expected utility of q relative only to the subdomain of possible worlds in which p holds. In other words, while probabilistic or decision-theoretic definitions of modals do not "quantify" in the narrow sense, they still utilize domains that can be restricted by *if*-clauses. Thus, our adoption of a best-worlds semantics is orthogonal to the core issues presented in this paper.

If modals are quantifiers that are restricted by *if*-clauses, the next obvious question is how, compositionally speaking, this domain restriction is effected. One possibility is to take inspiration from quantifiers in the nominal domain, such as *every*. Consider the sentence *Every student left*, in which *student* semantically restricts *every*, with the whole DP taking *left* as its scope.

(7) $\underbrace{\text{Every}}_{\text{Q}}\ \underbrace{\text{student}}_{\text{R}}\ \underbrace{\text{left}}_{\text{S}}$ Q = quantifier, R = restrictor, S = scope

In this sentence, *student* is by all relevant syntactic diagnoses the complement of *every*, and so its restriction can easily be effected via argument saturation. That is, if we define ⟦every⟧ in the standard manner as in (8), then the first argument (λP) can be saturated by ⟦student⟧, and the second argument ($\lambda P'$) by ⟦left⟧, generating the appropriate truth conditions.

(8) $[\![\text{every}]\!] = \lambda P \lambda P'.\ \forall x[P(x) \rightarrow P'(x)]$

It is tempting to say that *if*-clauses restrict modals in the same manner, i.e., by means of argument saturation. Thus, the structure of (6) at LF will be as in (9), with the *if*-clause occupying the complement position of *must*:

(9) $\underbrace{\text{must}}_{\text{Q}}\ \underbrace{\text{if Laura becomes a zombie}}_{\text{R}}\ \underbrace{\text{you shoot her}}_{\text{S}}$

Now suppose ⟦must⟧, evaluated at a context c and world w, is as in (10), where $\textsc{Cir}_{c,w}$ is the set of circumstantially accessible worlds from w as determined in c, and $\textsc{Best}_{c,w}(A)$ is the set of deontically best worlds in A as determined by the standards of c in w:

(10) $[\![\text{must}]\!]^{c,w} = \lambda p \lambda q.\ \forall w' \in \textsc{Best}_{c,w}(\textsc{Cir}_{c,w} \cap p)[q(w')]$

In this case, composing the LF in (9) will lead to the truth conditions in (11), i.e., true iff among those circumstantially accessible worlds in which Laura becomes a zombie, all of the best ones are worlds in which you shoot her:

(11) $[\![(9)]\!]^{c,w} = 1$ iff $\forall w' \in \textsc{Best}_{c,w}(\textsc{Cir}_{c,w} \cap \textsc{zombie})[\textsc{shoot}(w')]$

However, von Fintel (1994) argues based on a variety of evidence that *if*-clauses do *not* restrict their associated modals by means of argument saturation. With this

in mind, we will go a slightly different route. Suppose that modals only take one propositional argument, their scope. But each modal carries an index n, and the modal is restricted by whatever proposition the variable assignment g assigns to n. Thus, the denotation for *must* will look as in (12):

(12) $[\![\text{must}_n]\!]^{g,c,w} = \lambda q.\ \forall w' \in \textsc{Best}_{c,w}(\textsc{Cir}_{c,w} \cap g(n))[q(w')]$

When an *if*-clause attaches higher in the tree, with *if* coindexed with the modal, the ensuing interpretation is one in which $g(n)$ is replaced with the denotation of the antecedent. We will stipulate this through a semantic rule of composition, though no doubt there are more elegant ways of achieving this:

(13) $[\![\text{if}_n\ \Phi,\ \Psi]\!]^g = [\![\Psi]\!]^{g[n \to [\![\Phi]\!]^g]}$
where $g[n \to p]$ is the g' identical to g except that $g'(n) = p$

In this case, the LF for (6) will be as in (14):

(14) if_7 Laura becomes a zombie, you must_7 shoot her

Up to and excluding the *if*-clause, the denotation will be as follows:

(15) $[\![\text{you must}_7 \text{ shoot her}]\!]^{g,c,w} = 1$ iff $\forall w' \in \textsc{Best}_{c,w}(\textsc{Cir}_{c,w} \cap g(7))[\textsc{shoot}(w')]$

Thanks to the rule in (13), the interpretation of the clause as a whole is as follows:

(16) $[\![\text{if}_7 \text{ Laura becomes a zombie, you must}_7 \text{ shoot her}]\!]^{g,c,w}$
$= [\![\text{you must}_7 \text{ shoot her}]\!]^{g[7 \to \textsc{zombie}],c,w}$
$= 1$ iff $\forall w' \in \textsc{Best}_{c,w}(\textsc{Cir}_{c,w} \cap \textsc{zombie})[\textsc{shoot}(w')]$

We thus arrive at the exact same interpretation, albeit by different means: *if*-clauses effect their domain restriction through coindexation, rather than argument saturation. While we are skeptical of an argument saturation account for independent reasons, we adopt a coindexation account in this paper not for those reasons, and not because it is required for our own account (it is not). Rather, we adopt it because it allows for considerable syntactic flexibility in the relationship between modal and restrictor, since the restrictor does not need to appear as the complement of the modal at any syntactic level of representation. This in turn gives the restriction-based theories that we will be arguing against their best chance at success, since our evidence against them will not hinge on arbitrary syntactic constraints relating *if*-clauses to the operators they purportedly restrict.

As a final note before moving on to attitudes, it is well known that conditionals needn't come with overt modal operators, or overt quantificational operators of any kind, as illustrated in (17):

(17) If Hans hasn't eaten, then he is hungry.

In such cases, we will follow the commonly adopted view that there is a covert modal MODAL, which is a universally quantifying modal akin to *must* (or in some cases, the future modal *woll*), and which is restricted by the *if*-clause as in (18):

(18) if$_7$ Hans hasn't eaten, then MODAL$_7$ he is hungry

MODAL is restricted in its choice of modal flavors: it cannot, for example, be deontic in (17). We will posit a denotation of MODAL based on universal quantification over maximally likely worlds, as in (19), where $\text{ML}_w(A)$ is the set of most likely A-worlds from w:

(19) $[\![\text{MODAL}_n]\!]^{g,c,w} = \lambda p.\ \forall w' \in \text{ML}_w(g(n))[p(w')]$

Thus, the LF in (18) will receive the interpretation in (20), true iff among all of the most likely worlds in which Hans hasn't eaten, he is hungry:

(20) $[\![(18)]\!]^{g,c,w} = 1$ iff $\forall w' \in \text{ML}_w(\text{NO\_EAT})[\text{HUNGRY}(w')]$

It is worth noting, in accordance with a running theme of this paper, that adopting the particular denotation for MODAL in (19) is not crucial to any of the theories in this paper, including our own.

2.2 Attitudes

With our assumptions about the compositional semantics of conditionals out of the way, we next move on to our assumptions about the semantics of attitudes. We will mostly focus on *want* in this paper, so we will stick to that in this section, and discuss revisions for other attitude verbs where relevant.

We follow the long tradition of attitude semantics, dating back at least to the work of Hintikka (1969), in which the denotations of attitude verbs share core formal traits with those of modals. Thus, best-worlds analyses of modals are often paired with best-worlds (or similar) theories of attitudes, and decision-theoretic analyses of modals are generally paired with decision-theoretic analyses of attitudes. Much like in the case of modals and conditionals, for our purposes it does not really matter which path we take in terms of the lexical semantics of *want*, and so we will stick to a best-worlds approach. More specifically, we will adopt a somewhat simplified version of von Fintel's (1999) semantics for *want*, which more or less directly imports Kratzer's semantics for modals. This simplified denotation can be seen in (21), where $\text{Dox}_{x,w}$ is the set of worlds compatible with x's beliefs in w (x's *belief worlds* in w), and $\text{BEST}_{x,w}(A)$ is the set of best A-worlds by x's standards in w.

(21) $[\![\text{want}]\!]^w = \lambda p \lambda x.\ \forall w' \in \text{BEST}_{x,w}(\text{Dox}_{x,w})[p(w')]$

Thus, *Inge wants to leave* is true iff among those worlds compatible with Inge's beliefs, all of the best ones by her standards are worlds in which she leaves.[1]

Notice that by treating the denotations of attitudes as world-quantifiers, we invite the possibility that attitudes can themselves be restricted in the same manner as modals. Given our coindexation account of *if*-clause-imposed restrictions, a denotation for *want* that allows it to be restricted by an *if*-clause will look as in (22):

(22) $[\![\text{want}_n]\!]^{g,w} = \lambda p \lambda x.\ \forall w' \in \textsc{Best}_{x,w}(\text{Dox}_{x,w} \cap g(n))[p(w')]$

In this case, if a coindexed *if*-clause denoting q restricts *want*, then what *want* quantifies over is the set of x's favored worlds among those belief worlds of hers *in which q is true*. It is thus conceivable that *if*-clauses could restrict attitudes as well as modals, a possibility that is exploited by multiple of the restriction-based theories discussed in the next section.

3 Attitude conditionals: Three paths to a restriction-based account

In this section we will introduce what we take to be the most plausible restriction-based analyses for attitude conditionals like (1), repeated below.

(1) If Laura becomes a zombie, she wants you to shoot her.

The three proposals, which we refer to as **embedded conditional**, **attitude restriction**, and **double restriction** analyses, share in common the view that restricted readings arise due to the quantificational restriction of either the attitude itself or a modal in that attitude's complement; they differ in terms of which operators are restricted. In Section 4, we will provide evidence against all of these accounts.

3.1 Embedded conditional

On an **embedded conditional** analysis, the *if*-clause in an attitude conditional really makes its semantic contribution below the attitude, meaning that attitude conditionals are interpreted as simply making claims about attitudes toward conditional propositions. Thus, (1) will be interpreted approximately identically to (23):

(23) Laura wants it to be the case that if she becomes a zombie, you shoot her.

To our knowledge, this analysis of attitude conditionals has not been seriously argued for in the literature, though Drucker (2019) and Blumberg & Holguín

1 For alternative analyses in the narrowly quantificational tradition—that is, non-decision-theoretic denotations—see Heim 1992, Villalta 2008, Phillips-Brown 2018. For decision-theoretic proposals, see Levinson 2003, Lassiter 2011. For overviews comparing and contrasting these approaches, see Crnič 2011, Pasternak 2018. All of these are compatible with all of the analyses in this paper.

(2019) argue against versions of it based on the reasonable assumption that it is an intuitively plausible account. It also bears a *prima facie* resemblance to certain existing theories of non-attitude conditionals: while we assume that *if*-clauses restrict modal operators, there is a long history (for example, Broome (1999)) of instead positing that conditionals like (6) really involve bare conditionals inside the scope of the modal, as informally represented in (24):

(24) must [if Laura becomes a zombie, you shoot her]

A natural way of formulating an embedded conditional analysis is to posit the existence of a modal MODAL in the embedded clause, with the *if*-clause restricting this modal. Thus, the LF for (1) is predicted to be as in (25):

(25) if$_7$ Laura becomes a zombie, she wants MODAL$_7$ you to shoot her.

This will compose straightforwardly, with the *if*-clause restricting MODAL and generating an interpretation more or less identical to (23). More specifically, we predict the interpretation in (26), true iff all of Laura's ideal belief worlds w' are such that all of the most likely worlds from w' in which she becomes a zombie are worlds in which you shoot her.

(26) $[\![(25)]\!]^{g,c,w} = 1$ iff
$\forall w' \in \textsc{Best}_{x,w}(\textsc{Dox}_{x,w})[\forall w'' \in \textsc{ML}_{w'}(\textsc{zombie})[\textsc{shoot}(w'')]]$

We will call this the **low modal** analysis.

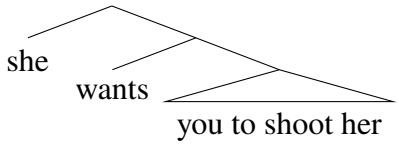
An alternate version of the embedded conditional approach, discussed in detail by Drucker (2019) and especially Blumberg & Holguín (2019), involves the *if*-clause making its semantic contribution in the matrix rather than embedded clause, with the attitude—and presumably, its external argument (the subject)—moving covertly in order to outscope the conditional. This is represented in (27):

(27) [she wants]$_1$ if$_7$ Laura becomes a zombie MODAL$_7$ t_1 you to shoot her

We will call this the **high attitude** analysis.

At least at first glance, a low modal analysis seems more syntactically plausible than a high attitude analysis. Notice that in the LF in (27), *she wants* moves together to a position that outscopes the conditional. However, *she wants* is generally not taken to be a constituent at any level of representation, as illustrated in (28):

(28)



Thus, in order for a high attitude analysis to work, it seems that one of two things has to happen. The first possibility is that *she* and *wants* move to separate positions that

outscope the conditional, in which case it is unclear how the compositional semantics is meant to generate the correct interpretation.

(29) she$_1$ wants$_2$ if$_7$ Laura becomes a zombie MODAL$_7$ t_1 t_2 you to shoot her

The second is that *you to shoot her* moves from its original position, with the matrix verb phrase—including *she*, *wants*, and the embedded clause's trace—then undergoing so-called *remnant movement* to a position that outscopes the conditional:

(30) [she wants t_2]$_1$ if$_7$ Laura becomes a zombie MODAL$_7$ t_1 [you to shoot her]$_2$

But once again, it's not entirely clear how the semantics is supposed to take this structure and return the correct interpretation. Either way, regardless of which version of the embedded conditional analysis is most desirable on independent grounds, in the next section we will provide ample evidence that both the low modal and high attitude versions of embedded conditional analyses miss the mark.

3.2 Attitude restriction

We next turn to **attitude restriction** accounts of attitude conditionals. Recall from the discussion at the end of Section 2 that because attitudes are world-quantifiers (in the broader sense), they can theoretically be restricted by *if*-clauses in much the same way that modals can. This is in fact the analysis that Pasternak (2018) proposes for attitude conditionals. We will tweak Pasternak's account to fit with the compositional assumptions detailed in Section 2, though the problems discussed in Section 4 extend equally well to his formulation.

Suppose that ⟦want⟧ is defined as in (22), repeated below:

(22) $[\![\text{want}_n]\!]^{g,w} = \lambda p \lambda x.\ \forall w' \in \textsc{Best}_{x,w}(\textsc{Dox}_{x,w} \cap g(n))[p(w')]$

Suppose in addition that in an attitude conditional like (1), *if* is coindexed not with an embedded MODAL, as in the low modal version of the embedded conditional account, but rather with the attitude itself, as in (31):

(31) if$_7$ Laura becomes a zombie, she wants$_7$ you to shoot her

The interpretation generated by the LF in (31) can be seen in (32):

(32) $[\![(31)]\!]^{g,w} = 1$ iff $\forall w' \in \textsc{Best}_{x,w}(\textsc{Dox}_{x,w} \cap \textsc{zombie})[\textsc{shoot}(w')]$

We thus predict (1) to be true iff, among those belief worlds of Laura's in which she becomes a zombie, all of her most preferred ones are worlds in which you shoot her. An attitude restriction analysis therefore represents the most direct translation of the intuition put forward in the introduction that attitude conditionals make claims about *restricted attitudes*.

3.3 Double restriction

In both the embedded conditional and attitude restriction analyses, exactly one world-quantifying operator is restricted by the *if*-clause: MODAL for the embedded conditional approach, and the attitude verb for the attitude restriction approach. Now suppose instead that we wish to posit that attitude conditionals are simply conditionals with attitude verbs in their consequents (just like they appear on the surface), while at the same time generating the desired restricted attitude interpretation. In this case, one may wish to adopt a **double restriction** analysis, in which the *if*-clause restricts both a higher MODAL in the matrix clause—hence why attitude conditionals look like conditionals with attitudes in their consequents—and the attitude itself, generating a restricted reading much like attitude restriction accounts.[2]

One means of implementing this, which we will refer to as **direct double restriction**, involves double coindexation: the *if*-clause is coindexed with, and thus restricts, both the higher MODAL and the attitude verb. This is represented in (33):

(33) if$_7$ Laura becomes a zombie, MODAL$_7$ she wants$_7$ you to shoot her

While we are not aware of any direct double restriction analyses of attitude conditionals proposed in the literature, we include it on the grounds that it is a seemingly plausible analysis of attitude conditionals.

Another implementation, put forward by Blumberg & Holguín (2019), is what we will call **indirect double restriction**. While Blumberg & Holguín do not explicitly adopt the view that *if*-clauses compositionally restrict modal operators (and in particular MODAL), we will translate their analysis into the present framework; this translation has no bearing on the nature of our arguments against their analysis. In short, Blumberg & Holguín propose that in attitude conditionals, MODAL is the only operator that is directly restricted by the *if*-clause through compositional semantic principles. Thus, attitude conditionals well and truly are just conditionals with attitudes in their consequents, with no other special compositional mechanisms at play. However, attitude verbs can be restricted through other, pragmatic mechanisms. Suppose that for any proposition q, $[\![\text{want}_q]\!]$ is defined as follows:[3]

(34) $[\![\text{want}_q]\!]^w = \lambda p \lambda x.\ \forall w' \in \textsc{Best}_{x,w}(\text{Dox}_{x,w} \cap q)[p(w')]$

The LF for (1) will then be as in (35):

(35) if$_7$ Laura becomes a zombie, MODAL$_7$ she wants$_{\text{ZOMBIE}}$ you to shoot her.

2 In terms of precedent, Kaufmann & Schwager (2011) consider what is essentially a double restriction account of conditional imperatives (e.g., *If I become a zombie, shoot me!*). They do not put forward a specific view about the precise nature of the mechanisms restricting the imperative's world-quantifier, instead rejecting double restriction on grounds that are independent of this issue.

3 Blumberg & Holguín (2019) use Heim's (1992) more complex semantics for *want*, rather than von Fintel's (1999). This difference is immaterial for our purposes, and the same criticisms apply.

Thus, while MODAL is directly restricted by the *if*-clause, *want* is pragmatically restricted by the proposition denoted by the antecedent clause, *Laura becomes a zombie*. (In contexts where attitudes are apparently not restricted, they are actually vacuously restricted by the necessarily true proposition $\top$.) Notice that in spite of the differences in compositional mechanisms used, the interpretations of attitude conditionals generated by (33) and (35) are predicted to be the same, since both MODAL and *want* are restricted by ZOMBIE.

The next step in their analysis is determining how structure constrains the choice of restriction. This is motivated in part by certain apparent ordering asymmetries in this regard. For example, the attitude disjunction in (4), repeated below, seems to permit a reading in which *regret* is restricted by $-$OUTSIDE (true iff not a lot of people are outside), hence why a reading is permitted that relates the speaker's regrets when only considering possible states of affairs in which not a lot of people are outside.[4] Meanwhile, this is not the case for (36), which we will call a *reverse attitude disjunction*, differing from (4) only with respect to the order of disjuncts.

(4) Blumberg & Holguín 2019: p. 380:

Either a lot of people are on the deck outside, or I regret that I didn't bring more friends.

(36) Blumberg & Holguín 2019: p. 391:

?? Either I regret that I didn't bring more friends, or a lot of people are on the deck outside.

Blumberg & Holguín (2019) implement a dynamic analysis to account for these observations; we will not go through the details here, and will instead take for granted that in their analysis, the antecedent of a conditional is available as a restriction for an attitude in the consequent, and the negation of the first disjunct is available as a restriction for an attitude in the second disjunct, but not vice versa.

One final thing worth noting about Blumberg & Holguín's (2019) analysis is that they never actually provide a complete compositional semantics for attitude conditionals in their paper, focusing only on which restrictions are available for attitudes in which syntactic environments. However, from their discussion it is clear that they view attitude conditionals as "regular" conditionals with restricted attitudes in their consequents, meaning that in our terminology they adopt a double restriction account. An important issue they do not address is how, given that restricted *want* is in the consequent of the larger conditional, we end up with an interpretation about actual Laura's restricted desires, rather than zombie-Laura's restricted desires along the lines of a would-attitude conditional. Since this issue is not directly relevant to

4 Here and throughout, we use $-p$ to indicate p's intensional negation (i.e., the complement of p in the domain of possible worlds/situations), reserving $\neg$ for boolean negation.

our arguments against their analysis, we will set it aside and assume that by some means or other they derive truth conditions for simple attitude conditionals that are roughly akin to Pasternak's (2018) attitude restriction analysis. We will assume the same of direct double restriction accounts.

3.4 Summary

To summarize, in this section we have introduced three broad categories of analyses for attitude conditionals built on quantificational restriction: **embedded conditional**, **attitude restriction**, and **double restriction**. Embedded conditional analyses can be further subdivided into **low modal** and **high attitude**, while double restriction approaches can be subdivided into **direct** and **indirect** double restriction.

In the next section we will provide arguments against all five of these (sub)proposals. In going through our counterarguments, it will be helpful to have the proposed LF structures for all of these potential theories gathered in one place for ease of reference; this can be seen in (37).

(37) a. **Embedded conditional, low modal**
if$_7$ Laura becomes a zombie, she wants MODAL$_7$ you to shoot her

b. **Embedded conditional, high attitude**
if$_7$ Laura becomes a zombie, [she wants]$_1$ MODAL$_7$ t_1 you to shoot her

c. **Attitude restriction** (Pasternak 2018)
if$_7$ Laura becomes a zombie, she wants$_7$ you to shoot her

d. **Direct double restriction**
if$_7$ Laura becomes a zombie, MODAL$_7$ she wants$_1$ you to shoot her

e. **Indirect double restriction** (Blumberg & Holguín 2019)
if$_7$ Laura becomes a zombie, MODAL$_7$ she wants$_{\text{ZOMBIE}}$ you to shoot her

4 A panoply of problems

In this section we will discuss several empirical problems that attitude conditionals (and to a lesser extent attitude disjunctions) pose for the restriction-based accounts discussed in Section 3. While our observations vary in terms of which of these theories they pose problems for, the cumulative result will be that none of them come out unscathed, and each faces multiple crucial problems in successfully analyzing the semantic interpretation of attitude conditionals. In Section 5 we will propose a novel, non-restriction-based theory of attitude conditionals, would-attitude conditionals, and attitude disjunctions that seeks to resolve all of these issues.

4.1 *Contra* embedded conditionals: Strange readings

As mentioned previously, Drucker (2019) and Blumberg & Holguín (2019) argue at length against an embedded conditional account of attitude conditionals, using various kinds of evidence. We will not rehash most of their arguments here, for two reasons. First, our own evidence similarly points away from an embedded conditional analysis, so repeating their arguments introduces unnecessary redundancy. And second, several of the arguments they offer only pertain to certain versions of an embedded conditional analysis; for example, Blumberg & Holguín exclusively address a very specific version of a high attitude embedded conditional account, noting some syntactic-semantic difficulties faced by such an approach. Ideally, of course, we would prefer arguments that rule out embedded conditional analyses as a broad class, rather than much narrower arguments targeting specific versions of an embedded conditional analysis. Fortunately, the simplest argument that Drucker (2019) and Blumberg & Holguín (2019) offer against an embedded conditional account also happens to be the one that best generalizes across different versions: simply put, it predicts strange and inaccurate truth conditions in a wide variety of cases.

Rather than go through the full argument, we will simply skip straight to the most extreme examples, i.e., the ones in which the embedded conditional analysis clearly makes the most absurd predictions. Consider the following example from Blumberg & Holguín (2019: p. 383), based on similar examples from Drucker (2019):

(38) If Bill is on a plane to Cuba, then that surprises me.

Operating under the reasonable assumption that *that* in (38) is a propositional anaphor referring back to the antecedent of the conditional, any version of an embedded conditional analysis should predict this to mean the same thing as the absurd (39):

(39) It surprises me that if Bill is on a plane to Cuba, then Bill is on a plane to Cuba.

Of course, (38) is a perfectly reasonable thing to say, and unlike (39) it can be true without any surprisal about tautologies. As another example, consider (40):

(40) If Bill moved to Cuba, then I'm surprised he went through with it.

On an embedded conditional account, (40) should have the same meaning as (41):

(41) I'm surprised that if Bill moved to Cuba, then he went through with it.

But once again, these two sentences are obviously not truth-conditionally equivalent, for more or less the same reason: with (41) one seems to be claiming surprisal about a (near-)tautology, but not so with (40). This distinction is unexpected on an embedded conditional account.

While Drucker (2019) and Blumberg & Holguín (2019) discuss other cases where the inaccuracies in truth conditional predictions are a little more subtle, examples like (38) and (40) make it plenty clear that embedded conditional approaches simply do not make the right truth conditional predictions in a wide variety of cases.

4.2 Donkey anaphora and scope

Our next empirical conundrum comes from donkey anaphora. In short, we will provide evidence strongly suggesting that the *if*-clause in attitude conditionals restricts some operator that outscopes the relevant attitude. This is problematic for any version of an embedded conditional or attitude restriction account, since on the former the *if*-clause only restricts a modal that scopes *below* the attitude, while on the latter it exclusively restricts the attitude itself. Meanwhile, double restriction analyses are (at least in theory) compatible with this observation, since the *if*-clause restricts an instance of MODAL that outscopes the attitude.

Donkey sentences feature an indefinite DP that is interpreted in the restrictor of some quantificational operator, as well as a pronoun in the scope of this operator that is anaphoric to that indefinite. (42) is an example where the quantifier is the determiner *every*, with the indefinite DP *a donkey* appearing in *every*'s NP restrictor:

(42) $\underbrace{\text{Every}}_{\text{Q}}\ \underbrace{\text{farmer who owns [a donkey]}_1}_{\text{R}}\ \underbrace{\text{beats it}_1}_{\text{S}}.$

If an *if*-clause restricts a modal operator whose scope is the consequent, then one should be able to construct similar examples with conditionals: an indefinite takes scope in the antecedent, and a pronoun anaphoric to that indefinite appears in the consequent. This is indeed the case, as illustrated in (43):[5]

(43) a. If Smith owns [a donkey]$_1$, he beats it$_1$.

b. $\underbrace{\text{if}_2\ \text{Smith owns [a donkey]}_1}_{\text{R}}\ \underbrace{\text{MODAL}_2}_{\text{Q}}\ \underbrace{\text{he beats it}_1}_{\text{S}}$

Donkey anaphora is a relatively complex phenomenon, and we will not provide an analysis in this paper. But one particular empirical observation about donkey anaphora in conditional environments will prove useful to us: namely, that if an indefinite DP is interpreted in the restrictor (i.e., antecedent) of some world-quantifying operator, then any donkey pronoun anaphoric to that indefinite must be interpreted in the scope

5 Crucially, donkey anaphora is about the indefinite *taking scope* inside the restrictor, not just being pronounced there. Thus, the relevant reading for (43) is the one paraphrasable as "If there is a donkey that Smith owns, he beats it", and not one in which the indefinite outscopes the conditional altogether, paraphrasable as "There is a donkey such that if Smith owns it, he beats it". This extends to all other examples of donkey anaphora in this paper.

(i.e., consequent) of that same world-quantifying operator.[6] For example, in (43), *it* must be in the scope of MODAL, since *a donkey* is in its restrictor. This observation will in turn allow us to use donkey anaphora as a litmus test for the scope of the world-quantifying operators restricted by *if*-clauses.

We first show that this observation actually holds. To do this we will use examples containing *just like*, such as (44):

(44) Leigh$_1$ is a genius, just like her$_1$ mother.

Sentences with *just like* can give rise to ambiguities of the sort seen in (45):

(45) I said that Leigh$_1$ is a genius, just like her$_1$ mother.

The two potential paraphrases for (45) can be seen in (46):

(46) a. I said that Leigh is like her mother in that she is a genius.
b. I am like Leigh's mother in that I said that Leigh is a genius.

Suppose that this ambiguity is a matter of scope: the narrow scope reading paraphrased in (46a) features *just like her mother* inside the embedded clause, while the wide scope reading paraphrased in (46b) involves higher adjunction in the matrix clause. Evidence in favor of this view comes from non-sentence-final uses of *just like*. Thus (47a), in which *just like her mother* can only be in the embedded clause, must have a narrow scope reading, while (47b), in which *just like her mother* must be in the matrix clause, only has a wide scope interpretation:

(47) a. I said that Leigh$_1$, just like her$_1$ mother, is a genius.
b. I, just like Leigh$_1$'s mother, said that she$_1$ is a genius.

Now consider what happens with sentence-final *just like* when the embedded clause is a donkey sentence, as in (48):

(48) I said that if [a child]$_1$ can solve this puzzle then she$_1$ is a genius, just like her$_1$ mother.

Here the ambiguity observed in (45) evaporates, and only a narrow scope reading is available: the child is like the mother in being a genius. This is readily explained if we adopt the previously mentioned constraint on donkey anaphora. The indefinite *a child* is interpreted in the antecedent within the embedded clause, meaning that it is inside the restrictor of an embedded MODAL:

6 This ignores so-called *modal subordination* (Roberts 1987, 1989, Stone 1997), such as (i):

(i) If [a zombie]$_1$ enters, you must shoot it$_1$. Otherwise, it$_1$ will eat me.

Clearly, the second *it* is in neither the restrictor nor the scope of *must*, and yet it can be anaphoric to *a zombie*. However, by all appearances the important examples to be discussed shortly are not cases of modal subordination—for example, several lack any additional modal like *will* in (i)—so modal subordination likely will not save an embedded conditional or attitude restriction account.

(49) I said that $\underbrace{\text{if}_7\ [\text{a child}]_1\ \text{can solve this puzzle}}_{\text{R}}\ \underbrace{\text{MODAL}_7}_{\text{Q}}\ \underbrace{\text{she}_1\ \text{is a genius}}_{\text{S}}$

As per the constraint on donkey anaphora, this means that any pronoun anaphoric to *a child* must be interpreted within the scope of MODAL, so *a fortiriori* it must be interpreted within the embedded clause. Since *her* in *her mother* is anaphoric to *a child*, and since a wide scope reading of (48) would be the result of *just like her mother* adjoining outside of the embedded clause, this reading is unavailable, and the ambiguity disappears. However, once the *just like* phrase no longer contains a pronoun that is anaphoric to *a child*, the ambiguity returns again, indicating that the lack of ambiguity is due specifically to the scope constraint on donkey anaphora:

(50) I said that if [a child]$_1$ can solve this puzzle then she$_1$ is a genius, just like the principal.

Thus it appears that our constraint on donkey anaphora is correct: if an indefinite takes scope inside an *if*-clause, then any donkey pronoun anaphoric to that indefinite must take scope below whatever operator that *if*-clause restricts. But what does this have to do with attitude conditionals? As it turns out, when an indefinite appears in the antecedent of an attitude conditional, a pronoun that outscopes the attitude can be anaphoric to that indefinite. Here is a first example:

> *Laura's desire is that you shoot her if she becomes a zombie and attacks a civilian. This is rooted in a belief that having an extra person around could be helpful to you. She does not care what your own reasons would be for shooting her.*

(51) If Laura becomes a zombie and attacks [a civilian]$_1$, then because he$_1$ could help you she wants you to shoot her.

Based on the context, the interpretation of (51) is not one in which the *because* clause somehow scopes below the attitude: Laura does not care what the addressee's motivations would be for shooting her. Rather, the *because* clause specifies the reason for Laura's desire, meaning that it is a matrix VP (or higher) modifier, and thus it is adjoined above the verb. (52), interpreted in the same context as (51), is similar:

(52) If Laura becomes a zombie and attacks [a civilian]$_1$, then for his$_1$ sake she wants you to shoot her.

Once again, Laura does not care about the addressee's motivations for shooting her, and her desire is what is for the sake of the hypothetical civilian, meaning that *for his sake* is once again an adjunct that outscopes the attitude.

The first two examples involve adjuncts that seemingly sit relatively low in the syntactic tree. But we can do the same thing with modifiers adjoined higher up the clausal spine, such as so-called "speaker-oriented" adverbials like *unfortunately*:[7]

> *Laura is the clear leader of the group: what she says, goes. She is also selfish and believes that a cure for zombie-ism will be developed soon, so no matter what she wants to be kept alive at all costs.*

(53) If Laura becomes a zombie and attacks [a civilian]$_1$, then unfortunately for him$_1$ she wants you to let her eat him$_1$.

For (53) it needn't be the case that Laura *wants* things to be unfortunate for her potential victim. In fact, she needn't even *think* it will be unfortunate for him: she might be a zealot who thinks it is in her victim's best interests to let her eat him. Rather, *unfortunately for him* indicates the *speaker's* belief that Laura's desire to be saved at all costs is unfortunate for her hypothetical victim. In other words, *unfortunately for him* does not somehow take scope below the attitude, but as with the other cases takes scope in the matrix clause, above *want*.

For an especially extreme case, consider (54):

(54) If one zombie attacks Laura and another attacks [a civilian]$_1$, then he$_1$ is the person that she wants you to save.

The anaphoric pronoun in (54) cannot possibly be interpreted in the scope of *want*: no upward movement or reconstruction could place the pronoun inside of the other DP (*the person*. . .), into an embedded clause in a relative clause that has no open arguments for *he* to saturate. But once again, the pronoun can be anaphoric to an indefinite in the antecedent of the attitude conditional.

Based on our observation about donkey anaphora in conditionals, the fact that donkey pronouns can outscope the attitude in attitude conditionals indicates that the *if*-clause must restrict some world-quantifying operator that scopes above the attitude. This runs directly counter to embedded conditional and attitude restriction accounts, since the former claim that the *if*-clause restricts a modal below the attitude, while the latter claims that the *if*-clause restricts only the attitude itself. Double restriction accounts, meanwhile, are potentially compatible with the observations in (51–54), since they posit that the *if*-clause restricts a higher MODAL that outscopes the attitude, in addition to the attitude itself.

7 For evidence that such speaker-oriented adverbials are merged at a relatively high position along the clausal spine, see Jackendoff 1972, Cinque 1999.

4.3 No counterfactual attitude conditionals with *wish* and *regret*

An important observation about attitude conditionals, first noted by Pasternak (2018), is that counterfactual attitude conditionals with *wish* and *regret* are deviant:

(55) a. # If Laura had become a zombie, she wishes you had shot her.
b. # If Laura had become a zombie, she regrets that you didn't shoot her.

By all appearances embedded conditional analyses seem to predict this fact, as evidenced by the oddness of the sentences in (56):

(56) a. # Laura wishes that if she had become a zombie, you had shot her.
b. # Laura regrets that if she had become a zombie, you didn't shoot her.

But as things currently stand, theories in which attitude conditionals involve restriction of the attitude itself—that is, attitude restriction and both direct and indirect double restriction—do not capture this observation. We will frame this discussion in terms of attitude restriction, but the same point extends to double restriction.

To see why these accounts fail to predict the unacceptability of the sentences in (55), we first must give a semantics for counterfactual attitudes like *wish* and *regret*. In our von Fintel (1999)-style semantics, the denotation for *want* takes the experiencer's belief worlds and universally quantifies over the ones that best meet the experiencer's bouletic standards. The transition from *want* to *wish* can be plausibly framed as follows: if *α wants p* is a claim about *α*'s desires relative to what they consider to be *currently* live possibilities, then *α wishes p* is a claim about *α*'s desires relative to what were *previously* live possibilities. Therefore, while we still use $\textsc{Best}_{x,w}$—we still care about the experiencer's ideal worlds with respect to some set of possible states of affairs—$\text{Dox}_{x,w}$ will be replaced with $\text{Dox}^{+}_{x,w}$, which is an expanded domain of possible worlds. While $\text{Dox}_{x,w}$ is the set of worlds compatible with what x currently believes to be possible, $\text{Dox}^{+}_{x,w}$ is the set of worlds compatible with what x believes to have previously been possible. Since anything that is currently possible was also previously possible—the passage of time closes doors, but does not open them—we can reasonably infer that $\text{Dox}^{+}_{x,w} \supset \text{Dox}_{x,w}$, though this needn't actually be the case in order for our argument to work.

With this in mind, ⟦wish⟧ can be defined as in (57). Note that in addition to the switch from Dox to Dox^{+}, a presupposition is added: *α wishes p* presupposes that *α* believes *p* to be false. That is, *Sandra wishes she had gone to the party* presupposes that Sandra believes she did not go to the party.

(57) $[\![\text{wish}]\!]^{w} = \lambda p \lambda x : \text{Dox}_{x,w} \cap p = \emptyset.\ \forall w' \in \textsc{Best}_{x,w}(\text{Dox}^{+}_{x,w})[p(w')]$

Naturally, if we would like to allow *wish* to be restricted by an *if*-clause, we can make the same revision as was made for *want*:[8]

(58) $[\![\text{wish}_n]\!]^{g,w} = \lambda p \lambda x : \text{Dox}_{x,w} \cap p = \emptyset.\ \forall w' \in \text{Best}_{x,w}(\text{Dox}^{+}_{x,w} \cap g(n))[p(w')]$

Therefore, on an attitude restriction account in which (55a) has the LF in (59a), the predicted interpretation is as in (59b):

(59) a. if$_7$ Laura had become a zombie, she wishes$_7$ you had shot her

b. $[\![(59a)]\!]^{g,w}$ is defined iff $\text{Dox}_{\text{laura},w} \cap \textsc{shoot} = \emptyset$. Where defined, $[\![(59a)]\!]^{g,w} = 1$ iff $\forall w' \in \text{Best}_{\text{laura},w}(\text{Dox}^{+}_{\text{laura},w} \cap \textsc{zombie})[\textsc{shoot}(w')]$

This interpretation is defined if Laura believes that you did not actually shoot her. As for the assertive content, recall that $\text{Dox}^{+}_{x,w}$ is the set of worlds that x believes (in w) were live options at some previous point. Thus, $\text{Dox}^{+}_{\text{laura},w}$ will contain not only worlds compatible with Laura's current beliefs, in which she never became a zombie and was never shot, but also worlds in which she did become a zombie and was shot, as well as worlds in which she became a zombie and was not shot. Now suppose she bouletically ranks worlds in the following fashion: the best are worlds in which she never became a zombie (and was not shot), the next best are worlds in which she became a zombie and was shot, and the worst are worlds in which she became a zombie and was not shot. Naturally, $\text{Dox}^{+}_{\text{laura},w} \cap \textsc{zombie}$ restricts $\text{Dox}^{+}_{\text{laura},w}$ to those worlds in which she became a zombie, and the best of these by Laura's standards are all worlds in which you shot her. An account based on attitude restriction thus predicts (55a) to be true in the context provided.

To summarize, an account of attitude conditionals built on restricting the attitude's world-quantification predicts (55a) to be true iff among those states of affairs that Laura thinks were previously live possibilities, the best ones in which she became a zombie are ones in which you subsequently shot her. This is a perfectly reasonable interpretation, so a restriction-based account seemingly fails to predict the infelicity of (55a). The same line of reasoning extends to the counterfactual *regret* conditional in (55b). Generally speaking, *α regrets p* is treated as more or less semantically identical to *α wishes* $\neg p$: the former presupposes that α believes p to be true (rather than false for *wish*), and asserts that the best worlds in the quantificational domain are worlds in which p is false (rather than true for *wish*):

(60) a. $[\![\text{regret}]\!]^{w} = \lambda p \lambda x : \text{Dox}_{x,w} \subseteq p.\ \forall w' \in \text{Best}_{x,w}(\text{Dox}^{+}_{x,w})[\neg p(w')]$

b. $[\![\text{regret}_n]\!]^{g,w} = \lambda p \lambda x : \text{Dox}_{x,w} \subseteq p.\ \forall w' \in \text{Best}_{x,w}(\text{Dox}^{+}_{x,w} \cap g(n))[\neg p(w')]$

8 One may wish to replace $\text{Dox}_{x,w} \cap p = \emptyset$ in (58) with $\text{Dox}_{x,w} \cap g(n) \cap p = \emptyset$, thus imposing the antecedent's restriction in both the assertive content and presupposition. This will have no effect on our point: either way, we wrongly predict the well-formedness of counterfactual attitude conditionals.

We thus clearly expect (55b) to have the same well-formed truth conditions as (55a), again contrary to fact.

Pasternak (2018) offers a relatively simple explanation for the ill-formedness of (55) within an attitude restriction framework, though his proposal may extend to double restriction approaches as well: these sentences are ill-formed simply because *wish* and *regret* are lexically incapable of restriction by *if*-clauses. Thus, *wish* can have the interpretation in (57) but not (58), and *regret* can have the interpretation in (60a) but not (60b). The fact that restricted readings do not arise in (55) then immediately falls out, and the fact that they are ill-formed and not simply false can presumably be traced to the semantically vacuous contribution of the *if*-clause.

However, Pasternak's proposal operates on the faulty assumption that because counterfactual attitude conditionals are ill-formed with *wish* and *regret*, it must be the case that *all* attitude conditionals are ill-formed with these verbs. But non-counterfactual attitude conditionals with *wish* and *regret* are completely fine, as illustrated in (61):

(61) I have no clue if it's raining or not, but if it is. . .
  a. . . . then I wish I had brought my umbrella.
  b. . . . then I regret leaving my umbrella at home.

We thus conclude that what (55) illustrates is not a general problem of attitude conditionals with *wish* and *regret*, but rather a problem that is specific to *counterfactual* attitude conditionals with these verbs.

Finally, one might wish to claim that the problem with counterfactual attitude conditionals is syntactic in nature, rather than semantic or pragmatic: whatever structural configuration generates an antecedent with a counterfactual-looking surface structure is incompatible with a consequent whose verb is in the simple present. That a syntactic account will not suffice can be seen through examples where the antecedent is expressed by *in that case* instead of an *if*-clause. As the would-attitude conditionals in (62) illustrate, *in that case* conditionals are compatible with both non-counterfactual and counterfactual interpretations:

(62) a. Laura might become a zombie. In that case, she will want to eat brains.
  b. Laura never became a zombie, but she could have. In that case, she would have wanted to eat brains.

But while restricted attitude readings can arise with non-counterfactual uses of *in that case*, as illustrated in (63a), with counterfactual interpretations restricted readings still do not arise, as evidenced by the oddness of (63b):

(63) a. Laura might become a zombie. In that case, she wants you to shoot her.

b. # Laura never became a zombie, but she could have. In that case, she {wishes you had shot her/regrets that you didn't shoot her}.

By all appearances, there is no relevant structural distinction between the second clauses in (63a) and (63b): the difference lies in whether *in that case* is interpreted as counterfactual or non-counterfactual in nature. Thus, the unacceptability of counterfactual attitude conditionals seems to be due to semantic or pragmatic deviance, rather than syntactic ill-formedness.

4.4 Relevance of real-world entailments

Another empirical problem posed by attitude conditionals comes from the ways in which real-world circumstances can affect truth conditions. To see what we mean, consider (64) in the context provided:

> *The department chair wishes to pursue a more aggressive anti-cheating policy, so she has asked the faculty to report any student misconduct. Prof. Johnston has scheduled his exam, and (unbeknownst to the chair) has decided that the students are not allowed to collaborate on it. He tells his TA:*

(64) If any students work together on the exam, the chair wants us to file a report.

Importantly, there is a reading of (64) on which it is true in the scenario provided, in spite of the fact that the chair does not know that working together constitutes cheating. This poses an apparent problem for all of the analyses of attitude conditionals discussed in the previous section, as they place the truth conditions for attitude conditionals "too far inside the head" of the attitude holder. We will once again illustrate using an attitude restriction analysis, though the same reasoning extends to the other accounts discussed in Section 3.

On an attitude restriction account, (64) will have the LF in (65a), leading to the truth conditions in (65b):

(65) a. if$_7$ any students work together on the exam, the chair wants$_7$ us to file a report

b. $[\![(65a)]\!]^w = 1$ iff $\forall w' \in \textsc{Best}_{\text{chair},w}(\text{Dox}_{\text{chair},w} \cap \textsc{collab})[\textsc{report}(w')]$

In other words, an attitude restriction account predicts (64) to be true iff among those belief worlds of the chair's in which students collaborate on Prof. Johnston's exam, all of her most favored ones are worlds in which a report is filed. But in this case we predict (64) to be false in the context provided, contrary to fact. After all, the chair does not know whether Prof. Johnston has disallowed collaboration on his exam, meaning that $\text{Dox}_{\text{chair},w} \cap \textsc{collab}$ will contain a mixture of worlds in which students collaborate and this is disallowed, and worlds in which students collaborate

and working together is fine. Therefore, her bouletically ideal worlds in this domain will consist of a mixture of worlds in which collaboration is fine and no report is filed, and worlds in which collaboration is disallowed and a report is filed. Thus, not all of the chair's bouletically ideal COLLAB-belief worlds are worlds in which a report is filed, so (64) is wrongly predicted to be false.

In short, then, the problem (64) poses is as follows. The interlocutors know that working together on the exam entails cheating on the exam. However, the chair does not know this, and as far as the accounts detailed in the previous section are concerned, the chair's beliefs are all that matters; the only relevant entailments are the ones that the attitude holder considers true, and not the ones that actually hold.

So what can be done about this conundrum? So far as we can tell, the only solution for the theories at hand is to posit that by some means or other, the antecedent of the conditional can be replaced with a contextually equivalent proposition, similar to ideas proposed by Schwager (2011) and Sudo (2014) regarding certain apparent *de re* readings of predicates. For example, if COLLAB_BAD is the proposition true in all and only those worlds in which collaboration has been disallowed, then since the context entails COLLAB_BAD any proposition p is contextually equivalent to $p \cap$ COLLAB_BAD. Thus, COLLAB is contextually equivalent to COLLAB $\cap$ COLLAB_BAD, and the interpretation in (65b) can be replaced with (66):

(66) $[\![(65a)]\!]^w = 1$ iff
$\forall w' \in \textsc{Best}_{\text{chair},w}(\textsc{Dox}_{\text{chair},w} \cap \textsc{collab} \cap \textsc{collab\_bad})[\textsc{report}(w')]$

Or, put another way, we expect (64) in the context provided to have an interpretation more or less as in (67):

(67) If any students work together on the exam *and working together is disallowed*, the chair wants us to file a report.

We thus correctly predict truth here: among those belief worlds of the chair's in which students collaborate and collaboration is disallowed, her favored worlds are all ones in which a report is filed.

Of course, such an analysis begs the question of what the grammatical mechanisms would look like that allow the replacement of a conditional antecedent with a contextually equivalent proposition. But more importantly, this approach runs into its own empirical problems. To see why, consider (68):

> *Jo the TA knows that the chair wants any misconduct reported, and that the chair has no clue if Prof. Johnston has disallowed collaboration on the exam. However, Jo is operating under the mistaken belief that Prof. Johnston has disallowed collaboration, when he has in fact decided to allow it.*

(68) Jo mistakenly believes that if any students collaborate on the exam, then the chair wants us to file a report.

In this context, (68) is clearly true: while Jo has a firm understanding of the chair's desires, her misunderstanding of the circumstances surrounding the exam have led to an inaccurate belief. However, we will see that an analysis of the effects of real-world entailments on attitude conditionals that is built around swapping in contextually equivalent antecedents does not make the correct predictions here.

First, let us see what goes wrong if we don't do any swapping out of antecedents at all: the embedded antecedent is simply interpreted as COLLAB. In this case, the denotation of the embedded clause in (68) will be (69):

(69) $\lambda w.\ \forall w' \in \textsc{Best}_{\text{chair},w}(\textsc{Dox}_{\text{chair},w} \cap \textsc{collab})[\textsc{report}(w')]$

The sentence in (68) can then be thought of as asserting two things: (I) Jo believes the proposition in (69) (the *belief claim*), and (II) the proposition in (69) is false in the actual world (the *inaccuracy claim*). The good news is that if the embedded clause denotes (69), then the inaccuracy claim of (68) is correct, since it is not the case that the chair's ideal COLLAB-belief worlds are all REPORT worlds, for the reasons stated above: the chair does not know whether collaboration is cheating. But unfortunately, if the embedded clause denotes (69) then the belief claim is false: Jo knows that the chair has no clue if collaboration has been disallowed, and so Jo knows that the proposition in (69) is false. After all, Jo's mistake is not a misunderstanding of the chair's beliefs and desires, but rather a misunderstanding of the actual rules surrounding the exam. We thus wrongly predict (68) to be false.

Now let's see what happens when we replace the antecedent with a contextually equivalent proposition. In determining whether two propositions are contextually equivalent we of course must choose a context, and for (68) we actually have two to choose from. The first is the *global* context, i.e., the common ground of the interlocutors, while the second is the *local* context, Jo's beliefs. But no matter which we choose, it seems that we can't capture the truth conditions of (68). Say we go with the global context. Since we the interlocutors know that Prof. Johnston has in fact allowed collaboration on the exam, COLLAB is equivalent to COLLAB ∩ COLLAB_FINE relative to the global context. If we swap this in, the interpretation of the embedded clause post-swapping will be (70) instead of (69):

(70) $\lambda w.\ \forall w' \in \textsc{Best}_{\text{chair},w}(\textsc{Dox}_{\text{chair},w} \cap \textsc{collab} \cap \textsc{collab\_fine})[\textsc{report}(w')]$

Much like with (69), this makes the right predictions with respect to (68)'s inaccuracy claim, since the proposition in (70) is clearly false: it is not the case that the chair wants a report to be filed if students collaborate and collaboration is fine. However, again like (69), it is also clear that Jo does not believe this proposition: Jo knows that the chair does not want spurious reports to be filed. We thus have the same problem as for (69): correct predictions about the inaccuracy claim, and incorrect predictions about the belief claim.

More interesting is the case in which we utilize the local context of Jo's beliefs. Since Jo mistakenly believes that collaboration has been disallowed, when we use the local context to determine contextual equivalence we can use the same trick as for (64): we can replace COLLAB with COLLAB ∩ COLLAB_BAD, since these propositions are equivalent in the context of Jo's beliefs. As a result, the embedded clause in (68) will have the denotation in (71):

(71) $\lambda w.\ \forall w' \in \textsc{Best}_{\text{chair},w}(\textsc{Dox}_{\text{chair},w} \cap \textsc{collab} \cap \textsc{collab\_bad})[\textsc{report}(w')]$

We now finally have as the denotation of the embedded clause a proposition that Jo genuinely believes is true. But here's the problem: this proposition *actually is* true. The chair's desires, restricted to possible states of affairs in which students collaborate and collaboration is disallowed, really *are* such that a report is filed. Thus, we once again wrongly predict (68) to be false, but this time for the opposite reason: the belief claim is true, but the inaccuracy claim is false. What's more, we wrongly predict (72) to be true: Jo believes the proposition denoted by the embedded clause, and that proposition is true in the real world.

(72) Jo correctly believes that if any students collaborate on the exam, then the chair wants us to file a report.

In summary, the effects of real world entailments on the truth conditions of attitude conditionals like (64) pose major problems for all of the theories outlined in the previous section, on the grounds that those theories put too much of the truth conditional burden of attitude conditionals inside the head of the attitude holder. Furthermore, if we try to get out of this problem by allowing antecedents to be replaced with contextually equivalent propositions, we make the wrong predictions about the truth conditions of embedded attitude conditionals such as (68).

4.5 Summary

In this section we have gone over a wide variety of problematic data that cumulatively rule out all of the analyses discussed in Section 3: embedded conditional analyses make strange truth-conditional predictions in several cases, the donkey anaphora facts run counter to embedded conditional and attitude restriction accounts, the lack of counterfactual attitude conditionals was problematic for every theory *except* embedded conditional, and the relevance of real-world entailments was problematic for all theories. These findings are summarized as a table in Fig. 1.

Having thoroughly problematized attitude conditionals and attitude disjunctions, we next turn to the task of providing an analysis of these constructions—as well as would-attitude conditionals—that overcomes these problems.

	no strange truth conditions	donkey anaphora	no counterfactual attitude conditionals	effects of real-world circumstances
emb. cond.	✗	✗	✓	✗
att. restr.	✓	✗	✗	✗
double restr.	✓	✓	✗	✗

Figure 1 Successes and difficulties for each candidate theory.

5 Faking restricted readings: A new approach

All of the analyses detailed and argued against in Sections 3 and 4 share one feature in common: the *if*-clause restricts, directly or indirectly, the world-quantification of either a modal operator in the scope of the attitude (embedded conditional), or the attitude itself (attitude and double restriction). The restricted reading that arises is an immediate result of this world-quantificational restriction effected by the *if*-clause. Blumberg & Holguín's (2019) analysis extends this principle to attitude disjunctions, but the core mechanism is the same: restricting the quantification of the attitude verb.

In this section we will propose an alternative according to which there is no quantificational restriction of the attitude whatsoever, nor of any modal that scopes below the attitude. Thus, in attitude conditionals, the only operator that the *if*-clause restricts—either directly or indirectly—is a wide-scope matrix-clause MODAL. In other words, structurally speaking, attitude conditionals are simply conditionals with attitudes in their consequents, and attitude disjunctions are simply regular disjunctions. The apparent restricted readings that arise in these constructions are then the result of other semantic mechanisms.

We start in Section 5.1 by refining the basic analysis of attitude semantics, adopting the view that attitudes are about situations (Barwise & Perry 1983, Kratzer 2002), where a situation is a partial possible world. In Section 5.2 we then turn to attitude conditionals, showing how a relatively simple situation-semantic analysis of conditionals, appropriately intertwined with our new situation-sensitive semantics of attitudes, generates the desired readings, while at the same time avoiding the pitfalls of restriction-based theories. In Section 5.3 we discuss would-attitude conditionals, deriving the differences in interpretation between attitude and would-attitude conditionals, as well as offering some tentative candidate explanations for why would-attitude conditionals should sometimes give rise to what look like restricted readings. In Section 5.4 we discuss attitude disjunctions, and in Section 5.5 we tentatively discuss the possibility of extending our fake restriction account to other, non-attitude conditionals.

5.1 Attitudes are about situations

We build our analysis on the premise that attitudes are in some sense about situations. We take situations to be partial possible worlds, or conversely, possible worlds are maximal situations. For situations s and s', $s \sqsubseteq s'$ iff s is a part of s', and $s \sqsubset s'$ iff s is a proper part of s' (i.e., $s \sqsubseteq s'$ and $s \neq s'$).

So how are these "about-situations" integrated into the lexical semantics of attitude verbs, and into the compositional semantics of clauses containing attitude ascriptions? On the lexical side, we will posit that attitude verbs simply have an extra argument that is saturated by their about-situation. We will also take the situation of evaluation—that is, the situation in which the wanting (or believing, wishing, etc.) is taking place—to be an argument of the verb. Thus, a verb like *want* takes four arguments, in the following order: (I) a proposition, (II) an about-situation, (III) an experiencer, and (IV) a situation of evaluation. This is indicated in (73):

(73) $[\![\text{want}]\!] = \lambda p \lambda s \lambda x \lambda s'.\ x$ wants p about s in s'

Intuitively, for x to want p about s in s' is for x's cognitive state in s' to be roughly as follows: of those ways in which x thinks that s could further develop, x prefers that s develop into a situation in which p is true. Thus, if s is a situation in which my favored sports team is playing, one could say that I want of s that my team win, meaning that I would prefer for s to develop into a situation in which my team wins. Of course, this is a very simple and imprecise analysis of the semantics of *want*, stripped of any quantification over possible worlds or situations. However, since attitudinal quantification goes unrestricted by any operators, for our purposes there is really no need to be any more precise than this, so long as we maintain the strong intuition that to want p of s is to in some sense want s to "become" a p-situation.

We next turn to the issue of how the denotations of attitude reports compose. We will use a relatively stripped-down syntax, keeping syntactic complexity to the bare minimum required to get our analysis off the ground. With this in mind, we take the LF of a simple attitude report like *Rivka wants you to leave* to be as in Fig. 2. The embedded clause *you to leave* denotes the proposition LEAVE. ABT is a situation-denoting pronoun that will provide the about-situation argument; it seems reasonable to suppose that ABT could be either deictic or bound by some other operator, though we will always treat it as bound. The head X above the subject *Rivka* represents the cluster of TAM (Tense-Aspect-Mood) heads higher in the clause. While TAM heads of course serve a variety of semantic purposes, for us most of those purposes are irrelevant. However, we take some combination of these TAM heads to provide the situation of evaluation for *want*; we will see that this is useful in explaining the difference between attitude and would-attitude conditionals. Finally, we follow Percus (2000) and others in positing the existence of a node λs_1 in the left periphery that binds the various situation arguments, thereby generating a final

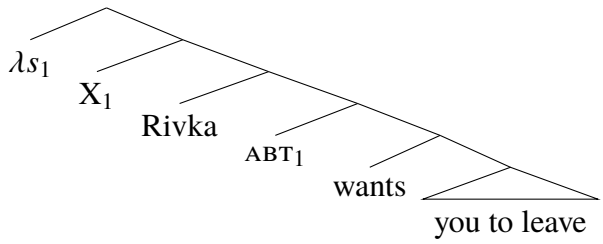


Figure 2 LF representation for *Rivka wants you to leave*.

interpretation that is a proposition (the characteristic function of a set of situations). When we get to more complex clauses, we will see that there may be multiple situation variable binders, with different options existing for which binders bind which situation variables.

Let's see what the LF in Fig. 2 gets us. First, ⟦want⟧ composes with ⟦you to leave⟧ in an unsurprising manner:

(74) ⟦want⟧(⟦you to leave⟧) = $\lambda s \lambda x \lambda s'$. x wants LEAVE about s in s'

ABT$_1$, being a (situation-denoting) pronoun, denotes a free variable: $[\![\text{ABT}_1]\!]^g$ denotes $g(1)$, where g is the variable assignment parameter. This naturally serves as the about-situation argument to *want*:

(75) $[\![\text{want you to leave}]\!]^g([\![\text{ABT}_1]\!]^g) = \lambda x \lambda s'$. x wants LEAVE about $g(1)$ in s'

This next composes with ⟦Rivka⟧, followed by $[\![X_1]\!]^g$; for simplicity's sake we take the latter to simply be $g(1)$, factoring out those ways in which the TAM field contributes to semantic interpretation outside of determining the situation of evaluation for the VP. Thus, the interpretation up to and including X is (76):

(76) $[\![X_1$ Rivka ABT$_1$ want you to leave$]\!]^g = 1$ iff
rivka wants LEAVE about $g(1)$ in $g(1)$

The lambda abstractor λs_1 then abstracts over index 1, leading to the final interpretation in (77), true of a situation s iff Rivka's desires in s about s are such that she leave:

(77) $[\![\lambda s_1$ X_1 Rivka ABT$_1$ want you to leave$]\!]^g =$
λs. rivka wants LEAVE about s in s

As mentioned above, the fact that the about-situation and the situation of evaluation are the same here is not a necessary feature of this analysis: it is equally plausible that ABT permits a deictic interpretation, in which case the ensuing proposition will be true of a situation s iff Rivka's desires in s about the contextually provided

situation are such that she leave. However, in our analyses of attitude conditionals, would-attitude conditionals, and attitude disjunctions, we will only make use of bound readings of about-pronouns and other situation variables.

5.2 Attitude conditionals

We next turn to our analysis of attitude conditionals. As mentioned previously, on a structural level we treat attitude conditionals simply as "regular" conditionals with attitude verbs in their consequents. But as is made clear in examples like (1), the attitudes of the experiencer's that we care about are her real ones: we care about Laura's actual desires, and not the desires of hypothetical zombie-Laura. This means that the attitude's situation of evaluation in attitude conditionals will be a situation in the actual world, not in the worlds quantified over by the conditional's modal. In other words, the attitude verb is essentially interpreted *de re*. Importantly, we will posit that it is possible to have actual attitudes toward non-actual situations, and that is what attitude conditionals convey: (1) constitutes a claim about Laura's actual desires toward non-actual situations in which she becomes a zombie.

In order to flesh this out more, we first need to revise our semantics for conditionals to suit our new situation-semantic framework. A natural way of formulating a conditional *if p, then q* in a situation-semantic framework is roughly as follows: in all of the most likely p-worlds w, the minimal p-situation in w extends to a situation in which q is true (cf. Berman 1987, Heim 1990, von Fintel 1994). This of course begs the question of what a minimal p-situation is, for a given proposition p. While attempts at formal definitions of this notion have been put forward, perhaps most thoroughly by Kratzer (2019), for our purposes it will suffice to stick to a relatively intuitive understanding of a minimal situation. Thus, a minimal ZOMBIE-situation includes, for example, Laura becoming a zombie, but not whether she is subsequently shot, or whether a cure for zombie-ism exists, or how she is treated post-zombification: it "just" includes her becoming a zombie.

With this in mind, we will redefine $[\![\textsc{modal}]\!]$ as in (78), where s is the situation of evaluation, $\textsc{MinSit}(p, w)$ is the minimal p-situation in w, and $g(n)$ will again end up being the antecedent of the conditional:

(78) $[\![\textsc{modal}_n]\!]^g = \lambda s \lambda q.\ \forall w \in \mathrm{ML}_s(g(n))[\exists s'[\textsc{MinSit}(g(n), w) \sqsubseteq s' \wedge q(s')]]$

As a result, we predict *if*$_7$ p *modal*$_7$ q to be true in a situation s iff the most likely p-worlds as evaluated in s are all such that the minimal p-situation expands into a q-situation.

We now have enough in place to put forward our analysis of attitude conditionals. Our proposed LF structure for (1) can be seen in Fig. 3. A few things are worth noting about this structure. The first is that there are now two situation-binding nodes, λs_1

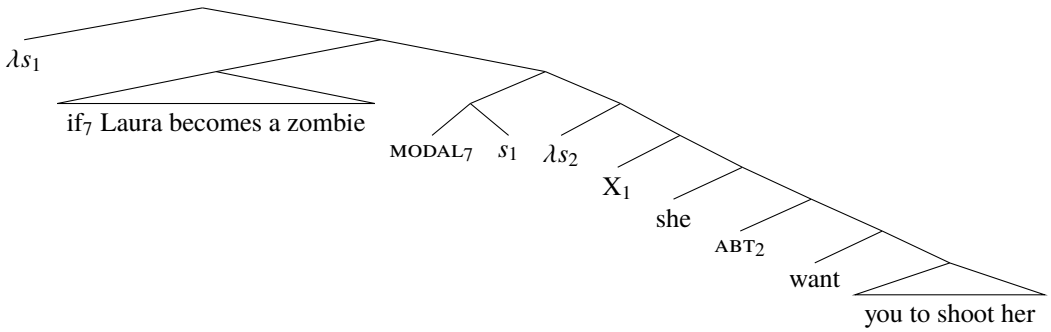


Figure 3 Proposed LF for *If Laura becomes a zombie, she wants you to shoot her*.

and λs_2, since one always appears in the left periphery, and one appears immediately below every intensional operator like MODAL.[9] Second, notice that X is coindexed with the highest lambda abstractor λs_1—*want* is interpreted *de re*—while ABT is coindexed with the lower lambda abstractor, since we care about Laura's desires toward non-actual situations in which she becomes a zombie. While we simply stipulate that X is coindexed with λs_1, it seems that a more promising route would be to postulate that whatever combination of tense, aspect, mood, and modal heads is responsible for the consequent's lack of conditional morphosyntax (we see *wants* and not *would want*) is responsible for *want*'s situation of evaluation being s_1 instead of s_2. Regardless of how this is actually to be captured, we will simply adopt the stipulation that in attitude conditionals X is coindexed with s_1 and not s_2.

We next turn to the compositional semantics. Up to and excluding the lower situation-abstractor λs_2, we get the interpretation in (79):

(79) $[\![\text{X}_1 \text{ she ABT}_2 \text{ want you to shoot her}]\!]^g = 1$ iff
laura wants SHOOT about $g(2)$ in $g(1)$

The next step is to lambda abstract over index 2 via λs_2:

(80) $[\![\lambda s_2 \text{ X}_1 \text{ she ABT}_2 \text{ want you to shoot her}]\!]^g$
$= \lambda s$. laura wants SHOOT about s in $g(1)$

This is then fed to MODAL (which has previously composed with the evaluation situation variable s_1), as seen in (81). Note that $g(7)$ will be a proposition (the denotation of the *if*-clause), while $g(1)$ will be a situation (lambda-abstracted by λs_1).

(81) $[\![\text{MODAL}_7\ s_1\ \lambda s_2\ \text{X}_1 \text{ she ABT}_2 \text{ want you to shoot her}]\!]^g = 1$ iff

9 Naturally, a third situation lambda-abstractor should appear below *want* for the same reason, but this does not concern us since we take the compositional semantics of the embedded clause for granted.

$$\forall w \in \mathrm{ML}_{g(1)}(g(7))[\exists s'[\textsc{MinSit}(g(7), w) \sqsubseteq s' \wedge$$
$$\text{laura wants } \textsc{shoot} \text{ about } s' \text{ in } g(1)]]$$

Finally, the *if*-clause replaces $g(7)$ with ZOMBIE, and λs_1 lambda-abstracts over index 1, leading to the final truth conditions in (82):

(82) $[\![(1)]\!] = \lambda s.\ \forall w \in \mathrm{ML}_s(\textsc{zombie})[\exists s'[\textsc{MinSit}(\textsc{zombie}, w) \sqsubseteq s' \wedge$
$\text{laura wants } \textsc{shoot} \text{ about } s' \text{ in } s]]$

We thus predict (1) to denote a proposition true of a situation s iff in all of the most likely worlds w from s in which Laura becomes a zombie, the minimal zombie-situation in w is contained in a situation about which Laura's actual desires are that you shoot her. If we accept that Laura can have desires about non-actual situations, then this seems to derive the correct truth conditions.

We will soon see that this account needs to be tweaked a bit, but in the meantime let's note a couple of positives about it straight off the bat. The first stems from our supposition, shared with double restriction analyses, that the *if*-clause in attitude conditionals restricts a modal that outscopes the attitude verb, with the difference being that for us MODAL is the *only* operator whose quantification is restricted. While a full analysis of donkey anaphora would send us too far afield for the purposes of this paper, an LF structure like that in Fig. 3 is fully compatible with donkey pronouns appearing in positions that outscope the attitude verb, unlike embedded conditional and attitude restriction analyses. After all, the constraint we observed with donkey anaphora was that a donkey pronoun must be in the scope of whatever operator the *if*-clause restricts; since the *if*-clause restricts an operator that outscopes the attitude verb, there should be no problem with a donkey pronoun outscoping the attitude.

Another benefit of the analysis as it currently stands is that it can account for the ways in which real-world circumstances can have an impact on the truth conditions of attitude conditionals, an observation that was problematic for other theories, as discussed in Section 4.4. Recall (64) in its associated context:

> *The department chair wishes to pursue a more aggressive anti-cheating policy, so she has asked the faculty to report any student misconduct. Prof. Johnston has scheduled his exam, and (unbeknownst to the chair) has decided that the students are not allowed to collaborate on it. He tells his TA:*

(64) If any students work together on the exam, the chair wants us to file a report.

Our analysis rightly predicts the existence of a true reading of (64) in the context provided. The denotation we assign to this sentence can be seen in (83):

(83) $[\![(1)]\!] = \lambda s.\ \forall w \in \mathrm{ML}_s(\textsc{collab})[\exists s'[\textsc{MinSit}(\textsc{collab}, w) \sqsubseteq s' \wedge$
$\text{the chair wants } \textsc{report} \text{ about } s' \text{ in } s]]$

That is, we predict truth in a given situation *s* iff in all of the most likely COLLAB-worlds from *s*, the minimal COLLAB-situation expands into a situation about which the chair wants a report to be filed. The context establishes that collaboration has been disallowed, so *a fortiriori* collaboration has been disallowed in each of the most likely worlds in which collaboration occurs. Thus, in each of the most likely worlds in which students collaborate, the minimal COLLAB-situation can be expanded into one in which students collaborate and collaboration has been established as disallowed, and the chair's desires about each of these situations are such that a report is filed. (64) is thus correctly predicted to be true, not through some recourse to swapping out antecedents for contextually equivalent propositions, but simply by virtue of the compositional semantics of attitude conditionals.

There is also no problem handling cases like (68), which was problematic on the antecedent-swapping account discussed above.

> *Jo the TA knows that the chair wants any misconduct reported, and that the chair has no clue if Prof. Johnston has disallowed collaboration on the exam. However, Jo is operating under the mistaken belief that Prof. Johnston has disallowed collaboration, when he has in fact decided to allow it.*

(68) Jo mistakenly believes that if any students collaborate on the exam, then the chair wants us to file a report.

If Jo is correct about the chair's desires, but incorrect about whether collaboration has been disallowed, then she indeed wrongly believes the proposition in (83) to be true. She incorrectly believes that all of the most likely collaboration-worlds are worlds in which collaboration has been disallowed, so she incorrectly believes that all of the minimal collaboration-situations can be extended to situations about which the chair wants a report to be filed. In terms of the discussion above, both the belief claim and the inaccuracy claim of (68) are rightly predicted to be true. Our analysis therefore correctly predicts the ways in which real-world circumstances can affect the truth conditions of both matrix-clause and embedded attitude conditionals.

However, additional constraints need to be placed on the analysis as it currently stands. To see why, consider (1) and (84) in the context provided:

> *A new treatment for zombie-ism has been developed, with the result being that if Laura becomes a zombie, she will still be able to lead a perfectly normal, happy existence. However, Laura does not know this: she still thinks that becoming a zombie is horrifying and terminal. She thus thinks that the best option upon her becoming a zombie is that she is shot.*

(1) If Laura becomes a zombie, she wants you to shoot her.

(84) If Laura becomes a zombie, she wants you to let her live.

Judgments are quite clear here that (1) is true in the context provided, while (84) is false. However, we currently seem to predict (84) to be true. The predicted truth conditions can be seen in (85):

(85) $[\![(85)]\!] = \lambda s.\ \forall w \in \mathrm{ML}_s(\textsc{zombie})[\exists s'[\textsc{MinSit}(\textsc{zombie}, w) \sqsubseteq s' \wedge$ laura wants LET_LIVE about s' in s]]

As per the context, for all of the most likely worlds w in which Laura becomes a zombie, the minimal zombie-situation extends to a situation in which Laura becomes a zombie, is cured, and lives a perfectly happy life. But surely Laura's desires about *this* situation would be that she is *not* shot, and that she is allowed to live. We thus apparently predict (84) to be true in the context provided, contrary to fact.

The problem can be paraphrased as follows. One of the benefits of framing things in terms of about-situations in maximally likely worlds, seen in our discussion of (64) and (68) above, was that the predicted truth conditions for attitude conditionals were somewhat divorced from the beliefs of the attitude's experiencer. This is why the chair's lack of knowledge that collaboration was disallowed did not get in the way of (64) being true. But (1) and (84) in the above context suggest that the truth conditions are currently *too* divorced from the experiencer's beliefs. We thus need to constrain the effects of real-world circumstances in a way that prevents non-existent true readings of (84), but that also continues to allow for the truth of (64) and (68).

We will capture this by imposing a requirement about what situations experiencers can and cannot have attitudes toward. More specifically, we will posit the following intuitively plausible principle: an individual can only have attitudes toward a situation s if s or some situation that they consider indistinguishable from s is compatible with that individual's beliefs. Or, put more informally, one only has desires about situations that one thinks are possible. To provide a more formally explicit definition of this principle, we start with the definitions in (86):

(86) a. $s \sim_{x,s''} s'$ iff s and s' are indistinguishable as far as x is concerned in s''.
b. $[s]_{x,s'} := \{w_{s''} \mid s'' \sim_{x,s'} s\}$, where $w_{s''}$ is the world containing s''.

For a given situation s, then, $[s]_{x,s'}$ will be the set of all worlds containing situations that, as far as x is concerned (in s'), are indistinguishable from s. Therefore, if $\mathrm{Dox}_{x,s'}$ is the set of worlds compatible with x's beliefs in s', then our new principle, which we call **Belief-Dependence of About-Situations**, states that in order for x to have an attitude about s in s', $[s]_{x,s'}$ and $\mathrm{Dox}_{x,s'}$ must have a non-empty intersection:[10]

10 Things get a bit more complicated when one considers that belief is itself an attitude, and thus must be about situations. We assume for simplicity's sake that a given individual has a set of all-things-considered belief worlds; for example, these might be the ones pertaining to that individual's beliefs about the world as a whole. There are presumably other, equally plausible ways of framing this.

(87) **Belief-Dependence of About-Situations**
For attitude verb V, individual x, proposition p, about-situation s and situation of evaluation s', $[\![\mathrm{V}]\!](p)(s)(x)(s')$ is defined only if $[s]_{x,s'} \cap \mathrm{Dox}_{x,s'} \neq \emptyset$

Note first that this principle does not prevent us from deriving the correct results with respect to the sentences in (64) and (68). Importantly, the chair is not sure whether collaboration has been disallowed for Prof. Johnston's exam. Thus, the existence of a minimal situation in which students collaborate and collaboration has been disallowed, while not entailed by the chair's beliefs, is compatible with her beliefs, meaning that she can have desires about such situations. These are precisely the situations that verify (64). Similar reasoning extends to the embedded attitude conditional case of (68). In other words, we still rightly predict that real-world circumstances can have an impact on the truth conditions of attitude conditionals.

However, Belief-Dependence of About-Situations does prevent us from erroneously ascribing truth to (84). Suppose that it is true that in all of the actually most likely worlds in which Laura becomes a zombie, she is subsequently cured of zombie-ism and leads a perfectly normal and happy life. However, Laura is unaware of this, and thinks that zombie-ism will inevitably lead to misery on the part of herself and/or others. In this case, situations in which Laura becomes a zombie and is subsequently cured are not compatible with her beliefs, meaning that these cannot be situations about which Laura has desires. Thus, in all of the most likely worlds in which Laura becomes a zombie, the maximal situation compatible with her beliefs that contains the minimal zombie-situation will not extend far past her simply becoming a zombie, and due to her misconceptions about zombie-ism Laura's desires about such situations will be that she is shot. So while we continue to correctly predict the impact of real-world circumstances on sentences like (64), we also rightly predict that (84) should be false and (1) true in the newly provided context.

Not only does Belief-Dependence of About-Situations appropriately constrain the ways in which actual circumstances affect the truth conditions of attitude conditionals, but it also points the way toward an account of the infelicity of counterfactual attitude conditionals with *wish* and *regret*, as well as the acceptability of indicative attitude conditionals with these same verbs:

(55) a. # If Laura had become a zombie, she wishes you had shot her.
b. # If Laura had become a zombie, she regrets that you didn't shoot her.

(61) I have no clue if it's raining or not, but if it is. . .
a. . . . then I wish I had brought my umbrella.
b. . . . then I regret leaving my umbrella at home.

We will go through the argument with *wish*, though the same point extends to *regret*. Suppose that to wish p about s presupposes believing not-p of s, and asserts that

one's counterfactual preferences would be to replace s with a p-situation (more or less). Once again this is relatively imprecise, but enough to illustrate how our account generates the right predictions.

In order to explain the ill-formedness of counterfactual attitude conditionals, we first need to have a basic semantics of counterfactuals. Recall that the distinction between *want* and *wish* was previously framed in terms of quantifying over worlds that are currently believed to be viable versus worlds that are believed to have previously been viable, with the result being that *wish*'s quantificational domain was a proper superset of *want*'s domain. The contrast between indicative and counterfactual conditionals can be framed similarly: while indicative conditionals quantify over currently circumstantially accessible worlds, counterfactual conditionals quantify over previously circumstantially accessible worlds, meaning that the domain for counterfactuals is a proper superset of the domain for indicative conditionals.[11] We will capture this in a relatively simple manner, by replacing the quantificational domain $\mathrm{ML}(A)$ with $\mathrm{ML}^{+}(A)$, the set of maximally likely counterfactual A-worlds.

With this in mind, the interpretation for (55a) is predicted to be as seen in (88):

(88) $[\![(55a)]\!] = \lambda s.\ \forall w \in \mathrm{ML}^{+}_{s}(\textsc{zombie})[\exists s'[\textsc{MinSit}(\textsc{zombie}, w) \sqsubseteq s' \wedge$ laura wishes $\textsc{shoot}$ about s' in $s]]$

In other words, the most likely counterfactual worlds in which Laura became a zombie are such that the minimal zombie-situation extends to one about which Laura wishes that you had shot her. However, Laura knows that she did not become a zombie, meaning that any situation in which she did become a zombie is incompatible with her beliefs. We thus necessarily run afoul of Belief-Dependence of About-Situations, since Laura cannot have attitudes toward situations in which she became a zombie. The ill-formedness of counterfactual attitude conditionals with *wish* and *regret* is rightly predicted.

Now consider the indicative (61a) in a context in which the speaker does not know whether it is raining. Since (61a) is an indicative and not a counterfactual conditional, we go from ML^{+} back to ML, predicting the truth conditions in (89):

(89) $[\![(61a)]\!] = \lambda s.\ \forall w \in \mathrm{ML}_{s}(\textsc{rain})[\exists s'[\textsc{MinSit}(\textsc{rain}, w) \sqsubseteq s' \wedge$ I wish $\textsc{umbrella}$ about s' in $s]]$

Since the speaker does not know whether or not it is raining, an interpretation like (89) does not violate Belief-Dependence of About-Situations. The predicted truth conditions are then that in all maximally likely $\textsc{rain}$ worlds w, the minimal $\textsc{rain}$ situation in w expands to a situation—namely, one that also includes the speaker's

11 For explicit discussion of this semantic parallel between the *want-wish* distinction and the indicative-counterfactual distinction, see Heim 1992. For similar discussion from a more morphosyntactically oriented perspective, see Iatridou 2000.

lack of an umbrella—about which the speaker wishes they had brought an umbrella. We thus have a potential explanation for why counterfactual attitude conditionals with *wish* and *regret* are ill-formed, but at the same time indicative conditionals with these attitudes are perfectly fine.

If Belief-Dependence of About-Situations is correct, the prediction seems to be that counterfactual attitude conditionals are ill-formed across the board. A tempting potential counterexample to this prediction would be sentences like (90). (A donkey pronoun outside of the scope of the attitude is included to show that this example cannot be reduced to an embedded conditional.)

(90) If one zombie had attacked Laura and another had attacked [a civilian]$_1$, then he$_1$ is the person that I think you would have saved.

If (90) is treated as a straightforward attitude conditional, by our account it should be ill-formed, since it requires that the speaker have beliefs about contrary-to-fact zombie attack situations. However, note the crucial inclusion of the embedded modal *would* beneath *think*. This modal is obligatory: (91) is ill-formed.[12]

(91) * If one zombie had attacked Laura and another had attacked [a civilian]$_1$, then he$_1$ is the person that I think you saved.

The fact that the modal is obligatory suggests an alternative analysis of (90). Suppose that the about-situation for *think* is not the situation introduced by the quantification of the matrix-clause counterfactual MODAL, but some actual situation—perhaps the one that the sentence as a whole is meant to characterize—in accordance with Belief-Dependence. That is, both the evaluation situation and the about-situation for *think* are real-world situations. However, the *if*-clause not only restricts the matrix-clause MODAL (hence the availability of donkey anaphora outside of the scope of *think*), but *would* as well, more or less akin to a double-restriction analysis:

(92) if$_7$ one zombie had attacked Laura and another had attacked [a civilian]$_1$, then MODAL$_7$ he$_1$ is the person that I think you would$_7$ have saved.

Thus, it would remain the case that counterfactual attitude conditionals are universally ill-formed, but sentences like (90) would not be attitude conditionals *per se*, despite sharing certain traits with attitude conditionals.

An alternative version of such an account could be framed as follows. Later we will suggest the possibility that fake restriction is a feature not only of attitude conditionals, but also of "regular" conditionals with overt modal operators. That is, modals, like attitudes, are about situations, and *If Laura becomes a zombie, you must*

12 (91) is well-formed on a distinct interpretation in which *had attacked* is a "regular" pluperfect in the antecedent of an epistemic indicative conditional, rather than the antecedent of a counterfactual. On this reading (91) reduces to a straightforward attitude conditional.

shoot her is a claim pertaining to actual obligations about hypothetical situations. In this case, the difference between (90) and "true" attitude conditionals would be that while the latter involve fake restriction of the attitude, the former only involve fake restriction of embedded *would*. In other words, while the about-situation of *think* is some actual situation, in accordance with Belief-Dependence, the about-situation for *would* would be the quantified-over non-actual situations. Since Belief-Dependence does not apply to *would*—hence why counterfactual conditionals are well-formed with *would* in the first place—we would predict no problem for sentences like (90). While such an account of either fake or real restriction of embedded *would* instead of *think* is currently only a sketch, rather than a fully fleshed-out analysis, we believe that it is a promising route to pursue for sentences like (90), and would be useful in explaining the obligatory presence of conditional morphosyntax in the embedded clause in such constructions.

To summarize, an analysis of attitude conditionals that defines their truth conditions around actual desires about non-actual situations generates the right predictions for a variety of cases that were problematic for the analyses discussed in Sections 3 and 4. Because the *if*-clause restricts a high-scope MODAL, donkey anaphora is not predicted to be constrained to the scope of the attitude. Because of the universal quantification over maximally likely worlds introduced by MODAL, the analysis rightly predicts that real-world circumstances can impact the truth conditions of attitude conditionals, though the effects of real-world circumstances had to be constrained by the intuitively plausible principle of Belief-Dependence of About Situations, which states that one can only have attitudes toward situations that one considers plausible. Once Belief-Dependence of About-Situations was adopted, the unacceptability of counterfactual attitude conditionals with *wish* and *regret*, as well as the acceptability of indicative attitude conditionals with these verbs, immediately fell out.

5.3 Would-attitude conditionals

Next we turn to would-attitude conditionals like (2), repeated below:

(2) a. If Laura becomes a zombie, she will want you to shoot her.
b. If Laura became a zombie, she would want you to shoot her.

Our analysis of attitude conditionals was that they constitute claims about actual desires toward non-actual situations. Viewed in this light, would-attitude conditionals become a simple matter of flipping *want*'s evaluation situation: they are claims about *non*-actual desires toward non-actual situations. Hence, the sentences in (2) make claims about the desires of zombie-Laura, and not present-Laura.

In our analysis of attitude conditionals, we posited that whatever constellation of tense, aspect, mood, and modal heads is responsible for the absence of conditional morphosyntax in the consequent (*want* instead of *would want*) was responsible for the evaluation situation of *want* being set to the matrix (real-world) situation, rather than the situation quantified over by MODAL. Conversely, suppose that in attitude conditionals whatever cluster of these heads leads to the *inclusion* of conditional morphosyntax is also responsible for the evaluation situation of *want* being MODAL's quantified-over situations, rather than the matrix situation. Beyond that, attitude and would-attitude conditionals are structurally identical, at least as far as the compositional semantics is concerned.

In this case, the predicted truth-conditions for would-attitude conditionals like (2) would be as in (93), with the sole difference between these truth-conditions and those for the attitude conditional (1) being in the choice of evaluation situation for *want*:

(93) $[\![(2)]\!] = \lambda s.\ \forall w \in \mathrm{ML}_s(\textsc{zombie})[\exists s'[\textsc{MinSit}(\textsc{zombie}, w) \sqsubseteq s' \wedge$ laura wants SHOOT about s' in $s']]$

Thus, the sentences in (2) are predicted to be true iff in each of the most likely worlds w in which Laura becomes a zombie, the minimal ZOMBIE-situation in w expands to a situation in which and about which Laura wants to be shot. On the assumption that zombie-Laura is incapable of desires for anything other than brains, or at least that zombie-Laura does not want to be shot, we rightly predict (2) to be false.

However, in the introduction it was noted that at least marginally, (2) does seem to allow for a restricted-like reading more or less on a par with (1), in which the interpretation pertains to Laura's present, actual desires. As things presently stand, our analysis cannot account for this fact, as the only desires that should matter for would-attitude conditionals should be zombie-Laura's. Moreover, if we are correct in tying the evaluation situation of *want* to the cluster of heads that leads to the pronunciation of either *would want* or simply *want*, then it cannot be the case that on "restricted" readings of would-attitude conditionals the evaluation situation for *(would) want* is simply set to the real-world situation, as it is with attitude conditionals. Thus, the present proposal needs to somehow be extended to account for both the observation that apparent restricted readings can be found with would-attitude conditionals, and the fact that these interpretations are marginal.

There are at least two intuitively plausible analyses for this marginal restricted reading of (2); we will not attempt to decide between these, but offer them as a potential path forward. The first starts with the observation that in certain matrix clause contexts, an attitude embedded under a modal seems to give rise to a reading quite similar to an unembedded instance of that attitude. This is illustrated in (94):

(94) STUDENT: I did my homework!
TEACHER: I certainly {would/should} hope that you did!

The truth conditions of the teacher's utterance in (94) seem to be (nearly) identical to the truth conditions were she to utter the same sentence without *would* or *should*. Such seemingly vacuous uses of modals are not restricted to attitude verbs; (95) illustrates an example with the speech verb *ask* and the future semi-modal *going to*, in which the interpretation is one of a present performative utterance, rather than a promise of a future instance of asking:

(95) I'm going to ask you to drop the weapon.

It is not obvious to us how or why these seemingly vacuous uses of modals should arise, but whatever the explanation is, it could lend itself to an analysis of our apparent restricted reading: while *would* generally requires the attitude's situation of evaluation to be bound by the modal's situation-quantification, there is a relatively narrow class of cases in which this can be avoided and a matrix-situation reading—or a reading that is practically truth-conditionally undifferentiable from one—can occur. Since such cases appear to be relatively infrequent in matrix clause uses like (94) and (95), this would also lead to the prediction that apparent restricted readings of would-attitude conditionals should be relatively marginal.

A second possibility that does not involve any seemingly vacuous *will* or *would* would be to posit that when one asserts (2) with an apparent restricted interpretation, one behaves as if when Laura becomes a zombie, there are in some sense really two Lauras: zombie-Laura and the "real", fully human Laura that somehow persists after her zombification. The pronoun *she* in the consequent of these would-attitude conditionals could refer either to zombie-Laura, as would seemingly be required for the true reading of (96) below, or to the metaphysically persistent human Laura, whose wishes would be that her zombie-self is shot.

(96) If Laura becomes a zombie, she will want to eat brains.

Alternatively, it could be that it is not the "real" Laura that persists, but just her desires; either way, some piece of pre-zombie Laura survives after she becomes a zombie, and it is this surviving piece of her that ensures that, after becoming a zombie, she wants her zombie self to be shot.

This may seem like a hefty ontological commitment, though it is worth emphasizing that we posit this as a (possibly transient) part of the "natural language ontology" that determines the metaphysical commitments of speech acts (Bach 1986), rather than as an actual metaphysical principle of the world itself. Viewed in this light, the potential ontological commitments are perhaps not so burdensome. In fact, it is fairly common to speak of objects and individuals that are gone as if they are still around. For instance, the sentences in (97) can be felicitously uttered at a reading of Sally's

will, in spite of the fact that Sally, being deceased, cannot be said to have any present desires at all:

(97) a. Sally's wishes are that we cremate her.
b. Sally wants half of her wealth to be donated to the Red Cross.

Of course, this doesn't change the fact that on this second proposal, someone who utters a would-attitude conditional on an apparent restricted interpretation is committing themselves to a somewhat quirky ontology for discourse purposes. Assuming that utterances that temporarily commit the speaker to an out-of-the-norm ontology are generally rare, this could once again explain the marginal nature of apparent restricted readings of would-attitude conditionals.

Before moving on to attitude disjunctions, it is worth considering the full set of conceivable ways in which the about-situation and evaluation situation variables can be bound in conditionals with attitudes in their consequents. On our analysis, attitude conditionals are interpreted as actual desires about non-actual situations, while would-attitude conditionals are interpreted as non-actual desires about non-actual situations. This suggests a possible 2×2 paradigm, where one variable is whether the about-situation is (non-)actual, and the other is whether the evaluation situation is (non-)actual. Our analyses of attitude and would-attitude conditionals cover the two cases in which the about-situation is non-actual, but what about the other two possibilities, namely actual desires about actual situations and non-actual desires about actual situations?

Let us start with actual desires about actual situations. This is precisely what happens in the case of unembedded attitude ascriptions, so clearly we predict there to be many possible cases in which the about-situation and the situation of evaluation for the attitude are both situations in the actual world. Moreover, we hypothesized about apparent well-formed counterfactual attitude conditionals like (90) that these involved actual attitudes toward actual situations, with the impact of the counterfactual antecedent being felt on the embedded modal *would* rather than on *think*. However, in simpler attitude conditionals like (1) we expect interpretations involving actual attitudes toward actual situations not to arise. To see why, consider what happens when we switch the about-situation in our analysis of (1) from the modal's quantified-over situation s' to the real-world situation s:

(98) $\lambda s.\ \forall w \in \mathrm{ML}_s(\textsc{zombie})[\exists s'[\textsc{MinSit}(\textsc{zombie}, w) \sqsubseteq s' \wedge$
laura wants SHOOT about s in s]]

In (98), the modal's situation-quantification does not actually bind anything. The sentence is predicted to be true iff (I) every maximally likely world in which Laura becomes a zombie contains a minimal situation in which Laura becomes a zombie, and (II) Laura currently, actually wants to be shot. But given the vacuity

of the first conjunct, we get the same truth conditions as if there were no *if*-clause or MODAL to begin with: both are entirely semantically vacuous. Since vacuous quantification is generally thought to lead to deviance, we predict that in conditional environments, interpretations involving actual attitudes toward actual situations should be constrained to cases in which the *if*-clause and MODAL are performing some other, semantically non-vacuous work.

What about non-actual desires about actual situations, i.e., sentences with the structure of would-attitude conditionals, but where ABT is bound by the highest (real-world) situation abstractor instead of the lowest (MODAL-quantified) one? One possible candidate for just such an interpretation is (99):

Kwame has passed away, and his wife Maria's friend is trying to persuade her to remarry.

(99) If Kwame were here, he would want you to remarry.

Notice that (99), which we will call a *were-they-here conditional*, is presumably not a claim pertaining to a hypothetically still-existing Kwame's desires about a situation in which he in fact still exists. After all, in this case he likely would not want Maria to remarry, since he would want to remain married to her if he is still around. Rather, (99) seems to be a claim pertaining to a hypothetically still-existing Kwame's desires about a situation in which he and Maria can no longer be married. This is precisely what we predict if were-they-here conditionals involve non-actual desires about actual situations, as in (100). (Notice that since (99) is a counterfactual conditional, we use ML^+ instead of ML; unlike with attitude conditionals, this is unproblematic for would-attitude or were-they-here conditionals.)

(100) $\lambda s.\ \forall w \in \text{ML}^+_s(\textsc{kwame\_here})[\exists s'[\textsc{MinSit}(\textsc{kwame\_here}, w) \sqsubseteq s' \wedge$
kwame wants REMARRY about s in $s']]$

Alternative accounts of were-they-here conditionals are conceivable: perhaps a metaphysical explanation in terms of the hypothetically existing Kwame in some sense being "less than fully existing", and thus incapable of being married to Maria. However, the existence of were-they-here conditionals suggests at least the possibility that conditionals with attitudes in their consequents can give rise to readings involving non-actual attitudes toward actual situations.

A table detailing the possible 2×2 paradigm of (non-)actual about-situations and attitude evaluation situations can be seen in Fig. 4.

5.4 Attitude disjunctions

Our final task is to provide an analysis of attitude disjunctions like (4):

	actual evaluation situation	**non-actual evaluation situation**
actual about-situation	unembedded, (90)	were-they-here conditionals
non-actual about-situation	attitude conditionals	would-attitude conditionals

Figure 4 Possible paradigm of (non-)actual about-situations and attitude evaluation situations.

(4) Blumberg & Holguín 2019: p. 380:

Either a lot of people are on the deck outside, or I regret that I didn't bring more friends.

We first start with a semantics for disjunction. We will use a very simple situation semantics for disjunction: *p or q* is true of s iff $p(s)$ or s expands into a situation s' such that $q(s')$. This is laid out more formally in (101); note once again that since the linearly second disjunct is the one with which *or* composes first, the order of arguments is reversed:

(101) $[\![\text{or}]\!] = \lambda s \lambda q \lambda p.\ p(s) \vee \exists s'[s \sqsubseteq s' \wedge q(s')]$

We take the syntactic structure for (4) to be as in Fig. 5; we exclude *either* for simplicity's sake. Notice that while the evaluation situation provided by the TAM cluster X is the one bound by the lower λs_3, the about-situation provided by ABT is the matrix situation bound by the higher λs_1, similar to our analysis of were-they-here conditionals (and an inversion of our analysis of attitude conditionals). However, it is worth noting that there is not as obvious a path to filling out the potential 2×2 paradigm for attitude disjunctions as there was for conditionals, for the simple reason that with disjunctions all of the situations involved are world-mates. As a result, there is no distinction between actual and non-actual evaluation situations or about-situations. This means that the predicted semantic differences based on how the evaluation situation and about-situation are assigned will be at best extremely subtle, and possibly non-existent. Thus, we see no reason to doubt that alternate bindings of the two situation variables are equally plausible and give rise to more or less the same truth conditions.

Given the structure in Fig. 5, the predicted truth conditions for (4) are as in (102):

(102) $\lambda s.\ \textsc{outside}(s) \vee \exists s'[s \sqsubseteq s' \wedge \text{I regret } \textsc{few\_friends} \text{ about } s \text{ in } s']$

Now suppose that s is the situation containing the party, as well as the events leading up to the party. If there are a lot of people outside, then the first disjunct in (102) is true of s, so (102) is predicted to be true of s. Meanwhile, if there are not a lot of people outside, then since s is a situation that includes woeful attendance of the party, s is a situation about which I regret not bringing more friends, meaning that

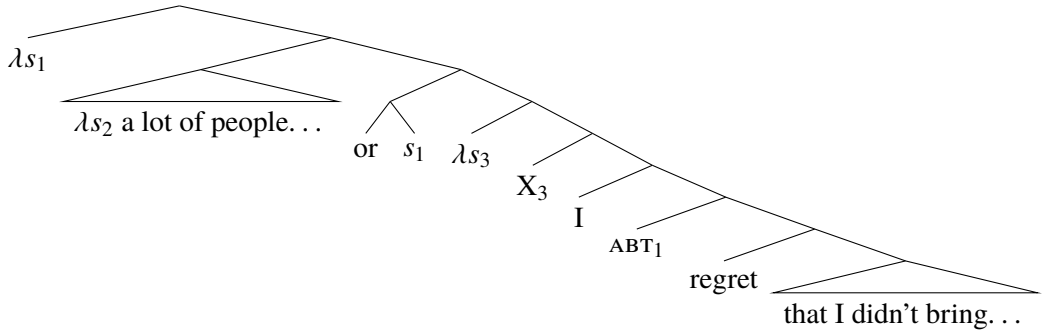


Figure 5 Proposed LF structure for (4).

the second disjunct is true, and (102) is again predicted to be true of s. Notice that this predicts the similarity in interpretation between (4) and its attitude conditional paraphrase (5), which is predicted to have the interpretation in (103):[13]

(5) If there aren't a lot of people on the deck outside, then I regret that I didn't bring more friends.

(103) $\lambda s.\ \forall w \in \text{ML}_s(-\textsc{outside})[\exists s'[\textsc{MinSit}(-\textsc{outside}, w) \sqsubseteq s' \wedge$
I regret FEW_FRIENDS about s' in $s]]$

Namely, each asserts that my attitudes toward situations that include minimal $-$OUTSIDE situations are such that I regret not bringing more friends to the party.

Two things are worth discussing about our proposed semantics for attitude disjunctions. The first is that as things stand, it does not predict the ordering asymmetry for attitude disjunctions observed by Blumberg & Holguín (2019) and illustrated by the oddness of the reverse attitude disjunction (36), repeated below:

36) Blumberg & Holguín 2019: p. 391:

?? Either I regret that I didn't bring more friends, or a lot of people are on the deck outside.

Blumberg & Holguín chalk up the infelicitous nature of (36) to the dynamic effects of disjunction, as a non-vacuous restrictor is permissible in the second disjunct, but not the first. However, we suspect that (36)'s oddness comes from a different source.

13 Defining a (minimal) situation of "not-p" is a well-known challenge in situation semantics. For example, a situation that just includes the potted plant in my office is not a situation of there being a lot of people being outside, but it is also clearly not a situation of there *not* being a lot of people being outside. For some discussion of this issue, see Kratzer 1989, 2019.

A common pragmatic use of uttering a disjunction *p or q* is to convey a direct, asymmetric causal or inferential connection between the falsehood of *p* and the truth of *q*. An example of this is given in (104):

(104) Zara will either pass her exams or get kicked out of the program.

While (104) seems to literally assert only that (at least) one of the two disjuncts is true, it also conveys that Zara's failure to pass her exams will be the direct cause of her getting kicked out of the program. Interestingly enough, reversing the order of disjuncts in such cases leads to an oddness that is very much of a similar sort to (36):

(105) ?? Zara will either get kicked out of the program or pass her exams.

In fact, to the extent that (105) is acceptable, it conveys a reverse causation: the only thing that will stop Zara from passing her exams is if she is kicked out of the program first. Notice that this inference is not simply due to some requirement that the second disjunct not temporally precede the first, as illustrated by the acceptability of (106):

(106) I will either go to Lidl tomorrow or swing by Aldi later today.

This is relevant because attitude disjunctions do in fact seem to convey precisely that the falsehood of the first disjunct is a direct cause of the truth of the second disjunct. Looking at (4) in particular and framing things in terms of our own analysis, if *s* is a situation of there not being a lot of people outside, then the fact that *s* has this trait is indeed the direct cause of *s* being a situation about which I regret not bringing more friends. In this case, whatever is the appropriate analysis of (105) should equally well predict the oddness of (36).

The second thing worth noting about our analysis is a particular prediction it makes about the natural language metaphysics of attitudes. More specifically, our analysis imposes a certain degree of *externalism*, in roughly the sense argued for by Drucker (2019): one can have an attitude without any of the cognitive phenomenology typically associated with that attitude. For example, suppose that Becky walks into the living room at a party and, seeing that it is practically empty but not knowing what the situation is outside, asserts (4). After making this assertion, Becky immediately suffers a heart attack and dies on the spot, without ever finding out if there were actually a lot of people outside. Now suppose in addition that there were not, in fact, a lot of people outside. Can we in this context truthfully assert that Becky regretted not bringing more friends to the party, even though her lack of knowledge about whether people were outside meant that she never felt the phenomenological experience of regret? On our analysis, we predict either that we can make this assertion truthfully—which may seem intuitively odd given that Becky never really *felt* the regret—or that Becky was incorrect in asserting (4) in the first place, which seems even odder since (4) essentially functions as an assertion about Becky's own

feelings, and surely she cannot be wrong about those. This same issue extends to attitude conditionals such as (5), the attitude conditional paraphrase of (4).

While there are a few different ways one could approach this issue within an analysis more or less along the lines of ours, perhaps the easiest way out is to follow Drucker (2019) and simply embrace this externalism, at least in the natural language ontology: for conversational purposes, Becky really *did* regret not bringing more friends to the party, even if she never "felt" that regret. Alternatively, a weaker version of this stance would be that interlocutors can temporarily adopt an externalist metaphysics in a given conversational context: while many features of the natural language ontology are generally assumed to be permanent fixtures that remain in place across conversations, presumably because they are tied to much broader features of human perception and cognition, it is not inconceivable that a concept as narrow as attitude externalism could sometimes be adopted for conversational purposes, and sometimes not. In other words, it is entirely plausible that we sometimes, but not always, speak like an externalist.

For some evidence that externalism—possibly only adopted temporarily for conversational purposes—may not be as bizarre as it sounds, suppose that Becky walks into a party, notices that no one is in the living room, and whispers (4) to her friend. Now suppose that another friend approaches Becky and asks her if she regrets not bringing more friends. In this context, it seems quite odd for Becky to reply with a simple and direct *No*, even though she is not experiencing the phenomenology of regret. A far more appropriate response would be *I don't know; are there a lot of people outside?*, which is precisely the sort of response the externalist would expect: Becky does not yet know whether she regrets bringing more people to the party, even if she knows that she is not actively feeling regret at that very moment.

Another piece of evidence in favor of (temporary) externalism may be what Blumberg & Holguín (2018), building on observations by Recanati (2012), refer to as *ultra-liberal attitude reports*, such as (107):

> *Grandmaster Caruana believes that any child can become a chess grandmaster if they start studying before the age of six. Sterling and Lana's daughter Jane, whom Caruana has never heard of, is turning five this month.*

(107) Caruana thinks that Jane can become a grandmaster if she starts studying before the age of six.

Clearly, GM Caruana has never experienced the phenomenology of believing that Jane can become a grandmaster, since he does not even know who she is, let alone that she is under six years old. And yet (107) can be truthfully asserted in spite of this. While Blumberg & Holguín (2018) posit that a seemingly true assertion of (107) is really an instance of what might be called "loose talk"—it is not literally true, but is treated as true—an alternative explanation that does not throw ultra-liberal

attitude reports into the "loose talk" wastebasket would be to posit some degree of attitude externalism: in some sense or another, as far as the discourse participants are concerned Caruana really *does* believe that Jane can become a GM, even if he does not know that he holds this belief (or why).

In summary then, the account of attitude conditionals and attitude disjunctions proposed in this paper apparently commits us to a certain degree of attitude externalism in the natural language ontology: whether or not one has a certain attitude can be entirely divorced from whether one experiences the phenomenology associated with that attitude, or whether one knows that one has that attitude. There may be ways in which one can seek to fix this feature of our account, but it may equally well be the case that there is nothing to fix, and that a certain amount of externalism, perhaps temporary and discourse-dependent, is required in order to capture the full range of linguistic data.

5.5 Are other conditionals fake restrictions too?

Before wrapping up, it is worth briefly discussing what our analysis of attitude conditionals might tell us about non-attitude conditionals like (6):

(6) If Laura becomes a zombie, you must shoot her.

Up to this point we have followed Kratzer (1981, 1991a,b, 2012) in supposing that for deontic conditionals like (6), the *if*-clause really does restrict the world-quantification of the modal *must*, in contrast to the fake restriction of attitude conditionals. However, our analysis of attitude conditionals invites another possibility, briefly mentioned previously in our discussion of apparent well-formed counterfactual attitude conditionals: namely, that conditionals with modals (and other operators that can be "restricted" by *if*-clauses) also involve fake restrictions. That is, if modals can be thought of as being about situations in the same way that attitudes are, (6) may have an LF more along the lines of Fig. 6, generating a reading of (6) pertaining to actual obligations about hypothetical situations.

In fact, if other conditionals involve fake restriction and the only operator that *if* genuinely restricts is the covert modal MODAL, then there is no longer any need to posit the existence of MODAL in the first place, and its semantic role can be shifted to *if*. That is, *if* can be assigned the interpretation in (108):

(108) $[\![\text{if}]\!] = \lambda s \lambda p \lambda q.\ \forall w \in \text{ML}_s(p)[\exists s'[\text{MinSit}(p, w) \sqsubseteq s' \wedge q(s')]]$

This definition loads the work of MODAL onto *if*, thereby allowing a simplification of the LF in Fig. 6 to that in Fig. 7, with an identical interpretation. The same reallocation of semantic work could be done for all of the attitude, would-attitude, and were-they-here conditionals discussed earlier in this section.

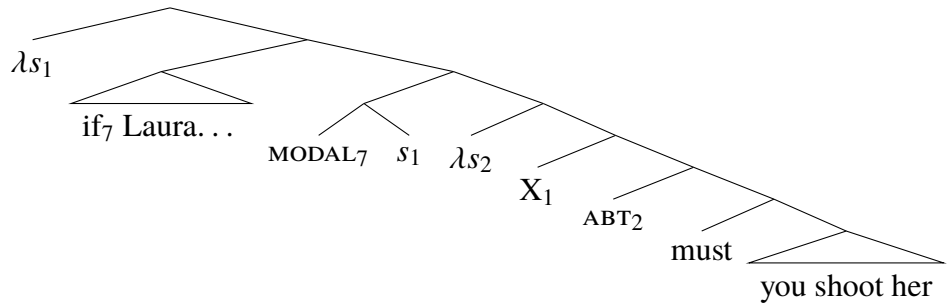


Figure 6 Potential fake restriction LF structure for (6).

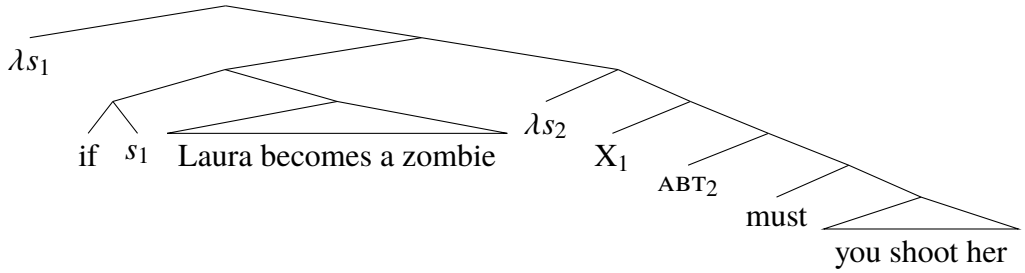


Figure 7 Alternate fake restriction LF structure for (6), with *if* providing world-quantification.

One benefit of such an approach is that it would allow for a unified treatment of attitude (and would-attitude) conditionals and other conditionals: much like how an attitude restriction approach reduces both modal and attitude conditionals to quantificational restriction by *if*-clauses, this new approach would reduce both to fake restriction through non-actual about-situations. Moreover, modal conditionals are known to exhibit a semantic contrast that bears a striking resemblance to the distinction between attitude and would-attitude conditionals, exemplified in (109) (see Frank 1996, Geurts 2004 among others):

(109) Geurts 2004: p. 1:

a. If Beryl is smoking marijuana, we have to report her to the police.

b. If smoking marijuana is illegal in this country, we have to report Beryl to the police.

(109a) looks like a regular deontic conditional: on a Kratzerian account, it is true iff in all of the deontically best worlds in which Beryl is smoking marijuana, we

report her to the police. But (109b) is slightly different, since the antecedent does not pertain to the circumstances of whether Beryl is smoking marijuana, but rather to what the laws actually are for this country. In other words, in (109a) the laws that inform the truth of the deontic claim are taken for granted, while in (109b) it is precisely the nature of these deontic rules that is called into question in the first place. This latter reading is also the most salient reading of the counterfactual (110):

(110) Frank 1996: p. 31:
If Luther hadn't brought about the Reformation, we would still have to pay indulgences.

(110) is clearly a claim about what our obligations would be not only if circumstances were different, but if the laws informing deontic modality were themselves different.

This distinction in interpretations could be captured on our account by framing the contrast between "have-to" and "would-have-to" conditionals in a manner similar to the contrast between attitude and would-attitude conditionals. Have-to deontic conditionals like (6) and (109a) would be interpreted analogously to attitude conditionals: they are claims pertaining to actual obligations about hypothetical situations. Meanwhile, would-have-to conditionals—that is, deontic conditionals pertaining to obligations if the laws were potentially different from how they actually are, such as (109b) and (110)—would be analogous to would-attitude conditionals: they pertain to non-actual obligations about non-actual situations.

While such a unification has its own appeal, it seems that firm empirical arguments in favor of extending our analysis of attitude conditionals to other types of conditionals are harder to come by than they were for attitude conditionals in the first place. For example, the impact of real-world circumstances on the truth conditions of attitude conditionals made for a compelling argument because with attitudes there is an important distinction between the belief-based domain of the attitude and the real-world-circumstance-based domain of MODAL (or, on the revised syntax-semantics outlined above, *if*). But with modals this is not the case: on practically any theory of deontic modality, the semantics of deontic *must* is itself sensitive to circumstances, meaning that even if the *if*-clause directly restricted the modal itself, we would expect real-world circumstances to affect the truth conditions of deontic conditionals. Therefore, while we think that a fake restriction account of conditionals other than attitude conditionals is a promising route to pursue, we set aside for future work the task of determining whether it has any meaningful empirical advantages (or disadvantages) when compared with a traditional account in which *if*-clauses restrict modal and other operators.

6 Conclusion

In this paper, we have provided an assortment of evidence against the proposal, put forward in a variety of forms in Section 3, that seemingly restricted readings of attitudes arise from a restriction of the world-quantification of either the attitude or some modal below the scope of that attitude. Instead we have argued that these apparent restricted readings are fake: they are a result of certain choices in fixing the (unrestricted) attitude's about-situation. This was shown to have several empirical advantages over restriction-based approaches, accounting for all of the observations that were shown in Section 4 to be problematic for such analyses. The analysis also poses wide-ranging questions about the compositional semantics of conditionals as a broader class, and more specifically, whether all cases of *if*-clause restriction are in fact instances of fake restriction.

Our analysis opens up several new lines of inquiry, some of which we have discussed over the course of the paper. In addition to these, an obvious place in which the present work needs to be extended is by incorporating a more robust lexical semantics of attitude verbs. For our specific purposes, the lexical semantics of attitude verbs was less relevant than the basic argument structure of these verbs. However, fleshing out the lexical semantics of these verbs will lead to further predictions about the semantic interpretation of attitude conditionals and other attitude constructions that were beyond the scope of this paper. A second potential area of future inquiry is the exploration of what further advantages can be gleaned from the use of about-situations in the interpretation of attitudes and, possibly, modals. For example, the ability to have attitudes about situations that are compatible with, but not entailed by, one's beliefs may be useful in analyzing the aforementioned phenomenon of ultra-liberal attitude reports (Blumberg & Holguín 2018). These and other matters are left for future research.

References

Bach, Emmon. 1986. Natural language metaphysics. In R. Barcan Marcus, G.J.W. Dorn & P. Weingartner (eds.), *Logic, Methodology, and Philosophy of Science VII*, 573–595. Amsterdam: Elsevier.

Barwise, Jon & John Perry. 1983. *Situations and Attitudes*. Cambridge, MA: MIT Press.

Berman, Stephen. 1987. Situation-based semantics for adverbs of quantification. *University of Massachusetts Occasional Papers in Linguistics* 12.

Blumberg, Kyle & Ben Holguín. 2018. Ultra-liberal attitude reports. *Philosophical Studies* 175(8). 2043–2062.

Blumberg, Kyle & Ben Holguín. 2019. Embedded attitudes. *Journal of Semantics* 36(3). 377–406.

Broome, John. 1999. Normative requirements. *Ratio* 12(4). 398–419. https://doi.org/10.1111/1467-9329.00101.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Crnič, Luka. 2011. *Getting* even. Cambridge, MA: MIT PhD dissertation.

Drucker, Daniel. 2019. Policy externalism. *Philosophy and Phenomenological Research* 98(2). 261–285.

von Fintel, Kai. 1994. *Restrictions on quantifier domains*. Amherst, MA: University of Massachusetts Amherst PhD dissertation.

von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148.

Frank, Anette. 1996. *Context dependence in modal constructions*. Stuttgart: Universität Stuttgart PhD dissertation.

Geurts, Bart. 2004. On an ambiguity in quantified conditionals. University of Nijmegen, Ms.

Goble, Lou. 1996. Utilitarian deontic logic. *Philosophical Studies* 82(3). 317–357.

Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13(2). 137–177.

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

Hintikka, Jaakko. 1969. Semantics for propositional attitudes. In *Models for Modalities*, 87–111. Dordrecht: Reidel.

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

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Kaufmann, Stefan & Magdalena Schwager. 2011. A unified analysis of conditional imperatives. In Ed Cormany, Satoshi Ito & David Lutz (eds.), *Semantics and Linguistic Theory (SALT)* 19, 239–256.

Kratzer, Angelika. 1981. The notional category of modality. In H.J. Eikmeyer & H. Rieser (eds.), *Words, Worlds, and Contexts: New Approaches in Word Semantics*, 38–74. Berlin: de Gruyter.

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

Kratzer, Angelika. 1991a. Conditionals. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik/Semantics: An International Handbook of Contemporary Research*, 651–656. Berlin: de Gruyter.

Kratzer, Angelika. 1991b. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik/Semantics: An International Handbook of Contemporary Research*, 639–650. Berlin: de Gruyter.

Kratzer, Angelika. 2002. Facts: Particulars or information units? *Linguistics and Philosophy* 25(5/6). 655–670.

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

Kratzer, Angelika. 2019. Situations in natural language semantics. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University summer 2019 edn.

Lassiter, Daniel. 2011. *Measurement and modality: The scalar basis of modal semantics*. New York, NY: New York University PhD dissertation.

Levinson, Dmitry. 2003. Probabilistic model-theoretic semantics for *want*. In Robert Young & Yuping Zhou (eds.), *Semantics and Linguistic Theory (SALT)* 13, 222–239. Ithaca, NY: Cornell University.

Lewis, David. 1975. Adverbs of quantification. In Edward L. Keenan (ed.), *Formal Semantics of Natural Language*, 3–15. Cambridge: Cambridge University Press.

Pasternak, Robert. 2018. *The mereology of attitudes*. Stony Brook, NY: Stony Brook University PhD dissertation.

Percus, Orin. 2000. Constraints on some other variables in syntax. *Natural Language Semantics* 8(3). 173–239.

Phillips-Brown, Milo. 2018. I want to, but. . . . In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of Sinn und Bedeutung 21*, 951–968.

Recanati, François. 2012. *Mental Files*. Oxford: Oxford University Press.

Roberts, Craige. 1987. *Modal subordination, anaphora, and distributivity*. Amherst, MA: University of Massachusetts Amherst PhD dissertation.

Roberts, Craige. 1989. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12(6). 683–721.

Schwager, Magdalena. 2011. Speaking of qualities. In Ed Cormany, Satoshi Ito & David Lutz (eds.), *Semantics and Linguistic Theory (SALT)* 19, 395–412.

Stone, Matthew. 1997. The anaphoric parallel between modality and tense. University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-97-09.

Sudo, Yasutada. 2014. On *de re* predicates. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 447–456. Somerville, MA: Cascadilla Proceedings Project.

Villalta, Elisabeth. 2008. Mood and gradability: an investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy* 31(4). 467–522.

Kai von Fintel
Department of Linguistics and Philosophy
Massachusetts Institute of Technology
Room 32-D808
77 Massachusetts Avenue
Cambridge, MA 02139, USA
fintel@mit.edu

Robert Pasternak
Leibniz-Center for General Linguistics (ZAS)
Schützenstraße 18
10117 Berlin, Germany
mail@robertpasternak.com