Discourse anaphoric otherwise: Information structure & modal subordination

Josh Phillips
Yale University

Hadas Kotek
Yale University

Abstract This paper provides a formal semantic/pragmatic analysis of the interpretation and meaning contribution of the English discourse anaphor otherwise. Otherwise is modeled as a discourse move (in the sense of Roberts 2012) which encodes an instruction to consider the complement of a set of worlds introduced in the clause preceding otherwise. Following Webber et al. (2001) and other authors, we take as key the observation that the identity of this antecedent clause to otherwise cannot be determined by the syntax alone. Instead, we argue that we must make crucial reference to the current information structure, and in particular to the current Question under Discussion, to determine the nature of the antecedent. We propose a dynamic semantic/pragmatic account for otherwise, which makes crucial use of modal subordination in order to both model otherwise’s flexible distribution, as well as previously unobserved limitations on its use.

1 Introduction

The work presented here develops an analysis of the English lexical item otherwise, drawing on tools from the dynamic semantics and information structural literatures. Our main focus is on otherwise’s use as a discourse ‘connective’ or ‘anaphor’ (e.g. Webber et al. 2001, Kruijff-Korbayová & Webber 2001), so named because of its apparent interpretive reliance on foregoing elements of discourse.1 This is demonstrated by the sentence pair (1), from Webber et al. (2001: 7). Each sentence is accompanied by a paraphrase that spells out its intended meaning.

(1) The discourse anaphor ‘otherwise’:
   a. If the light is red, stop. Otherwise go straight on.
      ≈ if the light is not red...

---
* For helpful comments and suggestions, we would like to thank Kai von Fintel, Francis Corblin, Simon Charlow, Kyle Rawlins, Ezra Keshet, Bob Beddor, as well as audiences at the LSA and NELS meetings in 2018.
1 For the purposes of this current paper, we restrict our attention to these “inter-clausal” adverbial uses. As we will discuss in §7, however, we anticipate that our account could be expanded to account for other uses as well.
b. If the light is red, stop. Otherwise you’ll get a ticket.
\[ \approx \text{if the light is red and you } \text{don’t stop...} \]

As example (1) makes clear, the question of how to determine the nature of the antecedent \( p \) is quite subtle. While the syntactic environment preceding otherwise is identical in both (1-a) and (1-b), it is clear that the proposition which is interpreted as the antecedent in each case is different. How, then, do speakers retrieve an antecedent to otherwise? In order to answer this, we take a view that emphasizes the flow of information in a discourse (see also Roberts 2012).

In a nutshell, we develop an analysis of otherwise which draws on existing dynamic semantic analyses of conditionals. We’ll argue that otherwise contributes a discourse move whose content is to predicate a subsequent proposition of the complement set of worlds made salient by the prior discourse. This accommodated antecedent may be a conditional, as in (1), but need not be:

(2) Otherwise’s antecedent may be a declarative, imperative, or (certain) interrogatives:
   a. She’s asleep, otherwise she would have come.
   b. Stop. Otherwise you’ll get a ticket.
   c. Do you have your car? Otherwise I’ll give you a lift.
   d. Do you want to get a beer at Three Sheets or Counterweight tonight? Otherwise you make a bloody suggestion.\(^2\)

A (declarative) otherwise statement, then, includes two components: the antecedent is put on the table as accurate to the best of the speaker’s knowledge.\(^3\) Otherwise then introduces an instruction to consider how the world must be if the speaker is incorrect about their initial assumption, or if their instruction to add the antecedent into the Common Ground is rejected. Its prejacent provides a description of such a world. Pragmatically speaking, this prejacent is often taken to represent a justification for the antecedent assertion, providing a discourse link between these two components.

The complement set of worlds under consideration by otherwise cannot be contributed strictly by the syntax, as we have just seen. Instead, we claim that it requires reference to the Question under Discussion (QuD) and the current Information Structure. We will show how this proposal can model cases such as (1) and, in addition, that it correctly predicts a previously unnoted interaction of otherwise with possibility modals as well as other restrictions on the choice of antecedent. We

\(^2\) To our ears, this question can be read as either a polar question or an alternative question.
\(^3\) That is, asserted, cf. Stalnaker 1979.
conclude the paper by briefly discussing an expansion of our proposal to nonclausal uses of *otherwise*. Two examples are given in (3):

(3)  *Intra-sentential uses of otherwise:*

a. The income they earn from it is likely to be the only source of cash to supplement their *otherwise* subsistence economy. (OED)
b. Amelia behaved well *otherwise*. (Flament-Boistrancourt 20114)

Before presenting our analysis in section 4, we first discuss in a bit more detail previous analyses and additional properties of *otherwise*, which our analysis builds on. We then present several novel observations about *otherwise* in section 3.

2  **Background: The meaning of otherwise**

As we have seen, *otherwise* acts as a discourse connective or anaphor, relating an antecedent sentence with a consequent sentence. A key example which we will concentrate on in this paper is the Red Light example, repeated here from (1) above. This example illustrates a key property of *otherwise*: that the continuation following *otherwise* appears to be discourse-dependent, and can’t be strictly calculated based on the syntactic material preceding *otherwise*. In (4), the same material appears before *otherwise*, but with different consequents. The nature of the consequent allows us to calculate what *otherwise* is operating on, as we spelled out in (1-a)–(1-b).

(4)  *The Red Light example:* = (1)

a. If the light is red, stop, *otherwise* go straight on.
b. If the light is red, stop, *otherwise* you’ll get a ticket.

A satisfactory approach to *otherwise*, then, requires a consideration of the structure and ‘flow’ of information in a given discourse context. Intuitively, the *otherwise* clauses in (4) have the semantics of conditionals: *Otherwise* targets a set of worlds in which some anaphoric proposition does not hold (i.e. converse nonimplication).

Two prior accounts of *otherwise* in Webber et al. (2001) and Kruijff-Korbayová & Webber (2001) adopt information-structural analyses of *otherwise*, which will inform our analysis below. In particular, Webber et al. (2001) argue for the existence of a “discourse adverbial” class (comprising lexical items including *then, nevertheless, otherwise*), and a distinct class of “structural connectives” (*or, and, but, because*). These authors appeal to an ‘anaphorically-derived contextual (eventive) parameter’ $e_i$ and an *inferrable relation* between two event descriptions (in the ab-

---

4 Here, and throughout, examples from both Flament-Boistrancourt 2011 and Inkova-Manzotti 2002 have been translated from the original French by the authors.
sence of an explicit structural connective). An example of these notions is given in (5):

(5) Two types of ‘inferrable relations’:
   a. If the light is red, stop, (but) otherwise go straight on.
   b. If the light is red, stop, (because) otherwise you’ll get a ticket.

Otherwise not only retrieves different antecedents, as we have already seen in (1), but also can encode different relations between two event descriptions (contrast in (5-a), and explanation in (5-b)). For Webber et al. (2001: 17), these effects are pragmatically derived, and are crucially unavailable to “structural connectives” which are restricted in the relations they can encode and the antecedents they can retrieve.5

This observation about the limited distribution of structural connectives has been independently made in the literature on conditional uses of or (‘pseudocoordination’, see Culicover & Jackendoff 1997, Klinedinst & Rothschild 2012, Biezma & Rawlins 2016, a.o.). As examples (6-a)–(6-b) show, the distribution of these uses is narrower than the equivalent use of otherwise. Although a conditional otherwise-like reading is available in (6-b), in (6-a), the conjoined imperatives ‘stop or go straight on’ must be interpreted as two options of what the addressee ought to do when the light is red. The otherwise-like reading that was available in (1-a)/(5-a) is infelicitous here.

(6) Conditional or has a more limited distribution:
   a. #If the light is red, stop, or go straight on.
   b. If the light is red, stop, or you’ll get a ticket.

Additional evidence that an adequate account of otherwise requires reference to a level of discourse representation comes from intra-sentential uses of otherwise, in cases such as (7). For Webber et al. (2001: 7), these examples necessitate an E-type anaphor.6 As these authors point out, this ‘suggests that discourse adverbials are accessing discourse entities (in particular, eventualities) rather than signaling a structural connection between clauses.’

(7) Intra-sentential otherwise:
   a. Every person selling “The Big Issue” might otherwise be asking for spare change. (Webber et al. 2001: 7)
   b. These moments give emotional ballast to what would otherwise be an exercise in wackiness.

5 A similar observation is made in Corblin (1994, 2002).
6 Although see our analysis below, an in particular section 6.1; for us, this move will not be required.
On the basis of data similar to the red light example (i.e. otherwise sentences with complex-clause antecedents), Kruijff-Korbayová & Webber (2001) model otherwise as a discourse connective that is sensitive to information structure in its retrieval of an antecedent. They assume that Logical Forms are partitioned into theme ($\theta_{is}$) and rheme ($\rho_{is}$) “phases”, which have the effect of updating a given discourse context. Following Steedman (2000), Kruijff-Korbayová & Webber (2001) assume that both $\theta_{is}$ and $\rho_{is}$ presuppose an alternative set (cf. Rooth 1985). Otherwise then updates the context with the complement of (a subpart of) either $\rho_{is}$ or $\theta_{is}$ with respect to the relevant alternative set.

Along similar lines, Inkova-Manzotti (2002) and Flament-Boistrancourt (2011) provide descriptions of the broad range of uses of French autrement ‘otherwise.’ Like English otherwise, the French particle requires use of context and pragmatics. Some examples are provided below.

(8)  
On peut se voir mardi. Autrement vendredi.
One can see Tuesday otherwise Friday
We’ll see each other Tuesday. Otherwise Friday.
(Inkova-Manzotti 2002: 114)

(9)  
Je pourrais faire une tarte. Je n’ai pas de farine. Autrement
I could make a quiche I NEG have NEG PART flour otherwise
j’ai tout ce qu’il faut
I have all DEM REL it necessary
‘I could make a quiche. I’m out of flour. Otherwise I’ve got everything needed.’
(Inkova-Manzotti 2002: 122)

In the analysis proposed below, we likewise acknowledge the importance of context and pragmatic computation in the use of otherwise. The existing analyses surveyed here suffer from the limitation that there are no constraints on the ‘range of things that can serve as antecedents’ (see Kruijff-Korbayová & Webber (2001) for an explicit discussion of this issue). Likewise, Webber et al. (2001), must make reference to complex event structures, and to yet another complex mechanism of E-type anaphora for examples such as (7). We will show in section 6.1 that these examples are naturally unified under our analysis, so that no additional assumptions must be made for intra-sentential cases as compared to inter-sentential cases.

Flament-Boistrancourt (2011) explicitly deals with distributional differences of French sinon and autrement (both are frequently translated as ‘otherwise.’) Francis Corblin (2002: 252; pers. comm) points out that sinon (lit. ‘if NEG’) admits of a compositional analysis and an identical distribution/use to si ce(la) n’est pas le cas... ‘if it is not the case that X...’
In the section that follows, we introduce several new observations regarding the distribution and use of otherwise, before spelling out a proposal which aims to capture these facts in section 4.

3 Other key properties of otherwise

We begin by laying out the key properties of otherwise that we set out to capture with our account. As we have seen in section 2, otherwise has a connective-like use. As example (2) showed, the antecedent of an otherwise sentence may be a declarative, an imperative, or an interrogative. Here we will concentrate on sentential cases, where otherwise connects two sentences, as in our Red Light example in (1). See section 6.1 for a brief discussion of intra-sentential cases such as (7).

We have also established that the content of the continuation which follows otherwise is discourse-sensitive, and cannot be computed solely based on the syntactic content of the antecedent. This has been an important guiding observation in prior work on otherwise, and one that we take up in our analysis as well. We highlight here several additional properties of otherwise that will become important for our analysis below.

3.1 Otherwise is an intensional operator

First, we argue that the notion of modality is crucial to the analysis of otherwise. Recall that Kruijff-Korbayová & Webber (2001) notice that the two components related by otherwise rely on an ‘inferrable relation’. We claim this relation follows from a view of otherwise as a modalized operator, admitting of different modal bases. We illustrate this in (10):

(10) Observation: otherwise admits different ‘modal flavors’:
    a. Jan is home, otherwise I don’t know where she could be.
    b. Jan is home, otherwise she’s breaking curfew.
    c. Jan does her homework, otherwise she won’t learn.
    d. Jan does her homework, otherwise she’d fail the class.

Example (10-a) requires an epistemic modal base for its interpretation, whereas the minimally different (10-b) is interpreted under a deontic modal base. Examples (10-c) and (10-d) are interpreted with a teleological and a bouletic modal base, respectively. Our account below will account for this flexibility of otherwise, unlike prior accounts.

In section 4.1 we additionally defend the claim that otherwise makes crucial use of modal subordination (Roberts 1989 et seq): a flexibility in the choice of
antecedents which will explain the two possible choices of antecedents in examples such as (11)—here, including or excluding the modal:

(11) Students are required attend the lecture, otherwise...

a. ≈ If ¬ (they ATTEND)...
   ...they’ll fail the class.

b. ≈ If ¬□ (they ATTEND)...
   ...it’ll be empty.

3.2 Non-emptiness

As our paraphrases above illustrate, otherwise asks us to consider what would be the case in the complement set of worlds to those introduced in its antecedent. That is, otherwise induces a partition over worlds into those that satisfy the conditions in the antecedent, and those that don’t. A crucial requirement on this partition is that both cells are non-empty.

To illustrate this, consider the contrast in (12):

(12) a. I must go to school, otherwise I’ll get in trouble.
    \[ \alpha = \{ w \in f_{\text{deontic}} \mid \text{I go to school in } w \} \]
    \[ \overline{\alpha} = \{ w \in f_{\text{deontic}} \mid \text{I don’t go to school in } w \} \]

b. #I can/am allowed go to school, otherwise I’ll get in trouble.

This judgment contrast emerges because the necessity modal must in (12-a) eliminates a set of worlds \( \alpha \) from the context set (viz. those in which I don’t go to school \( \overline{\alpha} \); otherwise is thus able to make a claim \( \beta \) about those worlds (namely, in all of them, I get into trouble: \( \overline{\alpha} \subset \beta \)).

Conversely, the circumstantial possibility modal can asserts the existence of an accessible world in which I go to school, but fails to exclude any worlds from consideration in (12-b). As a consequence, otherwise has no complement set available to operate on, and we correctly predict that an otherwise sentence is infelicitous in this case.

Compare this with the minimally different (13), which speakers judge as acceptable:

(13) I can go to school, otherwise I wouldn’t be able to get an education.
    \[ \alpha = \{ w \in f_{\text{circ}} \mid \text{I am able to go to school in } w \} \]
    \[ \overline{\alpha} = \{ w \in f_{\text{circ}} \mid \text{I am unable to go to school in } w \} \]

8 Notice that, like in (11), an alternative pragmatic reasoning could have led us to choose as prejacent the set of worlds in which I must go to school. A felicitous otherwise statement in such a case might be: “...otherwise (≈ if I didn’t have to go to school), I’d skip class and go to the park.”
Here *otherwise* is anaphoric on the entire modal claim. The resulting assertion is that — in all those worlds where it is not the case that I *can* go to school — I don’t receive an education. As a consequence, there is a non-empty complement set of worlds in which to evaluate the *otherwise* sentence (namely, all of those in which I *can’t* go to school). In (13), despite the presence of a possibility modal, we are still universally quantifying into the antecedent proposition.

Our account below will elegantly explain the felicitous use of *otherwise* in such sentences. We return to this non-emptiness constraint on the distribution of *otherwise* and its consequences in section 5.

### 3.3 An *otherwise* sentence is non-commutative

Another observation that will inform our analysis below is that *otherwise* is not a symmetric operator: *p otherwise q* is different from *q otherwise p*, even in cases where the two propositions related by *otherwise* are logically independent of one another — so that an ‘inferrable relation’ is difficult to establish.

(14) **Word order is important in an otherwise sentence:**

a. She’s in the living room. *Otherwise*, she’s in the bathroom.

b. She’s in the bathroom. *Otherwise*, she’s in the living room.

(15) a. She’s sick. *Otherwise*, she’d be here.

b. *She’d be here. Otherwise*, she’s sick.

Example (14) shows that even in such a case, speakers perceive a difference in the felicity conditions and contexts in which the two variants of the *otherwise* sentence will be appropriate. Roughly: my first guess is that she’s in the {living room/kitchen}; if it turns out that she’s not there, then she’ll be in the {bathroom, living room}. 9 Predictably, example (15) shows that when an ‘inferrable relation’ (here: causality) *is* present, changing the order of the two propositions connected by *otherwise* may lead to infelicitous consequences.

### 3.4 An *otherwise* sentence is conjunctive

The final crucial component of our analysis is the treatment of an *otherwise* sentence as a kind of asymmetric conjunction: the speaker puts the antecedent on the table for adoption, but also includes an explicit claim about how the world must be

---

9 See section 3.5 for more relevant discussion.
in case that it is rejected. In words, we might then say that otherwise asserts: $p$; and if not $p$, then $q$ will hold:  

$$(16) \quad \text{An informal description of the meaning of a } p \text{ otherwise } q \text{ sentence:}$$

$$p \land (\text{if } \neg p, \text{ then } \Box q)$$

If a sentence of the form $p$ otherwise $q$ has conjunctive semantics (as proposed in (16)), this ought to predict that its negation could be achieved by falsifying the first conjunct, the second conjunct, or the entire assertion. We show that this is the case in (17):

$$(17) \quad \text{Negating otherwise shows conjunction-like behavior}$$

A: Sam is always home by 6pm, otherwise little Susie has a tantrum.

$$p^{12} \text{ if } \neg p, \text{ then } \Box q$$

B: That’s not true…

(i) He often gets home late, and Susie’s just fine.
(ii) Susie would be just fine if he did ever get home later, although it’s true that Sam always get home on time.
(iii) He often gets home late, although it is true that little Susie indeed has a tantrum whenever that happens.

In (i), the speaker is negating both conjuncts: Sam isn’t always home on time ($\neg p$), but Susie doesn’t have a tantrum because of that ($\neg p \land \neg q$). In (ii), only the second conjunct is negated: we assert that the first conjunct is true (Sam is always home on time), but that the implication nevertheless doesn’t hold (Susie wouldn’t have a tantrum if Sam were late). In (iii) only the first conjunct is negated: we assert

---

10 Here we take the conditional to reflect the modal restrictor view, as in Lewis 1975, Kratzer 1981, 2012, Heim 1982, among others.
11 Note also the similarity of this treatment to ‘information parameter change’ readings of structural connective or as formalized by Klinedinst & Rothschild (2012: 155-6). Under their dynamic account an utterance of the form ‘$\alpha$ or $\beta$’ corresponds to $\llbracket \beta \rrbracket^{\text{c.in.$\alpha$}}$ (sc. an utterance of $\beta$ where the “information parameter” $\text{s}$ is updated with $\neg \alpha$ (the negation of the first disjunct)).
12 As with other examples we have seen, there are two possible antecedents to otherwise in this example: the sentence with always and the sentence embedded under always. We kept the antecedent constant in our examples for consistency.
13 Recall that an implication is falsified just in case that its antecedent is true and its consequent is false. In (16), the implication under consideration is if $\neg p$, then $\Box q$. It is falsified in case that $\neg p \land \neg \Box q$. Recall further that $\neg p \rightarrow \neg \Box p$. (If $p$ doesn’t hold, then must $p$ doesn’t hold.)
that Sam is late ($\neg p$); but the implication in the second conjunct holds: if Sam is late, Susie has a tantrum.\textsuperscript{14}

### 3.5 Modal weakening

Here we have claimed that in declarative sentences of the form $p$ otherwise $q$, a speaker asserts both $p$ and if $\neg p$ then $\Box q$. Notice that a consequence of this proposal is that it predicts the redundancy of otherwise-sentences with non-modalized antecedents like those in (18) below:\textsuperscript{15}

(18) \textit{Non-modalized antecedents should lead to infelicitous otherwise statements, given (16), but they are acceptable:}

a. Jan is home, otherwise she’s breaking curfew. (10-b)
b. Sam is always home by 6pm, otherwise little Susie has a tantrum. (17)

On the surface, both of these cases ought to be infelicitous: if I assert that, in the actual world, Jan is home, then asserting the conjoined proposition that \textit{If Jan isn’t home} (in the actual world) \textit{then she’s breaking curfew} ought to be judged as redundant. Similarly, if I’m willing to assert that \textit{Sam is always home by 6pm}, then the claim that Susie has a tantrum shouldn’t be verifiable in the actual world.\textsuperscript{16}

In both of these cases, the felicity of the otherwise clause appears to require the accommodation of a weakened $p$. For (18-a), note that in contexts where the speaker has direct perceptual access to the subject, the sentence is severely degraded. (18-a’) is infelicitous unless the speaker can be interpreted to have incomplete knowledge of where they are.\textsuperscript{17}

(18) a’. ??I’m home right now, otherwise I’m breaking curfew.

Consequently, we take it that while the speaker of a sentence like (18-a) is willing to confidently assert $p$, their addressee accommodates information about their evidence base for this assumption on the basis of their willingness to admit of an alternative.

---

\textsuperscript{14} Some speakers we have consulted find our example with ‘always’ difficult to process, and prefer a variant with ‘often’. The same point could be made with such an alteration, but we find our variant in the text even more striking. See section 3.5 for a relevant discussion.

\textsuperscript{15} We thank Kai von Fintel for discussing the context of this section with us, and Bob Beddor for pointing us toward some very helpful references.

\textsuperscript{16} This follows from a Stalnakerian view, where, by asserting $p$, we are proposing to eliminate all non-$p$ worlds from the Common Ground (e.g. Stalnaker 1979).

\textsuperscript{17} Compare von Fintel & Gillies’ 2010 treatment of epistemic must, the (evidential) use conditions of which are met in this scenario (viz. inference/indirect evidence.)
By virtue of a similar pragmatic mechanism, the interpretation of (18-b) involves accommodating a weakened assertion of $p$. The speakers we have consulted appear go about this in two different ways, paraphrased below:

(18)  
b’. Sam is **usually** home by 6pm, *otherwise* little Susie has a tantrum.

b''. **These days,** Sam is always home by 6pm, *otherwise in the past,* when he was sometimes late, little Susie **would have** a tantrum.

By weakening the quantificational force of the adverbial (18-b’) or restricting the domain to a stage-level predication (18-b’’), real-world alternatives to ‘Sam BE home by 6pm’ are made available. Both repairs allow for a non-empty complement set of worlds for *otherwise* to refer to, satisfying the non-emptiness requirement we discussed above.

Conversely, weakening is not necessary when we have an imperative or an interrogative antecedent, as both types of clauses by their nature always allow for a non-empty complement set of worlds: an addressee may fail to act on a command, admitting both worlds that satisfy the command and those that don’t; likewise, polar and alternative questions presuppose more than one possible answer, requiring a partition with non-empty cells.18 Along similar lines, when the prejacent to *otherwise* is counterfactual, the non-emptiness requirement can be satisfied without weakening the antecedent:

(19)  
I’m home right now, *otherwise* I’d be breaking curfew.

In the next section, we build on these observations about the nature of *otherwise* to develop an analysis rooted in dynamic semantics, and making use of the information structural notions of the Question under Discussion.

---

18 Notice that a prediction that follows from this discussion here is that *wh*-questions will not serve as felicitous antecedents for *otherwise*. Although constituent questions have been argued to impose a partition over the possible worlds in the context (e.g. Groenendijk & Stokhof 1984), there will not be a complement set for *otherwise* to refer to:

(i)  
  a. ??Who wants to present first in the seminar? *Otherwise* Max will.
  b. ??Where do you want to go? *Otherwise* we can stay home.

Unlike in the case of declaratives, we are not able to offer a repair such as modal weakening, and instead the examples are judged as marginal. For some speakers, the negations of these questions’ presuppositions — *viz.* ‘if there is noone who wants to present’ and ‘if there is nowhere that you want to go’ — are retrieved (sc. accommodated) with some amount of effort. We discuss other cases of infelicity due to a lack of a non-empty complement set in section 5.
4 Analysis

Here we draw on tools from the dynamic semantics and information structural literatures to model otherwise’s semantic contribution to a sentence. Section 4.1 introduces Discourse Representation Theory (Kamp 1981, Heim 1982), and in particular the notion of “modal subordination” (Roberts 1989, 1995, 2004, 2012). Section 4.2 lays out our proposal for the semantics of otherwise. It discusses previously unremarked limitations on the distribution of otherwise, and shows that they naturally follow from the modal subordination analysis we lay out. Finally, section 4.3 illustrates our proposal for the pragmatics of otherwise, and in particular how information structural notions (notably, the Question under Discussion) can be recruited to provide a treatment of otherwise as a discourse anaphor. An appendix to the paper provides a more detailed formal definition of modal subordination in the context of otherwise.

4.1 Background: Discourse representation & modal subordination

As we have seen, a key property of otherwise is its interpretational flexibility, which we have characterized as going beyond what is strictly contributed by the syntax of the utterance it is contained in. A number of authors have proposed dissociated syntactic and semantic notions of “subordination” (e.g. Yuasa & Sadock 2002, De Vos 2007, Culicover & Jackendoff 1997), noting the ostensible independence of these modules. In particular, Craige Roberts’ (1989) “modal subordination” formalism provides a way of capturing this dissociation, and consequently of explaining the different interpretations of otherwise in the Red Light sentences (1).

Roberts (1989) adapts Discourse Representation Theory (DRT), developed in Kamp (1981), in order to formally implement a notion of subordination which operates independently of the syntax (i.e. where even in the absence of a conventional trigger, the interpretation of some quantificational operator is restricted.)

(20) MODAL SUBORDINATION is a phenomenon wherein the interpretation of a clause \( \alpha \) is taken to involve a modal operator whose force is relativized to some set \( \beta \) of contextually given propositions. \( \) (Roberts 1989: 718)

In effect, modal subordination provides a way of understanding the relationship between sentence mood and the nature of an assertion in context. It operationalizes the insights of work on the structure of natural language quantification (i.e. the conception of modalized sentences as generalized quantifiers that relate ‘restrictor’ and ‘scope’). An illustrative example is provided in (21).

(21) An example of modal subordination in discourse:
a. If Edna forgets to fill the birdfeeder, she will feel very bad.
b. The birds will get hungry. (Roberts 1989: 683)

Notice that the birds need not get hungry (an entailment of (21-b), if it were to act as a standalone assertion) for the entire discourse to be true. Instead, (21-b) is *modally subordinate* to (i.e. its interpretation is dependent on) the conditional antecedent in (21-a). Therefore, only in a context in which the antecedent conditions in (21-a) are met must the consequent condition in (21-b) also be satisfied.

We take statements involving *otherwise* to rely on a similar logic. As we have seen, the syntactic form of *otherwise* sentences underdetermines their interpretation. Appealing to MODAL SUBORDINATION allows us to identify the relationship between the linguistic signal and its likely interpretation. Roberts (1989: 712–5) provides a formal syntax and semantics for modal interpretations of DRSs. The pertinent details are presented here.\(^1\)

Next we provide a basic overview of how to interpret the “box diagrammatization” of Discourse Representation Structures (DRSs), familiar from Roberts 1989, Kamp 1979, 1981, Partee 1984, a.o. These visualization conventions are associated with a formal language (the Discourse Representation Language), which is sketched in the appendix to this paper.

For a given DRS \(K\), \(K\) denotes a pair \(\langle X_K, C_K \rangle\), where \(X\) represents the *local domain* – a finite set of variables that represent discourse objects relevant in the context (including participants, eventualities, and times); and \(C\) is a finite set of ‘satisfaction conditions’ that eventually determine the truth value of a given proposition. For diagrams where a DRS \(K\) is represented as a box, the top of the box represents the variables \(X_K\) and the bottom represents the satisfaction conditions \(C_K\). For a simple discourse as in (22)–(23), we provide a DRS below. Notice that the indefinite is treated as a variable here, and is eventually existentially closed (Heim 1982). DRT allows us to continue to refer to a variable introduced in the prior discourse, as long as it is still accessible:

(22) A dog entered the room. (23) It barked.

\[
\begin{array}{|c|}
\hline
x \\
\hline
dog(x) \\
entered-room(x) \\
\hline
\end{array}
\begin{array}{|c|}
\hline
x \\
\hline
dog(x) \\
entered-room(x) \\
barked(x) \\
\hline
\end{array}
\]

\(^{1}\) An appendix to this paper provides some additional technical detail. The interested reader is referred to Roberts (1989) for a closer reading about the formal apparatus of modal subordination. See also Kamp (2017: 47–58) for a detailed formal presentation of a DRL that handles temporal relations.
Conditions of the form $P(x_1, \ldots , x_n)$ (where $P$ is an $n$-place predicate) are closed under the operations $\neg , \lor , \Rightarrow , \Box , \Diamond$. This is illustrated for a simple example below:

(24) If a dog is hungry, Pedro might feed it.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{dog}(x)$</td>
<td>$\text{Pedro}(y)$</td>
</tr>
<tr>
<td>$\text{hungry}(x)$</td>
<td>$\text{feed}(y, x)$</td>
</tr>
</tbody>
</table>

Crucial to the theory is the notion of an “accessible domain” $A_{K_i}$ – a superset of the local domain for any given $K_i$. As a discourse proceeds, the set of objects that can be referred to expands. The notion of ‘accessibility”, then, allows us to predict which objects can be referred to at a given stage in a discourse.

(25) The accessible domain $A_{K_i}$ contains all the variables that occur:
   a. In $K_i$’s local domain ($X_{K_i}$)
   b. In the domains of all DRSs that graphically contain $K_i$
   c. If $K_i$ is the right element of a (binary) modal condition ($\Rightarrow , \Box , \Diamond$), $A_K$ also contains all the elements of the antecedent’s (the DRS on the left’s) local domain.
      I.e. $K_{\ell} \Box K_i \rightarrow K_{\ell} \lesssim K_i$ where $\lesssim$ reads “is accessible from.”

In (24), we observe that the consequent box of the conditional makes reference to a variable introduced in the antecedent. We furthermore notice that the entire conditional statement is embedded inside a larger discourse, so that we are not committed to the existence of any dog in the context.

Based on the assumptions introduced in (25), a given DRS $K$ is *modally subordinate* to all those DRSs whose domains it has access to. Example (26) illustrates such a case, from Roberts (1989: 701). Here, the consequent clause is *modally subordinate* to the antecedent; the entire conditional is taken to assert that ‘you will find “John” at home reading a book’ *in those worlds* where “John” bought a book (further relativized to a Kratzerian context set – i.e. a modal base and ordering source). Like in (25), we need not be committed to the fact that John bought a book in the actual world; in other words, the entire statement is not a part of the matrix DRS $K$, but rather further embedded.
In (26), the DRS representing the consequent clause ($K_j$) is *modally subordinate* to its antecedent $K_i$ and, as a result, can access the individuals and eventualities in $K_i$ (i.e. $K_i \leq K_j$). Moreover, both $K_i$ and $K_j$ are subordinate to the matrix DRS $K$ (i.e. $K \leq K_i \leq K_j$); had any variables been introduced in $K$, they would have been accessible to $K_i$ and $K_j$.

### 4.2 Proposal: A dynamic semantics for *otherwise* and the role of discourse

With this background in mind, we are now ready to propose a semantics for *otherwise*. Specifically, we define an operator over DRSs $\ominus$ (and hence the condition $K_i \ominus K_j$) to represent the contribution of *otherwise*:

\[
(27) \quad \text{Proposal: A dynamic semantics for otherwise} \\
K_i \ominus K_j \iff (K_i) \land (\neg K_i \Box K_j) \\
\text{In words: } K_i \ominus K_j \text{ is satisfiable iff both } K_i \text{ and } (\neg K_i \Box K_j) \text{ are satisfiable.}
\]

This proposal can be paraphrased as claiming that: “the conditions in $K_i$ (should) hold”; however, in case that I am incorrect in my claim, the conditions in $K_j$ must then hold.” (Recall our discussion of *modal weakening* in section 3.5.) Notice that this treatment takes *otherwise* to be akin in its structure to a conditional, referencing our informal description in (16) that an otherwise statement can be paraphrased as $p \land (\neg p, \text{ then } \Box q)$. Moreover, we build an asymmetric conjunctive element into the analysis, building on the observations in section 3, especially surrounding example (17).

Notice additionally that we use Roberts’ necessity modal operator, building on our observation in (10) that *otherwise* appears to be a modalized operator. The se-
mantics of Roberts’ modal operator □ builds in a modal base m and ordering source o in order to capture the observations made by Kratzer (1981: §2.7) regarding different “flavors” of circumstantial modality.20

The satisfaction conditions (including the role to be played by a contextually retrieved modal base) for the condition in (27) are made explicit in the appendix. Although, as (27) makes clear, ⊖ can be expressed in terms of other defined operators (viz. ¬, □ and →), we use the ⊖ shorthand notation for convenience.

Notice further that the definition in (27) leads to the following accessibility relations:

\[(K_i \ominus K_j \in C_K) \rightarrow (K \leq K_i \leq K_j)\]

It follows from (28) that the condition \(K_i \ominus K_j\) entails that \(K_j\) is modally subordinate to \(K_i\) (and both are subordinate to the broader context that they are embedded in).

In section 4.2.1 immediately below, we show how our proposal for otherwise can model sentences like the Red Light example, and how it makes use of the notion of modal subordination. In section 4.2.2 we then show that adopting Roberts’ notion of the accessible domain into our proposal successfully predicts the range of possible antecedents for otherwise, and in particular previously unnoted limitations on its distribution. In section 4.3 we then discuss the role of discourse and pragmatics in calculating the antecedent to otherwise, and in the process introduce the notion of the Question under Discussion.

4.2.1 The proposal in action

We begin with the otherwise clauses in (1-a) and (1-b), repeated in (29). As in section 4.1, each DRS is represented as a box: its ‘local domain’ (the discourse objects that it introduces) at the top and its condition set at the bottom. These two examples differ in terms of the material that acts as the prejacent to otherwise. In (1-a), the DRS representing the prejacent to otherwise is modally subordinate to the entire conditional statement represented in the left box in (29-b). In (1-b), the DRS representing the antecedent to otherwise is modally subordinate to the antecedent of the if-clause.

20 For the sake of exposition, we abstract away from appeal to the ordering source, although this can easily be added back in following the semantics in Roberts (1989: 714). In effect, Roberts treats the material conditional as a modal with a singleton (i.e. “totally realistic”) modal base. See also Kratzer 1986 for another modalized analysis of conditionals, and e.g. Cantwell 2008, von Fintel 2011, Yalcin 2012 for arguments for preferring such an analysis to a material implication theory of conditionals.
(29) a. If the light is red, stop. Otherwise go straight on. = (1-a)
b. A DRT analysis of (1-a)

(30) a. If the light is red, stop. Otherwise you’ll get a ticket. = (1-b)
b. A DRT analysis of (1-b)

We provide a second illustration using example (11), repeated below. Again, the antecedent can lead to two distinct otherwise statements. The DRSs in (31) show how modal operators rely on the accommodation of material that is made available by preceding discourse elements. In the (31-a), the consequent clause is relativized to worlds where students attend the lecture. In (31-b), it is relativized to worlds in which students are obliged to attend the lecture (i.e. the entire modalized sentence is accommodated.)

(11) Students are required to attend the lecture, otherwise...

a. ≈ If ¬ (they ATTEND)... ...they’ll fail the class.
b. ≈ If ¬□ (they ATTEND)... ...it’ll be empty.
Different continuations mean that different antecedents are accommodated:

a. The non-modalized statement served as the prejacent to otherwise:

\[
\begin{array}{c}
\square \\
\begin{array}{c}
\text{x, y} \\
\text{students(x)} \\
\text{lecture(y)} \\
\text{attend(x, y)}
\end{array}
\end{array}
\oplus \\
\begin{array}{c}
\text{fail(x)}
\end{array}
\]

b. The modalized statement served as the prejacent to otherwise:

\[
\begin{array}{c}
\square \\
\begin{array}{c}
\text{x, y} \\
\text{students(x)} \\
\text{lecture(y)} \\
\text{attend(x, y)}
\end{array}
\end{array}
\oplus \\
\begin{array}{c}
\text{empty(y)}
\end{array}
\]

Once again, then, the choice of prejacent for otherwise varies between the two examples, and cannot be determined from the preceding syntax alone. Instead, the consequent clause plays a crucial role in the reasoning about the set of worlds under consideration in the evaluation of an otherwise-sentence, as we will discuss in detail in section 4.3.

4.2.2 Constraining the choice of antecedent

In this section, we further illustrate the utility of the notion of modal subordination in our analysis. To do so, we turn our attention to the available readings of otherwise. As we have seen, the choice of an antecedent for otherwise isn’t guided deterministically by the syntax. Nonetheless, its selection is not unconstrained. To
see this, consider example (32), which is judged as infelicitous with the intended reading of “If the light is red, stop. If it isn’t red, it will be green.”

(32)  #If the light is red, stop; *otherwise* it’ll be green.

    INTENDED ≈ If the light is *not* red...

This is crucially predicted by the modal subordination account described in this section. The “accessible domain” of otherwise will contain precisely those DRSs which can be felicitously accommodated as its antecedent. For sentences with conditionals such as the Red Light examples, the accessible domain will contain the content of either the consequent clause of the conditional or the entire conditional sentence — but there is no way to choose the antecedent of the conditional to the exclusion of the consequent. But this is precisely what would be required for (32) to be felicitous; its infelicity follows naturally from the use of modal subordination in our analysis.

To this point, compare the sentences in (33), adapted from Kruijff-Korbayová & Webber (2001: 76). Whereas (a) can be shown to encounter similar interpretation problems to (32), it is vastly improved in (b) when the relevant if-clause receives focus and associates with only.

(33)  a.  #If the light is red, stop; *otherwise* you’ll get rear-ended.

    INTENDED ≈ If the light is not red and you do stop...

b.  Only if the light is RED, stop; *otherwise* you’ll get rear-ended.

In fact, as shown by McCawley (1974), Barker (1993) and von Fintel (1994, 1997), the felicity of (33-b) follows naturally from a standard semantics for only, where only is taken to assert the negation of alternatives to its prejacent. These authors show that the truth-conditional content of only if can be derived compositionally (i.e. as a function of the standard semantics of only and if), where the assertive content of q only if p is modelled as \( \neg p(w) \rightarrow \neg q(w) \) (that is, q holds in no worlds other than those in which p does).

(34)  Presuppositional and assertive components of (33-b):

    Only if the light is RED, stop; otherwise you’ll get rear-ended.

    Presupposes: If the light is red, you stop.

    Asserts: If the light is {yellow, green}, you don’t stop. If you do stop, you get rear-ended.23

21 Speakers consulted frequently cited a reading where failing to stop at a red light would cause it to change color. This is predicted by our account.

22 Formally, only presupposes \( P(x) \) and asserts that for all alternatives \( y \) distinct from \( x \): \( P(y) \) is false.

\[
\text{[only]} = \lambda P . x : P(x) \land \neg \exists y (y \neq x \land P(y))
\]


23 Assuming here that red, yellow, and green are the contextually relevant alternatives to red.
This can then be schematized as in (35), where the otherwise clause is modally subordinate to all other content in the utterance, and interpreted as predicating into ‘the set of non-red-light worlds in which you stop.’

(35) A DRT analysis of the otherwise clause in (33-b)
If the light isn’t red, don’t stop; if you do stop, you’ll get read-ended.

This phenomenon is shown additionally in (36), where either all three conjoined clauses or the final conjunct can be easily accommodated as an antecedent proposition to otherwise. The other conjuncts are not accessible antecedents for otherwise in this context. Again, this is precisely what is predicted from the modal subordination account.

(36) You should have a snack, chill out for a bit, and then you should go to the gym, otherwise you’ll feel bad later on.

Having established that the choice of prejacent of otherwise is not syntactically determined but is rather constrained by the notion of the accessible domain, we now turn back to the role of context in determining the nature of the chosen prejacent of otherwise.

4.3 Proposal: otherwise as a discourse anaphor

As the preceding sections make clear, there is often more than one possible choice for the prejacent of otherwise. How is the prejacent chosen, then? We propose that the set of worlds that the complement set of which otherwise operates on is cal-
culated pragmatically from the prior discourse and the nature of the consequent clause.\textsuperscript{24,25}

By deploying the information structure notions proposed in Roberts (1996/2012), we can conceptualize of otherwise as representing a DISCOURSE MOVE (in effect, a stage in a given discourse), which adds to the QUESTION UNDER DISCUSSION in a given context.

(37) Two useful definitions:

a. The common ground is a set of mutually assumed background information. The cg is often modeled as a set of propositions, i.e. a set of sets of possible worlds (e.g. Stalnaker 1979).

b. The QuD is a partially structured set of questions which discourse participants are mutually committed to resolving at a given point in time. It is often modeled as a stack, consisting of ordered subsets of accepted question moves, the answers to which are not entailed by the cg (i.e. a set of “open” questions in the discourse at a given time.)

With these concepts, we have a means of representing the ‘flow’ of information and changes in the interlocutors’ information states over time. We take a sentence of the form \( p \) otherwise \( q \) to consist of (at least) three discourse moves. We propose that otherwise represents a discourse “setup” move with the effect of adding to the QuD.

(38) Proposal: the pragmatics of otherwise

Otherwise represents a discourse “setup” move with the effect of adding to the QuD stack a question about the COMPLEMENT of the set of worlds established elsewhere in the discourse.

The importance of this pragmatic aspect of our analysis is illustrated for example (39) below.

(39) \([\text{You must eat}]_{m_i}, \text{otherwise}_{m_j} [\text{you won’t grow!}]_{m_k}\)

\(m_i\) This clause represents a modalized assertion: in all worlds in some unspecified (here, likely teleological) conversational background \(f\), the addressee eats.

\[\forall w' \in \cap f : \text{EAT}(\text{Addressee})(w')\]

\(24\) This claim bears some similarity to the notion of a “anaphorically-derived contextual parameter” that features in the analysis of Webber et al. (2001: 14).

\(25\) Relatedly, Corblin (2002) notes the possibility of negative accommodation without otherwise in \textit{I didn’t buy the car. I wouldn’t have known where to put it (otherwise) and I should have accepted. I wouldn’t have been fired.} (our translations: 256, 258).
otherwise represents an instruction to consider the COMPLEMENT of some set of worlds established elsewhere in the discourse. This can be thought of as signaling the addition of a question to the QUd stack of the form:

what if we are in some \( w \in \text{COMP}(p) \)?

In other words, what if we are in a world in which the addressee doesn’t eat?

The consequent clause encodes the prejacent to otherwise. It is to be interpreted as proffering a (partial) answer to the current question under discussion by making a (modalized) assertion that the addressee won’t grow in the set of worlds picked out by otherwise (viz. the complement of the set of worlds in which the addressee eats).

\[ \forall w', w'' \in \text{COMP(} \text{EAT(Addressee)} \text{)} \rightarrow \neg \text{GROW(Addressee)}(w'') \]

As we know, the process of establishing the context set for a given otherwise sentence is underdetermined by the syntax of the sentence. We dub this the “RED LIGHT PUZZLE”, repeated in (40). Recall that the syntactic antecedents of the red light sentences are identical (hence \( m_i, m_j, m_k \) represent the same operation in each sentence), but they appear to constrain the interpretation of otherwise in markedly different ways, \( m_a \) vs \( m_b \).

(40) **THE RED LIGHT PUZZLE**

a. [If the light is red,] \( m_i \) [stop] \( m_j \) otherwise \( m_k \) [keep going!] \( m_a \)

b. [If the light is red,] \( m_i \) [stop] \( m_j \) otherwise \( m_k \) [you’ll get a ticket!] \( m_b \)

We provide an Information-Structure based analysis for (40-a) and (40-b). We first consider the two discourse moves in the if-clause, viz. \( m_i \) & \( m_j \).

(41) \( m_i \) The if-antecedent temporarily constrains the context set (Roberts 1989: 687). It adds a question to the QUd stack of the form:

what if we are in \( \{ w' \mid \text{RED.LIGHT} \in w' \} \)?

\( m_j \) Imperative stop represents an answer to QUd(\( m_i \)). As with the antecedent in (39), we treat it as a modalized proposition (again with some conversational background \( f \))\(^{26}\) which further restricts the domain established by \( m_i \).

\[ \forall w'', w'' \in \text{RED.LIGHT} \cap f \rightarrow \text{STOP(Addressee)} \in w'' \]

\( \)

\(^{26}\) See Portner (2007) a.o. for a modal treatment of imperative sentences.
As per our proposal, *otherwise* marks the addition of a question to the QU-D stack which considers what would happen if we were *not* in a world introduced in the prior discourse:

(42) The *otherwise* discourse move:

\[ m_k \]  

*Otherwise* represents an instruction to consider the complement of some set of worlds established elsewhere in the discourse. 

what if we are in some \( w \in \text{compl}(p) \)?

Given the salience of \( w' \) and \( w'' \) in (41), which have been added to the \( cg \) in \( m_i \), \( m_j \) respectively, both are possible candidates to form the set that *otherwise* builds on. The Addressee is thus required to infer which discourse move *otherwise* is anaphoric upon (*i.e.* its antecedent.), based on the content of the consequent. We dub this the *jeopardy!* effect: The addressee is given the consequent (=the answer) and must compute the correct antecedent (=question) based on it:

(43) The *JEOPARDY!* effect

\[ m_a \]  

*keep going* is interpreted as an answer to *what if we are in* \( \text{compl}(\text{red.light}) \)?

Here the propositional variable is saturated by the partition evoked in \( m_i \)

\[ \forall w''.w'' \in \text{compl}(\text{RED.LIGHT}) \rightarrow \text{KEEP.GOING}(w'') \]

\[ m_b \]  

*get a ticket* is interpreted as an answer to *what if we are in* \( \text{RED.LIGHT} \setminus \text{STOP} \)? (I.e. the complement of \( \text{STOP} \) relative to \( \text{RED.LIGHT} \))

Here a sub-partition (within the set of “red light worlds”) evoked in \( m_j \) saturates the propositional variable.

\[ \forall w''.w'' \in \text{RED.LIGHT} \cap \text{compl}(\text{STOP}) \rightarrow \text{GET.TICKET}(w'') \]

Our claim, then, is that computing the antecedent of *otherwise* is a pragmatic process, subject to reasoning by the addressee and depending on the given context in which the sentence is uttered.\(^{27}\) This follows from the pragmatic stipulation that, in a discourse, assertions represent ‘at least partial answers [...] to the question under discussion at the time of utterance’ Roberts (2012: 20–21).\(^{28}\) Broadly, the discourse contribution of *otherwise* can be understood as representing a “set-up move”: it signals to the addressee that its prejacent is to be understood as a modal claim, relativized to the complement of a set of worlds established elsewhere in the discourse context.

\(^{27}\) This makes predictions for online sentence processing — for example, that a given reading could be primed or ruled out by supporting contexts. We leave this for future work.

\(^{28}\) In fact, this effectively serves as a reformulation of Grice’s maxim of Relation, adapted for an information-structural framework.
5 Non-emptiness and possibility modals

Given that, by our analysis, otherwise requires reference to a set of “eliminated worlds” — the complement set of worlds to that introduced by the prejacent of otherwise — it follows that a sentence of the form $\alpha$ otherwise $\beta$ will be uninterpretable in discourses in which no worlds have been eliminated (i.e. where $\overline{\alpha} = \emptyset$). This principle is given in (44), and reflects the non-emptiness requirement we observed in section 3.2.

(44) **EXCLUSION: a felicity condition for otherwise**

The interpretation of otherwise $\beta$ depends on the retrieval of some antecedent discourse move $\alpha$ whose function was to eliminate a set of worlds $\overline{\alpha}$ from consideration.

Otherwise $\beta$ predicates $\beta$ of $\overline{\alpha}$.

In this section we show two consequences of this requirement for the interpretation of otherwise in modalized sentences.

5.1 Unambiguous scope

A sentence like Sam may not be a doctor is ambiguous between circumstantial and epistemic readings. This observation notwithstanding, the contrast between (45) and (46) further demonstrates the interpretive constraints that otherwise is subject to — namely, that it must be able to refer to a non-empty complement set of worlds, computed on the basis its antecedent and other components of the context. To illustrate this, consider the two contexts below, designed to support the circumstantial and epistemic readings, respectively, in the context of an otherwise statement:

(45) **CONTEXT.** Sam got horrible grades in school and is very clumsy

a. She may not be a doctor, otherwise... $\neg \gg \diamond_{\text{circ}}$

b. $\approx$ If she were (to become) a doctor... ...she might kill someone.

(46) **CONTEXT.** Sam works in a hospital and wears a white coat; I’m unsure what exactly it is that she does.

a. She may not be a doctor, otherwise... $\diamond_{\text{epist}} \gg \neg$

b. INTENDED $\approx$ If she is a doctor... ?...she’s probably a surgeon.

A crucial difference between the circumstantial (45) and epistemic (46) readings of the antecedent is the scope relation between the modal and negative operator. Just like in (12) above, otherwise is only licit if it can predicate into a non-empty set of worlds. In the $\neg \gg \diamond$ case, we can successfully achieve this result. But in the $\diamond \gg \neg$ case, where no worlds are eliminated, otherwise is unavailable. Given otherwise’s
observed infelicity with possibility readings of *may*, the epistemic reading is ruled out, leaving only the circumstantial one available.

5.2 Epistemic strengthening

A second, related result concerns so-called ‘weak necessity’ readings of possibility modals (Rubinstein 2012, von Fintel & Iatridou 2008).

The modals *ought* and *should* are described as encoding “weak” necessity, distinguishing them from other modal necessity expressions (e.g. *have to* and *must*.) Two examples demonstrating the relation between weak and strong necessities from von Fintel & Iatridou (2008: 117) are provided below.

(47) **Weak and strong necessity:**
   a. You *ought to* do the dishes but you don’t *have to*.
   b. #You *must* do the dishes but you don’t *have to*

(48) a. You *ought to* wash your hands – in fact, you *have to*.
   b. ?You *have to* wash your hands – in fact, you *ought to*.

Additionally, as with other modals, *ought* appears to admit of ambiguity between epistemic and circumstantial (e.g. deontic) readings, as shown in (49).29

(49) **Weak necessity and modal flavors:**
    Morris *ought to* be in his office. (von Fintel & Iatridou 2008: 116)

In view of the co-occurrence constraints on epistemic possibility modals with *otherwise*, compare the two sentences (both judged as acceptable) in (50) below.

(50) **A felicitous epistemic possibility modal with otherwise:**
   a. She *must* be sick, otherwise she’d be here.
      \[ \alpha = \{ w \in f_{epist} \mid \text{She is sick in } w \} \]
      \[ \overline{\alpha} = \{ w \in f_{epist} \mid \text{She is not sick in } w \} \]
   b. She *might* be sick, otherwise she’d be here.

The first clause of (50-b), uttered in isolation, asserts the existence of epistemically-accessible worlds in which the subject is sick (i.e. \( \exists w' \in f_{epist}, \text{She is sick in } w' \)). In such a case, as we have seen in the foregoing sections, a pure possibility reading is not available because no possible worlds have been excluded from consideration...
(that is, as far as the speaker is concerned, the subject may or may not be sick). Consequently, the felicity condition for otherwise as laid out in (44) is not met.

This problem is repaired here by strengthening the meaning of might, so that it is now interpreted as excluding possible worlds (requiring that it function as a universal quantifier: a hallmark of necessity modals). While the intended interpretation of (50-b) is weaker than that of its counterpart in (50-a), it can still be understood as quantifying universally over possible worlds, albeit over a more restricted set.

Following von Fintel & Iatridou (2008: 116), “while strong necessity modals say the prejacent is true in all of the favored worlds, weak necessity says that it is true in all the very best (by some additional measure) among the favored worlds.” With respect to the epistemic domain specifically, the difference could be understood as the difference between relativizing the prejacent to “hard and fast evidence” and “unreliable assumptions about the normal course of events.” Consequently, we propose the paraphrases below:

\[(51) \quad \text{With otherwise the possibility modal is strengthened to weak necessity:} \]

\[a. \quad \text{She must be sick, otherwise she’d be here.} \]
\[\approx \text{In all worlds consistent with what I know, if she is not sick, she’d be here.} \]

\[b. \quad \text{She might be sick, otherwise she’d be here.} \]
\[\approx \text{In all worlds consistent with my perception of her general behavior, if she is not sick, she’d be here.} \]

The finding that might/may — generally understood as encoding modal possibility — are encoding weak necessity suggests that the felicity conditions of otherwise coerce a non-canonical interpretation of these modals. This result follows from our proposal in section 4.2, that some non-empty set of worlds must be available for otherwise to predicate of.

\[\text{30 von Fintel & Iatridou (2008) and Rubinstein (2012) model weak necessity by appealing to at least one additional (“secondary”) ordering source. In the current case, the secondary ordering source might be described as stereotypical (g, e) following Kratzer (1981). Adopting this analysis, the relevant sets of worlds for (50-b) are:} \]

\[(i) \quad a. \quad \alpha = \{w \in \text{max}_{g, e}(f_{\text{epis}}) \mid \text{She is sick in } w\} \]
\[b. \quad \overline{\alpha} = \{w \in \text{max}_{g, e}(f_{\text{epis}}) \mid \text{She is not sick in } w\} \]

Where \(\text{max}_g(f)\) is a function that returns the “best” worlds in a set of worlds \(f\) as determined by an ordering source \(g\) (i.e. those worlds in \(f\) best conforming to the ideal contained in \(g\)).
6 Intra-sentential otherwise and complement anaphora

So far, the data we have focused on in this paper has comprised uses of otherwise that appear to signal a relation between clauses. We have claimed that, in these cases, otherwise adds a question of the form what if the antecedent proposition doesn’t hold? to the QuD stack. Nevertheless, as shown in section 1, intra-sentential uses of otherwise — namely, those which coordinate smaller structures — are also available. In this section, we briefly show how our analysis can be extended to account for such uses. We then relate our analysis to the phenomenon of complement anaphora, which has also benefitted from an analysis within a dynamic semantic framework.

6.1 Otherwise with donkey anaphors

A key advantage of DRT is in providing an analysis of so-called Donkey Sentences, such as in (52):

(52) Donkey anaphora:
    a. If a woman is rich, she owns a donkey.
    b. If a dog is hungry, Pedro will feed it. = (24)

Such sentences were famously used as a counter-examples to Montague’s formal analysis of quantification in natural language (1973), as they defy an analysis in first-order predicate logic.31 As we saw in section 4.1, DRT is able to provide a natural account, treating indefinites as variables rather than existential quantifiers (see Kamp 1981, Heim 1982). This is exemplified again in (53):

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>woman(x)</td>
<td>donkey(y)</td>
</tr>
<tr>
<td>rich(x)</td>
<td>owns(x, y)</td>
</tr>
</tbody>
</table>

One payoff of the approaches espoused by these authors is the conception of universal expressions as complex conditions of the form K_i → K_j, where K_i and K_j

31 A formula can be given, but only if the indefinite is translated using a universal quantifier — an arguably undesirable result.
are sub-DRSs representing the restriction and the scope of the quantified statement, respectively (Roberts 1989: 693-4).

Appealing to these same notions, we are able to naturally account for some intra-sentential uses of otherwise, as in (54) from Webber et al. 2001: 7, repeated here for convenience:

(54) Intra-sentential otherwise:

a. Every person selling “The Big Issue” might otherwise be asking for spare change.  

\[ X = \sum_{x} \text{person}(x) \]

b. 

\[
\begin{align*}
\text{sell.the.Big.Issue}(x) & \quad \Leftrightarrow \quad \text{ask.for.spare.change}(x)
\end{align*}
\]

c. In all worlds in which a person \( x \) isn’t selling the Big Issue, it’s possible that the person \( x \) is asking for spare change.

For Webber et al. 2001, example (53) requires the use of E-type pronouns. It thus receives a different analysis than inter-sentential uses such as (1). Our account, on the hand, doesn’t resort to any additional assumptions, and does not predict any distinction between such examples. We take this to be another advantage of our approach here.

6.2 “Intrapredicative” otherwise

Expanding on examples such as (53), in this section we investigate intra-sentential uses of otherwise (termed intra-prédicative by Flament-Boistrancourt 2011). We show how such cases can be united with the analysis presented above. The examples in (55) illustrate several relevant cases:
“Intrapredicative” otherwise:

a. I started meditating to find a bit of stillness in an otherwise hectic life.

b. The income they earn from [tea production] is likely to be the only source of cash to supplement their otherwise subsistence economy.  

(c) OED

c. Amelia behaved well otherwise. (Flament-Boistrancourt 2011)

d. She’s blonde. Otherwise she totally looks like her dad.  

(Inkova-Manzotti 2002: 124)

Observe that all of these uses are united insofar as they rely on processes of association (contextual retrieval of some domain set) and the exclusion of the complement of the prejacent from that set (see Webber et al. 2001).

For the intrapredicative uses shown here, then, otherwise can be understood to denote a relation that holds between properties \( P \subseteq D_{(x,(e,x))} \). Namely, where \( P \) is some accommodated property, otherwise \( Q \) can be understood as a property where if \( P \) didn’t hold of \( x \) in \( w \), then \( Q \) would. Building on our proposal in section 4, then, we would allow the (complement) set of worlds predicated of by otherwise to be constructed not only by considering a proposition (or set of propositions) and its negation, but also by considering a property (or set of properties) and its negation. In both cases, otherwise is to be understood to quantify over intensions. We leave the precise formulation of this extension to our analysis to future research.

6.3 Complement anaphora

Finally, we point out similarities between our analysis of otherwise and the phenomenon of complement anaphora, exemplified in (56) (Evans 1977, 1980, Nouwen 2003). Complement anaphora occurs in sentences where an anaphor appears to refer to the complement of a set of individuals introduced earlier in the discourse:

(56) **Complement anaphora:**

Few congressmen admire Kennedy.

a. They are (all) very junior.  

b. They think he’s incompetent.  

\[ A \cap \overline{B} \]

Some speakers struggle with the complement anaphora reading. The existence of complement anaphora was first extensively studied in a series of psycholinguistic experiments (Moxey & Sanford 1986, Sanford et al. 1994). These authors identify a small set of proportional determiners, including few, few, very few, not many, and hardly any, as allowing reference to the complement of a set of individuals introduced earlier in the discourse.
Moreover, although to our knowledge this has not been previously noted in the literature, we find similar effects in the temporal domain:\footnote{Such effects may be predicted by the discussion of ‘generalized discourse subordination’ effects of temporal quantifiers (Roberts 1989: 716ff, Corblin 1994: 8).}

\begin{enumerate}
\item \textit{Complement anaphora in the temporal domain:}
\begin{quote}
Senators rarely vote their conscience. They do what the Party tells them to.
\end{quote}

Building on Kibble 1997, \textit{Nouwen (2003)} develops a dynamic semantic analysis of complement anaphora, where reference to a complement set of individuals arises out of pragmatic constraints, key among them is the Non-Emptiness constraint:\footnote{See Corblin 1986 and Geurts 1997 for an alternative account whereby sentences described as involving complement anaphora in fact make reference to the \textit{maximal set}, and not truly to the complement set. \textit{Nouwen 2003} provides several arguments against this \textit{pseudo-reference} view.}

\begin{enumerate}
\item \textbf{NON-EMPTINESS:}
\begin{quote}
As the antecedent of an expression do not choose a set which is potentially empty, except when this set is the reference set of a quantificational sentence.
\end{quote}

Parallel to this proposal, we have argued that \textit{otherwise} picks out a complement set of worlds, and is subject to the exclusion felicity condition, (\textit{44}). We take \textit{otherwise} to lexically specify complement set reference, which is therefore not subject to the same pragmatic constraints as complement anaphora. We take (\textit{59}) to be a felicitous paraphrase of a sentence such as (\textit{56-b)}:

\begin{enumerate}
\item \textit{Complement anaphora with otherwise:}
\begin{quote}
Very few congressmen admire Kennedy. \textit{Otherwise} they (all) think he’s incompetent.
\end{quote}

\textit{Otherwise} encodes the instruction to consider a complement set of worlds as part of its semantics. As a consequence, \textit{otherwise} sentences are not marginal and are not subject to the same distributional restrictions as complement anaphora. This observation is similar to an observation \textit{Nouwen (2003: 109ff)} makes about the phrase ‘the others’:

\begin{enumerate}
\item \textit{Complement anaphora with ‘the others’:
\begin{quote}
Very few congressmen admire Kennedy. \textit{The others} (all) think he’s incompetent.
\end{quote}

As \textit{Nouwen} notes, \textit{the others} refers to the \textit{maximal set} of individuals which forms the complement to the set introduced in the antecedent sentence. This use is felicitous in cases where this complement set is necessarily non-empty. Again, the
resulting sentence, like in our otherwise examples, is then predicated of all individuals in this set. See also Corblin (1994, 2002) for a discussion of relativisations négatives (“negative accommodation”) in a modal subordination framework, which he takes as clear evidence of the need to appeal to some pragmatic phenomenon.

7 Conclusion & further work

In this paper we developed a formal semantic/pragmatic analysis of the interpretation and meaning contribution of the English discourse anaphor otherwise. The analysis was couched within the theory of dynamic semantics, and in particular relied on the notion of modal subordination for predicting the distribution of otherwise in English sentences.

We proposed that otherwise introduces a discourse move (in the sense of Roberts 2012) into the conversation, which encodes an instruction to consider the complement of a set of worlds introduced in the clause preceding otherwise. That is, otherwise introduces a modalized assertion which claims that proposition \( p \) holds (to the best of the speaker’s knowledge), and that in case that it didn’t hold, then some alternative proposition \( q \) must be true: \( (p) \land (\neg p \Rightarrow \Box q) \). We detail the intensional/modal-dependent property of otherwise, its asymmetric conjunctive behavior, and the weakening process affecting declarative antecedents in section 3.

An interesting consequence of our analysis is that otherwise imposes a restriction on the nature of its arguments; namely the NON-EMPTINESS of that complement set into which it predicates. In section 3.2, we empirically motivated this felicity condition; section 5 detailed a number of its consequences.

Following Webber et al. (2001) and other authors, we took as key the observation that the identity of the antecedent clause to otherwise cannot be determined by the syntax alone (the so-called Red Light puzzle). Instead, we argued that we

35 Ezra Keshet (pers. comm.) points out a related similarity between the others and otherwise. The others can pick up the members of the restrictor set not including the current individuals being quantified over:

(i) Few/Most boys ganged up on the others.
   (cf. *Few/Most boys ganged up on them)

In such configurations, otherwise is also available. In the examples below, otherwise picks up the worlds other than the winning or cheating worlds.

(ii) a. If you win, you’ll be happier than (you would have been) otherwise.
    b. If you cheat, you’ll always wonder if you could have succeeded otherwise.

This point is also addressed by Webber et al. (2001: 8).

36 For Corblin (2002: 260) the solution is found in relations from Rhetorical Structure Theory like EVIDENCE and JUSTIFY (apud Mann & Thompson 1988).
must make crucial reference to the current information structure, and in particular to the current Question under Discussion, to determine the nature of the antecedent. We dubbed this phenomenon the Jeopardy effect: the nature of the consequent to otherwise plays a crucial role in its determination of the antecedent.

This proposal allowed us to model both otherwise’s flexibility and previously unobserved limitations on its distribution. Specifically, we showed that otherwise can only be predicated of sentences that it is modally subordinate to, limiting the antecedents available to it in the discourse. We moreover introduce the exclusion criterion, whereby otherwise must refer to a non-empty set of worlds; in cases where the antecedent does not eliminate any worlds from consideration, an otherwise continuation is infelicitous.

Finally, we briefly showed how this dynamic account can naturally be extended to cases of reference to individuals, and in particular how it can be related to the phenomenon of complement anaphora.

Authors’ address

Josh Phillips & Hadas Kotek
370 Temple Street
New Haven, CT 06511
josh.phillips@yale.edu

Appendix: Modal subordination with otherwise: the formal mechanics

In this appendix, we provide further detail about the “discourse representation language” that formalizes the structures (and the satisfaction conditions for ⊓) presented above. Further, we show a complete derivation for an “otherwise”-sentence as a ‘proof-of-concept’ for our analysis.

As described in §4.1, formally a DRS $K$ is a pair $\langle X_K, C_K \rangle$. $X_K$ represents $K$’s local domain – a finite set of variables that are assigned to discourse objects at a given discourse stage. Consequently, each DRS can be thought of as introducing participants (represented by variables over the domain of individuals) as well as variables over eventualities and times (per Kamp’s (1979, 2017) treatment of temporal/aspectual phenomena, see also Partee 1984).

$C$ is a finite set of conditions that eventually determine the truth value of a given proposition. An atomic condition is of the form $P(x_1, \ldots, x_n)$ (where $P$ is an $n$-place predicate). Conditions are closed under the operations $\neg, \lor, \Rightarrow, \Box, \lozenge$.

Crucially, Roberts (1989: 713) also defines the notion of an “accessible domain” $A_K$ – a superset of the local domain for any given $K$. Accessibility is a partial order that obtains over DRSs such that for any $K$:
(61) Accessibility relations for operators and DRSs in DRT:
\[ K_i \lor K_j \in C_K \quad \rightarrow \quad K \leq K_i; K_j \]
\[ \neg K_i \in C_K \quad \rightarrow \quad K \leq K_i \]
\[ K_i \Rightarrow K_j \in C_K \]
\[ K_i \Box K_j \in C_K \]
\[ K_i \lozenge K_j \in C_K \]
\[ \rightarrow \quad K \leq K_i \leq K_j \]

The accessible domain of a given DRS, then, is given by the set union of all accessible DRSs’ local domains: \( A_{K_i} = \bigcup_{K \leq K_i} X_K \). As pointed out in §4.1, this relation is graphically represented in the box diagrams.

The primary payoff of this conceptualization is the epiphenomenal notion of modal subordination (Roberts 1989 et seq), where subordinate DRSs depend on access to objects introduced by (i.e. in the local domains of) those DRSs to which they are subordinate:

(62) Modal subordination is a phenomenon wherein the interpretation of a clause \( \alpha \) is taken to involve a modal operator whose force is relativized to some set \( \beta \) of contextually given propositions. (Roberts 1989: 718)

In (27), we defined the otherwise \( \ominus \) (and hence the condition \( K_i \ominus K_j \)) to represent the contribution of otherwise. In effect, \( \ominus \) can be expressed in terms of other defined operators (i.e. \( \lor, \neg, \Box \)). We repeat this proposal in (63).

(63) Proposal: A dynamic semantics for otherwise
\[ K_i \ominus K_j \leftrightarrow (K_i) \land ((\neg K_j) \Rightarrow \Box K_j) \]
In words: \( K_i \ominus K_j \) is satisfiable iff both \( K_i \) and \( ((\neg K_j) \Rightarrow \Box K_j) \) are satisfiable.

Consequently, \( K_i \ominus K_j \in C_K \rightarrow K \leq K_i \leq K_j \). As shown in §4.2.2, Roberts’ accessibility relation between DRSs successfully predicts the range of possible antecedents for otherwise.

In her extension to the discourse representation language, Roberts (1989: 714-5) provides a recursive definition of truth (i.e. verification in a model \( M \)) for DRSs. Given in (64), effectively, truth in a model is defined for a DRS \( K \) with respect to a world if there is some assignment function that satisfies all of the conditions in \( K \) in that world (recalling that \( K \) itself is a pair including a condition set \( C_K \)).

(64) \( \langle w, f \rangle \models_M K \leftrightarrow \forall c \in C_K (\langle w, f \rangle \models_M c) \)
A DRS \( K \) is verified (or “embedded”) in a model \( (\models_M) \) relative to a world \( w \) and assignment \( f \) iff all the conditions in \( K \) are satisfied (\( \models \)) by \( w \) and \( f \).
Phillips & Kotek

Roberts spells out a semantics for the satisfaction of all (atomic and non-atomic) conditions in $C_K$. Extending this, we can define a semantics for the $\ominus$ operator. The satisfaction conditions for $K_i \ominus K_j \in C_K$ are given in (65). These monotonically-growing assignment functions formally model the accessibility relation $\preceq$ described above, upon which modal subordination hinges. Effectively, they ensure that any modally subordinate DRS will be able to refer to (“access”) superordinate structures.

The formalism in (65) spells out the satisfaction conditions inherent in (63), assuming the notational conventions and adapting the proposals in Roberts (1989: 714). It makes use of a function $\text{max}$ (sometimes rendered $\text{best}$) which returns those worlds in a given set $m \subseteq \mathcal{W}$ (the modal base) which best conform to a given ordering source $o$ (i.e. contextually provided set of propositions inducing an order over $m$). Note that the notation $f'\langle X \rangle f$ reads: “$f'$ is exactly the same as $f$ except perhaps for the values it assigns to $X$” (implying that $f' \supseteq f$).

(65) **DRL formalization of $\ominus$ satisfaction conditions:**
\[
\langle w, f \rangle \models (K_i \ominus K_j) \iff \exists g [g \langle x_{K_i} \rangle f \land \langle w, g \rangle \models K_i] \land \\
\forall w', g'[g' \langle x_{K_i} \rangle f \land w' \in \text{max}_{o(w)} \{m(w) \cup \{w'' \mid \langle w'', g' \rangle \models \neg K_i\}\} \\
\implies \exists h [h \langle x_{K_i} \rangle g' \land \langle w', h \rangle \models K_j]\]
\]

A world $w$ and assignment $f$ satisfy the condition $K_i \ominus K_j$ iff:

- There is some assignment $g$ that satisfies $K_i$;
- If the worlds $w'$ — those in the modal base $m(w)$ best conforming to some ordering source $o(w)$ — verify the **negation of** $K_i$, then there will be an assignment $h$ that **verifies** $K_j$ in $w'$.

A DRT representation for one of the now-familiar red light examples is spelled out in (66). Alongside this representation, we list the set of satisfaction conditions introduced by the sentence.

37 The deployment of a function $\text{max}$ (given sometimes by other authors as $\text{best}$) significantly compresses the formalism given in Roberts (1989: 714, which follows Kratzer 1981). Given that an ordering source $o$ is modelled as a set of propositions which can induce an ordering $\leq_o$ ‘relative to $o$’, at least as good as’ over a given set of worlds. Consequently, $\text{max}_{o(w)}(m(w))$ returns $\{w' \in m(w) \mid \forall u \in m(w), w' \leq_{o(w)} u\}$.

38 In Roberts’ formalism, $f_{(X):g} \leftrightarrow \forall y (\neg (y \in X) \rightarrow f(y) = g(y))$ (1989: 714).
A formal DRT analysis of an otherwise sentence:
If the light is red, Jake will stop. Otherwise he’ll continue straight.

With the satisfaction conditions we introduced above, we can construct the truth-conditions that will verify the matrix DRS $K$:

(67) **Satisfaction conditions for (66):**

a. The matrix condition:

A world-assignment pair $\langle w, f \rangle$ verifies the entire DRS $K$ iff it satisfies the (complex) condition $K_i \ominus K_\ell$ :

$$\langle w, f \rangle \vDash (K_i \ominus K_\ell) \iff \exists g [g(x_{K_i})f \wedge \langle w, g \rangle \vDash K_i] \wedge$$

$$\forall w', g'[g'(x_{K_i})f \wedge w' \in \max_{o_{tel}(w)}(m_{circ}(w) \cup \{w'' \mid \langle w'', g'' \rangle \vDash (\neg K_i)\})]$$

$$\rightarrow \exists h (h(x_{K_\ell})g' \wedge \langle w', h \rangle \vDash K_\ell)$$

That is: $\langle w, f \rangle \vDash K$ iff:

- There is some assignment $g$ that verifies $K_i$ and
- If those worlds $w'$ in a circumstantial modal base $m_{circ}(w)$ that best conform to a teleological ordering source (likely one that contains Jake’s desires to both get where he needs to be and to be an upstanding road user) verify the negation of $K_i$, then there’ll be some assignment $h$ that verifies $K_\ell$

b. The antecedent to otherwise:

The antecedent $K_i$ is verified iff some world-assignment pair $\langle w, f \rangle$ satisfies the (complex) condition $K_j \square K_k$ :

$$\langle w, f \rangle \vDash (K_j \square K_k) \iff \forall w', g[g(x_{K_j})f \wedge w' \in \max_{o_{tel}(w)}(m_{circ}(w) \cup \{w'' \mid \langle w'', g'' \rangle \vDash K_j\})]$$

$$\rightarrow \exists h (h(x_{K_k})f \wedge \langle w', h \rangle \vDash K_k)$$

That is: $\langle w, f \rangle \vDash K_i$ iff for all $w'$ in a circumstantial modal base $m_{circ}(w)$ that best conform to a teleological ordering source $o_{tel}(w)$ if $w'$ verifies $K_j$ there’s some assignment $h$ that verifies $K_k$. 

35
c. **Simplex conditions:**

The remaining three DRSs — $K_j, K_k, K_\ell$ all contain only atomic conditions. Each of these DRSs is verified iff there is some world-assignment pair $\langle w, f \rangle$ which satisfies all of their respective conditions.

E.g.: $\langle w, f \rangle \models K_\ell \iff \langle w, f \rangle \models \text{continue}(y) \iff f(y) \in \text{\texttt{continue}}^w$

Notably, the $y$ is an unbound variable — however, because $K_\ell$ is modally subordinate to $K_i$, it has access to the local domain of this DRS ($X_{K_\ell} \supseteq A_{K_i}$). As a result, the assignment function ($h$ in (67-a) above) is able to assign to $y$ an individual introduced earlier in the discourse (namely ‘Jake’). We see, then, that our analysis is able to correctly model an *otherwise* statement, making crucial use of the notion of modal subordination and the tools of dynamic semantics to provide the truth conditions for the sentence.

**References**


Phillips & Kotek


38


