ROOTS OF MODALITY

A Dissertation Presented

by

AYNAT RUBINSTEIN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2012

Department of Linguistics
To Eshel, Imma, and Yanir
ACKNOWLEDGMENTS

This dissertation is a landmark in a journey that started far away, on a mountain that is always green on the other side of the ocean. I am grateful to my parents, who raised me with love and wisdom, and an unwavering encouragement to pursue my dreams. I am grateful for my brother, who has been a pillar of strength throughout this journey. And I am grateful to the Nirim for making sure to send Eshel to the Scouts. I would not have made it to this moment without Eshel, and I thank him for every year and every hour of being together.

I wish to express my deep and heartfelt gratitude to my advisor, Angelika Kratzer, and to Rajesh Bhatt, Kai von Fintel, and Robert Rothstein, my committee members. First and foremost, I thank Angelika – for being a wonderful teacher and guide, for challenging and encouraging me constantly, and for the many hours we spent together talking about modality. I will miss these dearly. I thank Rajesh for his optimism, and for making this study not just about English and not just about semantics. I thank Kai for his interest in this project and for inspiring discussions that taught me a great deal about how semantics should be done. Finally, I thank Bob for being such a careful reader, and for teaching me about Yiddish.

I am grateful to have been raised intellectually in the village that is the Department of Linguistics at UMass Amherst. I thank Chris Potts for being a source of inspiration and encouragement since my very first semester in Amherst and, later, across the country from the West Coast. I thank Lyn Frazier for being a true teacher and mentor, and for sharing her knowledge and enthusiasm whenever I came through her door. To Valentine Hacquard I am indebted for teaching an exciting proseminar on modality and for supporting my first steps
in this field. John Kingston, Kyle Johnson, Peggy Speas, and Tom Roeper are to thank for playing important roles in shaping my broader interests as a linguist.

I am especially grateful to Paul Portner – for reading parts of the dissertation and providing thoughtful comments on the material in Chapter 2, and for inspiring conversations about mood and modality. For stimulating discussions when they passed through South College, I thank Malte Zimmermann, Anna Szabolcsi, Philippe Schlenker, Roger Schwarzschild, María Arche, Pranav Anand, Floris Roelofsen, Hans Kamp, Chris Kennedy, Lisa Matthewson, Ana Arregui, and Kyle Rawlins.

My fellow students and colleagues are to thank for keeping my spirits up intellectually and otherwise during this journey. Thank you Amy Rose Deal, Annahita Farudi, Chris Davis, Karen Jesney, Kathryn Pruitt, María Biezma, and Masashi Hashimoto, for being such a great cohort. Thank you Cherlon Ussery, Andrew McKenzie, Florian Schwarz, Shai Cohen, Jan Anderssen, Ilaria Frana, Anne-Michelle Tessier, Keir Moulton, Mike Key, Helen Majewski, and Michael Becker for sharing your “higher class” and wisdom. Thank you Kathy Adamczyk, Tom Maxfield, and Sarah Vega-Liros for lending a hand and an ear on so many occasions. Thanks to Ruth Kramer for sharing the best coffee there is in D.C. as the final parts of the dissertation were written, and to Valentine Hacquard, Elena Herburger, and Graham Katz, who joined in. I am also fortunate to have crossed academic and personal paths with Paula Menéndez-Benito, Magda Oiry, E. Allyn Smith, Ofra Magidor, Reut Tsarfaty, Nora Boneh, Galit W. Sassoon, and Jennifer Rau. I hope our paths will cross again.

During my graduate studies, I had the fortune to participate in two ESSLLI summer schools. Courses with Denis Bonnay, Craige Roberts, and Judith Tonhauser were particularly important in broadening my horizons as a semanticist. In the Fall of 2009, I had the unique opportunity to visit the Academy of the Hebrew Language in Jerusalem and the Technion in Haifa. I thank Doron Rubinstein from the Academy and Alon Itai from the Technion for their hospitality and for providing me with excellent research environ-
ments throughout my visit. I gratefully acknowledge a fellowship from the Department of Linguistics at UMass that made this visit possible.

This work has roots in semantics and in discourse that can, and should, be traced back to Tel Aviv University. I am deeply grateful to Fred Landman and Rachel Giora, my teachers and mentors in Tel Aviv, and to Mira Ariel, Yoad Winter, Tal Siloni, Julia Horvath, Chuck Kisseberth, and Shuly Wintner, for supporting my entry into the world of linguistics.

It took my father less than ten minutes to realize that Amherst and Northampton come pretty close to being paradise on Earth. In the past years, I met wonderful people and made good friends in the Valley. Thank you to Tim and Caroline for a first home in the Hadley farmlands, to Tom Sadlowski for the best food and conversation, to Ram and Kim, Joe, Arun, Misato, Kathryn and Ryan, Annahita, and Lena for their friendship. Thank you, Cherlon, for always being there and for reminding me how to navigate this world with taste. Thank you to my parents’ friends, who are now my own, for coming to visit and being my family away from home. To Sam Bahn, Beth Latz, Ephraim, Sherry, and the extended Isaac-Rosen clan. And finally, Anat, Noa, Limor, Amir, Einat, Lilach, Dahlia, Nitsan, Naama, and Yuval – it’s been way too long. The next leg of this journey will hopefully bring us closer together again.
ABSTRACT

ROOTS OF MODALITY

SEPTEMBER 2012

AYNAT RUBINSTEIN

B.Sc., TEL AVIV UNIVERSITY
M.A., TEL AVIV UNIVERSITY
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Angelika Kratzer

This dissertation explores the interplay of grammar and context in the interpretation of modal words like ought, necessary, and need. The empirical foci of the discussion are patterns in the use of strong and weak necessity modals in conversation, and the interpretation of syntactically and semantically versatile modals like need in the various grammatical configurations they appear in across languages.

It is argued that a sensitivity to collective commitments in a conversation is necessary for understanding certain aspects of modal strength, in particular the traditional distinction between strong and weak necessity modals (exhibited by must and ought to in English). It is proposed that strong necessity modals can only reference priorities that are presupposed to be collectively committed to, whereas weak necessity modals are evaluated with respect to a mixed bag of priorities, crucially including ones that are presupposed not to be collectively committed to. A domain restriction approach to weak necessity is adopted, following a demonstration that it is superior to a number of probabilistic alternatives.
Modal verbs and adjectives that take both infinitival and nominal complements are shown to pattern alike across languages in requiring a teleological, or goal-oriented interpretation when their complements are not infinitives (but rather noun phrases or certain Complementizer Phrases). This limitation is lifted with infinitival complements, showing that transitive configurations of certain intensional verbs are not semantically equivalent to the infinitival configurations of the same verbs.

A result of this research is a fine grained analysis of the differences between closely related necessity modals and attitude verbs.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
</tbody>
</table>

## CHAPTER

1. **INTRODUCTION** ................................................................. 1
   1.1 What it means for a necessity modal to be weak ................. 1
   1.2 Microvariation among necessity modals ......................... 4
   1.3 Grammatical restrictions on the expression of priority modality .... 5

2. **WEAK NECESSITY MODALS AND MODAL DISCOURSE** ................. 8
   2.1 Grades of necessity .................................................. 8
   2.2 Theoretical approaches to weak necessity ...................... 10
      2.2.1 The comparative probabilistic approach and its limitations .... 14
         2.2.1.1 Comparing propositional alternatives ...................... 18
         2.2.1.2 Comparing alternatives as particulars .................. 22
         2.2.1.3 A decision theoretic analysis ......................... 26
         2.2.1.4 Summary and conclusions ............................. 32
      2.2.2 The domain restriction approach ............................. 36
         2.2.2.1 The challenge of distinguishing between priorities .... 39
         2.2.2.2 Cascades of priorities .............................. 48
         2.2.2.3 Summary ............................................ 50
   2.3 Motivating sensitivity to collective commitment .............. 51
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.1</td>
<td>Necessity modals and the expression of priorities in directives</td>
<td>55</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Necessity modals and rising declaratives</td>
<td>61</td>
</tr>
<tr>
<td>2.3.3</td>
<td>What is the collective whose commitments matter?</td>
<td>70</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Necessity modals and conversational backoff</td>
<td>73</td>
</tr>
<tr>
<td>2.4</td>
<td>Weak necessity as departure from collective commitment</td>
<td>77</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Formal proposal</td>
<td>77</td>
</tr>
<tr>
<td>2.4.1.1</td>
<td>Discussion of motivating examples</td>
<td>83</td>
</tr>
<tr>
<td>2.4.1.2</td>
<td>Inference patterns</td>
<td>86</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Performativity and strong necessity</td>
<td>88</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Secondary priorities versus conditional antecedents</td>
<td>91</td>
</tr>
<tr>
<td>2.4.4</td>
<td>Conditional application of priorities: a brief comparison</td>
<td>95</td>
</tr>
<tr>
<td>2.4.5</td>
<td>A different perspective on the modal landscape: negotiable norms,</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>non-negotiable norms, and facts</td>
<td></td>
</tr>
<tr>
<td>2.4.6</td>
<td>Conclusions and future directions</td>
<td>101</td>
</tr>
<tr>
<td>3.1</td>
<td>Root modalities and attitude predicates</td>
<td>104</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Attitude verbs of desire</td>
<td>104</td>
</tr>
<tr>
<td>3.1.2</td>
<td>A conditional comparative analysis: Heim (1992)</td>
<td>104</td>
</tr>
<tr>
<td>3.1.3</td>
<td>Multiple alternatives and less reliance on belief: Villalta (2006,</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>2008)</td>
<td></td>
</tr>
<tr>
<td>3.1.4</td>
<td>Refuting the argument for multiple alternatives</td>
<td>111</td>
</tr>
<tr>
<td>3.2</td>
<td>From verbs of desire to modal predicates</td>
<td>117</td>
</tr>
<tr>
<td>3.2.1</td>
<td>The comparative analysis extended to modals: Villalta (2006,</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>2008)</td>
<td></td>
</tr>
<tr>
<td>3.2.2</td>
<td>Necessary that there be a goal</td>
<td>120</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Want, necessary, and ought: similarities and differences</td>
<td>126</td>
</tr>
<tr>
<td>3.2.3.1</td>
<td>Necessary is a strong necessity modal</td>
<td>126</td>
</tr>
<tr>
<td>3.2.3.2</td>
<td>Preliminary lexical entry for necessary</td>
<td>127</td>
</tr>
<tr>
<td>3.2.3.3</td>
<td>Conclusion</td>
<td>130</td>
</tr>
<tr>
<td>4.1</td>
<td>A grammatical constraint on the interpretation of root modals</td>
<td>132</td>
</tr>
<tr>
<td>4.1.1</td>
<td>The basic observation: need to versus need</td>
<td>135</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Syntactic and semantic versatility</td>
<td>138</td>
</tr>
<tr>
<td>4.1.2.1</td>
<td>Deontic and teleological modalities</td>
<td>139</td>
</tr>
</tbody>
</table>
4.1.3 Experimental findings .............................................. 142
  4.1.3.1 Participants .................................................. 142
  4.1.3.2 Materials .................................................... 142
  4.1.3.3 Procedure ................................................... 146
  4.1.3.4 Results ...................................................... 147
  4.1.3.5 Discussion .................................................. 151

4.2 English (for)-to infinitives and the Expressive IP hypothesis ............ 154

4.3 Crosslinguistic data and the Lexical Ambiguity hypothesis ............... 157

4.4 Full CPs pattern with nominals ....................................... 168

4.5 Descriptive generalizations: subject orientation or complement types? ...................................................... 175
  4.5.1 The syntax of need .............................................. 176
  4.5.2 Subjectless constructions and the source of priorities ............ 183
  4.5.3 Conclusion ..................................................... 189

APPENDICES

A. A KRATZERIAN CONTEXT DEPENDENT POSSIBLE WORLDS SEMANTICS FOR MODALS .......................................................... 191
B. TRIPLE RELATIVITY: NEGOTIABLE NORMS, NON-NEGOTIABLE NORMS, AND FACTS .......................................................... 198

BIBLIOGRAPHY ............................................................... 200
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>A decision problem with two actions and ( n ) relevant possible worlds.</td>
</tr>
<tr>
<td>2.2</td>
<td>Decision problems and weak necessity</td>
</tr>
<tr>
<td>2.3</td>
<td>Presuppositions about commitment to priorities</td>
</tr>
<tr>
<td>3.1</td>
<td>Microvariation among necessity modals and attitudes of desire</td>
</tr>
<tr>
<td>4.1</td>
<td>Linear mixed model for Experiment 1</td>
</tr>
<tr>
<td>4.2</td>
<td>Results of Experiment 2’ (responses to comprehension questions)</td>
</tr>
<tr>
<td>4.3</td>
<td>Logit mixed model for Experiment 2’</td>
</tr>
<tr>
<td>4.4</td>
<td>Syntactic structures for <em>need</em> with nonfinite complements</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Four ways of getting from Cambridge to Amherst, Mass.</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>The contextual domain restriction approach to weak necessity</td>
<td>36</td>
</tr>
<tr>
<td>2.3</td>
<td>Domain restriction with sensitivity to collective commitments</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td><em>(preliminary)</em></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Accessible worlds in desire statements</td>
<td>115</td>
</tr>
<tr>
<td>3.2</td>
<td>Goal-oriented <em>necessary</em></td>
<td>121</td>
</tr>
<tr>
<td>4.1</td>
<td>Experimental results comparing <em>need</em> and <em>need to</em> in deontic and</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>teleological contexts</td>
<td></td>
</tr>
<tr>
<td>A.1</td>
<td>Ordering worlds with a single unstructured ordering source</td>
<td>195</td>
</tr>
<tr>
<td>A.2</td>
<td>Ordering worlds with a single unstructured ordering source (continued)</td>
<td>196</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

This dissertation is about modality, that signature property of human language that allows us to speak not about the here and now, but about the realm of mere possibilities.\textsuperscript{1} There are many words that allow us to engage in modal talk: must and have to, should and ought, necessary, want, need, and many others. How are the meanings of these modal words similar? How are they different? What are all the factors that influence how they are used in conversation?

This dissertation takes on these questions in the domain of necessity modals: weak and strong, performative, desiderative, and syntactically versatile. It offers an in-depth analysis of weak necessity and a novel criterion to explain how weak and strong necessity modals differ. The sum of these investigations is not only a more accurate description of the differences between individual modal words, but a deeper understanding of the semantic and pragmatic building blocks that language makes available for the construction of modal meaning.

1.1 What it means for a necessity modal to be weak

The entry point to the dissertation is the question of what makes necessity modals like ought weaker semantically than necessity modals like have to. It can be true, for instance, that You ought to rent a car, but simultaneously true that You don’t have to rent a car. If we replace ought to with must, the conjunction of these two sentences amounts to a

\textsuperscript{1} Hockett (1966).
contradiction. *Ought* is therefore said to be “weak.” This weakness has proven difficult to formally pin down.

A leading intuition in the literature has been that *ought* is weak because it makes a necessity claim that characterizes a subset of the possibilities that *have to* does. For example, if we are talking about your upcoming road-trip and identify that renting a car is one option among several you can pursue, it is not true that you have to rent a car. But it is true that you ought to, if for example an additional consideration of independence is taken into account. A smaller domain of quantification results in a weaker statement, assuming that both *ought* and *have to* effect universal quantification over possibilities.

This approach requires the considerations that necessity modals are sensitive to to come in two types. Some considerations should be relevant for both *have to* and *ought*, and there should be certain considerations that are only relevant for *ought*. Above, the additional consideration of independence that is behind the claim that you ought to rent a car should not be a consideration that *have to* is sensitive to (otherwise the claim that you have to rent a car would be true). So we have a question: what defines which considerations the different modals are sensitive to?

In my discussion, I argue that a satisfactory answer to this question has not yet been given in the literature. In particular, it doesn’t seem to be possible to make the split between considerations on conceptual grounds, and neither is it possible to rely on the linguistic signal to provide one type of consideration to the exclusion of the other. Often, considerations that form the background for the modal claim are entirely implicit and contextually determined, so as a rule it won’t do to look to the linguistic signal for clues. A new dimension for distinguishing between modal backgrounds is needed.

My goal in the second part of Chapter 2 is to propose a new dimension for characterizing the contextually determined considerations that weak and strong necessity modals are sensitive to. I propose that the relevant dimension concerns the presupposed commitment of the participants in the conversation toward a given consideration, or priority. Strong
necessity modals are strong because they take into account considerations that all the conversational participants are presupposed to be committed to, and weak modals are weak because they rely on additional considerations that not all the conversational participants are presupposed to be “on board with.”

As initial motivation for sensitivity to collective commitment as a way of distinguishing between priorities, consider the different necessity claims supported by each of the following two scenarios.

(1) [Rachel is coming to the United States next summer. It is now illegal not to have health insurance in the US.]

   a. [Rachel will be a graduate student in a respectable university.]
      She has to ($\l$ should) get health insurance.

   b. [Rachel will be employed illegally, selling cheap jewelry at a shopping mall.]
      She should ($\l^\text{False}$ has to) get health insurance.

In both scenarios, the consideration of following the law is salient and relevant. In (1a), it licenses a strong necessity claim. In (1b), it only licenses a weak one. Why is this the case? When we probe the judgment that a strong necessity claim is false in the scenario in (1b), the reason seems to be that it is not taken for granted that considerations of lawful conduct apply in the case of illegally employed workers. In contrast, such considerations are presupposed to guide the actions of those affiliated with a respectable university. We see that moving to a conversation in which there is no commitment to the crucial priority appears to affect the choice of necessity modal.

A number of independent probes of commitment status are argued to support the proposed split between priorities.

An intuitive way to think about the split is in terms of what is and is not negotiable in a conversation. The considerations that a strong necessity modal are sensitive to are presupposed to be collectively committed to, and thus not-negotiable. The weakness of an
ought claim, on the other hand, is due to the modal’s reliance on an additional consideration that is not committed to across the board in the conversation, hence still negotiable.

What are the consequences of this proposal for the way modal quantification domains are constructed? In the seminal theory of modality proposed by Kratzer (1981, 1991, 2012), there are two kinds of modal backgrounds that are used to construct these domains: a “modal base” and an “ordering source.” With the proposal that priorities come in two varieties, there are now three kinds of modal backgrounds to work with: the relevant circumstances, the relevant non-negotiable priorities, and the relevant negotiable priorities. The study of weak necessity thus offers a new perspective on the role of priorities in building modal quantification domains. This perspective may prove useful for the analysis of other modal expressions that are similarly sensitive to negotiable and non-negotiable priorities.

1.2 Microvariation among necessity modals

The family of necessity modals in English consists of additional members beyond the exemplary weak and strong necessity modals I opened the discussion with. In Chapter 3, I turn to examine the semantics of a third necessity modal, necessary, and an attitude verb of desire, want. The subtle differences between ought, have to, necessary, and want paint an intricate picture of how expressions of modal necessity may vary.

Necessary is a modal that is often thought of as a general purpose necessity modal, on a par with have to. This belief, which may be due to the specialized use of the modal in philosophical texts, is nevertheless incorrect as a characterization of how necessary is used in everyday language. Necessary seems instead to be primarily a teleological, or goal-oriented, modal; it demands a goal-oriented interpretation even when one is not readily supplied by context.

(2) a. I have to sneeze.

b. It is necessary for me to sneeze.
Thus it appears that *necessary* bears greater family resemblance to a bouletic attitude verb like *want*, or even a weak necessity modal like *ought*, than it does to a general purpose modal like *have to*, although both *necessary* and *have to* are strong (in the sense that they only take into account priorities that enjoy collective commitment in the conversation).

It is difficult to account for these similarities and differences with a toolbox that contains just “force” and “flavor” as parameters of variation between modals. The addition of “strength” as a third parameter of variation, understood in terms of sensitivity to collective commitment, may allow for a better understanding of the meaning of a larger class of necessity modals.

An informal summary of the microvariation in the domain of necessity modals that is discussed in Chapter 3 is shown in the Table below. *Need*, which is listed twice, presents a further challenge for classification. It is grouped here once with *necessary* and once with *have to*, for reasons that have to do with the type of modality it expresses in two different syntactic configurations it appears in.

<table>
<thead>
<tr>
<th>Modal</th>
<th>Presupposition</th>
<th>Truth conditional contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>have to, need</em>$_{IP}$</td>
<td>(no presupposition)</td>
<td>$\forall w': w' \in \max_{g(w)}(f(w)).q(w')$</td>
</tr>
<tr>
<td><em>necessary</em>, <em>need</em>$_{DP}$</td>
<td>new primary priority restricts the domain</td>
<td>$\forall w': w' \in \max_{g(w)}(f(w)).q(w')$</td>
</tr>
<tr>
<td><em>ought to</em></td>
<td>secondary priority restricts the domain</td>
<td>$\forall w': w' \in \max_{g(w)}(f(w)) \cap b.q(w')$</td>
</tr>
<tr>
<td><em>want</em></td>
<td>attitude holder’s desires restrict the domain</td>
<td>$\forall w': w' \in \max_{DES}(f(w)).q(w')$</td>
</tr>
</tbody>
</table>

*Obligatory teleological interpretation.

### 1.3 Grammatical restrictions on the expression of priority modality

The empirical focus of Chapter 4 is the expression of priority modality by necessity modals like *need*, in English and in other languages. In terms of the strength of necessity it expresses, *need* groups with strong necessity modals like *have to* and *necessary*. In other words, it is sensitive to priorities that are presupposed in the conversation to be collectively
committed to, or ones that are forced into collective commitment following acceptance of the modal claim. What makes it unique in comparison to other necessity modals is that its interpretation seems to vary depending on the syntactic configuration it appears in, or alternatively, the type of complement it takes.

The following scenario brings out the variable interpretation of *need*.

(3) [City regulations mandate that home owners put up fences between their properties. You and your neighbor get along very well without a fence. In fact, both of you object to a fence because it would have to go right on top of the beautiful flower beds that have been flourishing between your two properties. You say to your neighbor:]

a. False We need a fence.

b. False We need (for) there to be a fence.

c. True We need to put up a fence. (*according to city regulations*)

When its complement is a *to*-infinitive, the modal can express a wider range of necessities. In particular, it can express a necessity that is based on the content of a salient law. This is why (3c) is true in the scenario below. In the very same conversational context, the modal does not have access to this interpretation when its complement is a noun phrase, as in (3a), or a *(for)*-*to*-infinitive, as in (3b). The sentences in (3a)-(3b) cannot receive the interpretation that (3c) can, and are false in this scenario. Moreover, while adding a phrase like *according to city regulations* helps to bring out the deontic interpretation of (3c), this addition is somewhat awkward when *need*’s complement is a nominal, as in (4). Intuitively, the modal in (4) describes necessities that follow from our priorities, not necessities that follow from the what the law provides.

(4) *According to city regulations, we need a fence here.*

We thus say that *need* has a greater modal potential when its complement is an infinitival. Traces of this variability are characteristic of *necessary* as well.
The pattern in (3) is especially interesting because it uncovers a difference between two types of modal meaning that are conceptually very close: the deontic modality of what is possible or necessary given regulations or laws, and the teleological modality of what is possible or necessary given goals, priorities, or preferences. Both of these are traditionally grouped together as belonging to the class of priority modalities within the larger class of root, or non-epistemic modalities.²

It should not be set aside as a case of idiosyncratic ambiguity. The English pattern recurs in other languages (Hebrew and Hindi-Urdu are the main additional sources of data I present): necessity modals that are flexible in the types of complements they take in these languages are obligatorily teleological with nominal (as well as certain CP) complements, and exhibit a wider range of interpretations when their complement is an infinitival.

These data raise important questions about the syntax-semantics mapping in the domain of priority modals. How is the line between priority modalities drawn, and what makes deontic modality unavailable in certain grammatical configurations?

CHAPTER 2
WEAK NECESSITY MODALS AND MODAL DISCOURSE

We both know how this money should . . . nay, must be spent.

Bart, The Simpsons (1993)

This chapter proposes that differences traditionally labeled as differences in strength among necessity modals (exhibited by *must* and *ought to* in English) reflect a pragmatic distinction among the priorities these modals are sensitive to. Strong modals are argued to be sensitive to all and only the priorities that are presupposed to be collectively committed to in a conversation, whereas weak modals take into account additional priorities, ones that do not enjoy this status. The analysis builds on and improves on previous proposals, shedding light on the performativity of strong necessity modals and on the conceptual division of modal backgrounds representing negotiable and non-negotiable priorities.

2.1 Grades of necessity

Common among descriptions of modal systems of natural languages is the characterization of some modals as “weaker” than others in terms of the necessity or possibility they describe. In English, *must* and *have to* are called “strong” necessity modals. They are distinguished from their “weak” counterparts *ought to* and *should.*

The relationship between *must* and *should* is usually described as one of ‘degree’: both are said to express ‘obligation’, but this obligation is said to be weaker in the case of *should* than in the case of *must* (and weaker still in the case of *need*; cf. e.g. Leech 1971: 95). (Wierzbicka, 1987, p. 35)

1 There is some inconsistency regarding the classification of certain modals into these categories (compare, for instance, the contradictory claims regarding *need* in the two quotes in the text). It is necessary to explicate the notion of strength involved before any attempt at classification is made.

8
The expression of necessity may be one of a variety: beside *must* and *have* (to), *need*, *necessary*, *ought*, *should*, etc. Some, like *ought* and *should*, express a necessity that may seem weaker than *must.*

(Sæbø, 2001, p. 433)

The comparative strength metaphor for modals is rooted in scalar-like implications these modals give rise to. When entities are ordered along a dimension, the resulting scale supports certain well-known entailments and implicatures (Horn 1972, Levinson 2000). Higher elements on the scale are “informationally stronger” than lower elements on the same scale, so a higher element logically implies lower elements, and the negation of a lower element entails the negation of a higher element. A lower element is compatible with the negation of a higher and informationally stronger element. (5a) demonstrates this pattern with weak and strong necessity modals (and compare the contradiction resulting from substituting a strong modal in place of the weak one in (5b)).

(5) 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.

(von Fintel and Iatridou, 2008, 117(3-4))

A lower, or weaker, element in fact implicates the negation of higher elements. This is an implicature derived via Gricean reasoning, given the assumption that a cooperative and knowledgeable speaker would commit to the strongest possible entity on the scale if they could. As a conversational implicature, this inference is defeasible (6a).

(6) a. You ought to wash your hands – in fact, you have to.

b. ?You have to wash your hands – in fact, you ought to.

(von Fintel and Iatridou, 2008, 117(5))

Robust and widely assumed as it is, the perceived weakness of weak necessity modals has proven difficult to formally pin down. When the strength difference between them
is not at issue, weak necessity modals are simply analyzed as standard, strong, necessity modals.\(^2\)

### 2.2 Theoretical approaches to weak necessity

A fruitful intuition about the special meaning contribution of weak necessity modals originates in Aaron Sloman’s paper ‘Ought’ and ‘better’. Sloman (1970) made a connection between weak necessity and comparison, proposing that for a possibility \( p \) to be something that ought to be the case, \( p \) must be better than every alternative in a set of contextually determined possibilities. On the basis for comparing these alternatives, he wrote:

> In the simplest cases, the basis \( B \) will specify some condition, such as ‘promoting happiness’, or ‘getting to London within 2 hours’, and to say that something is better, or best, in relation to \( B \) will then be (roughly speaking) to say that it is more, or most, conducive to satisfaction of that condition. In more complex cases of comparison, ‘more conducive’ may be qualified by reference to other factors, e.g. ‘more conducive in respect \( R \), in circumstances \( S \)’.
> (Sloman, 1970, 389)

Discussions of weak necessity in the recent literature are couched within this comparative approach. They can be described as investigating the nature of the comparison invoked by weak necessity modals and, in particular, addressing the question of what makes one alternative “better” than another. In one line of work, von Fintel and Iatridou (2005b, 2008) capitalize on the idea that there is some condition (to use Sloman’s terms above) that separates the best alternative from the rest of the alternatives. On their approach, an alternative is better than another if and only if the first satisfies the relevant condition, and the second does not. We might call this a *qualitative approach* to comparison. In contrast, Finlay (2009, 2010) takes a quantitative approach. He explores the idea that alternatives are compared, roughly, based on how likely they are to lead to satisfaction of the relevant condition.

\(^2\) See, among many others, Arregui (2010, 243) on *should* (and references cited therein) and Cariani et al. (2011) on *ought*. 
I will refer to this second approach as the \textit{probabilistic comparison approach}. A variant of this approach assumes a quantitative representation of preferences in terms of utilities, in addition to a quantitative representation of likelihood (Goble 1996, Lassiter 2011b).

To clarify these two perspectives on weak necessity, consider the following example.

\textbf{(7) Driving to Amherst.} You are in Massachusetts, in Cambridge, and your goal is to get to Amherst. There are four ways to get from Cambridge to Amherst, all of which get you there in about 2 hours (the list is not ordered, route colors refer to Figure 2.1):

- Take the turnpike (I-90) all the way to I-91. (\textit{Green route})
- Take Route 9, through Worcester. Get to Route 9 via the turnpike. (\textit{Blue route})
- Take Route 9, through Worcester. Get to Route 9 without first driving on the turnpike. (\textit{Red route})
- Take Route 2, passing near the town of Concord. (\textit{Magenta route})

There is no teleological necessity in the scenario to take any one route in particular, thus it is false that you \textit{have to or must} do so. What about a weak necessity to take one of the four routes?
To go to Amherst, you ought to take Route 9.

Following Sloman (1970), taking Route 9 could be a weak necessity if it is more conducive than the other options for going to Amherst. Under a qualitative interpretation of “more conducive,” the ought claim would be true if there is a further condition that is relevant for the choice of route, and Route 9 satisfies that further condition. Under a quantitative interpretation, the ought claim would be true if taking Route 9 has a greater chance of getting you to Amherst than any of the other routes. Suppose it happens to be the case that Route 2 is not available on the relevant day (perhaps the road is closed for construction). Two out of the three remaining options involve taking Route 9, and presumably there is a probabilistic advantage for this option under these circumstances. On the other hand, Route 9 and the turnpike (the live options under the circumstances) have equal likelihood of getting you to Amherst, which means that there is no probabilistic advantage for any of them. The predictions depend on how precisely the probabilistic comparison is carried out.

The first part of this chapter is devoted to a detailed evaluation of the quantitative and qualitative approaches to weak necessity. In Section 2.2.1, I discuss the comparative probabilistic approach. I take as a starting point the motivating discussions in Finlay (2009, 2010), and present two implementations of his general approach. The two implementations result in two different lexical entries for ought, but both fall short of accounting for true preference-based weakness that is a hallmark of how the modal is used in conversation. In Section 2.2.1.3, I demonstrate that the empirical problems facing probabilistic comparison also pose a challenge for a decision theoretic approach to weak necessity of the kind proposed by Goble (1996) and developed recently by Lassiter (2011b).

I then present the qualitative approach and raise a conceptual challenge it faces regarding individuation of modal parameters (Section 2.2.2). The challenge can be stated succinctly as follows. If a weak necessity modal is evaluated with respect to a primary condition or priority \( c_1 \) and a secondary condition \( c_2 \), why is a strong necessity modal evaluated with respect only to \( c_1 \) and prevented from accessing both \( c_1 \) and \( c_2 \)? In order to
account for scalar inferences relating strong and weak necessity modals, there needs to be a motivated split of contextually salient priorities into those that are “primary” and those that are “secondary.” The basis for the division has not yet been explained in the literature.

My contribution in the second part of the chapter is to offer a way of overcoming this conceptual challenge. I argue that presuppositions about collective commitment towards a priority determine which strength of necessity the priority can support. Primary priorities are those that are presupposed to be collectively committed to; secondary priorities, in contrast, are presupposed not to be collectively committed to. Commitment itself is a modal notion, as it applies not just to the propositions representing priorities that an individual has actually endorsed in a conversation, but also to the propositions representing priorities he or she would endorse at that point, given a chance to do so. The proposal is then that weak necessity modals are sensitive to priorities for which lack of collective commitment is presupposed. More precisely, the secondary considerations that a weak modal is sensitive to are presupposed to be considerations that at least one conversational participant is not committed to at that point in the conversation.

Let me use an example to briefly illustrate the idea behind this proposal. In the conversation about driving to Amherst, (7), both you and I accept that your going to Amherst is a relevant priority. But other than getting to Amherst, no additional priorities that are relevant for deciding which road to take can be presupposed to be collectively committed to. Do you prefer nice scenery? Do you prefer to stop for lunch at a particular location on the way? Before such questions are answered, the scenario does not support strong necessity claims that single out any particular route choice. Intuitively, this is because the priority of going to Amherst can be achieved in multiple ways. The scenario nevertheless supports weak necessity claim, e.g., (8), that rely on exactly such priorities. Since I presuppose that

---

3 I believe that the locution “the priority is committed to” captures the modality in this description better than alternative locutions like “the priority is accepted” or “assumed” (which I have used in previous work; Rubinstein 2011, 2012). These latter phrases seem to imply actual acceptance, but actual acceptance is not necessary for commitment.
you are not committed to any priority that singles out Route 9 as your route choice, I use a weak necessity modal to convey that Route 9 is a necessary choice under some further assumption.

In Section 2.3, I present three arguments to support the hypothesis that collective commitment to priorities is implicated in the distinction between weak and strong necessity. I use constructions that have been argued in the literature to be sensitive to the presupposed commitments of participants in a discourse, independently of any discussion of modal strength. These include directives (self-directed ones and imperatives; Portner 2004, 2007), rising declaratives (following Gunlogson 2001, 2008), and conversational backoff with conditional questions (following Rawlins 2011).

In a qualitative framework, commitment toward priorities can be used directly to define the cutoff point between primary and secondary priorities (and thus also the cutoff point between primary and secondary ordering sources in a Kratzerian framework). I present a formalization of this kind in Section 2.4 for concreteness, although it is clear that the relevant distinction among priorities can be incorporated into analyses of necessity modals in other theoretical frameworks as well.

### 2.2.1 The comparative probabilistic approach and its limitations

Probability theory has entered into the semantics of modality through analyses of words like *likely* and *probable* (Frank and Kamp 1997, Swanson 2006, Yalcin 2007, 2010, Portner 2009, Lassiter 2011a,b, Kratzer 2012). These modal words have *epistemic* uses, in which they describe potentially non-actual possibilities that are compatible with some agent’s knowledge or beliefs, and *evidential* uses in which the relevant space of possibilities is determined by certain pieces of evidence.\(^4\) Probabilities provide the quantitative backbone

\(^4\) For descriptions of modality types, see von Wright (1951), Palmer (2001), and Portner (2009, §4.1) (which includes a comparison of classification systems in the literature). See Kratzer (2012, 21-22) for the characterization of epistemic modals as allowing evidence-based interpretations.
needed to interpret epistemic/evidential statements, such as (9), in which the reference to probabilities is explicit.

(9) There is a 60% probability that it’s raining.

(Portner, 2009, 73)

Complex expressions of probability and possibility, such as the numerical comparative statement in (10), also receive a straightforward analysis once probabilities are introduced into the semantics.

(10) If teaching assistants monitor exams, it is 31 percent more likely that students will cheat than if a professor is in the room[.]

A probabilistic analysis of the modal expressions in these sentences relativizes their interpretation to a *probability space*. A probability space is defined as a pair \(< B, Pr >\) of a set of possible worlds \(B\) and a probability distribution \(Pr\) defined over \(B\).\(^6\) \(Pr\) is a probability distribution over \(B\) if it maps all subsets of \(B\) to a number in the interval \([0,1]\) such that: \(Pr(B) = 1\), and \(Pr(p \cup q) = Pr(p) + Pr(q)\) for any disjoint subsets \(p\) and \(q\). Following Kratzer (1981, 1991, 2012), we can think of \(B\) as the set of accessible worlds determined by a *modal base*. An epistemic/evidential modal base \(f\) is a function from worlds to sets of propositions (where propositions are modeled as sets of possible worlds). It maps a world \(w\) to the set of propositions representing relevant bits of knowledge, belief, or evidence in \(w\). The worlds in the intersection of the propositions in \(f(w)\) are the worlds that are accessible from \(w\) in the relevant way. All the “probability mass” is distributed between these accessible worlds (Yalcin, 2010).

---


\(^6\) The set of worlds \(B\) is assumed to be countable (see, e.g., Benz et al. 2006), and sometimes, for simplicity, finite (see, e.g., Yalcin 2010).
Even when probabilities are not overtly specified, the probability space comes in handy for describing the meaning of *probably, probable* and their kin. Yalcin (2010) proposes the following probabilistic analysis for *probably*, based on the familiar idea that ‘probably’ means ‘more likely than not’:\footnote{7}

\begin{equation}
\text{p is probable iff } Pr(p) > Pr(\neg p) \text{ (equivalently: } Pr(p) > .5).\end{equation}

(Yalcin, 2010, 927)

In recent work, Finlay (2009, 2010) has proposed that weak necessity modals like *ought* are inherently – regardless of the type of modality they express in a given context – operators that effect probabilistic comparison. He has thereby extended the probabilistic analysis to encompass normative modalities (or *priority modalities*; Portner 2009) as well as epistemic modalities.\footnote{8}

Two options for analyzing priority modality within a comparative probabilistic theory are raised in Finlay’s work. The first option is to compare actions with respect to how likely they are to bring about a particular end (*e* below).\footnote{9}

Intuitively, I instrumentally ought to *ϕ* just in case it is *better*, with respect to bringing it about that *e*, that I *ϕ* than that I do any of the actions I could perform instead: ‘ought’ is essentially comparative. […] An action *ϕ* is an instrumentally ‘better’ means to an end than another, *ψ*, just in case *ϕ* is more likely to lead to that end than is *ψ* — or so I shall argue. (Finlay, 2009, 322)

In a later paper, Finlay argues that the means that is most likely to lead to the eventuation of an end can be identified with the means that is most likely chosen, under the assumption

\footnote{7 For an alternative probabilistic proposal and comparison with Yalcin’s proposal cited in the text, the reader is referred to Lassiter (2011b).}

\footnote{8 This attempt goes against the belief that only epistemic reasoning is based on probability and likelihood. For instance, Portner (2009) holds that “[w]hile it makes sense to understand epistemic modals in terms of probability, deontic, dynamic, and other modals cannot be analyzed in this way.” (pp. 76-77). Finlay provides the clearest and most compelling formulation I have encountered of the idea that non-epistemic modalities are amenable to a probabilistic analysis.}

\footnote{9 Finlay uses the term *end-relational* to refer to the analysis of priority modalities (including goal-oriented and desire-based modalities, also called *instrumental*, as well as rule-based, deontic modalities; see, e.g., Finlay 2010, pp. 78-79).}
that any dispositions an agent might have to pursue or neglect any of the alternative means are disregarded.

Therefore the base should involve an assignation of equal initial probability to the choice of every relevant potential means, by the principle of indifference. I will call this assumption of equal initial probability symmetry of choice. Given the subsequent base, i.e. that we have symmetry of choice and the end eventuates, the means most likely chosen is the means on which the end has the greatest likelihood of eventuating. (Finlay, 2010, 80-81)

In both of these descriptions, alternatives are invoked. Intuitively, alternatives to a prejacent \( p \) are “propositions (states of affairs, events, actions) such that they could obtain (occur, be performed) in place of \( p \)” (Finlay, 2009, 323). But what kinds of objects are the prejacent and its alternatives, formally? In the intensional typed lambda calculus I will be assuming, it is possible to characterize alternative actions as event particulars (objects of type \(<\text{ev},\text{t}>\) or \(<\text{s},\text{t}>\), respectively). When Finlay (2010) presents his formal analysis, he does not commit to any one of these interpretations in particular. It does seem clear that he treats alternatives on a par with (so, having the same type as) the prejacent, but a complete formalization of the proposal is not filled in.

In the following two sections, I discuss two such formalizations. One takes the prejacent and its alternatives to be of a property type (both of type \(<\text{s},\text{t}>\) or \(<\text{ev},\text{t}>\)) and is more faithful to Finlay’s (2010) implementation. The other takes the prejacent to be a property of events (type \(<\text{ev},\text{t}>\)) and the alternatives to be particulars (type \(\text{ev}\)) and is more in line with Sloman’s motivating discussion of \emph{ought} as a comparative operator, and to the earlier discussion by Finlay (2009). The two implementations give rise to different predictions and highlight the importance of formalization in testing hypotheses about the meaning of natural language expressions.
2.2.1.1 Comparing propositional alternatives

For the first implementation, we will assume that the prejacent and its alternatives are propositions, modeled as sets of possible worlds. The idea to be explored is that weak necessity modals compare the likelihood of a set of alternatives in an appropriately defined probability space, requiring that the likelihood of the prejacent top the likelihood of any other alternative. In a teleological context, ought will single out the action that is most highly represented among the set of accessible possible worlds in which a relevant end eventuates. This implementation of the comparative probabilistic analysis is fleshed out in (12), following Finlay’s (2010) formulation (ibid., pp. 80-81).

(12) (Following Finlay 2010)

(to p) ought to q is true in a world w with respect to a modal base f and a salient set of alternatives R iff \( Pr_B(q) > Pr_B(r) \) for all \( r \in R \),

where

\[ B = \bigcap f(w) \text{ (or } B = \bigcap(f(w) \cup \{p\}) \). \]

The set of worlds \( B \) is a subset of \( \bigcap f(w) \), the latter being the set of accessible worlds determined by a modal base \( f \) in an evaluation world \( w \). \( B \) may be narrowed down in a teleological context to include just those worlds in \( \bigcap f(w) \) in which the goal \( p \) obtains. (The ultimate subset of \( \bigcap f(w) \) identified in this way is the “subsequent base” in Finlay’s terms; I refer to it also as the set of accessible goal worlds.) The analysis assigns rationale clauses (phrases of the form (in order) to . . . ) the role standardly assumed for if-clauses by Kratzer (1981), namely that of “mak[ing] sure that a hypothesis is added to the modal base required by the modal expression to follow” (Kratzer, 1981, 317). The idea is that, semantically, goal-setting expressions conditionalize modal claims.

For purposes of exposition, Finlay (2010) defines probability as a proportion of a finite, non-empty possibility space. The probability of a proposition \( p \) relative to a base \( B \), written \( Pr_B(p) \), is thus defined “to be the proportion of \( B \)-worlds that are worlds in which \( p \) obtains” (ibid., p. 80):
The following example illustrates the theory at work.

Imagine that in the morning before school I know that Jorja hopes to skip school, and I also know that there are only two means to doing so: feigning illness, to which I assign a probability of success of 0.9, and feigning tiredness, to which I assign a probability of success of 0.1. I have no idea which she will choose, so (by the principle of indifference) relative to what I know, each choice has a probability of 0.5. Now imagine that in the evening I discover that Jorja somehow managed to skip school. It would be very natural for me to say, ‘She probably feigned illness, then’ (or: ‘She ought to have feigned illness, then’).

(Finlay, 2010, 81)

This example is problematic as an illustration of priority modality because it emphasizes the similarity between epistemic probably and ought, making it difficult to hold on to a goal-oriented interpretation of the weak necessity modal in the target sentence. The presence of a particle of inference (then) at the end of the sentence strengthens the bias toward an epistemic interpretation of ought. Presenting these two modal expressions side by side obscures the fact that only ought admits of the priority interpretation Finlay is after in this example. (13a) brings out the priority interpretation which probably and must have lack in (13b).

(13) a. She ought to have feigned illness (and that’s exactly what she did).

b. She {probably, must have} feigned illness (’and that’s exactly what she did).

Setting aside these differences between probably, must have, and ought to have, let us focus on the details of the proposed probabilistic comparison semantics for ought itself. A notable merit of the proposal is that it seamlessly incorporates goals into the semantics of probabilistic comparison priority modals. Comparative probability is simply conditionalized in the relevant cases on the attainment of a priority.

---

In case no one alternative outweighs the others among the accessible goal worlds, any weak necessity claim that singles out one of them is correctly predicted to be false. An example of this kind is provided in (14). The target sentence is false in this scenario, if we assume that equal probabilities are assigned initially to eating at each one of the four restaurants (so the scenario satisfies symmetry of choice).

(14) **A decent cheap lunch.** We are thinking where to send a student visitor for lunch in the area. Four restaurants in town serve decent lunches for under six dollars. All four are located in close proximity to each other. Three are owned by the Spageto company. The fourth one is called Green Side.

To eat a decent lunch here for under $6, the visitor ought to go to Green Side.

However, counterexamples for the implementation in (12) arise when contextually salient alternatives are “lumped” together such that the prejacent does outweigh the alternative(s) it is compared to in terms of probability. Still in the context of the lunch scenario, consider the target sentence (15).

(15) To eat a decent lunch here for under $6, the visitor ought to go to a Spageto restaurant.

According to (12), this sentence is true in the evaluation world with respect to a set of accessible goal worlds $B$ if and only if $Pr_B(\text{the visitor go to a Spageto restaurant}) > Pr_B(r)$ for all $r \in R$. Given the ownership of the restaurants, $Pr_B(\text{the visitor go to a Spageto restaurant})$ equals 0.75. Any alternative to eating at Spageto thus has a probability of at most 0.25 among the accessible goal worlds. This sentence is predicted to be true, although it is false and can be challenged by pointing out the lunch option at Green Side.

When the alternative to the prejacent is just its negation, this version of the probabilistic comparison analysis predicts a sentence *ought q* to be true just in case the probability of $q$ among the set of accessible (goal) worlds is greater than 0.5. These are the truth
conditions often assigned to the modal adverb probably in probabilistic frameworks (recall (11) above). In many uses of ought, however, the modal does not have a ‘more than half’ or ‘most’ semantics. Copley (2006) and von Fintel and Iatridou (2008) argue explicitly against analyzing goal-oriented weak necessity claims on analogy to most quantification over possible worlds.\footnote{Copley (2006) calls the ‘most’-analysis of weak necessity the “traditional view,” following a parallel drawn in Horn’s early work between modal quantification and quantification over individuals (Horn, 1972, 1989). Copley (2006, §2) and von Fintel and Iatridou (2008, 118) argue against this view. They allude to the technical obstacle of coherently defining quantifiers like most or more than half over infinite domains (von Fintel and Iatridou) and show that giving ought to/should less-than-universal quantificational force incorrectly predicts these modals to trigger versions of Moore’s paradox (Copley).}

Lumping, demonstrated by (15) above, is one challenge facing the implementation of the probabilistic comparison theory in (12). Another challenge is introduced by cases in which there is more than one alternative to the prejacent, and the alternatives are not equally represented among the accessible goal worlds. Consider the scenario in (16).

(16) **Raised rent.** Our friend’s landlord has raised the rent on him significantly with very short notice. As his friends, we identify four things he can do that are guaranteed to get him out of the tough situation: move back in with his parents, stay in the apartment and take a loan, stay and find another job, or stay and get a roommate.

Suppose we have a discussion about how the friend ought to act, and instead of assuming that the four alternatives are equally represented among the accessible goal worlds, suppose that moving back in with his parents is in fact twice as likely to take place than any of the other three alternatives. Thus, $Pr_B(he \ moves \ in \ with \ his \ parents) = 0.4$, and that $Pr_B(r) = 0.2$ for each of the other three alternatives (where $B$ is the set of accessible worlds in which the friend succeeds in making his upcoming payments). Together, the four alternatives exhaust the space of relevant possibilities. The implementation of the probabilistic comparison account in (12) predicts the teleological ought-claims in (17a) and (17b) to
differ in truth value in this scenario. (17b) is predicted to be true, while any weak necessity claim that focuses on a different alternative, e.g., (17a), is predicted to be false.

(17) One friend says, “In order not to default on his upcoming payments, he can do one of the four things we talked about.”

a. Someone remarks, “(In order . . . ) he ought to get a roommate.”

b. Another says, “(In order . . . ) he ought to move back in with his parents.”

Native speakers presented with these examples are inclined to accept them both as being potentially true. When they accept them as being potentially true, however, speakers provide explanations that show that they have to move beyond the priority stated in the rationale clause. They accept (17a) and say, “Well, yes, if he wants to save the legwork and the begging . . . ,” for example. They accept (17b) and say, “Sure, if he doesn’t mind not having a love life . . . .” The point is that the comparative likelihood of the different possibilities doesn’t affect the truth values of these sentences in the way predicted by (12). Either both sentences are true (with shifts to different sub-goals), or they are both false under a pedantic evaluation of the context, since there are multiple ways of achieving the relevant priority. The propositional probabilistic comparison analysis predicts an unattested split in truth values. It predicts (17b) to be true without a shift to any sub-goal, simply because the prejacent in this sentence describes more accessible goal worlds than any of the other three alternatives (and (17a) to be false for the same reason).

2.2.1.2 Comparing alternatives as particulars

A different interpretation of the probabilistic comparison analysis is not vulnerable to the problems raised in the previous section. This implementation is based on the assumption that alternatives are event particulars (type ev) and are not, in fact, “alternatives to the prejacent.”12 On this interpretation, the probabilistic comparison mechanism is sensitive to

---

12 In Finlay (2009, 323), this interpretation seems to be assumed. Thanks to Angelika Kratzer for clarifying this possibility.
whether or not a given alternative has the property described by the prejacent (the prejacent is assumed to be of type \(<ev,t>\)). It requires every alternative that \textit{is} a member of the prejacent to have a better chance of leading to a goal event than any alternative that isn’t a member of the prejacent. Rationale clauses do not function as conditionalizing the modal claim here, unlike in the previous implementation we considered. Formally, we entertain the analysis of \textit{ought} in (18).

(18)  \textit{(Another implementation of Finlay 2009, 2010)}

to \textit{p ought to q} is true in a world \(w\) with respect to a modal base \(f\) and a salient set of alternatives \(R\) iff for every \(r \in R\) such that \(q(r)\) and for every \(\bar{r} \in R\) such that \(\neg q(\bar{r})\), it is the case that \(Pr_{B_r}(p^w) > Pr_{B_{\bar{r}}}(p^w)\),

where
\[
B = \bigcap f(w),
\]
\[
B_r = \lambda w. B(w) \& \exists r' \in [r]. occurr(r')(w),
\]
\[
p^w = \lambda w \exists e. occurr(e)(w) \& p(e).
\]

In this implementation, \(B\) is the set of accessible worlds: those worlds in which all the modal base propositions are true. For any event \(r\), let \([r]\) be the set of worlds in \(R\) that are of the same event type as \(r\). For example, the set of events \([r]\) for an event \(r\) of eating at the first Spageto restaurant is the set of all eating-at-the-first-Spageto-restaurant events in \(R\). Let \(B_r\) be the subset of the accessible worlds in which an \(r\)-type event occurs. Finally, let \(p^w\) be the set of worlds in which the priority expressed in the rationale clause eventuates, i.e., the set of worlds in which a \(p\)-event occurs.

The problem of lumping does not arise in this implementation of the comparative probabilistic analysis. In the lunch scenario, for example, events described by the prejacent are events of the visitor going to any Spageto restaurant. These events do not have a probabilistic advantage of satisfying the relevant priority over alternative events of eating elsewhere: wherever the visitor chooses to eat, she will enjoy a decent cheap lunch with a probability of 1. Consequently, (18) correctly predicts the sentence in (15) to be false.
However, it is not immediately clear that this interpretation of the probabilistic comparison approach can yield a unified analysis of *ought* on all its modal interpretations. One concern is that it is unclear what would replace the priority *p* in a weak necessity claim with an epistemic interpretation. I set this question aside for now, and argue that concentrating just on goal-oriented modality already unearths a difficulty for the approach.

Consider what happens if no event in a salient set of alternatives *R* is guaranteed to lead to the eventuation of the priority *p*. It could be that the most reliable means is itself relatively risky. But if this risky alternative is more reliable than all other contextually salient alternatives, (18) predicts the corresponding *ought* sentence to be true. This seems incorrect, as the following scenario is designed to show.

(19) **Random muffins.** Two bakeries in town make muffins. Their selection is limited and varies by the day: the first bakery makes only chia or banana muffins; the second one makes only chia, banana, or anise muffins. The pastry chef of each bakery decides the unique flavor of the day each morning by rolling a fair die. My friend wants a banana muffin for breakfast.

In order to get a banana muffin today, you ought to go to the first bakery.

The event particulars that are compared in this scenario are events of my friend going to the first bakery (*bakery*$_1$ events) and events of her going to the second bakery (*bakery*$_2$ events). The probability that the first bakery will have banana muffins on any given day is greater than the probability that the second bakery will have them. More precisely, where $b^w$ is the set of worlds in which my friend gets a banana muffin for breakfast today:

\[
\Pr_{\text{Bakey}_1}(b^w) = \frac{1}{2}
\]

\[
\Pr_{\text{Bakey}_2}(b^w) = \frac{1}{3}
\]
Since going to the first bakery is more likely to lead my friend to a banana muffin, the analysis in (18) predicts the target sentence in (19) to be true. Crucially, the truth of the sentence follows from two things only: the relevant circumstances of the scenario, and the likelihood of a single priority (my friend’s getting the desired muffin) to materialize given those circumstances.

If these were the only ingredients that determined the truth of the modal statement, there would be no disagreement about what ought to be done. This is obviously not the case, however, as arguments are a very natural kind of discourse in which weak necessity modals are used (e.g., (20)). In such examples, there is a sense that the discourse participants agree about a primary preference (in this case, getting a banana muffin), yet they disagree about a secondary preference: conflicting secondary preferences are expressed in (20a) and (20b). Put another way, when someone says that such-and-such ought to be the case, it is not always clear what the basis for their claim is. What disagreements tell us is that there is more that *ought* is sensitive to than just the circumstances and the likelihood of the primary preference to materialize.

\[(20) \quad \begin{align*}
\text{a. A to C: In order to get a banana muffin today, you ought to go to the first bakery.} \\
\text{The odds are higher that you’ll get one there, so maximize your chances.} \\
\text{b. B to C: That’s not true. In order to get a banana muffin today, you ought to go to the second bakery. Your horoscope says that unexpected things will happen today, so follow that prediction.}
\end{align*}\]

The probabilistic analysis espouses a preference for greater likelihood as the sole secondary preference possible. To see this more clearly, consider the contrast between the two weak necessity claims in (21). (21a) is felicitous and true in the random muffins scenario; it expresses precisely the meaning that the probabilistic comparison analysis sets out to model. This meaning is not readily available for *ought* in (21b), however, in contrast to the predictions of (18). This sentence can be challenged by pointing out the possibility that
events that are not of the type described by the prejacent will lead to satisfaction of the primary priority.

(21) Both bakeries might have the muffins you want.

   a. In order to improve the odds of getting them, you ought to go to the first bakery.
   b. In order to get the muffins, you ought to go to the first bakery.

To summarize, chances are better that events described by the prejacent will lead to eventuation of the relevant priority, and a preference for improving the odds can make the use of ought felicitous and true if explicitly expressed, as in (20a) and (21a). The contrast in judgments between (21a) and (21b) shows that a probability-related preference for improving the odds is not contributed by the meaning of ought automatically, in every example. Therefore, sensitivity to greater likelihood should not be made an integral part of the semantics of weak necessity modals.

It is worthwhile to note that increasing the numerical value of the odds does not change the pattern of judgments reported in (21). Suppose that the odds of there being banana muffins for sale at one bakery are much higher than we first considered. Perhaps the first bakery uses a die that is biased towards this flavor: 9 out of 10 faces of the die say ‘Banana’ and only one says ‘Chia.’ Still, the target sentence in (19) is false under the relevant interpretation if there is a chance that the second bakery will have the same kind of muffin today (in addition to, or in place of, the first bakery).

2.2.1.3 A decision theoretic analysis

The random muffins scenario also serves to challenge a decision theoretic account of the weakness of ought. Before presenting a specific proposal of this kind – due to Goble (1996) and developed further recently by Lassiter (2011b) – let me briefly introduce some of the basic concepts of decision theory and the main ingredients of decision problems.

Decision theory provides a formal framework for describing how rational agents go about, or should go about, deciding between actions that are open to them. Its roots stretch
back at least three hundred years, with von Neumann and Morgenstern’s (1944) mathematic-
cal formalization of the connection between expectations and rational behavior leading to a
flurry of research on the economic, psychological, and social aspects of decision making.\footnote{I
follow Resnik (1987) and Goble (1996) in this exposition. Luce and Raiffa (1957) is a classic general
introduction to games and decision theory, and see Lewis (1969), Merin (1999), van Rooy (2003),
and Benz et al. (2006) for linguistic applications.}

In choosing between actions that might lead to multiple possible outcomes (such cases
are called \textit{decisions under risk}), the relative desirability of the actions and the probabili-
ties associated with the outcomes are assumed to guide the decision making process. The
“winning” action, namely the one that would be chosen by a rational agent, is the action
that strikes the optimal balance between what the agent prefers and the likelihood of her
preferences to materialize. In technical terms, each action is paired with its \textit{expected utility},
and the rule for solving the decision problem is to choose the action with the highest
expected utility among the actions under consideration.

Formally, a decision problem is a tuple $< W, A, \Pr, U >$, containing a space of possible
worlds $W$, a probability distribution $\Pr$ over $W$, a set of propositions $A$ (the actions
to choose from), and a utility function $U$ whose domain is $W$ and whose range is the real
numbers. The utility value $U(w)$ of a world $w$ is an abstract measure of the “goodness”
of $w$. It can be thought of in subjective terms, as modeling the particular preferences an
agent has among possible ways the world could be. It can also be given a more objective
interpretation which is not tied to the preferences of any particular individual. The prop-
erties of logics based on probabilities and utilities and the semantics of natural language
expressions that are interpreted using these logics can be safely investigated while setting
aside questions about the source of $U$ and $\Pr$ in actual conversation (Goble 1996, 353-354,
Lassiter 2011b, 159-160).

Decision problems can be represented graphically as in Table 2.1.\footnote{This representation is
inspired by Resnik (1987, 45ff.).} The rows of the
matrix list the actions in $A$, and the columns represent individual possible worlds in $W$
Table 2.1: A decision problem with two actions and $n$ relevant possible worlds.

(potentially infinitely many of them). Two bits of information are written in each cell: the utility value of the column world $w$, and the conditional probability of $w$ given the row action $a$. The conditional probability of a proposition $p$ given a proposition $q$ is defined as usual by:

$$ Pr(p|q) = \frac{Pr(p \cap q)}{Pr(q)} \text{ for } Pr(q) \neq 0 $$

Based on the utility values associated with individual worlds in the decision problem, the expected utility of any proposition, or set of worlds, can be calculated. The expected utility of a proposition $p$ is defined as the weighted sum of the utilities of the worlds in $W$, where the weights are the conditional probabilities of each world given $p$ (Goble, 1996, 329):

$$ EU(p) = \sum_{w \in W} Pr(\{w\}|p) \times U(w) $$

Table 2.2 is the decision problem (unique up to linear transformations of the utility function) that models the deliberation in the random muffins scenario. For simplicity, I assume a finite set $W$ of relevant possible worlds that differ according to two parameters: which flavors are served at the two bakeries on a particular day, and which bakery my friend goes to. (We can think of $W$ as determined by a modal base function, such that $W = \bigcap f(w)$.) The leftmost two columns represent worlds in which banana muffins are served at the first bakery and anise muffins are served at the second bakery. In $w_1$ these circumstances hold and my friend goes to the first bakery; in $w_2$ she goes to the second
Table 2.2: Decision problems and weak necessity. The random muffins scenario as a decision problem.

<table>
<thead>
<tr>
<th></th>
<th>Bakery 1</th>
<th>Bakery 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Banana,</td>
<td>Banana,</td>
</tr>
<tr>
<td></td>
<td>Chia</td>
<td>Chia</td>
</tr>
<tr>
<td>w₁</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>w₂</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>w₃</td>
<td>10/6</td>
<td>0</td>
</tr>
<tr>
<td>w₄</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>w₅</td>
<td>-1/6</td>
<td>0</td>
</tr>
<tr>
<td>w₆</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>w₁₁</td>
<td>-1</td>
<td>1/6</td>
</tr>
<tr>
<td>w₁₂</td>
<td>-1</td>
<td>0</td>
</tr>
</tbody>
</table>

bakery. The rightmost two columns represent worlds in which both bakeries serve chia muffins (in w₁₁ my friend goes to the first bakery and in w₁₂ she goes to the second bakery), and so on for the other possible combinations of flavors.

Utility values in the decision problem reflect my friend’s desire to have a banana muffin for breakfast today. There is a positive payoff for going to a bakery that serves the preferred flavor, and a negative payoff for going to one that doesn’t. Note that the payoff values for satisfying the relevant desire are the same regardless of the action that leads to them. This property of the decision problem represents the assumption that the bakeries are equal on every other potentially relevant consideration: they cannot be distinguished based on their distance to the house, their accessibility, their pricing, or any other criterion that would lead my friend to prefer one over the other.

Following the blueprint of Table 2.1, in the top leftmost cell we enter two numbers: the utility of going to the first bakery on a day in which the flavor line up at the bakeries is banana-anise (a positive utility of 10), and the conditional probability of that flavor lineup given that my friend goes to the first bakery. Since there is equal likelihood for each one of the six possible combinations of muffin flavors, this conditional probability is 1/6.

At this point, we can calculate the expected utilities of the actions under consideration in the scenario. We find that the expected utility of going to the first bakery is higher than the expected utility of going to the second bakery (EU(bakery₁) = 4.5, EU(bakery₂) =
This is so because it is more likely that the first bakery will have the desired flavor, and there are no additional desires or priorities that factor into the decision. According to Goble (1996), the action with the highest expected utility is the action that *ought* to be carried out.\(^\text{16}\)

What seems to be at work here is some good old-fashioned decision theory. What counts in determining what the player ought to do is not just the intrinsic values of the possible outcomes of the play, but rather the expected value attached to each play, and that takes probability into consideration. (Goble, 1996, 322)

The claim that both utilities and probabilities are essential for capturing the meaning of non-epistemic *ought* has recently been taken up by Lassiter (2011b).\(^\text{17}\) Lassiter proposes a variant of Goble’s “utilitarian” semantics that requires the expected utility associated with *ought*’s prejacent to be significantly greater than the expected utility of all the contextually available alternatives (*ALT* in the definition below) taken together.\(^\text{18}\) The lexical entry he proposes is given below.

\[
\text{(22) } \text{should/ought } q \text{ is true iff } EU(q) \geq \theta_{\text{should/ought}}, \text{ where } \theta_{\text{should/ought}} \text{ is a value significantly greater than } EU(\cup ALT(q)).
\]

(Lassiter, 2011b, 187)

The precise compositional contribution of rationale clauses like *In order to get a banana muffin today* is left implicit in the analysis, but we may assume that their utterance influ-

\(^\text{15}\) A sample calculation: \(EU(bakery_2) = 4 \times (-1 \times 1/6) + 2 \times (10 \times 1/6) + 6 \times (0 \times 10) = 16/6 = 2.67\).

\(^\text{16}\) If multiple actions tie in having the highest expected utility in a set of alternatives, their union defines the proposition that is a weak necessity. See Goble (1996, 343).

\(^\text{17}\) Cariani et al. (2011) have also incorporated decision problems into the semantics of necessity modals to account for their use in practical deliberations. Their use of the term “decision problem” is different, however, as it crucially does not involve a representation of utilities in the theory (ibid. §3.1). I do not discuss their proposal in detail here as it is not designed to account for the relative weakness of *ought* as compared to a strong necessity modal like *must*.

\(^\text{18}\) Lassiter (2011b) refrains from using the term “expected utility” and refers instead to “expectation” (\(E\)) or “probability-weighted preference.” I use \(EU\) here for consistency with the terminology introduced above.
ences, or at least matches, the preferences captured by the utility function in the relevant decision problem.

It seems plausible to assume that the utility function $U$ that is described in Table 2.2 faithfully represents the sole preference of getting a banana muffin in the random muffins scenario. The analysis in (22) thus predicts the truth of (19), repeated below, if the probability of there being a banana muffin at the first bakery is high enough for $EU(bakery_1)$ to count as “significantly greater than” $EU(bakery_2)$.

(19) In order to get a banana muffin today, you ought to go to the first bakery.

As argued in the previous section, this is an incorrect prediction. The decision theoretic analysis ties the meaning of weak necessity tightly to probabilistic reasoning, and as such falls short of accounting for cases like (17) and (19) in which what (arguably) ought to be done is not necessarily determined by the preference for maximizing expected utility.

In response to this argument, one might contend that my choice of the utility function was misguided (and thus, that these test cases are not true counterexamples to the decision theoretic approach). Indeed, if it is the case that additional preferences contribute to the evaluation of a weak necessity claim, above and beyond the primary preference modeled by $U$ in Table 2.2, a different utility function would be needed to represent them. In cases of disagreement, there could be a number of relevant utility functions, each representing distinct combinations of preferences.

This is a legitimate response, but it does not in fact support the decision theoretic approach to weak necessity. Recall that the idea behind introducing probabilities and utilities into the semantics of $ought$ was that maximization of expected utilities – this mechanism itself – is responsible for the weakness of the modal. The associated calculations provided the explanation for why it is that $ought$ expresses a weaker necessity than $must$ or $have to$ do. Indeed, keeping the parameters of the decision problem fixed (the utility function $U$, the probability distribution $Pr$, and so forth) is crucial for predicting the entailment.
relation between strong and weak necessity modals in this framework.\footnote{See Lassiter (2011b, 193). Similar remarks apply \textit{mutatis mutandis} to the probabilistic comparison approach. The denotations proposed by Finlay (2009) for strong and weak necessity modals are shown side-by-side in (23) and discussed below.} Thus, while it may be necessary to accept that \textit{ought} uses a different utility function than \textit{have to} does in order to uphold the decision theoretic analysis of these modals, it is important to realize that making this move without describing how such $U'$ would be different from the original $U$ begs the question of how the modals differ and thus provides little insight as to the semantic-pragmatic underpinnings of weak necessity.

If the burden of explanation is shifted from expectation maximization to properties of the utility function, it may well be the case that a (potentially complex) qualitative representation of preferences would be sufficient for analyzing weak necessity. Utility functions are, after all, nothing more than a quantitative representation of preferences.

### 2.2.1.4 Summary and conclusions

I evaluated two implementations of the probabilistic comparison approach to weak necessity in this section, following proposals by Finlay (2009, 2010). A related decision theoretic approach, pioneered by Goble (1996), was examined as well.

The two implementations of the probabilistic comparison approach embodied the idea that \textit{ought}’s prejacent somehow describes states of affairs that have a greater likelihood of occurring, given relevant circumstances and the eventuation of a goal, than any relevant alternative. This differs from strong necessity, which essentially requires the prejacent to have a probability of 1 in a probability space containing the worlds that are compatible with the relevant circumstances and goals. Indeed, Finlay assigns strong necessity modals like \textit{must} or \textit{have to} a standard analysis as universal quantifiers over possible worlds (23a) (and possibility modals, e.g., \textit{may}, are analyzed as existential quantifiers).
(23)  a. *must*: (In order that $e$) it must be the case that $p = \text{Every possible world in which circumstances/laws } C \text{ obtain (including its being the case that } e\text{), is such that } p$.

b. *ought*: (In order that $e$) it ought to be the case that $p = \text{It is more likely, given circumstances } C \text{ (including its being the case that } e\text{), that } p \text{ than that any other member of } R \text{ obtains.}

(Finlay, 2009, 319-323)

To test the predictions of the approach, I discussed the interpretation of *ought* statements in contexts in which the modal’s prejacent had a probabilistic advantage over other alternatives in bringing about a relevant priority, but, crucially, there was no indication that any of the alternatives were preferred over the others. These contexts turned out to be problematic for the approach. Despite the probabilistic advantage associated with some of the alternatives, native speakers’ intuitions were that *ought*-claims drawing attention to these alternatives were false unless additional preferences were taken into consideration. One contrast in judgments was particularly revealing. We saw that *ought* can be made sensitive to probabilities, as in (24a), but this is not what it does by default, hence the falsity of (24b) (see (21) above).

(24)  [The muffin selection of the day hasn’t been decided. No one bakery is guaranteed to make the flavor you want today, but the odds are higher that the first one will.]

   a. In order to improve the odds of getting the muffin you want today, you ought to go to the first bakery.

   b. In order to get the muffin you want today, you ought to go to the first bakery.

A preference for greater certainty is just one unprivileged preference that *ought* can be made sensitive to. There might be contexts which favor this interpretation, but there are many others which do not. This is part of the context dependency of weak necessity modals.
and therefore not something that should be hard-wired in their lexical entries. Finlay raises this point in his discussion.

It is plausible that the best means relative to a particular end is the most reliable means. Admittedly, we often evaluate means on other grounds, such as cost, pleasantness, moral permissibility, safety, etc. However these are plausibly evaluations of means relative to ends other than the ends to which they are means.[…] In ordinary conversation speakers are concerned with a variety of ends; this introduces complexities into our use of ‘ought’ that I will not attempt to address in this paper, but that would need to be accommodated for this account to be fully successful.

(Finlay, 2010, 81)

Indeed, evaluating means on grounds of reliability is precisely an example of evaluating means relative to ends other than the (primary) ends to which they are means. The challenge, if we want a truly unified semantics for weak necessity modals, is not to treat the cases in which increased likelihood matters (and similarly, greater expected utility) as a special case. While probabilistic reasoning is suitable for accounting for certain types of modal reasoning, our discussion indicates that it does not – at least not in ways discussed so far – give a satisfactory account of the full range of preference-based weakness that ought is used to express in conversation.

A central challenge in explaining the distinction between necessity modals of different “strengths” is to characterize the preferences (end, priorities) that these modals are sensitive to. It is a challenge facing not only the two quantitative approaches to weak necessity discussed in this section, but also the qualitative approach which we turn to next.

A final issue I would like to raise concerns the relationship between weak and strong necessity modals, and the question of whether or not there is such a thing as a basic concept of modal necessity (in English, and in natural language more generally). From eyeballing the denotations in (23), one would likely come away with the impression that weak and strong necessity modals express very different modal concepts: must is a universal quantifier, ought is not; ought makes reference to alternatives and involves comparisons, must does not. The fact that the two types of modals are paired with unrelated denotations does
not square well with what we know about their grammatical expression across languages. A recurring pattern discovered by von Fintel and Iatridou (2008), albeit mostly in languages that belong to the Indo-European language family, is that the expression of weak necessity is parasitic on, or derived from, the expression of strong necessity. Weak necessity modals in these languages are realized morphologically by applying counterfactual marking to a strong necessity modal (sometimes in addition to the availability of a dedicated lexical item that means ‘ought’). This is the case in Greek, Romance languages (French, Spanish), Slavic languages (Russian, Croatian), Germanic languages (Dutch, Icelandic, Swedish), and Hungarian. The following examples from Spanish illustrate this phenomenon. The conditional mood morphology that is found in consequents of counterfactual conditionals (as seen in (25)), also turns strong necessity modals to weak ones, as in (26).

(25) No esta borracho. Si estuviera borracho, gritaria más
Not is drunk if was.Sbjn drunk yell.Cond more
‘He is not drunk. If he were drunk he would yell more.’

(26) Debería (/Tendría que) limpiar los platos, pero no estoy obligado
Must.Cond (/have.Cond that) clean the dishes but not am obliged
‘I ought to do the dishes but I am not obliged to.’

(von Fintel and Iatridou, 2008, 122, simplified)

Why it is that counterfactual morphology has the power to make strong necessity modals weaker in some languages, and why this happens in some languages but not in others, is still an open question. Nevertheless, viewed through the prism of morphosyntax, weak and strong necessity modals have more in common than the probabilistic comparison analysis of ought seems to suggest.

The assumption that the two types of necessity modals share a semantic core is one of the basic tenets of the domain restriction approach to weak necessity, advocated by von Fintel and Iatridou (2005b, 2008).
2.2.2 The domain restriction approach

Here is an elegant summary of the domain restriction approach to weak necessity.

Our conception of weak necessity then makes [weak necessity modals] universal/necessity modals just as much as strong necessity modals are. What makes them weaker semantically is that they have a smaller domain of quantification: strong necessity modals say that the prejacent is true in all of the favored worlds, while weak necessity modals say that the prejacent is true in all of the very best (by some additional measure) among the favored worlds. (von Fintel and Iatridou, 2008, 119)

In a context which determines a modal base \( f \) and an ordering source \( g \), the favored worlds are the worlds compatible with the modal base that are most highly ranked by the ordering source (i.e., for an evaluation world \( w \), these are the \( g(w) \)-best worlds in \( \bigcap f(w) \); ibid., p. 117). According to von Fintel and Iatridou (2008), a strong necessity modal makes a claim about all of these worlds, whereas a weak necessity modal only makes a claim about a subset of these worlds.\(^\text{20}\) The result of shrinking the domain of a universal quantifier is weakening, since the universal quantifier is left downward monotone, or downward entailing in its restrictor argument. This analysis accounts directly for the scalar inferences that are observed to relate ‘have to’ and ‘ought.’

\[\text{Favored worlds}\]

\[\text{have to} \quad \text{By some additional measure} \quad \text{ought to}\]

Figure 2.2: The contextual domain restriction approach to weak necessity. Schematic representation following von Fintel and Iatridou (2005b, 2008).

\(^{20}\) Kratzer (1981, 312) analyzes the weakened necessity expressed by German soll in terms of domain restriction when she proposes that it requires a non-empty ordering source, unlike the stronger muß.
It should be intuitively clear which “additional measures” would be invoked by the analysis in the examples we have discussed so far. In the driving-to-Amherst scenario in (7), for instance, any consideration above and beyond reaching Amherst could count as an additional measure in interpreting the weak necessity claim *You ought to take Route 9*. Possible considerations would be avoiding toll roads, visiting Worcester on the way to Amherst, a superstition that nine is an auspicious number, and so forth. All of these additional measures characterize the subset of accessible Amherst-going worlds in which one takes Route 9. In the muffins scenario in (19), interpreting *ought* in a sentence like *You ought to go to the first bakery* could invoke the additional measure of avoiding risk, or “playing it safe.” The accessible worlds in which one gets the desired muffin flavor split in this case into two classes, according to which bakery the desired muffin was purchased at. Among the set of accessible goal worlds, the subset in which one behaves so as to minimize risk are worlds in which one goes to the first bakery and purchases the muffin there. The fact that any consideration – “chancy” or other – counts in principle as an additional measure is a point in favor of the domain restriction approach.

The two basic tenets of the domain restriction approach are, to recap, that weak necessity modals denote universal quantifiers, just like strong necessity modals, and second, that the domain of quantification of weak necessity modals is smaller than that of strong necessity modals (von Fintel and Iatridou, 2008, 119). As a way of formalizing domain restriction, von Fintel and Iatridou propose that there is a special ordering source that weak necessity modals use, which contains the “additional measure” they are sensitive to. This special ordering source is used to identify a subset of the favored worlds as the weak modal’s quantification domain.

While the standard Kratzer framework parametrizes the semantics of modals to two parameters (modal base and ordering source), we introduced a pair of ordering sources: (i) the primary one that is the only one that strong necessity modals are sensitive to and (ii) a secondary one which is the one that weak

21 See p. 33.
necessity modals use to refine the ranking of the worlds favored by the primary ordering source. We built that differential sensitivity into the lexical entries of must/have to and ought. (von Fintel and Iatridou, 2008, 136-137)

The authors’ detailed proposal is more elaborate than this, as we will see in Section 2.2.2.2, but the gist of the analysis is the same: weak and strong necessity modals differ in the number of ordering sources they are sensitive to. Strong necessity modals are characterized as being sensitive to one primary ordering source (or more, see below), while weak necessity modals are sensitive to at least one additional, secondary, ordering source.

Although my focus is on priority-type interpretations of weak (and strong) necessity modals, it is important to note that von Fintel and Iatridou’s proposal is not framed exclusively in the context of this subtype of modality. As such, it extends in principle to more epistemic uses of weak necessity modals, of the kind shown below. In explaining the weakness of ought in (27), von Fintel and Iatridou suggest that the modal takes into account not just the “hard and fast evidence” that is relevant, but also assumptions about “what is normally the case” (p. 119).

Let’s say you are on your way to Morris’s office, which is down the hall from mine, and ask me whether I think that Morris is in his office. Neither of us knows whether he is, in fact, there. Under those circumstances, I can utter [(27)].

(27) It’s 3pm. Given what I know about Morris’s habits, he ought to be in his office. Why don’t you go check?

(von Fintel and Iatridou, 2008, 126, renumbered)

This example raises an important question (the first among a number discussed below) about the proposed division between primary and secondary ordering sources. In order to explain why a weak modal is used in the given context, we would have to stipulate that the primary ordering source used cannot contain propositions describing what is normally the case (in other words, that it cannot be stereotypical in Kratzer’s 1981 sense). Presumably, we would assume that the primary ordering source is empty in the given scenario, and
therefore that a strong necessity claim corresponding to (27) is false (since Morris’s being out of the office is consistent with the evidence in the scenario). However, it is not the case that the primary ordering source can never contain information about what is normally the case. A variant of (27), in (28), shows that must can, in fact, take into account precisely this kind of information given slightly different contextual assumptions.

(28) [Let’s say you are on your way to Morris’s office, which is down the hall from mine, and ask me whether I think that Morris is in his office. Neither of us knows whether he is, in fact, there. I say to you:]

It’s 3pm. Since we can count on Morris sticking to his habits, he must be in his office. Why don’t you go check?

Why can the stereotypes influence the set of favored worlds in this latter case, but not in (27)? Providing an answer to this question (as it arises in corresponding priority-type necessity claims) is the center of my concern here. In the next section, I lay out the argument that in order to uphold the domain restriction approach to weak necessity, one must elucidate the distinction between “primary” and “secondary” priorities.

2.2.2.1 The challenge of distinguishing between priorities

The success of the domain restriction approach to weak necessity depends on an ability to make finer-grained distinctions than are standardly assumed between modal conversational backgrounds. Following Kratzer (1981, 1991, 2012), it is standardly assumed that conversational backgrounds are individuated based on the type of information they contain: epistemic backgrounds represent the available evidence, deontic backgrounds represent the relevant rules, bouletic backgrounds represent someone’s personal desires, teleological backgrounds represent priorities that are relevant, and so forth. In the analysis of necessity modals envisaged by von Fintel and Iatridou (2008), context is required to provide each of these backgrounds in two varieties. For example, it is assumed that some teleological backgrounds are primary (these a goal-oriented instance of have to is sensitive to, as well
as goal-oriented \textit{ought}, and others are secondary (only goal-oriented \textit{ought} takes these into account).

As mentioned in the discussion of the epistemic examples (27)-(28) above, without a criterion for determining the division between primary and secondary backgrounds, the meaning difference between \textit{ought} and \textit{must} cannot be captured by the analysis as it has been presented.\footnote{Silk (to appear) also notes this weakness of von Fintel and Iatridou’s analysis. His proposal, which also takes into account my previous work, is discussed in more detail in Section 2.4.4.} Let’s see why.

Let $f(w)$ be a set of propositions given by a circumstantial/factual modal base $f$ in a world $w$, and $g_1$ and $g_2$ two teleological ordering sources, such that $g_1$ is primary and $g_2$ is secondary. Suppose further that there are two propositions, $p$ and $r$, that represent all the salient priorities in $w$, and that neither is entailed by the propositions given by the modal base. Moreover, suppose $q$ is a necessity given these priorities (that is, in all accessible worlds in which $p$ and $r$ are true, $q$ is also true). If $r$ is included in a secondary ordering source, as in (i), the context supports a weak necessity claim \textit{ought $q$} according to the domain restriction analysis. If instead $r$ is included in a primary ordering source, as in (ii) or (iii), the context supports a stronger necessity claim, \textit{have to $q$}.\footnote{In (ii), neither priority takes precedence over the other, whereas a representation like (iii) reflects a cascading of the priorities such that the consideration represented by $p$ is more important, or takes precedence over, the consideration represented by $r$. See Appendix A for discussion.} Which of these contextual representations of the priorities (among others that are theoretically possible) is to be assumed?

(i) Context provides $f, g_1(w) = \{p\}$, and $g_2(w) = \{r\}$, supporting a weak necessity.

(ii) Context provides $f, g_1'(w) = \{p, r\}$, and $g_2'(w) = \emptyset$, supporting a strong necessity.

(iii) Context provides $f, g_1''(w) = \{p \cap r, p\}$, and $g_2''(w) = \emptyset$, supporting a strong necessity.

To make the dilemma a bit more concrete, consider the weak necessity claim \textit{You ought to take the turnpike}. This statement is true in the scenario of driving to Amherst, (7), if
alongside the priority of getting to Amherst ($p$-worlds), a secondary priority is taken into account that you drive 65 miles per hour the whole length of the way without breaking the law ($r$-worlds). Once this secondary consideration has been raised in a conversation, as in (29), $ought$ can pick up on it very easily. The final sentence in the exchange is true if I use a weak necessity modal, which suggests a contextual representation like (i) above. The sentence is false if I use a strong modal, implying that (ii) or (iii) are not possible representations of the context.

(29) [You ask about ways to get to Amherst, driving.]

You: I’m driving to Amherst. Are the different roads there all equally quick?
Me: No, the turnpike is the fastest.
You: I see.
Me: You {have to, ought to} take the turnpike.

Evidently, $have to$ does not have access to both $p$ and $r$ in this scenario, and the question is why not. What prevents the additional measure that $ought$ has access to from entering and being part of the primary teleological ordering source of the context? What is the criterion for “membership” in a primary ordering source that rules $p$ in and keeps $r$ out? Obviously, freely associating the contextually salient priorities to either primary or secondary backgrounds is problematic.

There has not been much discussion of this question in the literature, but two ideas, one linguistic and one conceptual, have been raised that deserve our attention. The first idea is that priorities that are expressed overtly in the linguistic signal are treated as primary priorities for purposes of interpreting teleological modal claims. The second idea is that there is a conceptual difference between goals and sub-goals, and that this difference determines the cutoff point between primary and secondary conversational backgrounds. I discuss these two ideas in turn, concluding that neither one offers a viable basis for distin-
guishing between priorities in the way that explains strength differences among necessity modals.

Let’s consider the linguistic approach first. In the case of teleological backgrounds, von Fintel and Iatridou (2008) make provisions for priorities expressed overtly, in rationale clauses or in antecedents of anankastic conditionals. They assume that the proposition designated by an *(in order) to* adjunct or an *if you want to* conditional makes it into the primary teleological ordering source, while propositions that describe ways of achieving the designated priority are mapped into the secondary ordering source (ibid., p. 119). As we saw in Section 2.2.1, rationale clauses play a similar role in Finlay’s (2009, 2010) comparative probabilistic analysis, where they are used to determine domains of quantification for strong necessity modals, and probability spaces for the interpretation of weak ones.

The idea that goal-setting adjuncts determine the content of primary ordering sources is at best a partial response to the problem of distinguishing between priorities, since it leaves open how propositions are assigned to ordering sources in the absence of overt goal-setting expressions. In ordinary conversation, where conversational backgrounds are typically implicit and are left for the conversational participants to infer, we have to worry about the principles governing where one conversational background ends and another begins.

Before turning to focus on these cases in the following sections, I would like to argue that even in the presence of a goal-setting adjunct, it is incorrect to assume that the proposition expressed by the adjunct determines the content of the primary teleological ordering source. The first problem is that strong necessity modals can, in certain contexts, take into account both the overtly expressed priority and a subsidiary consideration which relates to how this priority can be achieved. And second, a weak necessity modal can, in certain contexts, rely just on the overtly expressed priority, without taking any additional consideration into account (while a strong necessity modal is unable to pick up on the relevant priorities in the same context).
Examples of the first kind are ones in which an overtly specified priority provides “too little information” for a strong necessity claim. The Chinese train example in (30), attributed to Wolfgang Klein, is one such case. In this context, taking the Chinese train is not a necessary condition for reaching Vladivostok, because the Russian train goes there too. However, it is a necessary condition for reaching Vladivostok comfortably. While the overtly expressed priority is just that of getting to Vladivostok, von Stechow et al. (2006, 169) point out that speakers sometimes accept the sentence in (30) as true even if the modal in the main clause is a strong necessity modal like have to. This means that a subsidiary priority (in this case, the priority of comfort) can affect the interpretation of a strong necessity modal even when the modifying rationale clause does not express this priority.

(30) **Reaching Vladivostok.** Two trains serve the city of Vladivostok: a Russian train and a Chinese train. The Chinese train is much more comfortable.

To go to Vladivostok, you {have to, ought to} take the Chinese train.

The discussion in the literature following von Stechow et al. (2006) has treated differences in intuitions about (30) as a point of variation between speakers of English. I think it is more plausible that different contexts prompt speakers to reach different conclusions about the conversational parameters the modals are based on. While the judgment that the strong version of (30) can be heard as true is controversial, it is possible to increase the acceptability of this sentence by manipulating the context in a particular way. Consider the elaboration of the scenario in the following exchange:

(31) **You:** I’m planning to take a train to Vladivostok.

**Me:** Listen, I took the Russian train last year and the conditions on board were

---


25 Thus, even speakers that reject the strong have to in the Vladivostok scenario are likely to use it in other contexts in which the facts themselves do not license a strong necessity claim, e.g. You have to see this movie or You absolutely have to read this book! (Angelika Kratzer, p.c.).
horrible.

You: Wow. And you, unlike me, usually don’t care about comfort.

Me: I didn’t mind it too much, but I know you will.

To go to Vladivostok, you have to take the Chinese train.

The fact that the addressee’s comfort is assumed by both interlocutors to play a role in the deliberation seems to be relevant for licensing the strong necessity claim. This subsidiary priority is clearly in the background when the speaker insists that the Chinese train is necessary, i.e., that it is the only option available given the two priorities that were raised.

Next, I turn to examples in which the overtly specified proposition provides “too much information” for a strong necessity claim, because it specifies all the priorities that a following weak necessity modal picks up on. These examples are problematic for the idea that weak necessity modals are sensitive to unspoken priorities, i.e., priorities that are not expressed in rationale clauses and other goal-setting adjuncts. In the Vladivostok scenario, for example, (32) can be used very naturally to single out the Chinese train as the necessary mode of transportation, even if getting to Vladivostok and doing so comfortably are the only considerations assumed to govern your choice of train. There is no implication that additional preferences are at play.

(32) To go to Vladivostok comfortably, you ought to take the Chinese train.

Similar examples in the muffins scenario and in the scenario of driving to Amherst are given in (33) and (34), respectively. They are all examples of weak necessity claims in which an overt rationale clause mixes the primary and secondary preferences that ought is sensitive to.

(33) [There is a greater chance of getting a banana muffin at the first bakery.]

To get those muffins without taking unnecessary risks, you ought to go to the first bakery.

(34) [There are four different ways to get from Cambridge to Amherst by car.]
a. To get there going 65 mph all the way (legally!), you ought to take the turnpike.

b. To get there from the north, you ought to take Route 2.

c. To get there on a road with a lucky number, you ought to take Route 9.

These examples are problematic for the linguistic approach to individuating priorities. If, in order to accommodate data like (32)-(34), secondary ordering sources are allowed to be empty when *ought* is used, we lose the ability to distinguish between weak and strong modals in any principled way: given this assumption, there should be no difference between a strong necessity claim and a corresponding weak necessity claim in which the secondary source of preferences is empty. Yet the choice between *ought* and *must* in a sentence like (32) is intuitively meaningful even if no additional priorities are relevant beyond those stated in the rationale clause. (I begin probing into the factors that influence the choice of modal in (35)-(36) below.) Alternatively, under the assumption that *ought* requires a non-empty secondary ordering source in order to be used felicitously, (32)-(34) would be predicted to be true only if there is some other salient priority which can be retrieved from the context, above and beyond those expressed in the (in order) to adjuncts in those sentences. This is too strong. My intuition is that one can truthfully utter (32) even if only comfort is believed to govern your choice of train (and similarly in the other examples; no considerations above and beyond minimizing risk, driving quickly, etc., are relevant).

I conclude from this discussion that designated goals in construction with weak necessity modals are not always supplemented by additional unspoken preferences. From the perspective of the modal, I conclude that goal-oriented *ought* is not always sensitive to preferences above and beyond those stated in a modifying rationale clause.

What then determines the choice of modal when all the relevant priorities are stated overtly in a rationale clause? I will argue below that what is relevant is the commitment of the conversational participants to these priorities. The notion of commitment can be appreciated by observing the effects of contextual manipulations on the ability to use weak and
strong necessity modals in a conversation. Consider, for example, the two different elaborations of the Vladivostok scenario below. In both scenarios, Babushka publicly commits to wanting a comfortable train ride. But the identity of her interlocutor seems to have an effect on whether or not this desire ends up being collectively committed to. Babushka’s daughter can more easily accept her mother’s desire as a reason for ruling out the option of the Russian train. Accordingly, it would be somewhat artificial of her to use a weak necessity modal in her response in (35); ought would be too weak in this case (and also in (31)). The clerk, on the other hand, would be more hesitant to rule out options for Babushka, and more comfortable letting her call the shots. A strong necessity claim from the clerk in (36) would be too strong.

(35) [Babushka is preparing to take the Trans-Siberian train from Moscow to visit her daughter in Vladivostok.]
Babushka: I want the trip to be pleasant.
Daughter: To get here comfortably, you have to (/ought to) take the Chinese train.

(36) [At the train station, Babushka asks the clerk about the trains to Vladivostok.]
Babushka: I want the trip to be pleasant.
Clerk: To get there comfortably, you ought to (/have to) take the Chinese train.

The idea that collective commitment is implicated in the semantics of necessity modals of different strengths is motivated in Section 2.3 and formalized in Section 2.4.

In light of the problems facing the linguistic approach to the primary/secondary split among priorities, one might take a more conceptual approach. In the deontic domain, for example, there has been an intuition that weak necessity modals are sensitive to rules one might call “soft”: rules that can be ignored without leading to severe consequences (Bybee et al. 1994, 186). A corresponding distinction in the teleological domain could be made between important and less important goals, or goals and different ways of achieving them. This way of explaining the split among contextually determined priorities is suggested in the following quote from von Fintel and Iatridou (2008).
It should be noted that the choice of what is a primary ordering source and what is a secondary ordering source is presumably not an accident. In the goal-oriented case we have the designated goal and measures of ways of achieving it, in the epistemic case we have hard and fast evidence and guesswork based on unreliable assumptions about the normal course of events, and in the deontic case we have strict laws and less sanctionable codes of behavior. (von Fintel and Iatridou, 2008, 119-120, footnote 11)

While there is most certainly a connection between how strict a rule is and how strong the necessity is to follow it, the suggested conceptual classification of rules and other priorities cannot be the basis for identifying the content of secondary ordering sources. My main argument against this approach is based on the observation that, in certain circumstances, even relatively “soft” or insignificant priorities can form the basis for strong necessity claims. (31) above is a relevant example, since added comfort is arguably not a strong reason for ruling out what certain people would take to be a perfectly good way of achieving the primary priority of reaching one’s destination. In (37), whatever it is that makes purchasing new shoes a strong necessity is also only a secondary priority to what is plausibly the speaker’s primary priority of being happy. Finally, we have already seen that secondary priorities that describe ways of achieving a primary priority can be treated as primary priorities themselves, in the sense that they provide the basis for the utterance of strong necessity claims (as in (35)).

(37) i have to get those shoes, i keep seeing them everywhere and they’re so tempting!!!

Overall, we have seen in this section that the content of rationale clauses does not reliably correspond to the primary preferences in a discourse, i.e., those preferences that a strong modal like have to picks up on. I have also argued that it is crucial for the success of the domain restriction approach that the distinction between primary and secondary pref-

ferences be well defined. In the next section, I state this challenge in terms of the second, more detailed, implementation of the approach proposed by von Fintel and Iatridou (2008).

2.2.2.2 Cascades of priorities

Weak necessity is expressed in English by designated lexical items (*ought* and *should*), but crosslinguistically, von Fintel and Iatridou (2008) show that a more compositional method is found as well. In certain languages, the use of counterfactual morphology on a strong necessity modal results in the expression of weak necessity. In developing their analysis, the authors aim therefore for a compositional implementation that will be able to turn a strong necessity modal into a weak one by marking the former with counterfactual morphology. If strong necessity modals are normally sensitive only to one, primary, ordering source, how is it that they manage to “reach out” to a secondary ordering source in one particular morphosyntactic context?

The main idea von Fintel and Iatridou (2008) pursue to account for this fact is that modals are sensitive to a sequence of ordering sources, and that there is a bipartition – provided by the context – between the sequence that strong necessity modals are sensitive to and the sequence that weak necessity modals are sensitive to. In this setup, if a grammatical operation “promotes” an ordering source from the secondary sequence to the primary sequence, a strong necessity modal would appear to be interpreted in the same way that a modal like *ought* is.27 A more formal statement of the proposal is given in (38). It is the most detailed version of the proposal provided by von Fintel and Iatridou (2008), and thus

---

27 The alternative direction von Fintel and Iatridou (2008) mention is that of partial union of ordering sources. The idea would be that counterfactual marking triggers the union of the primary and the secondary ordering sources (presumably while keeping a copy of the former, in order to interpret non-counterfactual strong necessity modals correctly), but only propositions from the secondary ordering source that do not conflict with propositions in the primary ordering source would be added (ibid., p. 138). As the authors mention, this kind of union operation may not be deterministic, and so it is problematic. An operation of ordering source merger that seems more promising is investigated by Katz et al. (to appear), and thus it may be possible to pursue this direction further. Importantly, however, the representational issue of how to represent cascades of priorities, whether using multiple ordering sources or a single ordering source that is concocted in just the right way, is separate from the conceptual question I raise in these sections.
the one I will refer to with respect to the problem of distinguishing between priorities in what follows.

(38) The context provides for each modal, a modal base $f$ and a bipartitioned sequence of ordering sources $<<g_1, \ldots, g_i>, <g_{i+1}, \ldots, g_k>>$.

- Strong necessity modals say that the prejacent is true in all worlds in $\max_{g_i(w)}(\ldots(\max_{g_1(w)}(\bigcap f(w))))$.
- Weak necessity modals say that the prejacent is true in all worlds in $\max_{g_k(w)}(\ldots(\max_{g_{i+1}(w)}(\max_{g_i(w)}(\ldots(\max_{g_1(w)}(\bigcap f(w)))))))$.

(von Fintel and Iatridou, 2008, 138, footnote 35)

There may be multiple ordering sources in the first section of the bipartitioned sequence, and multiple ones in the second section as well. The function $\max_{g(w)}$ (for ‘maximally good according to $g(w)$’) retrieves worlds either from the set provided by the modal base $f(w)$ or, recursively, from the set of worlds that were found to be best according to a previously applied ordering source. Given a set of worlds and a set of propositions, the function outputs the subset of the worlds that are best according to the ideal represented by those propositions (the “better than” relation $\sqsubseteq_A$ is based on the standard Kratzerian ordering relation between worlds; see Appendix A). This is of course an analysis of weak necessity in terms of domain restriction, since weak necessity modals have access to more ordering sources than strong necessity modals do.

(39) For any set of worlds $X \subseteq W$ and set of propositions $A$:

$$\max_A(X) = \{w \in X : \neg \exists w' \in X. w' \sqsubseteq_A w\}.$$  

(Following von Fintel and Heim 2009)

As I argued in the previous section, the domain restriction approach to weak necessity is not complete without an explanation of how priorities are divided up between ordering sources. This challenge is exacerbated in the setup in (38). If before we had to worry
about just two ordering sources, now there are multiple ordering sources we would need to distinguish between. How is \( g_1 \) distinct from \( g_2 \) and the other ordering sources in the primary sequence? How are \( g_{i+1} \) and \( g_{i+2} \) to be distinguished in the secondary sequence? The prospects of making these distinctions on conceptual grounds seem grim, since the number of distinctions that need to be made is potentially very big, and worse still, it would seem to vary from context to context.

The challenge for a theory of weak necessity concerns the cutoff point between \( g_i \) and \( g_{i+1} \). For a given proposition \( r \) representing a salient priority in the context, we need a criterion for determining if it belongs in an ordering source on the primary side of the bipartition or in one on the secondary side. We need to know what defines the boundary between the two sections of the bipartition.

### 2.2.2.3 Summary

The domain restriction approach to weak necessity views weak necessity modals as universal quantifiers over possible worlds, with a domain of quantification that is restricted in comparison to that of strong necessity modals. A smaller domain results from the application of more ordering sources in weak necessity claims. *Ought to* uses all the ordering sources that *have to* does to retrieve relevant possibilities out of a modal base, plus at least one more. von Fintel and Iatridou (2008) assume a bipartition on ordering sources that determines a boundary between the first \( i \) ordering sources in the sequence, and all the rest. Only necessity modals like *ought to* take into account ordering sources from the second section of the bipartition.

My central concern in presenting the domain restriction approach was to highlight the importance of understanding what determines the cutoff point between the first and second sections of priorities. Using the choice of necessity modal as a guide, we saw that overt rationale clauses do not reliably do this job. Sometimes they express both primary and
secondary priorities; other times they express just a subset of the primary priorities. Most of the time there are no direct linguistic cues to rely on at all.

### 2.3 Motivating sensitivity to collective commitment

My goal for the remainder of this chapter is to answer the following central question: what property of the contextually determined conversational backgrounds is relevant for the choice of necessity modal in a given context? Assuming that some kind of sensitivity to this property is encoded in the lexical entries of weak and strong modals, the question can also be stated from the perspective of the language user: what in the semantics of weak and strong necessity modals guides a speaker to pick one over another in a given conversation? An answer to this question has implications for any theoretical approach to weak necessity. For the domain restriction approach discussed in Section 2.2.2, it amounts to a discovery of the organizational principle behind the bipartition of ordering sources suggested by von Fintel and Iatridou (2008).

In this section, I motivate an answer to this central question. I hypothesize that presuppositions about collective commitment toward a premise in a conversation determine which kind of necessity modal, weak or strong, can be sensitive to it. I show that strong necessity modals are only sensitive to prioritizing premises that the conversational participants are presupposed to be collectively committed to. At any given point in the conversation, these are the priorities (goals and sub-goals, norms, stereotypes, and so forth) that are assumed to form the ground rules of the modal deliberation: if any participant in the conversation were given the chance to defend these priorities, it is assumed in the context of the conversation that they would do so. Weak necessity modals take into account all these premises plus some more. For these additional premises, lack of collective commitment is presupposed. A presupposition of lack of collective commitment boils down to an assumption that there is a conversational participant who is not committed to the relevant premise, at that point in the conversation. Put another way, a speaker uses a weak modal when he or she believes
Table 2.3: Presuppositions about commitment to priorities. A proposition $q$ is presupposed to be collectively committed to if for every conversational participant (A, B), the set of propositions that participant is presupposed to be committed to includes $q$.

(perhaps mistakenly) that the secondary priorities it depends on are still up for discussion.

Table 2.3 presents my usage of the key terms graphically.

As initial motivation for this hypothesis, consider the two scenarios below.

(40) **Health insurance.** Rachel is coming to the United States next summer. It is now illegal not to have health insurance in the US.

a. [Rachel will be a graduate student in a respectable university.]

   She has to (/ should) get health insurance.

b. [Rachel will be employed illegally, selling cheap jewelry at a shopping mall.]

   She should (/False has to) get health insurance.

The consideration of following the law, which necessitates Rachel’s getting health insurance, is salient and relevant in both of these scenarios. In (40a), it licenses a strong necessity claim. In (40b), it only licenses a weak one. When we probe the judgment that a strong necessity claim is false in the scenario in (40b), the reason seems to be that it is

---

28 I thank Angelika Kratzer and Paul Smolensky for discussion of these examples. For more teleologically flavored examples, see the elaborations of the Vladivostok scenario in (35)-(36) above.
not taken for granted that considerations of lawful conduct apply in the case of illegally employed workers. In contrast, such considerations are presupposed to guide the actions of those affiliated with a respectable university. Thus, moving to a conversation in which there is no commitment to the crucial priority appears to affect the choice of necessity modal. Note that the personal preferences of the individual whose actions are under deliberation, Rachel in this case, are not directly relevant for the choice of modal. It could be that Rachel herself would prefer to follow the law and get health insurance, and perhaps she even made this preference of hers public. But this is not enough to make a strong deontic statement with have to true in a conversation in which (40b) is uttered, where whoever makes this claim and their addressee assume that Rachel will be illegally employed.

Converging evidence for the collective commitment hypothesis comes from the interaction of necessity modals with the expression of priorities in imperatives and self-oriented directives, with rising declaratives, and with conditional questions. These three linguistic phenomena have been linked to the representation of mutual presuppositions in discourse, independently of any discussion about the strength of necessity modals.

The first piece of evidence comes from the interaction of necessity modals with directives (including imperatives and self-directive utterances) that introduce priorities into a conversation. In the expression of priority type necessities, the choice of modal is seen to vary in relation to how the priorities it is relativized to were introduced. While accepting an imperative prepares the ground for a strong necessity claim based on the priority voiced in the imperative, a weak necessity claim that is relativized just to this priority is infelicitous. On the other hand, if a speaker’s directive fails to be accepted due to a clash with what another discourse participant is presupposed to prefer, only a weak necessity modal can take the priority expressed by the directive into consideration. This interaction can be

29 The focus of my discussion will be on weak necessity claims with priority root interpretations, although the analysis I present is intended to cover epistemic weak necessities as well. I will stick to ought and have to as representatives of weak and strong modals, respectively, in order to make root interpretations of the modals more easily available.
explained with the assumption that (“successful”) imperatives update a repository of mutually presupposed priorities (Portner, 2004, 2007). While strong necessity modals can only reference priorities that are presupposed to be collectively committed to, weak necessity modals that are keyed just to such priorities are infelicitous (Section 2.3.1).

Second, necessity modals split with respect to whether or not they license *rising declaratives* stating the priorities they are relativized to (Section 2.3.2). Weak necessity modals license certain rising declaratives, but strong necessity modals do not. Specifically, if $p$ denotes a priority that makes true a necessity statement *ought/have to* $q$, a subsequent utterance of $p$ with rising intonation (notation: $p?$) is felicitous if the modal is weak, but infelicitous if the modal is strong. Following Gunlogson (2001, 2008), if we assume that rising declaratives signal a request for confirmation of the content of the declarative, the felicity of uttering $p?$ after an *ought* claim can be interpreted as evidence that the speaker who uttered $p$ is not sure about its status as a priority in her conversation with the addressee. This means that at that point in the conversation it is not presupposed that the addressee would defend this priority, and thus the priority is presupposed not to be collectively committed to.

Third, the discourse dynamic of *conversational backoff* (Rawlins, 2011) diagnoses the modal backgrounds of weak necessity modals as a mixed bag in terms of collective commitment (Section 2.3.4). Only a proper subset of these assumptions can be raised by a conditional question to challenge the weak necessity claim without resetting the conversation. Following Rawlins’s (2011) analysis of conversational backoff, this means that the contextual assumptions in this subset are presupposed not to be collectively committed to by the discourse participants. When a strong necessity claim is challenged by a conditional question, and when certain other background assumption of a weak necessity claimed are challenged, the conversation is reset.
2.3.1 Necessity modals and the expression of priorities in directives

In searching for a criterion to distinguish between the priorities that different necessity modals are relativized to, it is useful to look at how exactly these priorities are introduced into a conversation. In this section, I focus on contrasts in the use of strong and weak necessity modals that correlate with the manner in which their background priorities are expressed. I observe a difference in the strength of necessity supported by self-oriented directives on the one hand, and less committal expressions of personal preference on the other hand. Of the two, only the first introduce priorities that support subsequent strong necessity claims. A presupposition that the speaker and the addressee are both committed to the priority is argued to be required for the use of a strong necessity modal, since self-oriented directives fail to support strong necessity claims precisely when the addressee is not committed to the priority, or is presupposed not to be committed to it. A similar interaction is observed with imperatives. The choice of necessity modal is argued to depend on whether or not the individual to whom the imperative is directed is assumed to accept the obligation placed on her. Only if the addressee is assumed to be committed to the priority voiced by the imperative is it possible to describe the necessities that follow from this priority as strong necessities. These facts support the hypothesis that strength of necessity for unembedded modals is sensitive to (pragmatic) presuppositions about collective commitments in the discourse context.\(^{30}\)

There are many ways that speakers can put the preferences they have for themselves into words. Two I will consider here are expressions like *such-and-such would be nice* or *it would be nice to such-and-such*, exemplified in (41a), and self-oriented directives, exemplified in (41b). Although the speaker expresses a preference for avoiding toll roads and big cities in both cases, the strength of the bouletic/teleological necessities that follow

---

\(^{30}\) In this work I focus on the interpretation of unembedded necessity modals. Extending the hypothesis to embedded instances of these modals would require careful investigation of which derived context, if any, is relevant for checking collective commitment in those cases.
from this preference are different in each case. *Ought* is felicitous and leads to a true statement in (41a), but it is too weak following the directive in (41b). *Have to*, on the other hand, is felicitous and leads to a true statement in (41b), but it is too strong and therefore infelicitous in the first dialog.

(41) *[Driving to Amherst.]*

a. You: Avoiding toll roads and big cities would be nice.
   Me: OK then, you ought to (/have to) take Route 2.

b. You: No toll roads and big cities please!
   Me: OK then, you have to (/ought to) take Route 2.

Intuitively, there is a difference in how “strong” the speaker’s preference is in these two cases. We might say that the preference itself is “weak” in the first dialog, and “strong” in the second. Another way to express this intuition is that in the first case the speaker comes across as being less committed to the preference, in the sense that she would give it up if it conflicts with another preference, or if it is directly argued against. In the second case, she would not. One might wonder, therefore, whether the contrast in (41) is simply a reflection of a qualitative difference between “strong” and “weak” preferences of the speaker.

By considering what happens when the priorities of the conversational participants clash, we can see that the level of personal commitment to a priority is not the correct characterization of the difference in modal strength that we are after. Consider the yoga scenario in (42).

(42) **Yoga or a play.** A friend and I are deciding what to do together after a doctor’s appointment that I have. We’ve narrowed down the options to yoga and a play. At the doctor’s office, I get a physical.

   Doctor (to me; friend is in the room, listening): You are not in good shape. Don’t skip those yoga classes!
a. Me (to friend): News flash... I have to go to yoga.

b. Me (to friend): I really should go to yoga.

The priority in this scenario is voiced forcefully, in an imperative, and the speaker (the doctor) is clearly strongly committed to it. However, this alone is not enough to license a strong necessity claim based on this priority. What matters for the choice of necessity modals is the addressee’s response to the speaker’s forceful expression of preference, and whether or not she is presupposed to be committed to the priority directed at her. Let’s consider two relevant sub-cases. In case the addressee is a cooperative patient who is known to take doctors’ orders seriously, there is successful uptake of the imperative. Consequently, commitment of the addressee to the doctor’s order is presupposed in the conversation (as well as commitment on the part of the friend, who is also in the room). In this case, a strong necessity claim like (42a) is felicitous and true. However, if the patient is known to be skeptical of doctors’ orders, the imperative seems to lose its force. The patient is not presupposed to be committed to the order in this case (she might not defend it), and concomitantly the intuition is that only should, a weak modal, can access the priority in this case. In this situation, (42b) would be used to express the necessity that follows from this priority, instead of (42a).

The choice of necessity modal is thus seen to correlate with what the conversational participants are presupposed to be committed to. In the case of a cooperative addressee, the priority the modal is sensitive to is presupposed to be committed to by all the conversational participants, and a strong necessity modal is chosen. In the case of a resistant addressee, the presupposition that at least one individual (namely, the addressee) is not committed to the priority results in a presupposition of lack of collective commitment. Only a weak necessity modal can access the relevant priority under such circumstances.

31 I thank an anonymous reviewer for this observation.
It is possible to construct examples in which the successful uptake of an imperative is not achieved directly, with the addressee willingly committing to the priority expressed by the imperative, but indirectly, due to the influence of a third party. Suppose that with me inside the doctor’s office is not just my friend, by also my mother. My mother is very much committed to my wellbeing, so much so that I will never defend in front of her any priorities of mine that conflict with my wellbeing. With my mother in the room, I am presupposed to be committed to the doctor’s orders, and (42a) is the only felicitous necessity claim I could make. If I utter the weak (42b) instead, I am in big trouble. This choice of modal on my part would suggest that I am not committed to do what the doctor told me to do, i.e., that I not committed to the priority of not skipping yoga.

Of course, weak necessity modals are not banned from following imperatives and being relativized to the priorities they express. In such cases, the modal is relativized not just to the priority contributed by the imperative (as argued by Portner 2004, 2007), but to some additional priority as well. (43) is a relevant example. Let’s assume that B is committed to following A’s order (he is a cooperative employee). If the 7 a.m. flight to New York were the only way to get to the meeting with the bankers, B would use a strong modal in his response; should would be too weak. B would also use a strong necessity modal if there are multiple ways to get to the meeting with the bankers on time, but company policy – to which both A and B are presupposed to be committed to – rules out all but the 7 a.m. flight. Indeed, any defensible further premise would allow B to present taking that particular flight as a strong necessity. Thus, intuitively, B makes a weak necessity claim when the further premise he has in mind is one that might be contested. This intuitive judgment is supported by the tests I present in the following two subsections.

(43) A: Go present this proposal to our bankers today!
   B: I should take the 7 a.m. flight to New York then.

(Portner, 2007, 353(2))
The intuitive difference between “weak” and “strong” preferences of a speaker, which was raised following (41) above, can also be explained in terms of how the speaker and her conversational partners are presupposed to construe the preferences she expresses. When the speaker uses a directive to express a self-directed priority, as in (41b), she commits herself to this priority. Unless the addressee has grounds for challenging a priority that is expressed this way, he or she will accept it and be committed to it too (in this particular example, this is the priority that “the speaker avoids toll roads and big cities”). In this kind of situation, the conversational partners are both presupposed to be committed to the preference, and a strong necessity modal must be used. However, the addressee need not accept the speaker’s self-oriented directive. When this is the case, as in (44), the addressee’s lack of commitment, which clashes with the speaker’s explicit commitment, requires the use of a weak necessity modal.32

(44)  You: No toll roads and big cities please!
   Me: What? No, you OUGHT TO take toll roads. The toll helps the battered state retirement fund.

Importantly, an addressee’s presupposed lack of commitment need not be real; it may be incorrectly presupposed. If you utter the directive in (44)/(41b) to a group of fellow travelers that includes officials from the state retirement fund, I could try and negotiate the choice of route with them by uttering (45), with ought to. This acknowledges the presupposed disagreement between the two of us on the one hand, and the rest of the conversational participants on the other. I wouldn’t use a strong modal like have to in this case. It doesn’t matter if I am mistaken about the actual commitments of the officials. The necessity claim must be weak in (45), even if the officials are actually committed to draining the retirement fund even more than it is, perhaps to secure some federal support for the little money that is still there.

32 I thank Angelika Kratzer for this variant of (41b).
[You, me, and a group of officials from the state retirement fund are going on a trip to Amherst.]

You: No toll roads and big cities please!

Me (toward fund officials): We ought/\^ have to take Route 2, but the toll definitely helps the battered state retirement fund.

Returning to less committal expression of personal preference, when the speaker in (41a) describes a merely-nice-to-have preference, she herself is presupposed not to be committed to this preference. If the issue came up and she had to defend the preference in a dialog like (41a), it is not presupposed that she would. In this case too, a weak necessity modal is used in a conversation in which it is presupposed that not all the conversational participants are committed to the priority the modal is relativized to (since at least the speaker is presupposed not to be committed to it).

In all the examples above, I argued for a connection between the strength of necessity a conversational context supports and what is presupposed in that context about the collective commitments of the conversational participants. More specifically, I argued for the following points:

- Strong necessity modals are relativized only to priorities that are presupposed to be collectively committed to.
- Weak necessity modals that are relativized just to such priorities are infelicitous.
- Weak necessity modals are relativized in part to priorities that are presupposed not to be collectively committed to.

Support for the first point, namely the connection between strong necessity and a presupposition of collective commitment, can be found in the literature on imperatives following Portner (2004, 2007).\textsuperscript{33} Portner proposes that imperatives update a component of

\textsuperscript{33} See especially Ninan (2005), Charlow (2011).
the context called the *To-do List*, which tracks commitments at the “mutually presupposed discourse level” (Portner, 2007, 359). In the examples above, the successful uptake of an imperative or a self-directive can thus be described as resulting in a presupposition that all the participants in the relevant conversation are committed to the priority denoted by the imperative. We saw that successful uptake is crucial for the felicitous use of strong necessity modals.

In the following two subsections, I provide additional empirical support for these points, focusing especially on the third one.

### 2.3.2 Necessity modals and rising declaratives

Weak and strong necessity modals split with respect to the licensing of subsequent rising declaratives, declarative sentences bearing rising intonation that function as questions. The first part of the pattern of interest is exemplified in (46a-b). In both dialogs, A utters a necessity claim and B and C follow by questioning the priorities that would make the necessity claim true. The follow-up declaratives by B and C are to be spoken with rising intonation, as indicated by the underlining and the sentence-final question mark. Following the weak necessity in (46a), B and C’s rising declaratives are felicitous. B’s utterance highlights the priority of passing through Worcester, whereas C’s utterance highlights the priority of taking a road with an auspicious number. Both priorities necessitate taking Route 9, as in the driving-to-Amherst scenario (see (7)). The situation with the strong modal is different. It is impossible to use a rising declarative to ask about the relevant priorities that *have to* in (46b) is relativized to, namely the priorities of driving to Amherst and not asking the aunt to borrow her car.

---

34 To avoid testing sentences that could receive parses as elliptical polar questions (e.g., *Do you want to see Worcester?*), the examples in this section all contain non-inverted auxiliaries (see Gunlogson 2008, 106-107).

35 I gloss over the complication that B’s utterance is not a direct statement of a priority, but a statement of A’s preference toward a priority. C’s utterance is similarly complex.
(46) **Family trip.** A family decides to take a road-trip to Amherst despite the fact that the family car is in the shop. An aunt has a free car she could lend, but the mother and her sister are not on speaking terms. On many previous occasions, the mother has made it clear that does not wish to speak to her sister ever again. There are four possible routes to Amherst.

a. A (mother): We should take Route 9.
   
   **B:** You’d like to pass through Worcester?
   
   **C:** An auspicious number for the road is important?

b. A (mother): We have to reserve a rental car.
   
   **B:** ?You’d like to go to Amherst?
   
   **C:** ?Not asking your sister is important?

In both exchanges, rising declaratives are used to raise a priority that would make the necessity claim true. Why do the priorities of the strong modal resist this kind of presentation, while those of the weak modal cohere with it?

The explanation cannot be that weak modals are simply good in this kind of discourse dynamic, because they are not always. Thus, alongside (46a), we have (47), where a priority that *should* is relativized to cannot be stated as a rising declarative following the modal. The same priority that was infelicitous in a rising declarative following the strong modal is infelicitous following the weak one.36

(47) A (mother): We should take Route 9.

36 We see this also in (i), which is the weak necessity variant of (46b). The two necessity claims do not form a “minimal pair,” however, since the contextual assumptions would have to change for a weak necessity claim to be felicitous in the first place (instead of the strong one). Similarly, there is no strong version of (46a) to consider; the strong necessity claim *We have to take Route 9* is false in the trip scenario.

(i) A (mother): We should reserve a rental car.

   **B:** ?You’d like to go to Amherst?
   
   **C:** ?Not asking your sister is important?
B: ?You’d like to go to Amherst?
C: ?Driving to Amherst is important?

The modal discourses in (46)-(47) diagnose a split in the priorities that weak necessity modals are sensitive to: some of these priorities are felicitous as rising declaratives, others are not. The ones that are not felicitous following should are also infelicitous as rising declaratives following have to. This split has not been previously observed, and it is difficult to explain if priorities are only divided according to “relevance” into those that are relevant and those that are not. All the priorities that should is relativized to in these examples are relevant, important, and contribute to its interpretation in context. The contrast between C’s utterances in (46a-b) is particularly puzzling. In both cases, C voices a priority that is ultimately a personal preference of A’s (an auspicious road number, not asking her sister). Why is there a contrast in C’s ability to raise these preferences in a rising declarative?

A clear split emerges when we examine the presupposed commitments of the interlocutors to these priorities. The priorities that are good as rising declaratives, in (46a), are priorities that have not been discussed previously in the conversation, and without special manipulation of the context, it is not taken for granted that the speaker of the modal claim is committed to them. Passing through Worcester is an example of such a priority. When C asks about this priority in response to the speaker’s should claim, he is trying to understand if this is the reason she has singled out Route 9 as a weak necessity. In contrast, the priorities that are bad in this construction, in (46b) and (47), are priorities that everyone (including the speaker) is presupposed to be committed to. Going to Amherst is one such priority, and not asking the aunt for a favor is another. The conversational participants are all presupposed to be “on board” with these priorities (which is why a strong necessity modal can be used when they are the only priorities that are relevant, as in (46b)).
This characterization of the difference between the priorities in terms of contextual commitments is supported by the analysis of rising declaratives in the semantic-pragmatic literature, to which I now turn.

A core observation in the recent literature has been that the distribution of rising declaratives is constrained by contextual factors (Bartels 1997, Steedman 2000, Gunlogson 2001, 2008, Nilsenová 2006, Poschmann 2008). The relevant factors have been argued to relate to the mutual beliefs of the participants in a conversation. Nilsenová (2006), for example, proposes that rising declaratives are semantically comparable to expressions of epistemic uncertainty. She analyzes the intonation of the final rise as contributing a meaning of epistemic possibility (comparable to that of it might be that), which is interpreted as a “test” on the content of the conversational participants’ mutual beliefs. Gunlogson (2008) allows a rising declarative to contribute to the mutual beliefs in the conversation like any declarative (i.e., no modalization is involved), but proposes that this contribution is not automatic. She argues that the final rise imposes a general condition of contingency which allows the update of mutual beliefs to go through only if the addressee ratifies the content of the declarative. I present the relevant components of Gunlogson’s proposal below and move on to explain how it relates to my discussion of rising declaratives following necessity claims.

First some background. Gunlogson (2008) assumes that when a declarative sentence is uttered, any declarative, it contributes information about the position taken by a conversational participant (by default, the speaker) in the discourse. In more technical terms, the utterance of the declarative updates the content of that participant’s commitment set: a contextual construct representing anything that participant “publicly treats as true for the

37 These studies are concerned primarily with English rising declaratives, with some attention to corresponding phenomena in Dutch and in French (see Beun 1989 and subsequent work and Nilsenová 2006, respectively).

38 Gunlogson (2008) concentrates on rising declaratives that are “not directly related to the content of a preceding utterance” (p. 102). This is a property that the modal discourses I have been discussing do not have. However, since Gunlogson’s formal proposals seem not to depend on this assumption, and since I believe her analysis provides insights into the discourse dynamic in the cases of interest, I will proceed under the assumption that it is does indeed apply to them.
purposes of the discourse” (ibid., p. 110). Note that this use of the term “commitment” is different from the use which is assumed in this dissertation. In particular, I have spoken about commitment to priorities, not just to beliefs, and have thus been able to describe individuals as being committed to propositions they do not necessarily treat as *true* in the context of a conversation, but merely desirable in that context. My application of the term can be seen as a generalization of the use which applies just to beliefs about “what is actually the case in the conversation.”

With this background, we turn to the main components of Gunlogson’s analysis of rising declaratives. The first component is a criterion for interpreting declarative sentences as questioning, the *Contingent Commitment Criterion* given in (48). The definition of contingency that is referred to in (48) is given in (49). Gunlogson proposes that, in general in English, rising intonation marks contingency. On a declarative, it marks that the speaker’s commitment to the content of the declarative is contingent upon the addressee’s ratification of this content. In terms of the formal definition in (49), the discourse move $\mu$ is a move of committing the speaker to the content of the declarative he uttered, and $\delta$ is the discourse condition that the addressee be committed to that content as well, based on her own judgment. Finally, there is an inequality between speaker and addressee that is inherent to the ratification process. The “context must cooperate in making evident the superiority of the addressee” in being able to affirm the content of the declarative (ibid., pp. 128, 132).

\begin{equation}
(48) \text{ Contingent Commitment Criterion:} \\
\text{An utterance of a declarative with content } \varphi \text{ is questioning to the extent that the speaker’s commitment is understood as contingent on the addressee’s ratification of } \varphi. \\
\text{(Gunlogson, 2008, 129(48))}
\end{equation}

Gunlogson (2008) notes that, “The commitment set is similar in many respects to Stalnaker’s (1978) notion of a context set, with the significant difference that commitment sets are relativized to individuals” (p. 107). Importantly, the Stalnakarian notion characterizes the information that all the conversational participants treat as true, and which they all believe they all treat as true, and so forth.
A discourse move $\mu$ by an agent $\alpha$ is contingent upon a discourse condition $\delta$ if:

a. $\delta$ does not obtain at the time of $\mu$.

b. It is inferable in the discourse context that the update effected by $\mu$ is to be retained only if $\delta$ obtains after the discourse move immediately succeeding $\mu$.

(Gunlogson, 2008, 128(45))

Tying the process of ratification to the addressee’s superior ability to affirm the content of the declarative allows Gunlogson to account for the observation that rising declaratives are odd “out of the blue” (e.g., (50b)). Out of the blue, there is no indication in this scenario that Harry has insights about the weather forecast for the weekend, and thus he is not in a better position than Gina to affirm the statement that the weather will be good. Polar interrogatives do not place similar preconditions on the context, and they may be felicitous out of the blue (e.g., (50a)).

(50) [Gina to her officemate Harry:]

a. Is the weather supposed to be good this weekend? (Polar interrogative)

b. ?The weather’s supposed to be good this weekend? (Rising declarative)

(Gunlogson, 2008, 103(6))

The analysis also explains why (50b) improves if the context is manipulated in such a way that puts the addressee in a superior position with respect to affirming the relevant content. Below, Harry is said to be “implicitly authoritative” when it is inferable in the context that he is able to judge whether the declarative is true or false without relying on input from Gina.

Say, for instance, that Harry is reading the local newspaper when Gina addresses him. Gina can see that he’s on the weather page, and although she is at some distance, she sees what appear to be several sunny-day icons in a row. In these circumstances Harry is implicitly authoritative; Gina’s imperfect view of the page is enough to make her a plausible source. The rising declarative [(50b)] is now fine and has the expected air of soliciting Harry’s ratification or rejection of Gina’s hypothesis.

(Gunlogson, 2008, 126)
In light of the first condition on contingency, (49a), the addressee must not be committed to the content of a rising declarative at the time the declarative is uttered. Indeed, if I believe that you believe the weather is supposed to be good this weekend, and you believe that I believe this, my asking you to ratify this information would be strange. In other words, at the time the declarative is uttered, there is a pragmatic presupposition that the addressee is not committed to the proposition denoted by the declarative. Thus, there is a pragmatic presupposition that not all of the conversational participants are committed to this proposition. (Upon the addressee’s subsequent ratification, however, both speaker and addressee are committed to the priority.) Therein emerges a tight connection between the analysis of rising declaratives proposed by Gunlogson (2008) and the distinction I have been motivating between weak and strong necessity modals.

When a statement of priority is uttered felicitously as a rising declarative, it follows that the conditions on questioning and contingency in (48)-(49) are met in the conversational context. Specifically, it follows that the priority is presupposed not to be collectively committed to (since the addressee is presupposed not to be committed to it, as per (49a)). In the modal discourses above, priorities of this kind were found only among the priorities that weak necessity modals are relativized to, e.g., in (46a).

What does the analysis allow us to conclude about priorities that are infelicitous as rising declaratives in the modal discourses above? Here, since there are multiple components to the analysis, there are multiple potential culprits. It could be that the infelicity results from some precondition for ratification not being met in the conversational context. For example, perhaps it is presupposed that the addressee is unable to commit to the priority (e.g., if she has committed to a conflicting priority), or perhaps she happens not to be better equipped than the speaker to commit to the priority (so she does not enjoy “superior authority”). It could also be that some precondition for contingency is not met. The obvious suspects here are that the addressee is already presupposed to be committed to the priori-
ity, or that there is no need for ratification, since the speaker and the addressee are both presupposed to be committed to the priority.

We can rule out some of these reasons as potential explanations of the modal data. For example, we can set aside the worry that the addressee is unable to commit to the priority, or lacks superior authority to do so. She is, after all, the one who issued the necessity claim which the rising declarative is a response to, and we assume that this claim – whether strong or weak – was not judged false or seen to contradict her presupposed position in the conversation (consisting of the beliefs and preferences she is presupposed to be committed to). If the problem is not in the potential to ratify commitment to the priority, it must be a problem of contingency. Assuming that infelicity reflects a failed precondition of contingency, we can conclude that in these cases the addressee is already presupposed to be committed to the priority stated in the declarative (that is, (49a) is not satisfied). A related source of infelicity would be failure of the Contingent Commitment Criterion, as when both the speaker and the addressee are presupposed to be committed to the content of the declarative. There is no need for questioning then, and no contingency to delay the commitment to the relevant content.

The infelicity observed in (46b)-(47), for example, is due to a colossal failure of the Contingent Commitment Criterion. B and C’s declarative sentences express priorities that are presupposed to be committed to by all the conversational participants. Only someone who wasn’t part of the conversation could be unsure about the status of these priorities. In contrast, the declaratives uttered by B and C in (46a) are genuinely presented to A for her ratification, and they are perfectly felicitous as questions. Not only are these statements perfectly felicitous following the weak necessity claim, the discourse as a whole imbues a sense of coherence. The questioners use the rising declaratives in these examples to turn to A and invite her to elaborate on the priorities she had in mind when uttering the necessity claim, priorities they are not yet privy to.
Before I conclude this discussion, I would like to provide a final example of the presuppositional restriction on the secondary priorities that weak necessity modals are relativized to. The crucial feature of this example, (51) below, is that it contrasts two priorities that are secondary to the main priority of the scenario (presumably, the priority of reaching a given destination; this was the priority shown to be infelicitous after should in (47)). The speaker says that finding a good map is a weak necessity in the context of a road-trip. While there may be many reasons for wanting there to be map in the car on a road-trip, the obvious reason, namely not wanting to get lost on the way, is not something that can be presented as a question to A in a rising declarative. B’s first utterance in (51) is infelicitous as a question; it implies that B is very seriously out of touch with priorities that are simply taken for granted in the conversation. On the other hand, the priority of shutting down one’s smartphone is a perfectly good priority to be raised in a rising declarative (B’s second utterance below). This priority, which would indeed necessitate bringing along a good old fashioned map, represents a dramatic action that A is not presupposed to be committed to. A may commit to this priority in a subsequent utterance, or she could reject it and present the reason she actually had in mind when she made her weak necessity statement.

(51) A: We should find a good map of the area.

B: You’d like not to get lost on the way? You’d like to be able to turn off your smartphone and its GPS for the weekend?

In this section, I used the felicity of rising declaratives in modal discourse to argue for a connection between strength of necessity and presuppositions about collective commitment to the conversational backgrounds of necessity modals. At the heart of the argument was Gunlogson’s (2008) analysis of rising intonation on declarative sentences as a vehicle for conveying contingent commitment. According to this analysis, rising declaratives contribute an inference that the speaker requests the addressee’s confirmation or rejection of the content of the declarative. A precondition for contingent commitment is a pragmatic presupposition that the addressee is not committed to the proposition denoted by the declar-
ative. Thus, priorities that can be raised felicitously as rising declaratives are priorities for which lack of collective commitment is presupposed. Using rising declaratives as a probe of the status of priorities in the discourse, I argued that:

- Secondary priorities that weak necessity modals are relativized to are presupposed not to be collectively committed to.

Primary priorities that weak necessity modals are relativized to, and all the priorities that strong necessity modals are relativized to, were argued to come with a presupposition of collective commitment. In Section 2.3.4, I provide additional support for the above characterization of secondary priorities. Before presenting this final piece of evidence, however, I discuss a number of apparent counterexamples to the proposed characterization of primary priorities. This discussion will also lead to a clarification of whose commitment is at stake for characterizing priorities as primary or secondary.

2.3.3 What is the collective whose commitments matter?

When a mother tells her child that he has to eat his spinach, as in (52), the child is arguably not committed to the priority behind this strong necessity claim. He is a participant in the conversation, however, so this use of have to looks like a counterexample to the characterization of strong necessity modals as relying on primary priorities, namely priorities that are presupposed in the conversation to be collectively committed to.

(52) Mother: You have to eat your spinach!
   Child: I don’t want to!

An initial response I would like to argue against is that the relevant collective for determining commitment is “society at large,” or some group representing an idealized version of society at large. Appealing to society at large provides an explanation of (52), if we assume that the priority that children eat healthy vegetables is collectively committed to
among the leaders of the community the mother and her child are members of. Perhaps appealing to the authority of this group is what allows the mother to make a strong necessity claim in (52).

While this explanation may work in the spinach example, it fails to provide a general explanation of the nature of primary priorities. First, it is clear that norms of the idealized community do not dictate what groups within the community take to form strong necessities. After all, radicalists utter perfectly felicitous necessity claims that are based on priorities only they and a handful of followers are presupposed to be committed to. The strong necessity claim in (53a) exemplifies this possibility in the context of a conversation between environmentalist city dwellers. Second, relativization to the priorities of a collective representing society at large is not an option that is freely available to strong necessity modals. (53b) is false when uttered in the conversation between the environmentalists, although the necessity it expresses follows from the presupposed commitments of society at large (and would be true if uttered by the mayor in a conversation with members of the city council, for example).

(53) **Green city living.** Environmentalists discuss plans for going green in the city.

a. We probably shouldn’t bring in cows or goats, but we have to start a chicken coop behind the buildings.

b. We have to use the cheapest chemical solution to eradicate the rats in the alley.

How do speakers justify making strong necessity claims in conversations that include participants that are not “on board” with the priorities they are assuming? The problem raised by (52) is not limited to conversations in which the speaker retains authority over the addressee, but arises more generally. (53a), for example, could be addressed to an outsider, a neighbor who came to talk to the environmentalists although she does not share their priorities. In such a conversation, one authoritative conversational participant addresses another in an attempt to convince the latter to adopt the priorities the necessity claim is
based on. This aspect of the use of strong necessity modals is the basis of my explanation of why they are felicitous in (52) and similar examples, despite the conflicting commitments of the conversational participants in these examples.

Strong necessity modals are rhetorical devices for persuasion, I propose, precisely because they come with a presupposition of collective commitment to the priorities they are based on. This presupposition implies that the speaker of the necessity claim (who may very well be aware of the actual “conflict of interests” with her interlocutor) is treating her interlocutor as part of a collective that is committed to the priority. In cases of conflict, the speaker conveys that she is operating under the assumption that her interlocutor will change his commitments so that they align with the relevant commitments of the collective. The result is an air of forced update of commitments which is related to the ability of strong necessity modals to place obligations on the addressee they are directed to. The latter ability is referred to in the literature as the performativity of strong necessity modals (see Section 2.4.2 for further discussion).

The proposal above is inclusive of participants: it assumes that all the conversational participants are included in the collective whose presupposed priorities matter. An alternative solution to the challenge presented by examples like (52) would be to allow for exclusion of participants from the collective; the hypothesis would be that there is flexibility in the identification of the set of participants whose commitments matter for the purpose of the necessity claim, such that one or two, or maybe a handful, can be ignored if their priorities conflict with the priorities of the rest of the participants. After excluding those participants, the presupposition of collective commitment would be satisfied and it would be possible to make a strong necessity claim based on the relevant priorities presupposed to be collectively committed to in that group. I believe that the explanation based on inclusion is superior to this alternative, first, because it offers insight into the performative abilities of strong necessity modals. Moreover, it seems that exclusion runs the risk of incorrectly predicting “too many” strong necessities. For instance, if A, B, C, and D are all
presupposed to be committed to taking a vacation in France, but A and B are committed to preferring Bordeaux over Paris, while C and D are committed to Paris over Bordeaux, an analysis of have to with a presupposition of collective commitment based on exclusion would predict the truth of the conflicting claims: *We have to go to Bordeaux* (because A and B are “collectively” committed to this priority) and *We have to go to Paris* (because C and D are “collectively” committed to this priority). Intuitively, neither of these claims is true in the scenario with conflict, although they can be used to persuade the dissenting parties to change their priorities.

A final important clarification concerns necessity claims about third persons, or persons that are not part of the conversation in the speech situation. These do not present a problem for my proposed generalization about collective commitment. When mother tells father that, “The boy has to eat his spinach!,” she uses a strong necessity modal because the priority that the child eat healthy food is presupposed to be committed to by the two parents, i.e., the conversational participants. Of course, these priorities may be relevant in situations quite removed from the speech situation, as is the case in (54), for example.

(54) Mother (to father): I don’t care if his friends are watching. He has to eat his spinach at summer camp too!

### 2.3.4 Necessity modals and conversational backoff

The connection between weak necessities and priorities that are presupposed to lack collective commitment is further supported by the pattern of responses to weak necessity claims exemplified by (55). We see in this dialog a split between the *what/even if*-questions that A can ask about priorities that are relevant for B’s weak necessity claim. The basic observation is that only a subset of questions with a conditional component constitute felicitous responses to the *ought* claim.

(55) [Driving from Cambridge to Amherst.]
A: Which road will he take?

B: He ought to take Route 2.

\[
\begin{aligned}
\text{A:} &\begin{cases}
\text{What if he doesn’t care about scenery?} \\
\text{What if he cares more about speed?} \\
\text{Even if he wants to save time?}
\end{cases} \\
\text{Possible responses}
\end{aligned}
\]

\[
\begin{aligned}
\text{A:} &\begin{cases}
?\text{What if he doesn’t want to go to Amherst?} \\
?\text{Even if he’s not driving?}
\end{cases} \\
\text{Impossible responses}
\end{aligned}
\]

The addressee of the *ought* claim, A, cannot challenge the priority of going to Amherst in her question, or the priority that the trip will be by car. She can, on the other hand, use a question with a conditional component to raise priorities of scenery or speed.

Rawlins (2011) analyzes utterances like those in (55) as a type of non-agreeing response that effectively allows addressees to ask for clarification of speakers’ assumptions.\(^40\) The discourse dynamic involved is quite complex. It begins with the speaker offering an answer to some question under discussion, and proceeds with the addressee not fully accepting this answer. The addressee then asks a conditional question, thereby accepting the attempted answer for cases where the conditional is not true, and “re-asking” the question for cases where it is true (Rawlins, 2011, 348). *Conversational backoff* is the term Rawlins uses to describe the partial acceptance triggered by the conditional component of the question, which allows the addressee to challenge the exhaustiveness of the answer they have been given. In (56) below, the addressee challenges the speaker’s claim that Alfonso is going to the party by raising the issue of Joanna’s presence there. According to Rawlins (2011), by asking a conditional question, A accepts B’s answer for those cases that Joanna is not going to the party, and further asks whether or not Alfonso will go if Joanna is going.

(56) A: Is Alfonso going to the party?

B: Yes, he is.

\(^{40}\) As such, these responses are most similar to *clarification requests* (see Rawlins 2011, §2.1).
The addressee in the dialog in (55) can use a variety of conditional questions to back off from the speaker’s claim that the driver ought to take Route 2. For example, asking *Even if he wants to save time?* would be a way of accepting the necessity of taking Route 2 in case the driver doesn’t care about saving time, but questioning this conclusion in case the driver does have such a preference. Raising a relevant consideration in the question (i.e., that the driver wants to save time) allows the addressee to ask the speaker indirectly to clarify which priorities they had in mind when uttering the weak necessity claim.

Rawlins (2011) argues that a precondition for conversational backoff is that the assumption raised by the conditional question be undecided in the discourse. More formally, the proposition raised by the conditional question must be unsettled in the common ground, which in turn is modeled as the intersection of the individual participants’ commitment sets (following Gunlogson 2001, 2008).\(^4\)

A central feature of all these cases is that conversational backoff can only target publicly omitted assumptions – it is not licensed if the crucial assumption has been committed to by a speaker in prior discourse.

A: Is Alfonso going to the party?
B: Yes, he is, even if Joanna might be there.

... 

A’: #Even if Joanna is there?

The assumptions in play during conversational backoff are purely implicit. (Rawlins, 2011, 350)

---

\(^4\) The proposition \(q\) signaled by the conditional question must be unsettled in \(cs_X^C \cap cs_Y^C\), where \(cs_X^C, cs_Y^C\) are the worlds compatible in a context \(C\) with the public commitments of the discourse participants \(X\) and \(Y\) (Rawlins, 2011, 359). A proposition \(p\) is unsettled in \(c\) if and only if \(c \cap p \neq \emptyset\) and \(c \cap p \neq c\) (cf. ibid.)
In the successful cases of conversational backoff following a weak necessity claim, we can conclude that the assumptions raised in the questions following the modal claim had not been previously committed to in the conversation. Thus, it is unsettled whether the driver wants to save time, or cares about speed, or doesn’t care for pretty scenery. Consequently, I assume, the conversational participants are not committed to any of the related priorities (that the driver save time, that the scenery be unremarkable, etc.) as priorities they are prepared to defend. Considerations of a leisurely drive or pretty scenery would make Route 2 a necessary choice among the four possible routes to Amherst, and they are precisely the kind of considerations that would be included in a secondary ordering source in von Fintel and Iatridou’s system.

Conversational backoff further distinguishes between two types of priorities that weak necessity modals are relativized to. The secondary considerations are felicitous as triggers of backoff, but the primary ones – that the trip will be a road-trip, and that Amherst is the destination – are infelicitous as triggers. Notably, the propositions raised by the second batch of questions in (55) are mutually assumed in the conversation. It is part of the contextual setup that the driver wants to go to Amherst and that he plans to drive, and we may assume that the conversational participants are consequently committed to the related priorities (that the driver go to Amherst, that it will be a road-trip). This description of the primary priorities as enjoying a presupposition of collective commitment is consistent with Rawlins’s analysis, although it doesn’t follow from it (there might be other sources for the observed infelicity, for example).

In sum, we have seen that discourses with conversational backoff can be used to diagnose a split among the priorities that weak necessity modals are relativized to. Only a proper subset of these priorities can be raised in a subsequent conditional question to challenge the extent of a weak necessity claim, and these are precisely the priorities for which lack of collective commitment is presupposed.
Taken together, the interaction of necessity modals with conversational backoff, rising declaratives, and directives supports the hypothesis that necessity modals differ with respect to presuppositions about collective commitment to the priorities they are relativized to. I presented converging evidence from these different sources that, first, strong necessity modals only reference priorities that are presupposed to be collectively committed to in the conversation; second, the evidence suggests that weak necessity modals are relativized in addition to prioritizing premises that are presupposed not to be collectively committed to.

The next step is to formally state this newly found sensitivity to collective commitments, keeping in mind the desideratum of accounting for the scalar-like inferences that relate the two types of necessity modals.

2.4 Weak necessity as departure from collective commitment

In this section, I formalize an analysis of weak and strong necessity modals that exposes their different presuppositions about the priorities they are relativized to. I use the analysis to explain the observations about modal discourse discussed in the previous section (Section 2.4.1.1) and show that it derives the scalar-like inferences that relate weak and strong necessity modals (Section 2.4.1.2). As a further application, I suggest a new perspective on the phenomenon of the performativity of strong necessity modals (Section 2.4.2). Section 2.4.3 highlights differences between secondary priorities and conditional antecedents which argue against analyzing the former as a silent analog of the latter. Finally, in Section 2.4.5, I offer an interpretation of the modal landscape embodied by the proposal (an interpretation elaborated in Appendix B). Section 2.4.6 concludes and presents directions for future research.

2.4.1 Formal proposal

This section offers a formalization of sensitivity to collective commitment within a domain restriction approach to weak necessity.
I assume a representation of the conversational context $C$ as a tuple containing at least a modal base $f$, an ordered sequence of ordering sources, $<g_1,\ldots,g_i>$, and individual records of the priorities that the participants in $C$ are publicly committed to. These records, called priority slates, are modeled on analogy to the individual records standardly assumed to represent the discourse participants’ public beliefs (Hamblin 1971, Gunlogson 2001, 2008, Farkas and Bruce 2010, Rawlins 2011).

**Priority Slates** A priority slate of a participant $X$ in $C$ is a set of propositions:

$$ps_X = \{p_{<s,t>} : X \text{ publicly endorses } p \text{ in } C \text{ as desirable}\}.$$ 

**Collective Commitment to Priorities** A priority $p$ is collectively committed to in a context $C$ iff $p \in ps_X \cap ps_Y \cap \ldots$, where $X,Y,\ldots$ are participants in $C$.

The priority slate of a discourse participant should not be confused with their To-Do List, in the sense of Portner (2004, 2007). A To-Do List of a conversational participant keeps track of properties that participant is expected to make true of himself, such as properties denoted by imperatives that are directed at that participant. A priority slate, in contrast, concerns not how the participant is expected to act, but what he or she is committed to. It may contain priorities that individual has concerning the actions of other individuals (“that my son eat his spinach”), concerning desirable future states of the worlds (“that there be world peace”), and so forth. What the constructs have in common is that they operate “at the public, mutually presupposed discourse level” (Portner, 2007, 359). If the speaker’s To-Do List contains the property of taking Route 9 to Amherst, the corresponding proposition (“that the speaker take Route 9 to Amherst”) is on the priority slate of every participant in the conversation.

---

42 The notion of public endorsement that is intended here is similar to the one standardly assumed in the literature on common knowledge and the common ground. Propositions in priority slates represent priorities that the conversational participants are committed to, believe that others are committed to, and believe others believe they are committed to.
Following Kratzer (1981, 1991, 2012), modal bases and ordering sources are functions of type \(<s, <st, t>>\), mapping possible worlds to sets of propositions. I define primary ordering sources as ordering sources whose content is collectively committed to in the sense defined above.

**Primary Ordering Sources** A function \(g\) from worlds to sets of propositions is a primary ordering source in a context \(C\) iff for every world \(w\) in the domain of \(g\) and every proposition \(p\) such that \(p \in g(w)\), \(p\) is collectively committed to in \(C\).

I propose that ordering sources that are represented in the conversational context contain only priorities that are collectively committed to.

(57) **Condition on Contextual Ordering Sources**

Every ordering source in \(<g_1, \ldots, g_i>\) in \(C\) is a primary ordering source.

Strong necessity modals are relativized to the primary ordering sources in \(<g_1, \ldots, g_i>\).

The lexical entry I adopt for them, in (58), is identical to the lexical entry proposed by von Fintel and Iatridou (2008) (see (38)-(39) above). Writing down this denotation requires introducing the assumption that there is always a set of worlds that come closest to the ideal represented by an ordering source (an assumption that is not made in Kratzer’s semantics for modals, following Lewis (1973); see Appendix A). Multiple ordering sources are applied one by one to retrieve the set of favored worlds: the worlds in the intersection of the propositions in the modal base which are best according to the ideals represented by the sequence of primary ordering sources.

(58) \[[\text{have to}]\]^C = \(\lambda q_{<s,t>}\lambda w.\forall w' : w' \in Fav^C(w).q(w')\).

**Favored Worlds** \(Fav^C(w) = max_{g_i^C(w)}(...(max_{g_1^C(w)}(\bigcap f^C(w))))\), where for any set of worlds \(X \subseteq W\) and set of propositions \(A\): \(max_A(X) = \{w \in X : \neg\exists w' \in X.w' \sqsubset_A w\}\).

(Following von Fintel and Iatridou 2008, von Fintel and Heim 2009)
Weak necessity modals are also analyzed as universal quantifiers over possible worlds, with a domain of quantification that is a subset of the domain that strong necessity modals quantify over. A restricted domain is arrived at by employing a secondary priority, $b$, which is defined to be a priority that is not collectively committed to in the context (see Figure 2.3). It is left open that different participants would use different secondary priorities to base their weak necessity claims on. To reflect this potential source of disagreement among participants, secondary priorities, in contrast to primary ordering sources, are not allotted a designated slot in the discourse context. To use Lewis’s (1979) metaphor of the “conversational score,” secondary priorities do not belong on the scoreboard (independently of being represented elsewhere in the context, e.g., in the speaker’s priority slate) until their status in the conversation has been clarified.

In the preliminary proposal presented (59), no further restriction is encoded about secondary priorities in the semantics.

\[(59) \quad (Preliminary \, proposal)\]

\[
\langle{\text{ought to}}\rangle^C = \lambda q_{<s,t>} \lambda w. \forall w' : w' \in \text{Fav}^C(w) \cap b.q(w'), \text{ where } b \text{ is a secondary priority in } C.
\]

**Secondary Priorities** A proposition $p$ is a secondary priority in a context $C$ iff $p$ is not collectively committed to in $C$. 


Worlds favored according to $f^C$ and $g^C_1, \ldots, g^C_i$

In a context $C$, a weak necessity modal quantifies over a subset of the worlds that are accessible and highly ranked based on priorities that are collectively committed to in $C$. The secondary priority it employs is not represented in the context.

The secondary priority it employs is not represented in the context.

The secondary priority it employs is not represented in the context.

The picture suggested by Figure 2.3 is that a secondary priority can be any proposition that is compatible with the set of contextually favored worlds (so that intersecting this proposition with the favored worlds results in the highlighted subset of the favored worlds as *ought*’s domain of quantification). This picture may capture the logical weakness of *ought* in comparison to *have to*, but it is too simplistic.

The main problem with the lexical entry in (59) is that it predicts far too many weak necessity claims to be true in a given context. For example, in a conversation about which route to take to Amherst, (59) incorrectly predicts the truth of *You ought to take Route 9 and have your night driving glasses handy*. How is this prediction derived? Obviously, if it is not a collective commitment in the context that you will drive through Worcester, so it definitely is not a collective commitment that you will drive through Worcester at night. The problem is that the priority of driving at night through Worcester counts as a secondary priority in the context, since there is a subset of the accessible worlds (those in which you drive to Amherst) in which you drive at night through Worcester. Relativizing the modal to this secondary priority incorrectly predicts the above sentence to be true. To fix this
kind of problem, we need secondary priorities to cut down on the set of favored worlds without introducing considerations that are more fine grained than those under discussion in the conversation (e.g., time of travel). Defining priorities in relation to a Question Under Discussion (Ginzburg 1996, Roberts 1996, Büring 2003) can solve this problem. In the conversation about driving to Amherst, the question under discussion is which route the driver will take. We can require the secondary priority to constitute a pertinent answer to this question (following Groenendijk 1999), and this will rule out propositions that are too fine grained from being treated as possible candidates for secondary priorities.

Following Groenendijk and Stokhof (1984, 213ff.), we view the denotation of a question as a partition of a set of possible worlds. A partition of a set $X$ is a set of non-empty and non-overlapping subsets of $X$ (called the cells of the partition) whose union equals $X$. I assume that the set of worlds that undergoes partitioning in modal discourse is the set of favored worlds, $\text{Fav}^C(w)$.

**Question Under Discussion** A context $C$ includes a question, $\text{QUD}$, whose denotation in a modal conversation evaluated at $w$ is a partition of $\text{Fav}^C(w)$.

**Pertinent Answer** A proposition $p$ is a pertinent answer to a question $q$ iff $p \cap \bigcup \llbracket q \rrbracket$ is the union of some but not all cells of $\llbracket q \rrbracket$.

(Following Groenendijk 1999)

In my final proposal for the lexical entry of *ought*, below, it is required that the secondary priority that derives the modal’s restricted domain also cohere with the question under discussion in the conversation.

(60) (Final proposal)

$$\llbracket \text{ought to} \rrbracket^C = \lambda q_{<s,t>} \lambda w. \forall w' : w' \in \text{Fav}^C(w) \cap b. q(w') ,$$

where $b$ is a secondary priority in $C$ and $b$ is a pertinent answer to $\text{QUD}^C$.

With this definition, only certain secondary priorities can serve to focus on a subset of the favored worlds. For example, when the question under discussion is which route
the driver will take to Amherst, only propositions that pick out one, two, or three out of the four possible routes can play this role. Within these limitations, the options are still essentially unlimited: the secondary priority for taking a route with an auspicious number determines the same domain of quantification – and the same weak necessities – as the secondary priority of passing through Worcester. Both make it a weak necessity to take Route 9.

2.4.1.1 Discussion of motivating examples

In this section, I revisit some of the examples that were used in the preceding sections to motivate a commitment based split between necessity modals.

Stereotypes  The first example I wish to revisit is also the first one that was mentioned in connection with the notion of commitment, in Section 2.2.2.1. I presented this example, the Vladivostok example and its variants, as a challenge for analyses of weak necessity that associate rationale clauses with the expression of primary priorities.

(61)  [Two trains serve Vladivostok: a Russian train and a Chinese train. Of the two, only the Chinese train offers a comfortable ride.]

To go to Vladivostok comfortably, you {have to, ought to} take the Chinese train.

In this scenario, taking the Chinese train is a necessary condition for reaching Vladivostok and doing so comfortably. There need not be any additional priority above and beyond these for the ought claim to be true. From the perspective adopted in this chapter, we expect the choice of modal to correlate with whether or not these priorities are collectively committed to. This expectation is borne out by the elaborated contexts below (repeated from (35) and (36)).

(62)  [Babushka is preparing to take the Trans-Siberian train from Moscow to visit her daughter in Vladivostok.]
Babushka: I want the trip to be pleasant.

Daughter: To get here comfortably, you have to (/ought to) take the Chinese train.

(63) [At the train station, Babushka asks the clerk about the trains to Vladivostok.]

Babushka: I want the trip to be pleasant.

Clerk: To get there comfortably, you ought to (/have to) take the Chinese train.

In both scenarios, Babushka publicly commits to the priority of taking a comfortable train ride. The daughter, who knows her mother and her situation intimately, may thus presuppose that her mother is indeed committed to paying for a comfortable ride, and commit to this priority herself. Since there is a presupposition of collective commitment to the priority, the priority has the status of a primary priority and have to is relativized to it. The clerk at the train station, in contrast, does not know the woman who stands in front of him and says she wants a pleasant trip. He cannot judge how strong her desire is and whether she is willing to back it up, and in order to be on the safe side, he does not automatically endorse the priority himself. In this case, the priority does not have the status of a primary priority. Since it is a pertinent answer to the question of which train the passenger will take (it rules out an entire cell in the partition of the set of accessible goal worlds, according to nationality of the train), it is a priority that ought can be relativized to.

The expression of preferences in directives Next, let us return to the yoga scenario, repeated in (64), to examine the interaction between imperatives and varieties of necessity. We observed that the choice of necessity modal in this example varies with certain background assumptions about the speaker. Assuming that the speaker is cooperative, the doctor’s utterance of the imperative causes her to commit to the priority it denotes (namely, her not skipping yoga). It is further presupposed that this priority is on the priority slates of the doctor and the friend in the room, so collective commitment is presupposed and a strong necessity modal is used. If the imperative is rejected by the patient, for example if she chooses not to cooperate with the doctor, the relevant priority is not added to her
priority slate. Although the priority is not represented in a primary ordering source in this kind of context, it nevertheless qualifies as a secondary priority: the doctor’s order would rule out one of the two options the girls are discussing, so it can be used as a secondary priority to make a weak outht claim about going to yoga.

(64) [A friend and I are deciding what to do together after a doctor’s appointment that I have. We’ve narrowed down the options to yoga and a play. At the doctor’s office, I get a physical.]

Doctor (to me; friend is in the room, listening): You are not in good shape. Don’t skip those yoga classes!

Me (to friend): I have to/should go to yoga.

**Rising declaratives** In Section 2.3.2, I observed that certain statements of priorities are felicitously uttered with rising intonation following outht/should claims, while certain others are not. This contrast is seen in the example below (repeated from (51)).

(65) [Getting ready for a road-trip.]

A: We should find a good map of the area.

B: You’d like not to get lost on the way?

You’d like to be able to turn off your smartphone and its GPS for the weekend?

In (65), A utters a weak necessity claim based on some secondary priority she has, and B asks questions to discover what that priority is. Priorities that are collectively committed to, like the priority not to get lost, are not possible candidates for being secondary priorities. B’s first rising declarative is infelicitous because it violates the condition for contingency that the rising intonation contributes (following Gunlogson 2008). In her second utterance, B asks for A’s confirmation or rejection of a priority she does not know if he is committed to. Rising intonation on a statement of this priority is the perfect way to raise this question.
Conjunctions of incompatibles  A final issue I wish to discuss concerns conjunctions of weak necessity statements. The analysis I have proposed allows for the truth of conjunctions of the form $\text{ought } p \text{ and ought } \neg p$ (assuming that neither $p$ nor $\neg p$ are presupposed to be collectively committed to in the conversation), and this may seem like a problem: how can one and the same individual present two conflicting weak necessities as being both true? I would like to argue that it is possible for such conjunctions to be true, under certain circumstances, and thus this flexibility of the analysis is welcome. First, an example showing the problem raised by conjunctions. Assuming that it is not an option to rest and work out on the same day, (66) sounds like a contradiction.

(66)  [An athlete consults her coach about how to spend the last day before the competition. It is not possible to rest and work out on the same day.]

Coach: You ought to rest today and you ought to work out today.

The crucial point is that the sentence sounds like a contradiction specifically when coming from the coach. If we imagine the athlete consulting with a friend, her friend could utter the same sentence without any air of contradiction. Intuitively, there is an added pragmatic presupposition in (66) that the coach has an opinion about which priority prevails: the one which necessitates rest, or the one which necessitates a workout, while the friend is allowed to hypothetically reason through conflicting options, leaving the resolution of the conflict to the athlete herself.

2.4.1.2 Inference patterns

The domain restriction approach to weak necessity accounts directly for the inference patterns that relate weak and strong necessity modals. Strong necessity claims logically entail corresponding weak necessity claims with the same prejacent, since the domain of quantification of the former is a superset of that of the latter. As an implementation of

43 See Section 2.1 on additional scalar-like inferences.
this approach, the analysis proposed in Section 2.4.1 accounts for these patterns. I would like to comment on an alternative statement of the sensitivity to collective commitment that I have rejected since it does not deliver these predictions.

The idea would be, along the lines of (67), to encode a presupposition of collective commitment just in the lexical entries of strong necessity modals (not as a general requirement on ordering sources in C, as I have done), and to associate weak necessity modals with a lexical entry devoid of any presupposition. With a pragmatic principle such as Maximize Presupposition (Heim, 1991), the use of weak necessity modals would be predicted in cases that the necessity claim relies on priorities that are presupposed not to be collectively committed to.

(67) (Alternative rejected)

a. \([\text{have to}]^C\) is only defined if \(<g_1, \ldots, g_i>\) in C are collectively committed to. If defined, \([\text{have to}]^C = \lambda q \lambda w. \forall w': w' \in Fav^C(w).q(w')\).

b. \([\text{ought to}]^C = \lambda q \lambda w. \forall w': w' \in Fav^C(w).q(w')\).

This analysis is problematic because it does not allow strong modals to ignore ordering sources representing priorities that are not collectively committed to. If \(g_1(w)\) contains the collectively committed to priority of driving to Amherst, and \(g_2(w)\) contains the secondary priority of driving through Worcester, the weak claim You ought to take Route 9 is correctly predicted to be true by the analysis, but the strong claim You have to drive would result in presupposition failure (due to the offending \(g_2(w)\)).

Now, if in order to solve this problem we allow strong modals to “pick and choose” the ordering sources they are relativized to, we lose the ability to account for the entailment relation between the modals. Suppose \(g_1(w)\) and \(g_3(w)\) contain priorities that are collectively committed to, and \(g_2(w)\) does not, as before. A modal which is relativized just to \(g_1\) and \(g_3\) is not guaranteed to have a domain of quantification that includes the worlds that are most ideal according to all three of \(g_1\), \(g_2\), and \(g_3\).
It is possible to avoid such problems by splitting the ordering sources in C to primary ones and secondary ones (as in the proposal by von Fintel and Iatridou 2008), according to the criterion of collective commitment. However, the representations in (67) would no longer be appropriate in such a setup, since the quantification domains for strong and weak modals would have to be made sensitive to different parts of the bipartition. The difference between my proposal and the alternative just sketched then boils down to a question of whether or not secondary ordering sources should be represented in the context. I believe that disagreements about secondary priorities argue against representing them in the context as ordering sources, but I will not provide a detailed argument to support this position here.

2.4.2 Performativity and strong necessity

As a first result of my proposals, I would like to suggest a connection between modal strength and the phenomenon of performativity of strong necessity priority modals. This connection has also been noted independently, and worked out in more detail, in recent work by Silk (to appear).

Following up on an intuition in the descriptive literature on modality, Ninan (2005) argues that deontic must is different from deontic should in having imperative-like force. This imperative-like contribution wins must the status of a performative modal, since “in uttering a simple must sentence, the speaker is attempting to initiate action – she’s either trying to get someone to do something or committing herself to doing something” (ibid.)

Performativity is said to be responsible for two patterns that distinguish between must and should in their uses as deontic modals. First, an assertion of must/should q permits a denial of q if the modal is should, but not if it is must (68). Second, when must takes past/perfect complements it cannot be interpreted deontically, but should can (69).

(68) a. Sam should/ought to go to confession, but he’s not going to.

---

44 See Ninan (2005) and Portner (2009, §4.3.3) for a survey of the descriptive and theoretical literature on the topic.
b. #Sam must go to confession, but he’s not going to.

(69) a. Sam should/ought to have gone to confession. \textit{(Deontic interpretation possible)}
b. Sam must have gone to confession. \textit{(Epistemic interpretation mandatory)}

(Ninan, 2005, §1)

I will not go into the details of how the two patterns are explained by assuming that \textit{must} has an imperative-like meaning (the reader is referred to Ninan 2005 for discussion); I wish only to raise the question of whether this is indeed a separate meaning component of the strong necessity modal, as has been assumed in the previous literature.

Ninan (2005) explicitly addressed the question of whether performativity and modal strength (\textit{practicality} and \textit{uniqueness} in his terms) are related. Considering the domain restriction analysis of weak necessity that was around at the time (due to von Fintel and Iatridou 2005b), Ninan concluded that performativity and strong necessity are not related. The gist of the analysis should be familiar: \textit{must} was analyzed as being sensitive to one normative parameter, while \textit{ought} was analyzed as being sensitive to two normative parameters. Formally, it was proposed that only weak necessity modals have access to an ordering source.\textsuperscript{45,46}

I mention this difference between \textit{should} and \textit{must} only to set it aside. I don’t think it will solve our puzzle, because it only serves to make \textit{must} stronger than \textit{should} by allowing \textit{must} to ignore the ordering source. But it is

\textsuperscript{45} This predecessor of the analysis presented by von Fintel and Iatridou (2008) is discussed in Section 2.4.5.

\textsuperscript{46} Ninan (2005) presents the empirical contrast in (i) as a further argument that the puzzle in (68) is not related to the strength of the modals involved. The argument is that while the modals in (ia) and (ib) are equally strong, one can be followed by a denial of its prejacent, but the other cannot. (An analogous argument is made regarding the puzzle in (69).)

(i) a. Religious duty absolutely requires John to go to confession, but he’s not going to go.
b. #Sam must go to confession, but he’s not going to.

This contrast points to a difference between \textit{require} and \textit{must} in the type of modalities they allow, a difference that could be lexically or grammatically determined (e.g., given the different types of syntactic complements they take).
difficult to see how this could solve either of our two puzzles, since *must* and *should* have the same basic semantics, except that the domain of worlds that *must* quantifies over is a superset of the domain of worlds *should* quantifies over.

(Ninan, 2005, §3.1)

In light of the updated implementation of the domain restriction approach, I would like to suggest that we revisit this conclusion. The possibility to be explored is that the performative aspect of strong necessity modals can be derived from their semantics and the constraints on ordering sources, without introducing a separate imperative-like component to their meaning (as proposed by Ninan 2005).

Strong necessity modals are sensitive to conversational backgrounds which all carry a presupposition of collective commitment. When a speaker utters a strong necessity claim, despite being aware that her addressee is not committed to the priorities on which that claim is based, she conveys her expectation that the addressee will revise his priorities. Depending on the authority of the speaker and other considerations, this expectation may lead the addressee to endorse the necessary priority (and revise his other commitments as appropriate). The result is an air of forced update of commitments which is related to the forced update triggered by imperatives, as in the analysis of imperatives by Portner (2004, 2007). Assuming an analysis of imperatives along these lines, performative modals and imperatives have in common a potential to update the context with priorities that become collectively committed to.

Performativity in the domain of priority modals emerges as a phenomenon that is intimately related to strong modality, although the details of this connection still need to be elaborated and fleshed out in detail (see Silk to appear for a proposal). Although I do not undertake this task here, I expect that such an analysis will be free of some of the problems facing Ninan’s (2005) speech act-based approach to the performativity of *must*. Answering the question of how priority modals are interpreted when embedded is particularly

---

47 See ibid. §5.3 “Complications” and §6.1 “Embedding.”
urgent, given that *must* “loses its performativity” when embedded under attitude verbs and in antecedents of conditionals (Ninan, 2005). Obviously, this is an important question to which I owe an answer as well.

2.4.3 Secondary priorities versus conditional antecedents

The function of the not-collectively committed to proposition \( b \) in (60) is to restrict the quantification domain of a modal. This is exactly what an *if*-clause modifying the modal would do, according to the view pioneered by Lewis (1975) and Kratzer (1981, 1986). Recently, a conditional analysis of weak necessity has been proposed, in work by Alex Silk (to appear). In this section, I discuss differences in persistence that distinguish between secondary priorities of weak necessity modals and conditional antecedents, taking up a more detailed evaluation of the particular conditional analysis proposed by Silk (to appear) in the following section.

From the parallel in function between them, secondary priorities emerge as the silent analog of a conditional antecedents. However, there are two basic properties that distinguish conditional antecedents from the extra assumptions accompanying weak necessity modals. First and foremost, the latter are not reliably conveyed by linguistic content, whereas *if*-clauses are always included in the overt linguistic signal. Relatedly, \( b \) does not persist in the discourse unless it is introduced overtly (e.g., by a rationale clause), whereas the content of an *if*-clause normally has the option to persist.

Conditional antecedents do not only affect the context in which their consequent is interpreted, but they may also shape the context for the interpretation of subsequent sentences. Roberts (1989) discusses this property of *if*-clauses, exemplified in (70), in the broader context of the phenomenon of *modal subordination*. The discourse as a whole can be true, Roberts notes, even if in actuality the birds do not get hungry.

---

a. If Edna forgets to fill the birdfeeder, she will feel very bad.

b. The birds will get hungry.

(Roberts, 1989, 683)

Dynamic treatments of modal subordination allow the hypothetical fact raised by the antecedent to affect the interpretation of subsequent sentences in the discourse (Frank 1996, Stone 1999). The if-clause is assumed to establish a temporary context against which multiple sentences may be evaluated, and thus is said to “persist.”

The infelicitous exchange in (71) is designed to show that the additional assumption that accompanies a weak necessity claim – if implicit – does not persist in the discourse. Suppose that the first ought claim in this exchange rests on a preference speaker A has for seeing pretty scenery (b'). This preference is inaccessible to speaker B. It therefore cannot, and does not, circumscribe the domain for a modal claim in the sentence that follows. This is the case even if the subsequent claim itself is merely a weak necessity.

(71) A: How does one get to Amherst from here?

B: There are four routes; the ones you see on this map.

A: I ought to take Route 2 then. (unspoken b' = enjoy the scenery)

B: Yes, and you ought to leave early. (unspoken b = avoid the afternoon fog that obscures the view)

Note that the additional assumption does not persist even if its content is easily inferable in the context. Thus, in (72), even though B outlines all the possible sub-goals for A, and A can be taken to make his or her choice between these sub-goals, B’s second ought claim cannot be modally subordinated to the context created by the first. The final sentence in (72) is just as infelicitous as it is in the exchange (71).

A: How does one get to Amherst from here?

B: There are four routes; the ones you see on this map. Choosing between them depends entirely on whether you prefer to visit Worcester or whether you prefer scenery. Route 2 is the most scenic, but it doesn’t pass near Worcester. The other three routes do.

A: I ought to take Route 2 then. (unspoken b' = enjoy the scenery)

B: ?Yes, and you ought to leave early. (unspoken b = avoid the afternoon fog that obscures the view)

Of course, if the preference for pretty scenery is collectively committed to (so a strong modal like have to has access to it), it necessarily circumscribes the domain of possibilities for modals in the discourse. This is what we see in (73).

(73) A: What is the prettiest road to Amherst from here?

B: Route 2, I would say.

A: I have to take Route 2 then. (gC = {enjoy the scenery})

B: Yes, and you ought to leave early. (unspoken b = avoid the afternoon fog that obscures the view)

Conditional antecedents differ from implicit assumptions that accompany ought claims not only in terms of persistence, but also in terms of identifiability. If-clauses are clearly packaged in the linguistic signal, whereas the content of a b-proposition is harder to pin down. In the absence of an anaphoric relation to a salient proposition in the discourse, preferences the speaker has – about himself or others – seem to determine the content of b in unembedded weak necessity claims. Perhaps it is not surprising, therefore, that when the secondary priority is raised explicitly, e.g., in a rationale clause modifying the modal claim, it has the option to persist.

(74) A: How does one get to Amherst from here?

B: There are four routes; you can see pictures I took on each one in this album.
A: To enjoy scenery on the way, I ought to take Route 2 then. \( (b' = \textit{enjoy the scenery}) \)

B: Yes, and you ought to leave early. \( (\text{unspoken } b = \textit{avoid the afternoon fog that obscures the view}) \)

It is not clear, however, that what we see in (74) is persistence of a temporary assumption in the discourse. One thing that is clear is that this “persistence” is not automatic. When multiple secondary priorities are enumerated in a list, as in (75), the last one does not have a privileged status over the first few in delimiting the favored worlds for a following modal. A plausible explanation of the dynamic in (74) is therefore that persistence of \( b \) in B’s final utterance reflects a change in the contextual commitments (such that there is presupposition of collective commitment to the priority of scenery at that point, triggered by A’s preceding weak necessity claim). In (75), in contrast, speaker A is not presupposed to be committed to any one of the priorities she enumerated. The presupposed collective commitments at the end of the conversation are just as they are in the beginning: A’s goal is to reach Amherst, and there are multiple ways to get there.

(75) A: How does one get to Amherst from here?
   B: There are four routes; you can see pictures I took on each one in this album.
   A: To enjoy scenery on the way, I ought to take Route 2 then, and to visit Clark University in Worcester, I ought to take another route.
   B: Yes, and you ought to leave around noon. \( (\text{unspoken } b = \textit{avoid the morning rush hour traffic to Worcester}) \)
2.4.4 Conditional application of priorities: a brief comparison

In this section, I briefly comment on a recent analysis of weak necessity by Silk (to appear) which includes a conditional component and responds to my previous work on weak necessity.\(^{50}\)

Silk (to appear) advances an interesting proposal about the representation of ordering sources, which he uses to develop an analysis of weak necessity within the domain restriction approach. Silk’s idea is that a priority (e.g., a desire, a goal, a norm) is usually not categorical, but rather comes with a specification of its applicability conditions: the conditions under which, and only under which, it is to be pursued (ibid. §3). These applicability conditions are used to determine the restricted quantification domains of weak necessity modals, as well as the priorities that necessity modals in general – weak and strong – are sensitive to. For example, if the value of you going to Amherst is established in the conversation (for Silk, this means representation in the common ground), the ordering source representing this priority will contain a biconditional in which the priority of you going to Amherst is paired with the applicability condition that “there are no more important countervailing values” (ibid. (10)):

\[(76) \quad g(w) = \{\text{You go to Amherst} \leftrightarrow \text{there are no more important countervailing values}\}\]

(Silk, to appear; adapted to the driving to Amherst scenario)

There is just one ordering source which is used to make both strong and weak necessity claims, and priorities end up influencing the ordering of accessible worlds differently, depending on whether or not their applicability conditions are satisfied. What makes a weak necessity modal weaker is that it uses some (potentially all) of the applicability conditions in the ordering source to ignore some of the accessible worlds determined by the modal base: the worlds in which those specific applicability conditions are not satisfied. The lex-

---

\(^{50}\) The manuscript cited here as (Silk, to appear) has not yet been published, and I thus refer to example and section numbers instead of page numbers in this discussion. Silk (2011) and Silk (2012) also provides relevant discussion of the semantics of weak necessity modals.
ical entries for weak and strong necessity modals proposed by Silk (to appear) are given in (77) (some of the notation has been adapted to fit the conventions I have been using).

\[(77)\]
\[\begin{align*}
\text{a. } & \llbracket \text{must } \phi \rrbracket^{c,w} = 1 \text{ iff } \forall w' \in \text{max}_{g(w)}(\bigcap f(w)).[\llbracket \phi \rrbracket]^{c,w'} = 1. \\
\text{b. } & \llbracket \text{should } \phi \rrbracket^{c,w} = 1 \text{ iff } \forall w' \in \text{max}_{g(w)}(\bigcap (f(w) \cup C_{g(w)})).[\llbracket \phi \rrbracket]^{c,w'} = 1,
\end{align*}\]

where \(C_{g(w)}\) is a (properly improper) subset of the set of applicability conditions for each of the premises in \(g(w)\).

(Silk, to appear, §5.1, notation adapted)

In the two examples of ordering sources Silk provides, the applicability conditions for priorities are superlative and related to the priorities they are paired with. That is, we have conditions like “there are no more important countervailing values,” or “you don’t prefer to do anything else more” (ibid. ex. (6)), which means that the contents of such ordering source are propositions of the form “\(p \iff \text{there are no more important countervailing values than } p\).” This is crucial, as I will now argue. Nothing in Silk’s account seems to prevent other applicability conditions from being paired with priorities, however. Consider (78), for example.

\[(78)\]
\(g'(w) = \{\text{You go to Amherst } \iff \text{there is one more important countervailing value}\}\)

This ordering source seems to represent a somewhat strange priority or goal, e.g., a priority that you go to Amherst if and only if Amherst is the least preferred of two possible destinations. A biconditional like this in the ordering source will generate the prediction that the strong necessity claim \(\text{You have to drive}\) is true, assuming that all the circumstentially accessible worlds in which you go to Amherst are worlds in which you drive. This is an incorrect prediction if the circumstentially accessible worlds are diverse with respect to the destination you go to, and the more preferred destination is one you reach by plane, for instance. In a situation in which it is the case that there is one other destination that you prefer more than Amherst, this aspect of the situation would be treated as one of the relevant circumstances in the modal base (ibid. discussion of ex. (6)). Thus, the applicability
condition in (78) would be satisfied throughout the contextually determined circumstantial modal base. The modal would quantify (universally) over all the accessible worlds in which the priority of going to Amherst is satisfied, and the strong necessity claim that you have to drive is predicted incorrectly to be true.

I take this to show that it is not enough to pair priorities with any applicability condition to explain the semantics of strong necessity. Those priorities that are relevant are those for which, I have suggested, there exists a presupposition of collective commitment in the conversation. Collective commitment depends on presupposed commitment of every conversational participant, and commitment of each participant implies their endorsement of the priority over its alternatives. If you haven’t decided whether going to Amherst is something you prefer over going to Hartford, for example, then you are not committed to any one of these priorities in the sense that is relevant for licensing a strong necessity claim. We might say then that a primary priority $p$ is a priority for which “there are no more important countervailing values than $p$” in the conversational context. I have argued that this characterization of primary priorities is an important ingredient that distinguishes primary from secondary priorities, and it seems crucial for the analysis proposed by Silk (to appear) based on ordering sources with applicability conditions as well.

A potential problem for the implementation of the conditional analysis proposed by Silk (to appear) arises in scenarios with cascades of priorities. These kinds of scenarios, which provided important evidence in arguing for a distinction between primary and secondary priorities, are not analyzed in detail in Silk’s paper. The problem for the conditional analysis seems to be that applicability conditions that are mutually exclusive result in one priority at most being operative as a positive priority, when in fact multiple ones should be. Let’s step through an example. Suppose we are in the context of driving to Amherst, where it is a collective commitment that you go to Amherst. Suppose, moreover, that you mention that a scenic drive and a quick trip are also important, but I assume that you have not yet committed to either one. According to my understanding of Silk’s discussion, the fact that
you expressed a preference for going to Amherst, for pretty scenery, and for a quick trip means that these priorities are all represented in the ordering source, with the applicability conditions shown below.

\[(79)\quad g(w) = \{ \text{You go to Amherst} \iff \text{you don't prefer to do anything else more}, \]

\[\quad \text{You take a scenic road} \iff \text{you don't prefer to do anything else more}, \]

\[\quad \text{You take a quick road} \iff \text{you don't prefer to do anything else more} \} \]

Now, the weak necessity claim *You ought to take Route 2* is true in this scenario, if pretty scenery, for example, functions as a secondary priority. This is not what we get according to the analysis in (77b). The sentence is true, according to (77b), if and only if you take Route 2 in all the \(g(w)\)-best worlds in the subset of the modal base in which it is true that you do not prefer anything else more than a scenic road. There are circumstantially accessible worlds in which you take a scenic road to a different destination, however, not to Amherst. At least in some of those worlds, you prefer a scenic road more than anything else, but, importantly, you do not take Route 2. Indeed, in the restricted set of circumstantially accessible worlds that *ought to* would be quantifying over, worlds in which you do not go to Amherst are better according to this \(g(w)\) than worlds in which you do go to Amherst. This seems problematic.

I believe these considerations merit closer examination of the conditional approach proposed by Silk (to appear) and detailed comparison with the commitment based approach I have argued for here (and which is free from certain problems of its predecessor, Rubinstein 2011). I must leave these tasks for another occasion, however.

### 2.4.5 A different perspective on the modal landscape: negotiable norms, non-negotiable norms, and facts

My proposal for the semantics of weak necessity modals makes use of “modal backgrounds” of three kinds: modal bases, primary ordering sources, and secondary priorities. In this section, I would like to suggest a *triply relative* perspective on modal parameters that
distinguishes these three kinds of backgrounds. From this perspective, modal bases and primary ordering sources represent non-negotiable (consistent and potentially inconsistent) information, and together they determine a simple, unstructured, set of live possibilities that all modals are sensitive to. Secondary priorities represent negotiable information. Thus, they do not eliminate possibilities from consideration, but only add structure to the set of live possibilities by ordering it, allowing for comparison of alternatives.

How are secondary priorities related to the familiar kinds of conversational backgrounds, namely modal bases and ordering sources? Ignoring the simplification of using a single proposition instead of a set of propositions to represent relevant secondary priorities in a context, secondary priorities are similar to ordering sources in that they “may themselves be inconsistent or else be inconsistent with the established facts” (Kratzer, 1981, 307). In contrast, sets of propositions determined by a modal base are always consistent, at least in the case of circumstantial or factual modal bases (ibid.)

There is another sense, however, in which secondary priorities are distinct from primary ordering sources. According to my proposal, secondary priorities represent negotiable priorities, priorities that are still up for discussion. In contrast, the priorities in a primary ordering source are non-negotiable. They serve to eliminate possibilities from consideration in the same way, intuitively, that the relevant facts delimit what is possible and not possible in a certain context. Sloman (1970) conveys this intuition clearly when he describes the class of possibilities to which ought’s prejacent is compared as reflecting prior decisions, principles, and commitments:

Although the above is not an exhaustive or systematic classification of types of comparison class, it does illustrate the variety of ways in which the possibilities under consideration may be restricted. Of course, some race of supermen might always consider all possibilities when trying to decide what ought to be the case, or what ought to be done. But our considerations are limited, not only by the bounds of our own imagination and knowledge, but also by the structure of the practical situations in which we are faced with such questions: our prior decisions, principles, commitments, etc., impose restrictions or constraints on the possibilities we can consistently take into account. (Sloman, 1970, 388)
From this second perspective, modal bases and primary ordering sources have something in common. Jointly, these conversational backgrounds determine the set of live possibilities that weak necessity claims are focused on. Secondary priorities carry out the true purpose of an ordering source in a Kratzerian semantics: that of comparing possibilities and ordering them according to how well they approach an ideal.

An early proposal by von Fintel and Iatridou (2005b) reflects this intuition about the way that quantification domains for necessity modals are constructed. According to this proposal, given in (80), weak necessity modals are relativized to a modal base, a designated priority, and an ordering source, while strong necessity modals are relativized only to a modal base and a designated priority.

(80) a. *to p, ought to q* is true in a world *w* with respect to a modal base *f* and ordering source *g* iff all the *g(w)*-best worlds in \( \bigcap f(w) \) where *p* is true are *q*-worlds.

b. *to p, must q* is true in a world *w* with respect to a modal base *f* iff all the worlds in \( \bigcap f(w) \) where *p* is true are *q*-worlds.

(von Fintel and Iatridou, 2005b, ex. (24), notation slightly adapted)

Despite the intuitive appeal of the proposal, it does not generalize automatically to other priority type interpretations. To accommodate deontic necessity claims, in particular, the designated priority would need to be replaced by an ordering source, and one that might contain an inconsistent set of propositions.

The technical challenge of dealing with inconsistencies has been overcome in a different context, in work by Anette Frank. Frank (1996) defends the idea that facts and norms may jointly determine unstructured modal domains, and develops the necessary tools for joining consistent sets of propositions (modal bases) with potentially inconsistent ones (ordering sources).\(^{51}\) While Frank proposes to dispense with ordering sources altogether, I suggest that we conceptualize their role a bit differently – and keep them. Specifically, I suggest

---

\(^{51}\) Inspiration for some of these tools derives from Gazdar (1979) via Kasper (1992).
that secondary priorities be used to order the unstructured domain of accessible worlds, or live possibilities, that the facts and the primary priorities jointly determine. A formal sketch of this suggestion is found in Appendix B.

In summary, the true power of ordering sources is in comparing possibilities, and comparisons are made when negotiable priorities are at stake. In this study, the connection between ordering and lack of collective commitment, and hence negotiability, emerged through the analysis of weak necessity. I suspect that it will play a role in the analysis of other comparative modal expressions as well.

2.4.6 Conclusions and future directions

With this chapter, I hope to have shown that understanding certain aspects of modal strength requires sensitivity to presuppositions about collective commitment in a conversation. I motivated a split between two kinds of priorities: ones that are presupposed to be collectively committed to, and ones that are presupposed not to be collectively committed to. I used independent probes of commitment status to show that weak necessity modals are sensitive to priorities of both kinds, while strong necessity modals are sensitive only to priorities of the first kind.

What makes weak necessity modals such useful modal words is that they allow interlocutors to entertain priorities that may be outright contested, controversial, or simply still up for discussion. Different participants might have different opinions about the relative goodness of different priorities and, in normal conversation, disagreements about what should or shouldn’t be done reflect this variety of opinions.

Since it provides a criterion for the split between priorities, I argued that sensitivity to collective commitment solves a conceptual challenge facing the domain restriction approach to weak necessity. I proposed that this criterion is superior to criteria that have previously been entertained in the literature (a split on conceptual grounds, or a split based on the content of rationale clauses). Having argued that the domain restriction approach
developed by von Fintel and Iatridou (2008) is superior to a number of probabilistic alternatives, I formalized sensitivity to collective commitments within this approach. The crucial priorities that fix domains of quantification for weak necessity modals were defined in relation to the Question Under Discussion in the conversation as priorities for which lack of collective commitment is presupposed.  

A first application of my proposals concerns the performativity of strong necessity modals, a phenomenon which may be analyzed as reflecting the presupposition of collective commitment that is required for the felicitous utterance of these modals.

The questions that remain most intriguing for me at this point concern the extent of commitment-based weakness that one might expect to find in modal systems of natural languages. As von Fintel and Iatridou (2008) have shown, languages differ in whether or not, and how, they express weak necessity. What is the explanation for these differences? What are the relevant properties of languages that develop weak necessity modals from strong ones? A further obvious question to ask, given the influence of presupposed commitments on the expression of modal necessity, is whether similar influences can be detected with possibility modals as well. The null hypothesis would be that they do, and that we might be able to find an existential counterpart to ‘ought’ in languages that grammatically distinguish between weak and strong necessity.

Higher items on my to-do list, however, are two open issues regarding the composition of priority modal meaning. The first concerns the analysis of weak necessity in embed-
ded contexts. The second concerns the compositional contribution of rationale clauses in teleological discourse. While there have been proposals to treat rationale clauses and *if*-clauses on a par (von Fintel and Iatridou 2005b, von Stechow et al. 2006, Finlay 2010), reconciling the differences in persistence between the two is reason to think afresh about the relationship between them.
CHAPTER 3
ROOT MODALITIES AND ATTITUDE PREDICATES

The mountain, oh mountain that is always green
I still dare dream and dwell
To breathe in your winds as in the begin’
To lie in your shade, Carmel.

Yoram Taharlev, ‘The Mountain that is Always Green’, (1972)

Attitude verbs like want have a preference-based meaning, and like weak necessity modals, they have been given modal, comparative, and probabilistic analyses. In this chapter, I critically review recent developments in the semantics of preference-based attitudes, focusing on comparative and modal accounts stemming from Heim (1992). I show that the introduction of focus alternatives to the semantics of desire predicates and modal predicates like necessary is not necessitated by the evidence amassed so far (contra Villalta 2006, 2008). I propose a modal analysis of want and necessary and compare them to the strong and weak modals analyzed in Chapter 2. Necessary is argued to be a strong necessity modal that is nevertheless always priority-oriented.

3.1 Attitude verbs of desire

3.1.1 A conditional comparative analysis: Heim (1992)

Leading proposals in the semantic literature on want and other desire predicates are designed with the idea that such predicates involve comparison of alternatives. This idea is summarized succinctly in the following quote from Stalnaker.

[... ] wanting something is preferring it to certain relevant alternatives, the relevant alternatives being those possibilities that the agent believes will be realized if he does not get what he wants.

(Stalnaker, 1984, 89)
Heim (1992) uses this characterization as motivation for proposing a conditional comparative semantics for desire verbs.

An important feature of this analysis is that it sees a hidden conditional in every desire report. A little more explicitly, the leading intuition is that *John wants you to leave* means that John thinks that if you leave he will be in a more desirable world than if you don’t leave.

(Heim, 1992, 193)

The conditional paraphrase is then used to implement a semantics for desire verbs using a version of the semantics proposed for conditionals (indicative and counterfactual) by Stalnaker (1968) and Lewis (1973). In this tradition, a conditional sentence of the form *If* \( \phi \), then \( \psi \) *is true in a world* \( w \) *if and only if the consequent* \( \psi \) *is true, not necessarily in* \( w \), *but in the worlds most similar to* \( w \) *in which the antecedent* \( \phi \) *is true. For any proposition* \( p \) *and world* \( w \), *the* \( p \)-*worlds most similar to* \( w \) *are abbreviated as* \( \text{Sim}_w(p) \).

(81) \[ \text{Sim}_w(p) = \{w' \in W : w' \in p \text{ and } w' \text{ resembles } w \text{ no less than any other world in } p \} \]

(Heim, 1992, 195, 197)

According to the conditional paraphrase, *a wants p* means, roughly, that *a* considers the \( p \)-worlds she can think of to be better than the \( \neg p \)-worlds she can think of. As von Fintel (1999, 119) points out, there are two orderings of worlds at play in Heim’s analysis. One is the above ordering based on similarity. The other is an ordering (of worlds and propositions) based on desires or preferences of an individual:

(82) a. For any \( w, w', w'' \in W \),

\( w' <_{a,w} w'' \) iff \( w' \) is more desirable to \( a \) in \( w \) than \( w'' \).

b. For any \( w \in W, X \subseteq W, Y \subseteq W, \)

\( X <_{a,w} Y \) iff \( w' <_{a,w} w'' \) for all \( w' \in X, w'' \in Y \).

(Heim, 1992, 197)

As hinted by the informal paraphrases, desires are evaluated in light of one’s beliefs, not in light of what is actually the case. They are thus easily rooted in beliefs that do not
match reality, as shown by (83). Patrick may have a desire to sell a cello which is not in fact in his possession, but which he thinks is.

(83) Patrick is under the misconception that he owns a cello, and he wants to sell his cello.

(Heim, 1992, 183(2))

Additional motivation for the connection with belief comes from examples like (84). Let’s assume that worlds in which I don’t teach at all next semester are more desirable to me than ones in which I get a good teaching assignment. But if teaching next semester is a given, I believe it is not possible for me not to teach. That is, in all the worlds that represent my beliefs in the evaluation world (these worlds are called my doxastic alternatives in the evaluation world, or the worlds doxastically accessible to me in the evaluation world, or simply my belief worlds), I teach next semester. Since (84) is typically judged true in the situation described, it seems that the desirability comparison cannot reach out to worlds that are not doxastically accessible to the subject in a given situation.

(84) I want to teach Tuesdays and Thursdays next semester.

(Heim, 1992, 195(35))

Assuming that comparisons of desirability are restricted by the subject’s doxastic alternatives, Heim (1992) defines the semantics of want as in (85). $Dox_a(w)$ here is a function that retrieves the set of possible worlds that match the beliefs of an agent $a$ in a world $w$. It maps an individual to their doxastic alternatives in that world.\(^1\)

---

\(^1\) This use of the term doxastic alternatives to refer to doxastically accessible worlds is standard practice following Hintikka (1962). Note that it contrasts with the use of the term by Lewis (1986) to refer to individuals: “[…] a class of possible individuals – call them the believer’s doxastic alternatives – who might, for all he believes, be himself. […] These individuals are the believer’s doxastic possibilities. But they are not different possible ways for the world to be; rather, they are different possible ways for the individual to be, and many of them may coexist within a single world.” For Lewis, “The doxastic alternatives determine the doxastically accessible worlds, though not conversely: a world is accessible iff at least one of the alternatives inhabits it.” (ibid., pp. 28-29).
Heim’s motivation for crafting an analysis of desire predicates is to account for Karttunen’s (1974) observations about presupposition projection from complements of attitude verbs. Although the original observations have not gone uncontested, Heim’s (1992) analysis of want has become the benchmark for much subsequent work on the semantics of attitude predicates (e.g., von Fintel 1999, Geurts 1998). A particularly interesting adaptation of the theory has been proposed recently in work by Elisabeth Villalta.

### 3.1.2 Multiple alternatives and less reliance on belief: Villalta (2006, 2008)

Villalta (2006, 2008) proposes a semantics for a family of predicates, among them predicates of desire, which makes comparison of alternatives the driving force of the analysis. Her proposal brings to light certain subtleties regarding the doxastic component of desire predicates. It handles naturally cases in which there is more than a binary choice between alternatives (although we will see below that the additional machinery it uses for this purpose is unnecessary). Finally, it applies not only to want and a handful of other desire verbs, but to a number of classes of predicates unified in Spanish, the language analyzed, by being predicates that select the subjunctive mood in their embedded clauses. Among these are desire predicates and (adjectival and verbal) modal predicates.3

The first of Villalta’s contributions concerns the doxastic component of desire predicates. In the version of Heim’s conditional semantics in (85), she finds that “the desires of the subject are too tightly connected to his/her beliefs” (Villalta, 2008, 478). Since the desirability of $p$-worlds (over $\neg p$-worlds) is compared only in worlds that match the beliefs

\[
\forall w' : w' \in Dox_a(w). Sim_{w'}(Dox_a(w) \cap p) <_{a,w} Sim_{w'}(Dox_a(w) \cap \neg p).
\]

(A static rendering of the dynamic proposal by Heim 1992, 197(39))

\[\text{want}(a)(w) = 1 \text{ iff }\]

\[\forall w' : w' \in Dox_a(w). Sim_{w'}(Dox_a(w) \cap p) <_{a,w} Sim_{w'}(Dox_a(w) \cap \neg p).\]
of the subject of the sentence, the analysis in (85) predicts that if someone believes that two propositions are equivalent, there is no way he or she can desire one to a different degree than the other. This is incorrect, since in fact, “someone may believe that two propositions $p$ and $q$ are true in the exact same set of worlds, and at the same time want $p$ without wanting $q$” (Villalta, 2008, 479). Let’s refer to this as the Doxastic Problem for Heim’s analysis.

(86)  

a. I want to teach Tuesdays and Thursdays next semester.

b. I believe that I will teach Tuesdays and Thursdays next semester if and only if I work hard now.

c. *Invalid inference:* I want to work hard now.

(Villalta, 2008, 478(33))

It should be noted that not all versions of Heim’s conditional semantics for want face this problem. In particular, the truth-conditional version of her analysis, given in (87), does not make this incorrect prediction. This is the version Heim (1992) considers before transposing her proposal into the framework of context change semantics.

(87)  

$$ [[\text{want}]][p](a)(w) = 1 \text{ iff }$$

$$ \forall w' \in \text{Dox}_a(w) : \text{Sim}_{w'}(p) <_{a,w} \text{Sim}_{w'}(\neg p).$$

(Following Heim 1992, 193(31))

In (87), the conditional component (which is based on the similarity ordering of possible worlds) is not limited to worlds that are included in the subject’s doxastic alternatives. Therefore, worlds that are not doxastically accessible to the subject can also enter the desirability comparison. Their existence is enough to ensure that invalid inferences like (86) do not go through.

However, when Villalta claims that the invalid inference is in fact derived, she remains true to the spirit of Heim’s analysis. Heim (1992) states quite explicitly that worlds outside
the subject’s doxastic alternatives should not enter the desirability comparison (see below, emphasis in the original). Formally, while the truth-conditional version of her analysis in (87) does not deliver this result, the version inspired by dynamic semantics, in (85), does.

[(85)] implies that all the desirability comparisons that enter into determining the truth of a want-sentence are entirely among the subject’s belief worlds. If a belief world w’ has φ true in it, it must be more desirable than otherwise similar belief worlds where φ is false, and if a belief world has φ false in it, then it must be less desirable than otherwise similar belief worlds where φ is true. The desirability of non-belief-worlds never has any bearing on the truth of a want-report. (Heim, 1992, 197)

Villalta’s argument regarding (86) is therefore valid. If doxastic alternatives constrain the possibilities which are compared for desirability, then:

[...] from the truth of [(86a)] it follows that in $\text{Dox}_I(w)$ all the worlds in which I teach Tuesdays and Thursdays next semester are more desirable than maximally similar worlds in which I don’t teach on Tuesdays and Thursdays next semester. From the truth of [(86b)] it follows that in $\text{Dox}_I(w)$ the worlds in which I teach Tuesdays and Thursdays next semester correspond exactly to those worlds in which I work hard this semester. Hence, from [(86a)] and [(86b)] it follows that in $\text{Dox}_I(w)$ all worlds in which I work hard this semester are more desirable than maximally similar worlds in which I don’t work hard this semester. (Villalta, 2008, 478-479)

Villalta’s solution to the Doxastic Problem is to relax the connection between desire and belief in the semantics of want. The subject’s beliefs are relegated to a definedness condition of the predicate and no longer restrict the desirability comparison. According to the proposal in (88), wanting $p$ requires $p$ and its alternatives to be represented in the
subject’s doxastic alternatives, but the desirability comparison is not restricted to worlds in this set.\textsuperscript{4,5}

(88) \[[\text{want}_C]^{g}(p)(a)(w)\] is defined iff \(\forall q \in g(C) : \text{Dox}_a(w) \cap q \neq \emptyset\).

If defined, \([[\text{want}_C]^{g}(p)(a)(w) = 1 \text{ iff } \forall q : q \neq p \& q \in g(C) : p <_{\text{DES},a,w} q.\]

(Villalta, 2008, 480(37), notation adapted)

A virtue of this analysis is that it avoids the universal quantification over belief worlds in Heim’s proposal. Concerns about this aspect of the original comparative desire semantics have been voiced in the literature. Levinson (2003), for example, argues that according to Heim (1992), “a person wants something only if he/she believes that it will necessarily improve the situation in any possible case. […] this excludes many actual cases of wanting” (p. 230).

Notice that the embedded proposition \(p\) is compared here not to its negation \(\neg p\), as in (85) or (87), but potentially to many propositions in a set of alternatives \(g(C)\). This claim about the comparisons carried out by the attitude verb is a central claim put forth by Villalta (2006, 2008). \(C\) in (88) provides access to the relevant set: it is a variable that is anaphoric to a contextually determined set of propositions (and receives its value from the variable assignment \(g\); cf. Rooth 1985, 1992, von Fintel 1994, Westerståhl 1985). I come back to this innovation shortly.

\textsuperscript{4} Villalta (2008) adopts Kratzer’s (1991) definition of \textit{better possibility} as the basis for comparing the desirability of sets of worlds (as opposed to individual worlds). The following thus replaces Heim’s definition of \(<_{a,w}\) in (82b) above.

(i) a. For any \(w,w',w'' \in W, w' <_{a,w} w''\) iff \(w'\) is more desirable to \(a\) in \(w\) than \(w''\).

b. For any \(p \subseteq W, q \subseteq W, p <_{\text{DES},a,w} q\) iff \(\forall w'' \in q \exists w' \in p\) such that \(w' <_{a,w} w''\), and it is not the case that \(\forall w' \in p \exists w'' \in q\) such that \(w'' <_{a,w} w'\).

(Villalta, 2008, 479(35), notation adapted)

\textsuperscript{5} In what follows, I limit myself to discussion of the pre-final version of Villalta’s (2008) proposal, ignoring the decompositional degree-based refinement she introduces (see Villalta 2006, Villalta 2008, \S\S8.3).
The invalid inference in (86) is no longer derived in the new setup. There are two desire statements being considered: one concerns my wanting to teach Tuesdays and Thursdays next semester, the other concerns my wanting to work hard now. The Tuesday-Thursday schedule is compared to a Monday-Wednesday-Friday schedule, and preferred over it. The contextually determined alternatives to which the second statement is compared are different, and include the alternative that I don’t work at all and the alternative that I work only a little (Villalta, 2008, 480). All the alternatives mentioned are represented in my belief worlds, so the doxastic definedness condition is satisfied. However, since the desirability comparison is not restricted to worlds the subject believes to be possible, the comparisons may yield different results in (86a) and (86c). If it is the case that the most desirable worlds are Tuesday-Thursday-worlds in which I do not work hard, (86c) would be false.

3.1.3 Refuting the argument for multiple alternatives

The semantics proposed by Villalta (2008) for desire predicates in (88) includes a treatment of alternatives which marks a departure from Heim (1992) in several respects. For Heim (1992), there are exactly two alternatives that enter the comparison, $p$ and $\neg p$, where $p$ is the proposition complementing the desire predicate. Villalta (2008) argues that it is necessary to compare $p$ to a set of contextually determined alternatives, not just to $\neg p$. Furthermore, she argues that this set is to be derived based on $p$’s focus structure. My aim in this section is to discuss the first of these two innovations. I point out that Heim’s analysis, without any change, can deal with a purported counterexample advanced by Villalta (2008). Thus, moving from a singleton set of alternatives to a potentially larger set is not necessitated by the evidence amassed so far. I will leave the details of the focus-based analysis for another occasion.6

6 Thus, I will ignore for the sake of this discussion what happens when $p$ does not contain any focus-marked constituent in it, as well as some inaccurate predictions of the theory regarding association with focus (see Villalta 2008, footnote 13, for one such challenge).
Villalta’s argument for generalizing the definition of alternatives in the semantics of desire predicates is based on examples in which there are more than two contextually available alternatives \((p)\) and at least two more) that seem to be compared. The picnic potluck scenario is such an example.

Sofía has promised to bring a dessert to the picnic. Victoria believes that there are three possibilities for what she may actually do. She could prepare a chocolate cake, even though Victoria considers that extremely unlikely because it represents far too much work. She might bring an apple pie, which Victoria considers very likely since she can just buy it at the bakery nearby. Or Sofía might bring ice-cream, which seems most likely to Victoria, since she usually has some in her freezer. Victoria prefers the chocolate cake over the apple pie and the apple pie over the ice-cream.

\[(89) \text{False} \quad \text{Victoria wants Sofía to bring an apple pie.}\]

In this scenario, Victoria’s preferences run counter to what she believes is likely to happen. She would prefer Sofía brought a chocolate cake, but this is also the most unlikely thing that could happen, in her mind. Rather, she thinks it is most likely that Sofía will bring ice cream, the least desirable option of the three as far as she is concerned. It is clear that the sentence in (89) is false in this scenario.\(^7\)

Contrary to intuitions, Heim’s conditional semantics for the verb \textit{wish} predicts that this sentence should come out as true in this scenario. Under Heim’s proposal, the sentence [(89)] is true iff for all worlds \(w’\) that are compatible with Victoria’s beliefs, the worlds in which Sofía brings the apple pie are more desirable to Victoria than all the minimally different worlds in which she doesn’t. This is true, since, in the given scenario, the closest worlds in which Sofía doesn’t bring an apple pie are worlds in which she brings ice-cream. This set doesn’t contain worlds in which she brings chocolate cake (these are not minimally different for Victoria since she considers them extremely unlikely).

\[(Villalta, 2008, 477)\]

\(^7\) Villalta (2008, 477(30)) uses a version of this sentence with \textit{wish} instead of \textit{want}, but it is clear that she is making an argument about \textit{want}.
This conclusion is not warranted. The problem is that the argument ignores worlds in which Sofía does bring a chocolate cake, “against all odds.” Such worlds exist in the set of Victoria’s doxastic alternatives, $Dox_V(w)$, because she considers the possibility that Sofía bring a chocolate cake unlikely, but not impossible. Crucially, the desirability of these worlds according to Victoria is higher, not lower, than any minimally different world in which Sofía brings apple pie. The analysis of (89) based on comparison of just the prejacent and its negation, in (90), captures this fact. For any world $w''$ in $Dox_V(w)$ such that $w'' \in \llbracket Sofía brings chocolate cake \rrbracket$, it holds that $Sim_{w''}(\neg \llbracket Sofía brings apple pie \rrbracket)$ simply equals $\{w''\}$ (a world is always most similar to itself). But such a world $w''$ is ranked higher by Victoria’s desirability ordering, $<_{V,w}$, than any world minimally different from $w''$ in which $\llbracket Sofía brings apple pie \rrbracket$ holds. Therefore, (89) is predicted correctly to be false in the given scenario.

(90) $\llbracket want \rrbracket(W)(V)(w) = 1$ iff $\forall w' \in Dox_V(w) : Sim_{w'}(Dox_V(w) \cap \llbracket Sofía brings apple pie \rrbracket) <_{V,w} Sim_{w'}(Dox_V(w) \cap \neg \llbracket Sofía brings apple pie \rrbracket)$.

In sum, the picnic scenario does not provide an argument for replacing $\neg p$ in the semantics of want with a set of alternatives, even though the proposition $\neg p$ represents a number of finer grained alternatives to $p$.

Villalta’s approach nevertheless provides a solution to the Doxastic Problem she identified in Heim’s theory.

Can we apply the solution proposed by Villalta (2006, 2008), namely relaxing the dependency of want on belief, in order to come up with a comparative proposal that only has $p$ and $\neg p$ as the alternatives compared? A first stab at such a lexical entry is (91).

---

8 Additional empirical arguments Villalta (2008) uses to support her proposal do not seem to depend on the cardinality of the set of alternatives (e.g., the gradability of some of these predicates and their ability to appear in certain comparative constructions).
(91) (First try)

\[ [[\text{want}]](p)(a)(w) \text{ is defined iff } \text{Dox}_a(w) \cap p \neq \emptyset \text{ and } \text{Dox}_a(w) \cap \neg p \neq \emptyset. \]

If defined, \[ [[\text{want}]](p)(a)(w) = 1 \text{ iff } p <_{\text{DES}_a,w} \neg p. \]

This implementation is reminiscent of Heim’s truth-conditional semantics for \text{want} in (87), minus the similarity constraint built into that proposal, and for this reason it fails to account for the truth conditions of desire statements that seem to be circumscribed by what is believed to be possible. In particular, the naive semantics for \text{want} in (91) incorrectly predicts \text{I want to teach Tuesdays and Thursdays next semester} to be false in a situation in which I would really prefer not to teach at all. Worlds in which I don’t teach are irrelevant in a context in which my teaching is a given, but we lose the ability to keep them out of the desirability comparison by comparing \(p\) to all of \(\neg p\).

We need a way of circumscribing the possibilities \text{want} has access to so that irrelevant possibilities do not impact the desirability comparison. A straightforward way to do this is to use a contextually determined modal base to allow certain background assumptions to choose subsets of \(p\) and \(\neg p\) for the desirability comparison. This is the function \(f\) in (92).

(92) (Second try)

\[ [[\text{want}]](f)(p)(a)(w) \text{ is defined iff } \bigcap f(w) \cap \text{Dox}_a(w) \cap p \neq \emptyset \text{ and } \bigcap f(w) \cap \text{Dox}_a(w) \cap \neg p \neq \emptyset. \]

\[ \text{If defined, } [[\text{want}]](f)(p)(a)(w) = 1 \text{ iff } \bigcap f(w) \cap p <_{\text{DES}_a,w} \bigcap f(w) \cap \neg p. \]

In order to avoid the Doxastic Problem, the modal base should not be hard-wired to equal the subject’s doxastic alternatives (or the superset of the doxastic alternatives derived by ignoring beliefs the subject has about their own future actions, which is what Heim ends up proposing).\(^9\) The modal base \(f\) can be circumstantial, as is arguably the case in

---

\(^9\) “[…] we don’t take into account all [the subject’s] beliefs, but just those that he has about matters unaffected by his own future actions” (Heim, 1992, 199); see also von Fintel (1999). In the teaching scenario, my teaching next semester would be a belief I have about my future actions; nevertheless, it is not ignored for the purpose of the desire claim \text{I want to teach a Tuesday-Thursday schedule}.\]
the teaching scenario. The intuition that no-teaching worlds do not annul my desire to get a particular teaching schedule is modeled by allowing \( f \) to be a circumstantial modal base that contains the proposition \([I \text{ teach next semester}].\) *I want to teach Tuesdays and Thursdays next semester* is predicted to be true, and *I want to work hard now* is predicted to be false given my preferences in the scenario (see Figure 3.1). The modal base \( f \) can also be a doxastic modal base, if the relevance of someone’s belief state is made explicit (e.g., *Patrick is under the misconception that . . . , and he wants . . . ,* Heim’s example (83) above).\(^{10}\)

![Figure 3.1: Accessible worlds in desire statements. Doxastic alternatives may be a subset of the worlds determined by the modal base. In the teaching scenario of (86), the modal base worlds are circumstantially accessible worlds in which I teach next semester (left side of the figure, in blue). Only a proper subset of these are the belief worlds of the subject. The larger set is relevant for evaluating *I want to (not) work hard* (and perhaps also *I want to teach Tuesdays and Thursdays next semester*, although focusing just on the doxastic alternatives would also predict the truth of the sentence since the MWF schedule is less desirable than the TuTh schedule in general).](image)

\(^{10}\) Note that if the facts are that Patrick does not own a cello, the unmodified desire statement *Patrick wants to sell his cello* is odd.
Finally, Heim (1992) notes that reports of unrealistic desires with *want*, e.g., (93), are problematic for a semantics of desire that is tied to belief. This “loose end” for the Heimian analysis is more easily accommodated by an analysis that leaves the identity of \( f \) open.

(93) I want this weekend to last forever. (But I know, of course, that it will be over in a few hours.)

(Heim, 1992, 199(42))

What such examples show is that the possibilities that are relevant for a desire statement may be possibilities that are circumstantially accessible, yet doxastically inaccessible. Although it is impossible for any Sunday to last more than 24 hours, it *is* possible – circumstantially – that my work engagements on Monday will be canceled and my “weekend” will last one day longer. If doxastic alternatives can, but need not, restrict what is desired, a third stab at the semantics of *want* would be (94). The truth-conditional part of the denotation can also be given a quantificational implementation as in (95), following von Fintel (1999).\(^{11}\) The function \( \text{max}_g(f) \) retrieves the \( g(w) \)-best worlds among a set of accessible worlds \( \bigcap f(w) \) (which, I have argued, need not match the attitude holder’s set of doxastic alternatives).\(^{12}\) In the case of *want*, the attitude holder’s desires in the world of evaluation contribute the ordering source.

(94) \((\text{Third try})\)

\[
\llbracket \text{want} \rrbracket^f(p)(a)(w) \text{ is defined iff } \bigcap f(w) \cap p \neq \emptyset \text{ and } \bigcap f(w) \cap \neg p \neq \emptyset.
\]

If defined, \( \llbracket \text{want} \rrbracket^f(p)(a)(w) = 1 \text{ iff } \bigcap f(w) \cap p <_{\text{DES}_{a,w}} \bigcap f(w) \cap \neg p. \)

(95) \((\text{Quantificational implementation})\)

If defined, \( \llbracket \text{want} \rrbracket^f(p)(a)(w) = 1 \text{ iff } \forall w' \in \text{max}_{\text{DES}_{a,w}}(f(w)).p(w'). \)

---

\(^{11}\) See Crnić (2011) for recent discussion of monotonicity properties of desire predicates in relation to the Heimian similarity-based analysis and the quantificational analysis due to von Fintel (1999).

\(^{12}\) See (39) in Chapter 2 for the definition of this function.
This denotation for *want* is similar in some ways to the denotation for weak necessity *ought to* developed in Chapter 2. Both predicates require their prejacent *p* to be unsettled in a set of accessible worlds (the definedeness condition), and both bring in an additional consideration which favors *p*-worlds over ¬*p*-worlds among the accessible worlds. What is special about *want* is that the additional consideration is restricted to desires and linked to a particular individual, namely the individual denoted by the verb’s subject. Are considerations of personal desire also similar to *ought*’s secondary priorities in terms of presupposed commitments in the discourse? Comparing *ought* not just to *want*, but also to *necessary*, is helpful in answering this question.

### 3.2 From verbs of desire to modal predicates

One of Villalta’s (2008) central proposals is that the comparative analysis of desire verbs extends to a class of predicates in Spanish that select for subjunctive-marked embedded clauses. Among these are the modal predicates *es necesario* ‘is necessary,’ *es probable* ‘is probable,’ *es posible* ‘is possible,’ and *necesitar* ‘need’ (ibid., p. 519).

In this section, I discuss the analysis of one predicate in this class: the necessity modal adjective *necessary*. Exposing the inherently goal-oriented nature of this modal as a clausal operator, I argue that *necessary* does not contribute a likelihood-oriented modality (as argued by Villalta 2008) and neither is it the general purpose necessity modal operator it is often considered to be (Sections 3.2.1-3.2.2). However, I observe that *necessary* is a strong modal, like *have to*, and thus propose that it places different presuppositions on the priorities it is sensitive to than a weak necessity modal like *ought* does. *Want*, on the other hand, is similar to *ought* in being sensitive to prioritizing information that is presupposed not to be collectively committed to (Section 3.2.3).
3.2.1 The comparative analysis extended to modals: Villalta (2006, 2008)

In Spanish, both *querer* ‘want’ and *es necesario* ‘is necessary’ are predicates that select for subjunctive mood in their embedded clauses. Villalta (2008) proposes that this shared trait means that ‘want’ and ‘necessary’ both have a semantics that is based on comparison of alternatives. The lexical entries she associates with the two predicates are shown below, side by side (the first is repeated from (88)).

(96) \[ [\text{want}_C]^g(p)(a)(w) \text{ is defined iff } \forall q \in g(C) : \text{Dox}_a(w) \cap q \neq \emptyset. \text{ If defined,} \\
[\text{want}_C]^g(p)(a)(w) = 1 \text{ iff } \forall q : q \neq p \& q \in g(C) : p <_{\text{DES}_{a,w}} q. \]

(97) \[ [\text{be necessary}_C]^g(p)(w) = 1 \text{ iff } \forall q : q \neq p \& q \in g(C) : p <_{\text{LIKELY}_{w}} q. \]

(Villalta, 2008, 482(42), notation adapted)

A modal predicate like *be necessary* differs from *want* in three respects according to these proposals. First, it does not have an individual argument. Second, it does not include a precondition that restricts alternatives to ones that overlap with a set of beliefs. And third, it compares alternatives based on likelihood instead of desirability. The notion of likelihood that is used to discriminate between alternatives is “based on comparative similarity (closeness to the actual world, cf. Lewis 1986)” (ibid., p. 483).

(98) a. For any \( w, w', w'' \in W \), \( w' <_w w'' \) iff \( w' \) is closer to \( w \) than \( w'' \).

b. For any \( p \subseteq W, q \subseteq W, p <_{\text{LIKELY}_w} q \) iff \( \forall w'' \in q \exists w' \in p \) such that \( w' <_w w'' \), and it is not the case that \( \forall w' \in p \exists w'' \in q \) such that \( w'' <_w w' \).

(Villalta, 2008, 482(42), notation adapted)

An immediate worry that arises is that this analysis creates an entailment relation between *be necessary* and *be likely*, a relation that does not exist in everyday English (in contrast to modal logic). First, what is necessary may be unlikely:

(99) a. Marijuana reform necessary, but unlikely.\(^\text{13}\)

b. It is necessary, but unlikely that the countries who need such a central clearing-house would actually take advantage of it.14

Second, what is likely may be unnecessary. Villalta’s picnic potluck scenario (see above, p. 112) illustrates this point. Since the most likely event in that scenario is that Sofía will bring ice cream, the semantics of *be necessary* in (97) seems to predict that *It is necessary that Sofía bring ice cream* is necessarily true in the context. This is not the case, however, even if Sofía does bring ice cream to the picnic in actuality.

Independently of the meaning of *necessary*, the definition of likelihood in (98) is problematic for the reason that it equates likelihood with greater similarity to the evaluation world. As we know from good detective stories, the real world can be host to highly unlikely events:

> There are certain worlds among the accessible worlds which are more far-fetched than others. [...] Far-fetched in respect to what? In respect to what is the case in the real world? This can’t be true, since it seems quite natural to say that something which was almost impossible, turned out to be the case. Actually, it is things like this which usually happen in detective stories. The most unlikely candidate is the murderer.

(Kratzer, 1981, §3)

An additional concern regarding the semantics in (97) has to do with the type of modality conveyed by *necessary*. Villalta’s proposal was made specifically in the context of teleological interpretations of *be necessary*. This is not apparent in the denotation in (97), but is somewhat more apparent in a proposal by Krasikova (2008), on which Villalta’s semantics for *necessary* is based. Krasikova (2008) proposes that the semantic contribution of purpose clauses, whether overtly present or supplied by context, is to provide a standard of comparison to which the likelihood of the proposition embedded under the modal is compared. The purpose clause is understood as setting a lower bound for the likelihood

---

of the proposition embedded under the modal. According to this view, modals like *necessary* “relate propositions to degrees of their comparative possibility in a given world” (ibid., p. 346). I will not at this point delve deeper into the specifics of this proposal. This would require at the very least clarifying how the likelihood of propositions is assumed to be calculated, something which is not immediately clear from Krasikova’s presentation. Some of the challenges facing a comparative probabilistic account of teleological modality were discussed in Chapter 2 in connection with Finlay’s (2009, 2010) recent work (see Section 2.2.1).

Based on English data, I argue instead that *necessary* is a priority type necessity modal which, at least as a clausal operator, is primarily teleological.

### 3.2.2 *Necessary that there be a goal*

A rough guide to the semantics of *necessary* are its synonyms, a cohort which includes the modal adjectives *essential* and *crucial*. Tracing the diachronic development of *essential* and *crucial* in English, Van linden et al. (2008) explain how over time these adjectives have developed what we would call teleological, or goal-oriented interpretations.¹⁵ The fact that *necessary* is often provided in definitions of these two modals and offered as their synonym (especially *essential*, see, e.g., the Oxford English Dictionary) suggests that it too is a teleological modal, at least in some of its uses.

I will argue in what follows that *necessary* always depends on a priority to get its domain of quantification. In certain syntactic configurations (in particular, when it takes untensed *that*-clause complements or *(for)-to*-infinitives), this results in strictly goal-oriented interpretations of the modal. This claim is new and surprising. It contradicts what seems to be a common, if mostly silent, assumption that *necessary* is a general purpose (or poly-

---

¹⁵ They also conclude that the adjectives have undergone a further semantic change and are able to express deontic modality as well. See Section 4.4 for arguments against this conclusion.
functional) necessity modal which can be used deontically and even non-circumstantially in appropriate contexts.

Intuitively, it is impossible to decide if something is (or is not) necessary without first specifying what it is necessary for. Suppose we had in front of us a few drops of lemon juice. We could run a series of tests on the juice to measure some of its properties: the degree to which it is yellow, transparent, heavy, or sour. We could not, however, measure the degree to which the juice is “necessary.” The only way to make sense of a claim that lemon juice is necessary, as in (100), is to evaluate if it follows from the achievement of some priority or goal. In this example, the relevant priority is to prepare poppy seed tea.

(100) Is lemon juice necessary when preparing poppy seed tea/wash?16

Being necessary is not a property that entities or events have independent of a particular priority. Sometimes the priority is made explicit in a rationale clause, as in the notice below.

Figure 3.2: Goal-oriented necessary. Here in construction with a rationale clause to keep the bagshare up and running (August 2010, Serio’s Market, Northampton, Mass.).

A surprising finding is that deontic interpretations of necessary are ruled out when the modal takes an untensed that-clause complement. This is surprising on the traditional view,

where all root (or non-epistemic) modality types are grouped together in one natural class, generating the expectation that a polyfunctional modal will allow different root interpretations in different contexts of utterance. Consider (101). Rationing food is a teleological necessity if one’s priority is to survive a famine. It can also be a deontic necessity, e.g., if there are rules mandating the rationing during that time. Both (101b) and (101a) would be true in such a situation, but for different reasons. Whereas (101a) describes the content of a law, (101b) describes survival during difficult times.

(101) During the famine, …
   a. … it was mandated by law that food be rationed.
   b. … it was necessary that food be rationed (and retired laboratory animals eaten).

Next, consider (102). This example shows that necessary that \( q \) cannot convey the deontic necessity that mandated by law that \( q \) does. The context in this example supports food rationing as a deontic necessity (and therefore (102a) is clearly true). But since the priority of survival can be reached without food rationing, the teleological necessity based on survival is false.

(102) **Years of plenty.** After the famine ended there was plenty of food again for everybody, but the law hadn’t changed for a while.

   During that interim period, …
   a. … it was mandated by law that food be rationed.
   b. … it was necessary that food be rationed.

   Why are we still hesitant to say that (102b) is false – full stop – in this context? The reason is that necessary depends on a priority, and a different priority could make the sentence true. In particular, since the prejacent to the modal matches what is deontically necessary in the scenario, the preference of abiding by the law would do the trick. Imagine that (102b) is uttered by a lawmaker who is an adamant believer in upholding and enforcing
the law. (102b) is true in the context of (102) if necessary means “necessary in order to behave in accordance with the law.” To the extent that we are willing to judge (101b)/(102b) as “deontic,” it is not because the modal describes the content of the law in these examples (this necessary cannot do), but rather because it is restricted by a priority that exists because the law exists, as a side-effect if you will.\footnote{See Chapter 4 for further discussion.}

Another kind of apparent non-teleological use of necessary is common in the writings of logicians (103), who use this modal in an alethic sense. Alethic modality (from the Greek word for ‘truth’) is the philosophical modality of modal logic, the modality used to talk about what is “necessarily true” (Gamut, 1991, Vol. 2, 30). In its alethic use, is necessary can often be paraphrased as is a necessary fact.

\begin{enumerate}
\item It is necessary that gold has atomic number 79.\footnote{Halbach and Welch (2009, 72).}
\item If something is necessary, then it is necessarily so.\footnote{Gamut (1991, Vol. 2, 21).}
\item It is possible that you don’t understand me, but it isn’t necessary.\footnote{Gamut (1991, Vol. 2, 21).}
\end{enumerate}

How do these uses fit the generalization that necessary is teleological in nature? First, note that the modal’s prejacent in (103) (a, and elided in c) is tensed.\footnote{Brennan (1993) notes that epistemic interpretations of necessary that are found only when the modal’s complement is tensed (p. 88, note 19). Portner (2009) groups alethic modality in the same class as epistemic modality (suggesting factual modality as a term that covers the two types and perhaps others; p. 135).} This type of complement is not natural in everyday non-technical English. (In the OED, alethic-type interpretations are associated specifically with uses of the adjective in philosophical and theological discourse; s.v. necessary, adj., senses III 6. a. and III 6. b.) When we give the modal an untensed that-clause complement, we see that necessary, the adjective, is degraded as an
alethic modal in comparison to the adverb *necessarily*. Consider (104), a pair inspired by the example in (103a).

(104) If this particle is Helium, then . . .

   a. . . . necessarily it has atomic number 2.

   b. . . . it is necessary that it have atomic number 2.

   The sentences in (105) make the same point for the culinary inclined. Using *necessary* in the consequent clause invites the interpretation that containing olives is somehow useful (for the spread, for reaching some goal). But we have no clue as to what this usefulness is about, and the result is that (105b) sounds strange and is degraded with respect to (105a). Furthermore, the comparison between (105b) and (105c) highlights the fact that a teleological interpretation (where the goal is provided in the *if*-clause) restores the acceptability of *necessary* in the main clause.

(105) a. If this spread is a tapenade, then it necessarily contains olives.

   b. . . . If this spread is a tapenade, then it is necessary that it contain olives.

   c. If you want to call this a tapenade, then it is necessary that it contain olives.

   From a purely logical perspective, there are reasons to treat phrases in which the adjective *necessary* is interpreted in an alethic sense as standing in for (perhaps as a shorthand for) phrases including *necessarily* and a predicate of truth. The argument for considering this substitution comes from quantificational statements involving alethic *necessary* (e.g., *All theorems of Peano arithmetic are necessary*), which cannot be formalized in a first-order logic augmented with a necessity modal operator, □ (“box”). Halbach and Welch (2009) argue that replacing *is necessary* by *is necessarily true* in such statements solves the formalization problem and is furthermore technically feasible.

   Let’s set aside alethic uses of *necessary*, then.

   The goal-oriented generalization still falls short of explaining the interpretation of *necessary* in other syntactic configurations, and as an attributive modifier in particular. Mod-
ifying nouns like result and consequence, the modal is used to qualify events as being “inevitably determined or produced by a previous state of things” or “occurring as the next logical step in a sequence of events” (OED, s.v. necessary, adj., sense III 6. c.)

(106) a. The Sitwells are known to everyone who has even a casual acquaintance with modern literature . . . This is a necessary consequence of the Sitwellian methods of publicity. (OED 1927)

b. Along with references to the traditional rise and fall of fortune, there is a tendency to regard the rise of Bullingbrook as a necessary result of Richard’s negligence. (OED 1992)

Descriptively, when necessary modifies nouns of result and consequence, it evaluates accessible possibilities of the kind described by the complement to the noun. One might argue that even in this case the modal does its ordinary job, except that the “goal” it is oriented to is of a circumstantial nature. I leave the details of this argument for future work.

The teleological generalization thus applies primarily to the interpretation of necessary when it takes tenseless clauses as complements. This interpretation is evident even in examples like (107), which are designed to encourage a dynamic, dispositional-type, reading of the modal (see Brennan 1993, 88, note 19). There is nevertheless a hint of goal-orientation even here: the sentence conveys that signing is good for Peter, or sits well with his personality, not merely that he is disposed to do it (compare to dispositional will in In view of his personality, Peter will sign whatever he’s presented with).

(107) It is necessary, in view of his personality, for Peter to sign whatever he’s presented with.

(Brennan, 1993, 88)

---

22 This meaning is arguably found also in Gamut’s (103c) above, especially once we consider this paraphrase of that sentence: It is possible that you don’t understand me, but it isn’t (a) necessary (result).
Finally, necessary is goal-oriented where a more polyfunctional necessity modal like have to can be purely circumstantial. Have to can be used to express the circumstantial necessity that non-parallel lines intersect (108a), or that I can’t help sneezing when I’m out in the sun (109a). In the corresponding sentences with necessary, the modality is teleological and the sentences invite a What for? follow-up. (109b) cannot be used to express an incontrollable urge to sneeze, for example, but suggests that sneezing is a necessity in light of achieving some goal, e.g., making it difficult to hear what someone next to me is about to say.

(108)  
a. If these lines are not parallel, they have to intersect at some point.

b. If these lines are not parallel, it is necessary that they intersect at some point.

(109)  
a. I have to sneeze.

b. It is necessary for me to sneeze.

3.2.3 Want, necessary, and ought: similarities and differences

A lexical entry for necessary must capture the observation that the modal is teleological, and thus different from a polyfunctional necessity modal like have to. What complicates the task of writing down a lexical entry is that both of these modals are strong necessity modals (in the sense of Chapter 2), as we will now see.

3.2.3.1 Necessary is a strong necessity modal

How does necessary compare in strength to other necessity modals in the inventory of English? Is it strong, like have to, or weak, like ought?

Entailment patterns that have been used in the literature to establish the weakness of ought fail to detect weakness with necessary. For example, (111) is just as contradictory as (110b), suggesting that must, have to, and necessary all express strong necessity.

(110)  
a. You ought to take the train, but you don’t have to.
b. #You must take the train, but you don’t have to.

(Following von Fintel and Iatridou 2008; see (5) in Chapter 2)

(111) #It’s necessary that you take the train, but you don’t have to.

There is an interesting caveat to the claim above. A sentence of the form necessary q can be true while the corresponding sentence with have to is false, but only when the modals wear their restrictions on their sleeves, so to speak. Thus, (112a) may be a conjunction of true teleological claims, but take away the phrases that describe the different backgrounds for the modals and you end up with a contradiction (112b).

(112) a. You don’t have to take the train (you can also drive there or fly), but it’s necessary for reducing (if you want to reduce) your carbon footprint.

b. #You don’t have to take the train, but it’s necessary.

3.2.3.2 Preliminary lexical entry for necessary

The claim that necessary is a strong necessity modal raises the following analytical puzzles:

i. Assuming that teleological modality is analyzed as quantification over a restricted domain of accessible worlds (namely, those in which certain priorities are achieved), why does domain restriction lead to weakness in the case of ought, but not in the case of necessary?

ii. How does the fact that necessary is primarily teleological relate to its being strong? How does the fact that ought is weak relate to the fact that it admits epistemic interpretations?

I address these puzzles by referring to the distinction between negotiable (secondary) and non-negotiable (primary) priorities defended in the previous chapter, and add a twist. The basic idea is to relativize necessary just to primary ordering sources, namely those
which are represented in the context and contain priorities that are presupposed to be collectively committed to in the conversation. This will account for the observation that the modal is strong.

To explain the inability of *necessary* to receive pure circumstantial and dispositional interpretations, I propose that the modal places a further condition on the context, requiring it to provide a set of favored worlds (determined by \( f^C \) and \( < g_1, \ldots, g_i > \) in \( C \)) that is a subset of the worlds determined just by the modal base (i.e., by \( f^C \) alone). This boils down to a requirement that the ordering sources in \( < g_1, \ldots, g_i > \) be non-empty and non-trivial. *Have to*, in contrast, places no restrictions on the contextual parameters it draws on and is thus able to expresses necessities that characterize the set of modal base worlds in its entirety, as in the case of pure circumstantial and dispositional modalities.

The condition placed by *necessary* cannot be as simple as \( < g_1, \ldots, g_i > \) being non-empty, however. After all, nothing prevents a non-empty, non-trivial, stereotypical ordering source (implicated in an epistemic claim) or an ordering source representing relevant rules (implicated in a deontic claim) from being represented in the context as primary ordering sources. But even in such contexts, *necessary* is interpreted in relation to a priority, and that priority goes beyond what the stereotypical or deontic ordering source provide. The sentences in (99) are relevant examples of the first kind: the author that claims that *Marijuana reform necessary, but unlikely* is clearly attuned not just to the possibilities opened up by the relevant circumstances, but also to what is likely and unlikely to take place. His necessity claim is thus sensitive to a modal base, a stereotypical ordering source, and some additional priorities that further restrict the possibilities that count as “favored” (in all of these, there is marijuana reform). *Necessary* adds these additional priorities to the context. I thus propose to incorporate the following additional ingredient into *necessary*’s lexical entry: an update of the sequence of primary ordering sources with a new ordering source, last in the sequence, that represents the additional priorities the modal is sensitive to.
These ideas about *necessary* motivate the proposal of the preliminary lexical entry in (113). According to this proposal, the modal has the same truth conditional contribution as *have to*, but different presuppositions. Accepting the necessity claim amounts to confirmation of the contextual update effected by *necessary*, which is the addition of a new final primary ordering source to the context.

I assume that propositions in this primary ordering source are relevant to the question under discussion, just as *ought*’s secondary priorities are.

\[
\text{(113) (Preliminary proposal)}
\]

\[
[[\text{necessary}]^C] = \lambda q_{<s,t>} \lambda w. \forall w' : w' \in \text{Fav}^C(w). q(w'),
\]

where \(b\) is a pertinent answer to \(\text{QU}_{113}\) and \(g\) such that \(g(w) = \{b\}\) is a new final ordering source in \(<g_1, \ldots, g_i>\).

A prediction of this analysis is that *necessary* is a performative modal, and moreover a performative modal whose performativity originates in the semantics, not in the pragmatics of conversation. This aspect of its meaning is predicted to be the main factor distinguishing *necessary* from *ought*. This seems intuitively correct, as the contrast between the pair in (114) suggests: the organizers of the bagshare chose to say (114a), not (114b). They use a teleological modal that can force their priorities into the “conversation” with their patrons.

\[
\text{(114) [The bagshare scenario of Figure 3.2.]}\]

a. It is necessary for the bags to come back.

b. The bags ought to come back.

Taking stock of what we have explained about the different necessity modals, note that there is still a piece that needs to be filled in. The observation that *necessary* is primarily teleological, an important aspect of the modal’s meaning that I have argued for, is unexplained by the lexical entry in (113). Specifically, it is unexpected that a modal with such a

---

23 I thank Paul Portner for discussion of this point.
Table 3.1: Microvariation among necessity modals and attitudes of desire. Three types of necessity modals and *wanted* compared.

<table>
<thead>
<tr>
<th>Modal</th>
<th>Presupposition</th>
<th>Truth conditional contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>necessary</td>
<td>$g^C_{i+1}(w) = {b}$, $b$ is pertinent to $QUD^C$</td>
<td>$\forall w': w' \in Fav^C(w) \cdot q(w')$</td>
</tr>
<tr>
<td>have to</td>
<td>(no presupposition)</td>
<td>$\forall w': w' \in Fav^C(w) \cdot q(w')$</td>
</tr>
<tr>
<td>ought to</td>
<td>$b$ is secondary, $b$ is pertinent to $QUD^C$</td>
<td>$\forall w': w' \in Fav^C(w) \cap b \cdot q(w')$</td>
</tr>
<tr>
<td>want</td>
<td>$g = DES_{a,w}$, $q$ isn’t settled in $\bigcap f(w) \cap Dox_a(w)$</td>
<td>$\forall w': w' \in max_q(f(w)) \cdot q(w')$</td>
</tr>
</tbody>
</table>

lexical entry would not be able to express epistemic and deontic modalities. No problem, one might suggest; simply add a lexical specification of this restriction to the preliminary denotation in (113). While this move might be appropriate for *necessary* – perhaps this modal is simply lexically restricted to one particular type of modality – I choose not to adopt this solution. In light of the evidence to be presented in Chapter 4, I believe that this aspect of the adjective’s meaning will receive a grammatical explanation.

3.2.3.3 Conclusion

I conclude the chapter with a side-by-side comparison of the key players in our discussion so far: the strong necessity modals *necessary* and *have to*, the weak necessity modal *ought*, and the desire verb *want* (Table 3.1).

Three types of necessity modals have emerged from our discussion: ones that place no restrictions on the possibilities they quantify over (e.g., *have to*), ones that require context to highlight a way of restricting the domain of relevant possibilities (e.g., *necessary*), and ones that rely on extra-contextual assumptions to achieve their restricted domain (e.g., *ought to*).

Attitudes of desire are unique among this group of predicates in relying on an attitude holder to acquire their ordering sources. However, the propositions in these ordering sources are similar in nature to the secondary priorities that weak necessity modals are rel-
 ativized to. An individual’s desires, like secondary priorities, are typically not part of the collectively committed to information.
CHAPTER 4
A GRAMMATICAL CONSTRAINT ON THE INTERPRETATION
OF ROOT MODALS

The optimist says, “The glass is half full.”
The pessimist says, “The glass is half empty.”
The rationalist says, “This glass is twice as big as it needs to be.”

Thomas Cathcart & Daniel Klein, Plato and a platypus walk into a bar . . . (2007)

An interesting perspective on the relationship between modals and attitude predicates is provided by versatile modal words – words that compose with multiple types of syntactic complements and express a variety of modality types. One can need a car and need to have a car in English, and a common assumption is that these phrases mean the same thing. This chapter substantiates a semantic argument against the belief that transitive configurations of intensional verbs are equivalent semantically to the infinitival configuration of the same verbs. Based on experimental and crosslinguistic data, I show that different subtypes of root modality are expressed in different syntactic configurations of versatile modals. When versatile modals take nominal complements, they are shown to allow teleological interpretations and to resist deontic interpretations. When their complements are verbal, this limitation is lifted and the modal is free to pick up on a range of priorities and deontic sources that are available in the context. The recurrence of the phenomenon across languages motivates a new look at the syntax-semantics mapping of root modals and the limits of their context dependency.

How fine-grained is the sensitivity of natural language to different varieties of modality types (modal “flavors”)? Obviously, our lexicons keep track of very fine-grained distinctions when it comes to modals and attitude predicates. We can talk about wanting things, craving them, hoping for them, yearning for them, and having a host of other subtly different attitudes of desire towards them. We have no trouble keeping track of these subtle
differences, which means that they must be encoded in our knowledge of our native language. The question about the sensitivity of the grammar to such distinctions is much more interesting. Is there corresponding evidence that the grammar is sensitive to distinctions between subtly different types of modality? If so, what is the level of granularity with which it allows us, the language users, to make distinctions among modal possibilities?

A test case that addresses one aspect of these broad questions is the distinction between deontic and goal-oriented modalities. These sub-varieties are usually only distinguished on conceptual grounds: the former is the modality of possibilities determined by rules and norms; the latter is the modality of possibilities determined by goals and preferences more generally. Portner (2009) suggests the term priority modalities to refer to deontic, teleological, and desire-based modalities. Conceptually, these are the modalities that allow us to talk about what would be possible or necessary if some priority (be it a desire, a directive, or a goal) were to materialize. There are modals that are not restricted in the range of priority modalities they can express in context, e.g., need to below.

(115) Do we need to install a power outlet in the bathroom?

   a. Yes, to get the house up to code. (deontic interpretation)

   b. Yes, to put a digital picture frame in there. (teleological/bouletic interpretation)

In the literature, deontic and teleological varieties of priority modality are sometimes not distinguished at all (Palmer 2001 is a notable example). If they are, the difference between them is usually not tied to semantic properties of the word introducing the modality. For Kratzer (1981, 1991, 2012) and much of the formal literature on modality, both deontic and teleological modals are analyzed as circumstantial modals that employ an ordering source to rank their accessible worlds. Since the choice of ordering source is open to contextual pressure, the resolution of modal flavor is not made in the grammar. A grammatical split among priority modals is not acknowledged even in theories that grant the grammar a role in the calculation of modality types. Root modals in Cinque’s “Universal Hierarchy of
Clausal Functional Projections” come in three varieties: ability, obligation, and permission (Cinque 1999, and especially Cinque 2006, 92). The latter two seem to cover both deontic and teleological modalities, distinguishing modals only according to force. Hacquard (2006, 2009) offers an inspiring approach to understanding the syntactic-semantic basis of modality types, but she too does not draw a distinction between teleological modals and the conceptually similar class of “ought-to-do” deontics. Existing theories generate the expectation that a modal that is not lexically restricted to expressing a specific priority modality type should be open to contextual influence on its interpretation. If the modal can be interpreted teleologically in some syntactic configuration, it is expected to give rise to deontic interpretations in that configuration as well, given appropriate contextual support.

The goal of this chapter is to show that the context dependency of root modals is restricted grammatically, since the syntactic configuration a modal appears in influences its ability to express the full range of priority modalities. The argument is based on the extent to which need and modals like it across languages can express deontic and teleological modalities. I observe that need can express both sub-varieties of priority modality when its complement is an infinitival clause, and that one particular sub-variety, namely deontic modality, cannot be expressed when the modal’s complement is a noun phrase. This basic observation is presented in Section 4.1. Section 4.1.2 sharpens the distinction between teleological modality and deontic modality and specifies the properties of contexts that can be used to tease them apart. The relevant types of contexts are then used in two controlled experiments to investigate the interpretation of need with infinitival and DP complements (Section 4.1.3). The results of these experiments show a clear difference in the modal potential of need depending on the syntactic category of its complement.

---

1 See also Brennan (1993), Barbiers (1995, 2002).

2 I assume that desires are a special type of goal (one that can be associated with a desiring agent) and will thus not distinguish between teleological and bouletic modalities in the discussion. These two sub-varieties of priority modality are hard to tease apart in many examples and should not be thought of as mutually exclusive categories (Portner, 2009, 185).
The second part of the chapter extends the empirical basis of the generalization to additional syntactic configurations (Section 4.2) and to additional languages (Section 4.3). The basic correlation between type of modality and type of modal complement is observed in Hindi-Urdu and Hebrew (with suggestive data from Czech). These languages also show clearly that tenseless complementizer phrases (CPs) pattern with noun phrases in disallowing deontic interpretations of a modal they are in construction with (Section 4.4).

The empirical findings are stated as a generalization about the syntax-semantics mapping of modals and attitude verbs (Section 4.5). Several hypotheses about the source of this generalization are described and evaluated in the course of the discussion, although the theoretical basis for the generalization remains undetermined in the present work.

4.1 The basic observation: need to versus need

The empirical focus of the chapter is the expression of priority modality by necessity modals like need. In terms of the strength of necessity it expresses, need groups with strong necessity modals like have to and necessary.\(^3\) If my proposal in Chapter 2 is on the right track, this means that it is sensitive to priorities that are presupposed in the conversation to be collectively committed to by the conversational participants, or ones that are forced into collective commitment following acceptance of the modal claim:

(116) [Mother to her stubborn child:]
   a. You have to eat your spinach!

---

\(^3\) Both need to and need with a DP complement pattern with strong necessity modals on the following diagnostic of strength (see discussion of necessary in Chapter 3, p. 126):

(i) a. You ought to take the train, but you don’t have to.
   b. *You must take the train, but you don’t have to.

(ii) a. *You need to take the train, but you don’t have to.
   b. *You need the A train, but you don’t have to take the A train.
b. It’s necessary that you eat your spinach!

c. You need to eat your spinach!

Things get a bit more interesting when we investigate the types of modalities the modal expresses. When it takes infinitival complements, need is closer in its modal potential to have to, but in other syntactic configurations it is more similar to teleological necessary. What looks like a grammatically conditioned variability in modal potential will be the focus of this chapter.

In English, we can appreciate the grammatical influence on need’s interpretation by comparing sentences in which the modal’s complement is a to-infinitive and minimally different ones in which the modal’s complement is a noun phrase. Compare the modal question in (115) above with the corresponding question in (117) (A’s first utterance in the dialog). The crucial difference between these question is the type of complement that need takes in each. In terms of their meaning, need with a nominal complement is a strictly teleological modal. If there is no reason to put in an outlet, then B is justified in concluding that they don’t need an outlet. The modal is resistant to deontic interpretations in this configuration. Even after A brings up a relevant rule, this rule is not immediately available to serve as a background for interpreting the phrase need an outlet. In contrast to need with a DP complement, need to has less problem receiving the missing deontic interpretation (as in (115) above and in A’s final utterance in (117)).

(117) **Outlets.** There is a tiny bathroom in A and B’s garage. It is never used.

A: Do we need a power outlet in the bathroom?

B: No. I can’t think of any reason for putting one in there.

A: Doesn’t the building code require that there be an outlet in every bathroom?

B: Maybe, but it makes no sense. We don’t need an outlet down there.

A: Still, we need to install one.
Note that need to in A’s final utterance in (117) gives rise to a necessity statement that is almost as weak as the corresponding claim with ought to in (118). In both cases A bases the final necessity claim on a priority related to the content of a regulation that is not endorsed by her interlocutor, but only the need to statement has the performative effect of forcing collective commitment to the regulation, postfactum. Consequently, if it is the case that we need to install an outlet, it is also the case that we ought to do so, but the reverse does not hold.

(118)  
A: Doesn’t the building code require that there be an outlet in every bathroom?
B: Maybe, but it makes no sense. We don’t need an outlet down there.
A: Still, we ought to install one.

The dialog could also have evolved in a different direction after A brought up the rule pertaining to power outlets. In (119), B accepts the relevance of the rule and uses the phrase need an outlet to describe a necessity that follows from it.

(119)  
A: Doesn’t the building code require that there be an outlet in every bathroom?
B: Okay, so maybe we do need an outlet down there.

Isn’t this an example of the modal taking a nominal complement and expressing deontic modality? I will maintain that the modality in this kind of example remains teleological, despite the reference to the content of the law. B’s utterance of okay, so . . . before the need claim signals acceptance of the law as relevant for the deliberation, and commitment to it. This move triggers a change in the priorities that are presupposed to be collectively committed to, which in turn results in a change in the truth value of the modal statement in the scenario. I will call this a pseudo-deontic interpretation: a goal-oriented interpretation with a goal of abiding by the law. Changes in the committed to priorities bring out the differences between deontics and pseudo-deontics in contexts like (117) (see Section 4.1.2).

The basic observation, to summarize, is that when the rules conflict with the priorities that are presupposed to be collectively committed to in a context (i.e., the primary prior-
ities), a phrase like *need a power outlet* seems restricted to describing necessities in line with the contextual priorities. In the very same context, *need to install a power outlet* has access to the primary priorities as well as to contextually salient rules and regulations, even if these are not presupposed to be collectively committed to before the modal claim is uttered. The modal verb *need* is thus able to express more varieties of priority modality when its complement is an infinitival; we may say that it has “greater modal potential” in this configuration. Importantly, these examples show that the ban on deontic interpretations is not a lexical property of the modal. Quite the contrary, speakers of English have plenty of evidence that *need* can express deontic modality in certain configurations. It will be my working hypothesis that the restriction to teleological modality in construction with nominal complements is a fact to be given a grammatical explanation.

### 4.1.1 Syntactic and semantic versatility

Modals that are of interest for testing grammatical constraints on the expression of priority modality are what I will call *versatile modals*. Versatility of a modal is both syntactic and semantic, as follows.

i. *Syntactic versatility*: the modal takes verbal complements (e.g., infinitival phrases, or IPs) as well as nominal complements (e.g., noun phrases, or DPs, and in some languages also CPs).

ii. *Semantic versatility*: the modal allows teleological, deontic, and epistemic/circumstantial interpretations with verbal complements.

English *need* is a versatile modal according to this definition. It takes both verbal and nominal complements and exhibits the full range of modal interpretations with verbal complements.

(120) *Syntactic versatility of need:*

a. I need to have a cup of coffee. (*verbal complement*)
b. I need a cup of coffee. (*nominal complement*)

(121) *Semantic versatility of need:*

a. The soup tastes like water. There needs to be more salt in it. (*teleological*)

b. There needs to be more than one way to enter and exit each room in the building. (*deontic*)

c. According to the map, there needs to be a bus stop around the corner. (*epistemic*)

d. I really need to sneeze. (*circumstantial*)

It is noteworthy that epistemic interpretations of *need to* are rare and atypical (Smith, 2003), but even so, English speakers have evidence that *need* can accommodate many different types of modality. It is thus all the more surprising that in one specific syntactic configuration, the *nominal complement construction* of (120b), this richness disappears. Note that I will refrain from calling this construction “transitive;” since some of the modals we will discuss outside of English do not exhibit transitive *syntax* in the relevant configuration. That is, in some languages, the subject of the modal ‘need’ that takes nominal complements does not receive nominative case and the nominal complement itself does not receive accusative case (Harves 2008, Harves and Kayne 2012). Indeed, one result of this chapter is that the modal potential of versatile modals with nominal complements is restricted in exactly the same way crosslinguistically, regardless of the case marking on the subject and object noun phrases in construction with the modal.

### 4.1.2 Deontic and teleological modalities

In the previous chapter, I drew a distinction between the sub-varieties of root modality that are expressed by the modal adjectives *mandatory* and *necessary* (see (102)). While

---

4 The possibility of using *need to* when the modality is keyed to some body of evidence seems to depend on temporal orientation. Typical examples of epistemic reasoning are past-oriented and cannot be expressed using *need to*. 

139
the existence of a law is enough to make a statement with *mandatory* true, it is not always enough to make a corresponding statement with *necessary* true. The example we looked at is repeated in (122): since an abundance of food means that there are multiple ways for people to survive without rationing, a teleological claim oriented toward the priority of survival is false in the context of (122). However, *necessary* can come quite close to expressing the rule-based meaning that is conveyed by *mandatory*. This happens when the priority it orients to is a priority of conforming to the law. For example, it is necessary that food be rationed even in times of plenty if a desire for political order and lawful conduct is an operative priority. I concluded in Section 3.2.2 that *necessary* is a teleological modal, in the sense that it is always interpreted with respect to a priority (either contextually determined, or specified overtly, as in a rationale clause).

(122)  [For a certain period of time in the aftermath of a famine, rules mandating food rationing were still in effect although food was already in large supply.]

During the interim period, it was \{mandated by law, necessary\} that food be rationed.

The ability to use the law as a source for the formation of priorities creates an affinity between deontic and teleological modalities that makes it difficult to distinguish between them in many cases (von Fintel and Iatridou 2005a, §5.5, Hacquard 2006). Hacquard (2006) makes this point with respect to “ought-to-do” deontics (Feldman, 1986): deontic modals that affirm a connection between an individual and a state of affairs.

In fact, many instances of *ought-to-do* deontics can be reduced to goal-oriented modality, where the goal is unpronounced and often deals with avoiding some kind of punishment. The following sentence is not so much a statement about what the laws of this world look like, but rather a statement of what I need to do, given a certain body of laws and regulations in effect in the actual world, *in order for me not to get a ticket*: […] I have to take out the trash on Wednesdays.

(Hacquard, 2006, 41)
The famine context in (122) and the power outlet scenario in (117) are examples of a type of context that is particularly useful in teasing apart deontic and teleological modalities. They present scenarios in which an original set of primary priorities is in conflict with what the law provides in the scenario. (In the first scenario, there are ways to survive without rationing food; in the second, there is no reason to follow protocol and install a power outlet in a room that nobody uses.) A contextual shift in the primary priorities is required in such scenarios for a teleological modal to be made sensitive to a priority that derives from the law. Detecting a shift in priorities prior to the utterance of the modal is an indication that the modal has received a pseudo-deontic interpretation, since a preparatory priority shift is not required for the expression of true deontic modality.

Scenarios in which deontic and teleological modalities can be distinguished will be called \((+\text{content}/-\text{preferences})\) scenarios, referring to the tension between what is lawfully prescribed in them (the content of the law) and what the primary priorities are taken to be (the preferences). In such scenarios, the set of favored worlds varies depending on whether the law or other priorities are used to order the worlds determined by the relevant circumstances.

Shifts in the contextual priorities are often not recorded in the speech signal, but sometimes they are. A shift is overtly indicated in (119), for example, where an acceptance move (Okay, so . . .) follows the mention of a rule that is relevant for the deliberation. A plausible interpretation of B’s acceptance move in this dialog is that it indicates a change in the priorities he endorses. This effects a change on what the collectively committed to priorities are presupposed to be in the context.

All else being equal, if there is no indication of a shift in priorities, a strictly teleological strong necessity modal is expected to be somewhat infelicitous in a \((+\text{content}/-\text{preferences})\) scenario. This infelicity is subtle, but it can be detected by direct comparison of strictly teleological and strictly deontic necessity modals, as we have done with necessary and
mandatory in the famine scenario. In the next section, we will see that it is also psycholinguistically real.

Section 4.1.3 presents experimental evidence that need in its nominal complement construction is infelicitous in (+content/-preferences) scenarios in comparison to need taking verbal complements.

### 4.1.3 Experimental findings

Two experiments were designed to test the correlation between the syntactic configuration of need and its semantic interpretation in different contexts.\(^5\) The results obtained provide quantitative evidence, first, that the nominal complement construction of need is naturally used to express teleological modality, and, second, that need to is significantly more open to expressing deontic modality than need complemented by a noun phrase.

#### 4.1.3.1 Participants

40 undergraduate students from the University of Massachusetts Amherst participated in the two experiments reported here. The students were all native speakers of American English. They were compensated with course credit for their participation.

#### 4.1.3.2 Materials

Two experiments were designed to test the correlation between the syntactic configuration of need and the range of modal meanings it allows. Each experiment consisted of twelve test items which took the form of a short paragraph followed by a target sentence, as shown in (123). The complement to need in the target sentence varied between an infinitive (the IP condition: needs to take the coconut cake) and a noun phrase (the DP condition: needs the coconut cake).

---

\(^5\) This study was presented at the 23rd Annual CUNY Conference on Human Sentence Processing (Rubinstein, 2010).
In both experiments, the verb in the complement to *need* was *have* in half of the items and *take* in the other half. This variability resulted in stimuli that sounded natural yet were able to express similar semantic relations, as described below.

(123) This weekend the local community center is organizing a cake exchange event. Everyone who comes brings a homemade cake and gets someone else’s cake in return. The exchange is determined randomly in a drawing. When Sharon saw what she had just won she was very upset: it was the coconut cake – the flavor she hates most of all!

   a. Sharon needs to take the coconut cake. (*IP condition*)
   b. Sharon needs the coconut cake. (*DP condition*)

Stimuli were based on only two meanings that the polysemous verbs *have* and *take* can express: the meaning of possession (as in *needs to take the coconut cake*) and the light-verb meaning of undergoing or experiencing an event (as in *needs to take a nap*). Items in the DP condition were able to express the same lexical relations of possession and undergoing as their IP condition counterparts.

6 The observation that the full polysemry of lexical meanings associated with overt *have* (*take*) is also available when the light verb seems to be missing has been used to argue for the existence of a covert ‘have’ (*take*) in the nominal complement construction of verbs like *want*. For example, Harley (2004) argues that, “[…] *John wants Mary or Jim wanted the apple pie* is multiply ambiguous in exactly the same way that *John wants to have Mary or Jim wanted to have the apple pie* is, making the notion that *want* introduces *have* into the interpretation of [*want DP*] very plausible” (ibid., p. 259). More on this below.

Contexts in Experiment 1 described (+content/-preferences) situations. Like (123), these were scenarios in which certain external rules or regulations were in conflict with a salient priority of an individual in the scenario. There were a total of twelve items in this experiment, six with *have* in the IP condition and six with *take*. Half of the *have* items
described a possession relation and half described an undergoing relation. Among the take items, two described an undergoing, and four described (grabbing into) possession.\footnote{There is no good reason for this imbalance of senses except for the fact that I was guided by the naturalness of the target sentences in the design of scenarios. In Experiment 2, the relations were equally represented for take, but not for have (there were 4 examples of undergoing and 2 of possession).}

Participants were asked to rate how well the target sentence in each item fit with the preceding short story, on a scale of 1 (Poorly) to 5 (Very well). We expected that sentences would be judged as congruent with stories in the IP condition, and incongruent in the DP condition.

Experiment 2 was used to further probe the interpretation possibilities of need in its two syntactic configurations. Items took the same format as in Experiment 1, namely a short story and a target sentence followed by a rating question about the fit between the story and the sentence. However, the stories in Experiment 2 described scenarios of a different kind, ones in which external rules or regulations were in agreement with the salient priorities (124). Scenarios of this kind, which we might classify as (+content/+preferences), support both a deontic and a teleological interpretation of need. The ratings of fit between stories and target sentences were thus expected to be uniformly on the side of congruence in this experiment, with nominal complement need picking up on the preferences given in the story, and need to picking up on either the preferences or the content of the rules at play.

(124) It is a rule at Sarah’s house that everyone naps for at least one hour between 2 and 4 in the afternoon. It’s 2pm, and Sarah just got back from a long day trip with her friends. She is particularly tired right now because she also didn’t get much sleep last night.

a. Sarah needs to take a nap. (IP condition)

b. Sarah needs a nap. (DP condition)

(125) Why does Sarah need to take a nap?

a. Because of the rules of the house. (Deontic reason)
b. Because she is tired. \textit{(Teleological reason)}

As a window into the interpretation actually assigned by participants to target sentences, every rating question in Experiment 2 was followed by a forced choice comprehension question (Experiment 2’). (125) provides an example. The version of the question shown is from an item in the IP condition, and one should keep in mind that the syntactic configuration of \textit{need} in the follow up question always matched the one in the sentence rated for fit. Items in the DP condition were expected to allow only an interpretation which was based on the salient priority in the scenario, while those in the IP condition were expected to be consistent with either this priority or the content of the rule that was described. The option describing rules or regulations was always presented first in order to make sure that it was read and considered in the IP condition even though the teleological reasons would have been an appropriate choice there too.

Target sentences in both experiments were inflected with present tense and had a bisyllabic feminine proper name as subject.

In addition to the experimental items, 10 filler items were included that were similar in form and length to the items in the two experiments. The fillers were created with two requirements in mind. First, they were to show that a verbal complement to \textit{need} may be incongruent with a preceding scenario. Since, in both experiments, items in the IP condition always resulted in congruent discourses, it was necessary to deter participants from using this incorrect generalization to inform their responses. The scenarios in these filler items would be characterized as (-content/-preferences). Four fillers had this property. As for \textit{need} with a nominal complement, items from Experiment 1 will have provided evidence for the possibility of a negative rating of congruence in the DP condition. A second group of fillers was designed for the forced choice task that, unlike items in Experiment 2, were not compatible with both answers to the question. In particular, the contexts given in these fillers were incompatible with needs that were based on rules or regulations. Two were of the (-content/-preferences) type, and two were of the (-content/+preferences) type. These
fillers were included in order to verify that participants were not picking the first answer across the board, i.e. the answer based on a deontic interpretation. The remaining four fillers featured the verb want in its transitive construction, and were similar in form to items in the DP condition.

Each participant responded to a total of 24 experimental items and 10 fillers in Experiment 1 and Experiment 2 together. Items within each experiment, filler items, and items from three unrelated experiments were mixed and randomized for each participant. The choice of items for a participant was determined based on their number in the experiment: Subject 1 was given the first condition of item 1, the second condition of item 2, the first condition of item 3, etc. Subject 2 was given the second condition of item 1, the first condition of item 2, and so forth.

The experiments were implemented using the block reading template in the software package Linger (version 2.94).

4.1.3.3 Procedure

Participants read items on a computer screen and used numbers on a keyboard to indicate their responses. For any given item, the scenario, the target sentence, the question introducing the rating scale, and the rating scale appeared all at once on the screen. For those items that included a comprehension question, this question along with its answers were displayed immediately after the rating choice was recorded, on a fresh screen (no longer containing the short paragraph describing the scenario). An additional keystroke was required in order to advance from item to item and a 500 ms delay was inserted between items.

A practice session familiarized the participants with the structure of the experiment. It included a total of 5 items, one of which demonstrated the setup of the rating task in the experiments (the other four items were practice items from unrelated experiments). The ‘target’ in the practice did not contain the modal verb need, and participants were not given
feedback on their choices during the practice. Participants were instructed to judge the fit between the sentence and the story based on the information in the story. They were also told that they would be asked a follow-up question after some items and that they would be able to take a break midway through the experiment. Each session lasted approximately 30 minutes including the break.

### 4.1.3.4 Results

Sentences in which need’s complement was a noun phrase received lower ratings of fit than corresponding need-IP sentences in the scenarios of Experiment 1 (Figure 4.1). In Experiment 2, where stories were consistent with both a deontic and a teleological interpretation of the modal, need-IP sentences and need-DP sentences both received high rating scores: 4.43 and 4.40 respectively (grand mean over all items and all subjects). Experiment 1 featured contexts which were only consistent with a deontic reading of need. Here, need-DP sentences received a rating of 2.94 on average, whereas need-IP sentences continued to be highly rated, 3.75 on average.\(^8\)

To test whether the difference between the two types of modal complements was significant in the (+content/-preferences) scenarios of Experiment 1, we conducted a mixed effects regression with crossed random effects for subjects and item (Table 4.1). The effect of modal complement on the ratings was found to be significant \((t = 7.155, p < 0.001)\). As a further test of the main effect of complement type, we built a second model for the ratings with random slopes for this treatment variable both by subject and by item.\(^9\) This model embodies the assumptions that different subjects may have differed in their sensitivity to need-DP and need-IP constructions and that different short stories brought out the difference between the conditions more clearly than others. In the more specific model, the type

---

\(^8\) All calculations and visualizations were carried out with the R open source software for statistical computing and graphics (R Development Core Team, 2008).

\(^9\) In R, the formula used to describe this model was \(\text{Rating} \sim \text{ModalComplement} + (1 + \text{ModalComplement}|\text{Subject}) + (1 + \text{ModalComplement}|\text{Question})\).
Figure 4.1: Experimental results comparing *need* and *need to* in deontic and teleological contexts. Need-DP sentences received significantly lower ratings than need-IP sentences in Exp. 1 ($t = 7.155, p < 0.001$). In Exp. 2, the ratings in the two conditions did not differ ($t = 0.43, p = 0.66$). The models fitted were linear mixed models with crossed random effects for subjects and items. Between experiments, a $t$-test revealed that the difference among the need-IP sentences was significant as well ($t = 6.45, p < 0.001$).
of modal complement still had a significant effect on participants’ ratings ($t = 3.628$). With more degrees of freedom, this model is more complex, yet a likelihood ratio test shows that it is justified ($\chi^2 = 18.435, p = 0.001015$; see Baayen 2008, p. 253, for discussion).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2.9417</td>
<td>0.1771</td>
<td>16.607</td>
</tr>
<tr>
<td>ModalComplement = IP</td>
<td>0.8083</td>
<td>0.1130</td>
<td>7.155</td>
</tr>
</tbody>
</table>

**Random effects** (40 subjects, 12 questions)

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>(Intercept)</td>
<td>0.26217</td>
<td>0.51202</td>
</tr>
<tr>
<td>Question</td>
<td>(Intercept)</td>
<td>0.22128</td>
<td>0.47041</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>1.53159</td>
<td>1.23757</td>
</tr>
</tbody>
</table>

$N = 480$, Log likelihood = -817.9

Table 4.1: Linear mixed model for Experiment 1. Summary of a linear mixed model fit to the raw data from Experiment 1 using R’s `lmer` function (Bates et al. 2008). The formula used in the call to `lmer` was `Rating ~ ModalComplement + (1 | Subject) + (1 | Question).

We conclude from this that when need’s complement is a nominal, the modal does not readily have access to modal interpretations that are based on deontic content. These interpretations are nonetheless freely available to it when it combines with a to-infinitive.

Next, we turn to the findings of Experiment 2. Recall that scenarios in this experiment supported both deontic and teleological readings of need, and the grand mean ratings of sentences in the two conditions were correspondingly quite similar: 4.43 for the IP condition, 4.40 for the DP condition (see the right half of Figure 4.1 above). Statistical analysis reveals no effect of modal complement on the ratings in this experiment (the estimated increase of 0.033 in mean ratings of IP complements was associated with a $t$-value of 0.43, $p = 0.66$).

More interestingly, speakers’ responses to the forced choice comprehension question that followed each item (Experiment 2’) give us direct insight as to the interpretations they assigned to the modal in the two syntactic configurations tested. Recall that, for each item,
participants chose between two possible answers to a *why* question about the item (as in (125) above). This response was taken to reflect the interpretation they assigned to the modal in light of the information provided in the short story. The deontic response was the one that was based on rules or regulations mentioned in the story. The teleological response was the one that was based on the preferences described in the story.

The number of responses indicative of deontic and teleological interpretations in each syntactic configuration is shown in Table 4.2. Although there was no overall difference in the proportions of the two types of response (the column totals are similar), the breakdown by condition shows that items in the DP condition contributed the majority of teleological interpretations \((141/248 = 56.85\%)\), whereas items in the IP condition were the main source of deontic interpretations \((133/232 = 57.33\%)\).

<table>
<thead>
<tr>
<th></th>
<th>Deontic response</th>
<th>Teleological response</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP condition</td>
<td>133</td>
<td>107</td>
</tr>
<tr>
<td>DP condition</td>
<td>99</td>
<td>141</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>232</strong></td>
<td><strong>248</strong></td>
</tr>
</tbody>
</table>

Table 4.2: Results of Experiment 2’ (responses to comprehension questions). A question like *Why does Tammy need (to have) a laptop?* was presented with two possible answers: one based on the content of a rule or regulation (*Because it’s company policy*) and one based on preferences or desires (*To watch her DVDs*). A significant tendency for content-based deontic interpretations to be paired with the IP configuration of the verb \((2.16\) times higher odds than in the DP configuration, \(p < 0.001\)) was found in a mixed logit model.

In order to model the random effects of subject and item while also respecting the categorical nature of outcomes in this task, we applied a mixed logit model to these data.\(^{10}\) The properties of the resulting model are shown in Table 4.3. They confirm the significance of the effect that syntactic configuration had on the interpretation of *need* in context.

The estimate for the fixed effect of complement type tells us that when the complement to *need* was an infinitive, the odds for a deontic response (which was the one coded as

---

\(^{10}\) See Jaeger (2008) for a detailed critique of ANOVA-based approaches to data which are proportions of categorical outcomes.
1 in this analysis) were 0.7695 log-odds higher than when the complement was a noun phrase. That is, the odds for a deontic interpretation were $e^{0.7695} \approx 2.16$ times higher in the IP condition than in the DP condition. As seen in the table, this difference between the conditions is significant ($p < 0.001$). That is, even though the ratings for *need* in the two conditions were essentially indistinguishable, they were clearly distinct in terms of their interpretation.

**Fixed effects**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>z value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.4525</td>
<td>0.3709</td>
<td>-1.220</td>
<td>0.222533</td>
</tr>
<tr>
<td>ModalComplement = IP</td>
<td>0.7695</td>
<td>0.2122</td>
<td>3.627</td>
<td>0.000287</td>
</tr>
</tbody>
</table>

**Random effects** (40 subjects, 12 questions)

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>(Intercept)</td>
<td>0.90501</td>
<td>0.95132</td>
</tr>
<tr>
<td>Question</td>
<td>(Intercept)</td>
<td>1.10807</td>
<td>1.05265</td>
</tr>
</tbody>
</table>

$N = 480$, Log likelihood $= -288$

Table 4.3: Logit mixed model for Experiment 2’. Logistic regression was carried out by specifying `family = binomial` in the call to `lmer`. The formula that was used in the call was `WhyResponse ~ ModalComplement + (1|Subject) + (1|Question)`.

### 4.1.3.5 Discussion

The results of these experiments establish a clear distinction between the interpretations available to modal *need* in two of its syntactic configurations in English. Experiment 1 showed a contrast in the acceptability of *need* depending on the type of its complement in contexts that supported deontic but not teleological necessity claims. Sentences in which the modal was complemented by a DP received significantly lower acceptability ratings in these contexts, suggesting that the modal does not readily allow deontic interpretations in this syntactic configuration. In Experiment 2 we saw that when rules and preferences were in agreement, the difference in ratings between the two conditions disappeared. *Need* in the nominal complement construction was appropriate and true in these scenarios because
salient priorities, to which it is keyed, sanctioned a universal modal claim. Importantly, the variable modal potential of *need* was manifest in this kind of scenario as well, in the answers participants gave to the comprehension question that followed each item. Echoing the finding from Experiment 1, we found a significant tendency for deontic readings of the modal to be paired with its IP configuration. In summary, while both deontic and teleological modalities are expressible by the modal verb *need*, it is mainly teleological modality that is available to it with nominal complements.

One must at this point wonder about two findings in the experiments that challenge this generalization. First, if deontic modality is not a possible interpretation for nominal complement *need*, why were there nonetheless traces of this interpretation in the data? And second, if *need to* can accommodate both types of modality considered here, why does there seem to be a difference in the ratings of sentences in the IP condition in the two experiments? I address these two questions in turn.

The line that separates deontic modality from teleological modality is blurred by pseudo-deontic interpretations, and therefore easy to overlook. The fact is that a deontic obligation can be the basis for the formation of a priority, if following the rules is considered to be a goal in itself. Following the rules entails at least fulfilling the latent goal of being obedient, and it is advantageous to be obedient, especially if the rules in question and one’s own interests are not in conflict. The items we included in Experiment 1 tried to make pseudo-deontic interpretations unlikely, by portraying situations in which following the rules was disadvantageous to an individual and not something they would be expected to be committed to (for example in (123), taking home a cake one cannot eat).

The benefits associated with following the rules are not easy to shake off, however, and an example of how they might raise the acceptability of sentences with *need* in the nominal complement construction is demonstrated by (126) below. This item received the highest positive adjustment to the intercept among the twelve items in Experiment 1 and is useful in making this point. If we focus just on Nicole’s preferences and purposefully
ignore the new regulation at her workplace, it is reasonable to say that Nicole does not need an assistant. But although she doesn’t need one, it would probably be more harmful for her to object about the new policy than to swallow the unwelcome change and deal with a personal assistant. Native speakers would likely attach little significance to the shift in contextual assumptions that is required for Nicole needs an assistant to be true in the context. They would have no difficulty construing the situation in a way that licenses a pseudo-deontic interpretation of the modal in the DP condition.

(126) Nicole is resourceful and independent. She always runs errands herself, manages her bank accounts herself, and doesn’t like strangers interfering with her business. That’s why she hates the new policy at her workplace: every manager in Nicole’s rank or higher is required to hire a personal assistant starting next month.

a. Nicole needs to have an assistant. (IP condition)

b. Nicole needs an assistant. (DP condition)

Some items were better than others in the sense that they were more resistant to contextual manipulation of this kind. The very fact that items differed from one another in this respect reminds us of the highly interactive nature of modal disambiguation in conversation. Depending on what aspects of a situation one chooses mentally to focus on, a modal statement they heard will receive subtly different interpretations. These choices are influenced by many factors, some of which are linguistic and concern the presentation of considerations in the scenario, and some which are not linguistic. Often, the context can only narrow down the options without resolving this indeterminacy completely. The responses we see in our experiments reflect the complex nature of the phenomenon under discussion.

The second question I raised concerns what looks like an effect of context type on the rating of sentences in the IP condition. Sentences in this condition received lower ratings in the (+content/-preferences) scenarios of Experiment 1 compared to the (+content/+preferences) scenarios of Experiment 2. A simple independent observations t-test
shows that the difference we perceive visually between the two IP columns in Figure 4.1 is also statistically significant ($t_{424.293,0.975} = 6.4509, p < 0.001$). This is a bit surprising, because one would think that need to would be completely felicitous as a deontic modal in the scenarios of Experiment 1. However, since need to can (grammatically) accommodate a variety of priority interpretations, participants may have considered a teleological interpretation of the modal in these sentences, in addition to a deontic interpretation. Since settling on this interpretation of the modal would result in falsity, the sentences may have been given a lower rating of fit overall. In other words, lower ratings may reflect intrusion from an unsuccessful resolution of the modality as teleological in a (+content/-preferences) scenario.

The experimental evidence presented in this section disconfirms an expectation from the literature that deontic and teleological modalities would be equally possible as root interpretations of a versatile modal. We saw that need can express only a proper subset of the priority root modalities in the nominal complement construction; it does not readily acquire rule-based interpretations in this configuration, even in contexts in which the existences of the rules is salient.

At a more general level, these results substantiate a new argument against a semantic equivalence between modal verbs that combine with nominals (or “deceased clauses,” to use McCawley’s 1974 figurative term) and those that combine with infinitivals.

In the next section we will see than not all nonfinite complements allow the semantic flexibility we observed with to-infinitives.

### 4.2 English (for)-to infinitives and the Expressive IP hypothesis

A plausible hypothesis regarding the greater modal potential of need with infinitival complements is that something about giving a versatile modal a nonfinite complement is necessary and sufficient for allowing a deontic interpretation of the modal.
(127) *(The Expressive IP hypothesis)*

Given enough contextual support, a versatile modal may receive a deontic interpretation if and only if its complement is a nonfinite clause.

Sentences like (128) (with *need*) constitute a crucial test case for this hypothesis. These are examples of what is traditionally called the *Accusativus cum Infinitivo* or “accusative and infinitive” construction (AcI, Bresnan 1972, pp. 149ff.)\(^\text{11}\) In this configuration, a modal or attitude verb is followed by an accusative marked noun phrase (*him*, optionally preceded by *for* in the case of *need*) and an infinitival clause (*to be in Amherst*).

(128)  We *{need (for), want}* him to be in Amherst.

Assuming that AcI complements count as nonfinite clauses, the Expressive IP hypothesis predicts that they should pattern with smaller *to*-infinitives in allowing deontic interpretations of *need*. If the larger variety of modality types is somehow related to the presence of a *to*-infinitive in a modal’s complement, we expect *(for)-to*-infinitives to have access to this variety as well. This prediction of the Expressive IP hypothesis is not met.

Consider the (+content/-preferences) scenario in (129), where the law is once again in conflict with the committed to priorities.

(129) **Fences.** City regulations mandate that home owners put up fences between their properties. You and your neighbor get along very well without a fence. In fact, both of you object to a fence because it would have to go right on top of the beautiful flower beds that have been flourishing between your two properties. You say to your neighbor:

a. False *We need a fence.*

b. True *We need to put up a fence. (according to city regulations)*

---

\(^{11}\) See Section 4.5.1 for a review of the distinguishing syntactic properties Bresnan (1972) associates with “*want* type AcI.”
The basic contrast between nominal and infinitival complements recurs: a nominal complement to *need* in (129a) prevents the modal from receiving the deontic interpretation that is more freely available in (129b). What is surprising is that the *(for)-to*-infinitives in (130), which seem to differ insignificantly from the infinitival complement in (129b), nevertheless cannot automatically express the necessity that follows from abiding by the law. In other words, (130) cannot naturally receive a deontic interpretation in this context, although the modal’s complement contains an infinitive.

(130) a. False We need (for) a fence to be here.
     b. False We need (for) there to be a fence.

Moreover, while adding a phrase like *according to city regulations* helps bring out the deontic interpretation of (129b) in this scenario, this addition is somewhat awkward when *need*’s complement is a noun phrase or a *(for)-to*-infinitive, as in (131). Intuitively, the awkwardness in these examples is due to a clash between the teleological interpretation that *need* is forced into by the syntactic configuration it is in, and the deontic background contributed by the adjunct.\(^{12}\) The city regulations do not set our priorities (i.e., the things we need), they only determine what counts as lawful conduct (i.e., the things we need to do).

(131) a. According to city regulations, we need a fence here.
     b. According to city regulations, we need (for) there to be a fence.

I conclude from these examples that giving *need* a nonfinite complement does not by itself secure a deontic interpretation for the modal in (+content/-preferences) contexts. AcI complements also show that *size* of the modal’s complement does not determine the range of interpretations the modal will allow. It is clearly not the case that any complement that is at least a *to*-infinitive allows for deontic interpretations of the modal. Only com-

\(^{12}\) I thank Kai von Fintel for this point.
plements that are exactly to-infinitives – nothing more and nothing less – seem to make non-teleological modality available.

Turning to other languages in the next section, we will see the English pattern recur: modals with noun phrase complements are obligatorily teleological, and semantic versatili-
ity is only found when the modal’s complement is an infinitive. Moreover, in languages
that allow full (subjunctive/future-marked) CPs as complements to versatile modals, these
constructions pattern with nominal complement constructions in terms of modal potential.

4.3 Crosslinguistic data and the Lexical Ambiguity hypothesis

Since we see need appearing in different syntactic configurations with different mean-
ings, a simple hypothesis is that it is simply a lexically ambiguous word. Ambiguity or polysemy are not foreign concepts when it comes to need. As Loureiro-Porto (2009) notes, the verb has traditionally been classified as having homomorphs belonging to three different lexical categories (auxiliary as in He need not go, full verb as in He needs a car, and catenative verb need to; ibid. §1.2, Jacobsson 1974).13 Why not say that need that takes nominal complements is lexically specified as a teleological modal, while the catenative verb need to is lexically specified more generally as a priority root modal?

The Lexical Ambiguity hypothesis is not easily falsifiable. However, it seems implausi-
ble that the lexical ambiguity involved is of the kind invoked to account for idiosyncrasies in lexical meaning. Blind lexical ambiguity generates expectations that are not met about the pairing between the particular root meanings that verbs meaning ‘need’ express crosslin-

13 It is usually only in its guise as an auxiliary that need is referred to as a true modal in discussions of the English modal inventory (e.g., Quirk et al. 1985, 138, Smith 2003, 245, Stowell 2004, 621), but it is clear that this use of the modal is disappearing. Jacobsson (1974) surmised that: “[...] observers on both sides of the Atlantic are agreed that in the spoken language catenative need is gaining ground at the expense of the auxiliary, which is often felt to be bookish and is sometimes described as a decaying modal” (ibid., p. 63). The decay of auxiliary need was complete in spoken English by the end of the twentieth century (Leech, 2003), whereas need to has shown a remarkable rise in frequency (ibid., p. 230). A number of studies have converged on the conclusion that need to is replacing auxiliary need in all contexts in present day American English as well as in British and Australian English (Smith 2003, 248ff., Loureiro-Porto 2009, 8).
guistically in the various configurations they appear in. In addition, positing such lexical ambiguity is not an explanatory move, and it presents challenges for learnability (a classic subset problem, since the modal meanings allowed in one configuration are a subset of those allowed in another).

One unfulfilled expectation of a blind lexical ambiguity account is that there would be languages with versatile modals in which the pairing between meaning and form is exactly the opposite of what we found for English need. That is, there should be languages with modals that express a larger variety of (say, root) modalities when they take nominal complements, and a subset of these modalities when complemented by an infinitival. An initial survey I have conducted on versatile necessity modals in other languages (Hindi-Urdu, Indic; Hebrew, Semitic; Czech, Slavic, see below) has instead revealed a pattern that closely resembles the pattern in English. In Hebrew, the pattern extends beyond the lexical item tsariχ ‘need’ to other versatile necessity modals. One of them, χayav ‘must, have to,’ derives from a root that typically contributes deontic meanings of debt, obligation, and requirement. Nonetheless, it is exclusively teleological in (+content/-preferences) scenarios when it takes nominal complements.

**Hindi-Urdu**  
The modal word caahiye in Hindi-Urdu subsumes the meanings of ‘want, need, should.’ It is morphologically related to the verb caah ‘want.’ -iye is an archaic participle with ‘want,’ but is productively used to make polite imperatives elsewhere in the language (e.g., aa-iye ‘come!’, a polite imperative; Rajesh Bhatt, p.c.)

There are a number of syntactic configurations in which caahiye can appear. When it takes noun phrase complements, caahiye gives rise to the dative-nominative case marking pattern. In terms of Davison’s (2004) classification of Hindi-Urdu polyvalent verbs, this modal/attitude verb belongs to Class C: the class of primarily psychological verbs that

---

14 Echoing Hacquard’s (2010) concern with respect to the epistemic-root polysemy of modal auxiliaries, “these solutions, which rely on separate, interpretation-specific entries, ultimately fail to explain why, crosslinguistically, epistemics and roots are expressed by the same lexical items” (p. 81).
mark their subjects with dative case and their direct object with nominative case (ibid., pp. 201-202). An example of an experiencer verb belonging to Class C is (132).15

(132) mujhe ek upaay suujh-aa.

I-Dat one means.M.SG.Nom see-Pfv.M.SG

‘A solution came to my mind, I saw a solution.’

(Davison, 2004, 207, quoted from Bahri, 1992)

Modal caahiye gives rise to different interpretations depending on its syntactic configuration.16 Consider the scenario in (133).

(133) Water drinking. Ram is participating in an educational game designed to encourage kids to drink water in summer instead of soft drinks. In this game, the winner of a round is required to drink a glass of water. Ram has been winning many rounds and consequently drinking lots of water. Oh, no! He just won another round . . .

a. ram-ko ek cup aur pii-naa caahiye.

Ram-Dat a cup more drink-Inf should

‘Ram needs to drink another cup.’

b. ram-ko ek cup aur caahiye.

Ram-Dat a cup more should

‘Ram needs/wants another cup.’ (deontic impossible)

Matching the judgments about their English translations, a difference in truth values emerges between the sentences in this scenario. (133a) is true under a deontic interpretation (because of the rules of the game). (133b) is judged false and an odd thing to say (because Ram will suffer from additional liquid intake and doesn’t want to drink any more).17 Thus,

15 I have slightly adapted the transcription of certain vowels in Davison’s examples to fit the convention used elsewhere in this section. Also, note that in (132), suujhaa would be better glossed as an inflected form of ‘strike’ (Rajesh Bhatt, p.c.)

16 Judgments about the Hindi-Urdu data presented below are due to Rajesh Bhatt unless noted otherwise.

17 Caahiye can also be interpreted as a desire verb meaning ‘want’ in the dative-nominative construction. To form a desire statement with a verbal complement, the verb caah ‘want’ must be used:
when *caahiye* has a (dative) subject and a (nominative) nominal complement, its modal potential is restricted in comparison to when its complement is an infinitive. (133b) can be teleological but not deontic in this configuration.

The (+content/-preferences) scenario in (134) shows the same contrast. *Caahiye*-IP in (134a) is judged true, referring to the order issued in the context. It contrasts sharply with *caahiye* in the nominal complement construction (134b), which is judged false in the same context. The latter cannot receive the deontic interpretation that makes the former true. (These judgments hold for the English translations as well.)

(134) **A day off cut short.** Ram is a school teacher whose classes are off on a field trip. He took the day off and was having a great time resting, until a phone call from the principal ruined it all a few minutes ago. Ram was told to show up at school in an hour for a meeting with the school gardener. He couldn’t care less about this meeting, but the principal explicitly ordered him to come.

a. *ram-ko maalii-se mil le-naa caahiye.*
   Ram-Dat gardner-with meet take-Inf should
   ‘Ram needs to meet with the gardener.’

b. *ram-ko maalii-ke saath meeting caahiye.*
   Ram-Dat gardner-Gen with meeting should
   ‘Ram wants/needs a meeting with the gardener.’

**Hebrew** Contemporary Israeli Hebrew provides further insights concerning grammatical constraints on the interpretation of modals. It has three versatile necessity modals (*tsariχ* ‘need,’ *muχraχ* ‘obliged, must,’ and *χayav* ‘must, have to’), and we will see that all three

(i) *ram ek cup aur pii-naa caah-taa hai*
   Ram a cup more drink-Inf want-Hab is
   ‘Ram wants to drink another cup.’
are restricted to teleological modality with nominal complements.\textsuperscript{18} The inability of \textit{χαυαυ} to express deontic modality with noun phrase complements is particularly interesting. The \textit{χαυαυ}-DP construction is a recent innovation in Hebrew, and one that is considered colloquial speech (Choueka et al. 1997, s.v. \textit{χαυαυ}, adj., sense 6). Given that the modal is homomorphomic with an adjective that means ‘owe,’ one might have expected that deontic modality would be its default interpretation. Instead, it emerges as a teleological necessity modal in the nominal complement construction.

Semantic versatility of \textit{ţαριχ} ‘need’ is demonstrated by the full range of modality types expressed when the modal takes verbal complements. The need expressed in (135a) is oriented to the goal of winning a soccer match (a teleological interpretation). (135b) is a deontic example: the necessity follows from what is decreed in a legal document (the road map). Finally, (135c) is an example of circumstantial or logical modality: the prejacent describes a conclusion necessitated by the application of a mathematical formula (calculating the average height of a population).

\begin{verbatim}
(135) a. le-da‘at-i kol eχad me-ha-saχkan-im tsariχ latet ‘et misχak to-mind-1.SG every one of-the-player-PL need give.Inf Acc game.CS χaya-v be-yom ḫabat biforn ḫe-netse ‘im nitsaχon life-3.M.SG in-day Saturday so that-leave.Fut.1.PL with victory mi-tedi
from-Tedi

‘[...] I think that each player needs to play the game of his life on Saturday so that we leave Tedi Stadium victorious.’\textsuperscript{19}
\end{verbatim}

\textsuperscript{18} Other necessity modals in the language are not versatile according to my criteria (either syntactically or semantically, see p. 138). Ambar (1989) provides a detailed overview of Hebrew modal expressions, and see Sima’an et al. (2001) and Netzer et al. (2007) for categorizations of Hebrew modals from a computational point of view.

\textsuperscript{19} Forum for soccer fans, \texttt{4everblue.co.il}, November 2006. Last accessed on September 22, 2009.
b. lefi mapa-t ha-dray-im ha yi-ta tsariχ la’asot b-a-faza according to map-CS the-road-PL be.Past-2.M.SG need do.Inf in-the-phase ha-riʃon-a ’et ha-davar ha-ba the-first-F.SG Acc the-thing the-next ‘According to the road map, you should have done the next thing in the first phase.’

c. ze ’omer fe-ha-memutsa be-dor ha-tʃe’ir-im tsariχ lihyot 1.80 this says that-the-average in-generation.CS the-young-PL need be.Inf 1.80 ‘This means that the average height in the younger generation must be 1.80 (meters).’

Nominal complements are grammatical with tsariχ (as well as with muχraχ and χayav), but taking such complements limits the adjective’s modal potential. The (+content/-preferences) scenario of water drinking shows that in Hebrew too, verbal complements allow the modal to receive a deontic interpretation (136a). (136b) is judged false, since Sam’s wellbeing is a priority that is presupposed to be collectively committed, taking precedence over a priority like winning the game. A nominal complement thus forces a teleological interpretation of the modal.

(136) [Drinking too much water, as in (133).]

a. sem tsariχ/χayav liʃtot od kos.

Sam need/must drink.Inf another cup

‘Sam must drink another cup.’

---


22 Judgments about Hebrew data that are reported in this section are my own.

23 This is the case for the adjective tsariχ as well as the verbal modal yiʃtareχ ‘need’ that is derived from the same root (yiʃtareχ is the future form of the root ts. r.χ √NEED conjugated in the hitpael binyan).
Sam needs another cup.

The same pattern emerges in sentences that do not have overt subjects, showing that the overt presence of a subject does not determine tsariχ’s modal potential (this point is taken up in Section 4.5.2). Recall the fences scenario described in (129). This scenario is “+content,” because there is a law that mandates a certain action (putting up a fence). It is also a “-preferences” scenario, since the action is dispreferred: it goes against the presupposed commitments of the participants in the conversation. When the modal takes a verbal complement, as in (137a), it can express the deontic interpretation just described, but in (137b) it cannot. The second sentence is judged false in the fences scenario. (It should be noted, however, that the complement to the modal is a small clause in (137b), not a noun phrase. Small Clauses have been proposed to be the underlying structure of intensional transitive need by several researchers, as I discuss in Section 4.5.1.)

(137) [The fences scenario of (129).]

a. tsariχ/χayav-im livnot kan gader.
   need.M.SG/must-M.PL build.Inf here fence
   ‘A fence needs to be built here.’

b. tsariχ/χayav-im kan gader.
   need.M.SG/must-M.PL here fence
   ‘We need a fence here.’

A variant of (137a) with an existential prejacent, (138), shows that ‘need’ can also receive an epistemic interpretation when it is in construction with an infinitival. For example, (138) can be used to communicate the results of a survey about fence locations in town. (137b) cannot be used in this way. Note that in contrast to (137a), a modal with an exis-
tential prejacent agrees in number and gender with the single noun phrase in the existential prejacent (the noun ‘fence’ is grammatically feminine).24

(138) \textit{tsr}a\textit{γ}-a/\textit{hayev}-et \textit{lihyot kan gader}.  
need-F.SG/must-F.SG be.Inf here fence(F.SG)  
‘There need to be fences here.’

“Weather predicates” are useful for testing how a versatile modal like ‘need’ expresses epistemic necessities. The drought scenario in (139) invites discussion of both teleological and epistemic possibilities. However, only a verbal complement allows the epistemic, forecast-based interpretation to emerge. (139b) has a reading similar to that of \textit{It will probably rain} in English, where the modal reports on what the weather is expected to be like.25 The absence of this interpretation is truly striking in (139a) (and in the corresponding example where the modal’s complement is a future-marked clause; see Section 4.4). The modal in (139a) can only be interpreted teleologically: the sentence is judged true because in order for the drought to end, there must be rain.

(139) \textbf{Drought}. After a prolonged drought, there are finally forecasts of rain.

\begin{itemize}
  \item a. \textit{tsariγ gefem}. 
  
  need rain 
  ‘Rain is needed.’
\end{itemize}

\footnote{Agreement with a plural form of the noun is shown in (i). Agreement is obligatory, as evidenced by the ungrammaticality of the non-agreeing or partially agreeing forms in (ii). See Section 4.5.2 for additional discussion of the non-inflecting form of ‘need’ in (137a).}

\begin{itemize}
  \item (i) \textit{tsr}a\textit{γ}-ot/\textit{hayev}-ot \textit{lihyot kan gder-ot}.  
  
  need-F.PL/must-F.PL be.Inf here fence-F.PL  
  ‘There need to be fences here.’
\end{itemize}

\begin{itemize}
  \item (ii) *\textit{tsr}a\textit{γ}-a/*\textit{hayev}-et/*\textit{tsariγ}/*\textit{hayev}-im \textit{lihyot kan gder-ot}.  
  
  need-F.SG/must-F.SG/need.M.SG/must-M.PL be.Inf here fence-F.PL
\end{itemize}

\footnote{Note that the more similar word-by-word counterpart of the Hebrew, \textit{It needs to rain}, seems incompatible with such an epistemic forecast-based reading. This points to there being subtle differences in the restrictions placed by English \textit{need} and Hebrew \textit{tsariγ} on their domains of quantification.}
Czech In Czech, the modal *muset* ‘must’ is a versatile modal that takes both infinitival and Prepositional Phrase (PP) complements (marginally also nominal complements, see below). It allows both epistemic and root interpretations in its infinitival construction, as shown in (140). Like other modals in the language, *muset* ‘must’ is ungrammatical with clausal complements (Kyncl, 2008, 122).

(140) Jan musí být v knihovně.
     Jan must.3.SG.Pres be.Inf in library
     ‘It must be true that Jan is in the library.’ (epistemic)
     ‘Jan is obliged to be in the library.’ (deontic)
     (Kyncl, 2008, 260)

Example (141), from Kyncl (2008), is suggestive that the modal *muset* resists deontic interpretations with a nominal complement (‘hamburger’ below). Kyncl glosses the meaning of the modal as ‘liking’ and notes explicitly that a (deontic) ‘is required to’ interpretation is not allowed.

(141) Jan musí hamburger každý den.
     Jan must.3.SG hamburger.Acc every day
     ‘Jan likes a hamburger every day.’ (‘is required to’ impossible)
     (Kyncl, 2008, 185(fn. 125))

Since the sentence is presented out of context, however, it is difficult to judge exactly what its modal potential is.\(^\text{26}\) The translation using ‘like’ also raises the question of whether

---

\(^{26}\) The Czech possibility modal *moci* ‘can’ presents further complications. Kyncl (2008) cites examples such as (i) as exemplifying a deontic interpretation of the modal in the same configuration as (141).
the modal allows a teleological, ‘need’-type, meaning in this configuration in Czech in addition to, or in place of, the ‘want’-type desire meaning that *like* has in English. What is crucial in the context of the present discussion is the fact that a deontic interpretation is not available.

Variability in modal potential was found when the modal was given PP complements instead of infinitivals. The scenario in (142) was used to test the range of interpretations that *muset* can receive in these syntactic configurations. In this scenario, two priorities are pitted against each other: a priority for a home delivery, and a priority for monetary support. Due to the legal situation described, these two priorities are in conflict, but only the first priority is presupposed to be collectively committed to in the context (we assume that the interlocutors in this conversation would both defend the mother’s preferences). What we find is that *muset* has a wider range of possible (teleological) interpretations when its complement is verbal. The sentence in (142a), with an infinitive, is true in the scenario (based on the necessity of going to the hospital for claiming the stimulus money). On the other hand, (142b) is false, because it is sensitive to the priority that Sarah give birth at home.\(^{27}\)

(142) **Birthing stimulus.** There is a law in Israel that only recognizes deliveries completed in hospitals as eligible for a governmental birthing stimulus, as opposed to deliveries completed at home. Sarah strongly prefers to give birth in the comfort of her home, and she made all the necessary arrangements to do so. An hour ago the delivery started as planned at home, and everything seems to be going smoothly.

\(^{27}\) The judgments I report are due to Radek Šimůk.
a. musí jet do nemocnice
must.3.SG.Pres go.Inf to hospital
‘She needs to get to the hospital (to receive the benefit).’

b. musí do nemocnice
must.3.SG.Pres to hospital
‘She needs (to get to) the hospital.’

This means that the modal can access certain priorities but not others when its complement is a PP, and that it is less restricted when its complement is a verbal. Note that (142b) becomes true if there is a change of circumstances, for instance, if there start being complications in the delivery and Sarah needs to reach the hospital for treatment. This accords with must-PP’s depending on the primary priorities, since staying alive in the case of complications is a priority that is assumed (implicitly) to take precedence over any subsidiary considerations of comfort.

Although these data are preliminary and inconclusive regarding the deontic/teleological divide, it is tempting to tentatively conclude that Czech must is a versatile necessity modal that has a limited modal potential when its complement is a PP (and perhaps also a noun phrase, as suggested by (141)). The nature of this restriction is the same as is found in the nominal complement configurations of versatile necessity modals in the other languages in our sample.

Kyncl (2008) concludes in light of the contrast between (140) and (141) that Czech ‘must’ is lexically ambiguous.28 I believe that the crosslinguistic emergence of goal-oriented modality as a default in the nominal complement construction does not have the mark of a lexical ambiguity, and that it calls for a principled grammatical explanation.

28 See ibid. §5.3.4, “Explanation using ‘double lexical entries’. ”
English  English itself provides some evidence that the restriction to teleological modality is not a property specific just to *need* (and *necessary*, as we will see again below). Consider the meanings contributed by the verb *require* in the two sentences in (143).

(143)  a. The dentist required her assistant to do a new filling.
       b. The dentist required a new filling.

*Require* is a verb that very naturally expresses deontic necessity, as in (143a). But in (143b), the verb has primarily a teleological interpretation, whereby getting a new filling is described as necessary in order to better the dentist’s situation. The dentist is the patient, not the one giving orders in the relevant interpretation of (143b). The teleological interpretation (143b), and only this interpretation, can be paraphrased by substituting *need* for *require* in the sentence.

Additional examples are given in (144), exhibiting the teleological interpretation of *require* in what looks like a nominal complement construction. In (144a), the priority that gives the verb its teleological flavor is expressed in an accompanying purpose clause (*to open door*). These examples cannot receive deontic interpretations. For instance, (144a) cannot be used to report on the content of a new regulation about door knobs.

(144)  a. Ergonomically Designed Door Knob of the Future **Requires** Only 1/2in Movement to Open Door.\(^{29}\)
       b. A lie **requires** a fresh coat of paint often.\(^{30}\)

### 4.4 Full CPs pattern with nominals

The versatile necessity modals of Hindi-Urdu and Hebrew are grammatical not only with noun phrase complements, but also with certain Complementizer Phrases (CPs). In

---


this section, we will see that modals in this configuration are restricted in terms of their modal potential in exactly the same way that they are when complemented by a DP. English facts about necessary follow the same pattern. Thus, clausal complements pattern as nominals in forcing a teleological interpretation of a modal they are in construction with, showing that teleological modality is not just the modality of “small” modal complements.

**Hindi-Urdu**  
*Caahtye* ‘want, need, should’ is complemented by a subjunctive CP in the sentence in (145). The prominent interpretation of the modal/attitude verb is ‘want’ in such examples, but it is possible to show that a teleological ‘need’-type interpretation is possible as well. The scenario in (145) is useful for this purpose, since it is incompatible with a bouletic/desire interpretation of *caahiye*. The subject’s desires are described as conflicting with the need to increase the family income. The sentence is nevertheless judged true, under a teleological interpretation that is oriented toward fulfillment of the salient economic priority.

(145) **Sending daughters to work.** Ram doesn’t want his daughters to be out of the house working. He has conservative views about such things. The problem is that he needs the extra money that their work will bring in.

ram-ko caahiye [ki us-kii beTiyaaN kaam kar-eN].

Ram-Dat should that he-Gen.F daughters work do-Sbjn.3.PL

‘Ram wants that his daughters work.’ or ‘Ram needs his daughters to work.’

The above scenario contrasts desires with teleological needs, but it does not address the question of whether or not a deontic interpretation of the modal is possible in the CP complement configuration. To answer this question, we need to consider a (+content/-preferences) scenario, like (146).

---

31 I thank Rajesh Bhatt for discussion and for the judgments reported in this section.
Work required by law. Ram is a very wealthy businessman who emigrated with his family to a new country for political reasons. He doesn’t want his daughters to be out of the house working (he has conservative views about such things), and the family’s economic situation is wonderful without anyone else getting a job. But the community that accepted the family in the new country has very strict rules about employment: every adult must get a job, and this applies to Ram’s daughters as well.

Getting a job is presented as a legal obligation in this scenario (hence, “+content”), but one that runs counter to the accepted preferences of the individuals involved (hence, “-preferences”). The sentence in (145) is judged false in this scenario, the reason being that neither Ram’s desires nor his needs dictate that his daughters get a job.\(^{32}\) I take this as evidence that caahiye-CP cannot receive a deontic interpretation in the context of (146).

For completeness, is should be noted that not every subjunctive CP complement correlates with a teleological interpretation of the modal (Rajesh Bhatt, p.c.). If the embedded subject is nominative, it has to be different from the matrix subject for teleological modality to be imposed (as in (145)). Otherwise, when the subject of the embedded subjunctive CP is marked with nominative case and is also co-indexed with the matrix subject, caahiye contributes the meaning of a necessity modal like ‘should.’ We see this in (147), which should be evaluated in the context of the day-off scenario of (134). Caahiye-CP can receive a deontic interpretation in this scenario: (147a) is judged to be true, and is comparable to the infinitival construction in (147b). This particular CP complement construction notably diverges from the DP complement construction exemplified by (134b).

\[(147)\] [A day off cut short, see (134) above.]

\[a. \text{ram-ko (yeh) caahiye [ki vo waqt-pe school pahunc jaa-e].} \]

\[
\text{Ram}_{1-\text{Dat this should that he, time-on school reach go-Sbjn.3.SG}}
\]

\(^{32}\) “The daughters are required to work by law but this is not Ram’s need/desire” (speaker’s comment).
‘Ram should reach school on time.’

b. ram-ko waqt-pe school pahunc jaa-naa caahiye.

Ram-Dat time-on school reach go-Inf should
‘Ram should reach school on time.’

Finally, subjunctive CP complements with dative subjects pattern like (145). In this configuration, exemplified in (148), *caahiye* is restricted exclusively to teleological modality. (148) is true on a ‘need’ (or ‘want’) interpretation in the depressed artist scenario: the sentence is true from the perspective of cheering Ram up, although even his best of friends would admit that Ram doesn’t deserve a medal and shouldn’t get one. The same sentence is false in a scenario like (149), which only supports the existence of a deontic necessity, not a teleological or a bouletic one.

(148) **Depressed artist.** Ram is a so-so artist who doesn’t deserve a medal for his artistic achievements. He has been depressed recently and getting a medal is sure to cheer him up. A close friend of Ram’s says the following to a mutual acquaintance, out of concern:

ram-ko caahiye [ki us-ko medal mil-e].

Ram-Dat should that he-Dat medal receive-Sbjn.3.SG
‘Ram wants to get a medal.’ or ‘Ram needs a medal.’

(149) **Arrogant artist.** Ram is a highly acclaimed award winning artist. He is so tired of winning prizes for his work, in fact, that he has developed an aversion to any form of public appreciation. Last week he was notified that he won the prestigious Padma Vibhushan medallion. His mother made it clear that she won’t forgive him if he doesn’t accept this particular prize. She told a friend: “Out of respect for me, *Ram-ko caahiye ki us-ko medal mil-e!’*”

Accounting for the full range of the semantic and morphosyntactic facts we find in Hindi–Urdu is beyond the scope of this chapter. The constructions I will focus on are those
in which the versatile modal is complemented by a CP (like in (145) and (148)) that is in a sense “different enough” from a simple infinitival.

**Hebrew** Hebrew modals *tsariχ* ‘need,’ *muχraχ* ‘obliged, must,’ and *χayav* ‘must, have to’ are grammatical with verbal and nominal complements, as we saw in Section 4.3. A third type of complement they allow is a full clausal complement, specifically, a Complementizer Phrase with obligatory future inflection.

Landau (2004) argues that the future marking in CP complements of *tsariχ* ‘need’ is actually the realization of subjunctive mood (pp. 820, 859-860). This makes the Hebrew need-CP construction and the corresponding Hindi-Urdu construction similar in their grammatical makeup.  

The interpretation of the modals in construction with a future marked CP is exclusively teleological. Thus, in the same scenarios used above to argue for a difference in the interpretation of need-DP and need-IP constructions, CP complements patterns with nominal DP complements and not with infinitivals. (150) is false in the water drinking scenario, (151) is false in the fences scenario, and (152) in the drought scenario is exclusively teleological and cannot be interpreted as a forecast about the likelihood of rain.

(150) [Drinking too much water, see (136) above.]

\[
\text{sem } \text{*tsariχ*/χayav } [\text{fe-yitnu } l-o \text{ liftot od kos}]. \\
\text{Sam need/must that-give.Fut.3.PL to-him drink.Inf another cup}
\]

‘Sam needs that he be given another cup to drink.’

(151) [Fences, see (137) above.]

\[
\text{*tsariχ*/χayav-im } [\text{fe-tihye kan gader}]. \\
\text{need.M.SG/must-M.PL that-be.Fut.F.SG here fence}
\]

---

33 ‘Need’ and related predicates select the subjunctive mood in Spanish as well (Villalta 2008, 519, and see Chapter 3).
‘A fence is needed here.’

(152) [Drought, see (139) above.]

ţariχ [je-yered  gefem].
need that-fall.Fut rain

‘Rain is needed.’

**English**  In Chapter 3, I documented the teleological interpretation of the modal adjective *necessary* when it is complemented by (tenseless) *that*-clauses and *(for)*-*to*-infinitives (both CPs). Integrating these facts into the present discussion, I will show that the modal in these nominal configurations contrasts has a more limited modal potential than when it is complemented by a *to*-infinitive. Intuitions are not crystal clear, but they point in the expected direction: a CP complement forces a teleological interpretation of *necessary* more strongly than a *to*-infinitive does.

Consider the trial scenario is (153). This is a (+content/-preferences) type scenario: the law requires deporting a certain witness, but the witness is absolutely critical for putting the villain behind bars. (153a) with a CP complement is false in this scenario. In contrast, (153b) can, marginally, receive a deontic interpretation and be judged true.34

---

34 *Essential* and *crucial* are less flexible than *necessary* in this respect. (ia) is false just like (ib) in the trial scenario, contrasting with deontic *need to* in (ic).

(i) a. **False** It is essential (/crucial) to deport the only remaining witness.
   b. **False** It is essential that the only remaining witness be deported.
   c. **True** The only remaining witness needs to be deported.

I thus disagree with Van linden et al.’s (2008) claim that *essential* can express deontic modality (ibid. §2.3). Perhaps it is not a coincidence that the (teleological) example these authors use to motivate their conclusion that this interpretation is possible is one in which the modal’s complement is a *to*-infinitive.

(ii) The Anglo-Catholics consider it *essential* to be ordained by bishops receiving their appointment in regular succession from the apostles. (OED 1842)
Mobster trial. In a big trial against the head mobster, all the witnesses on the side of the prosecution have slowly been lost. Some mysteriously disappeared. Others fled the country. There is now only one remaining witness that can testify. But his status is highly problematic since his visa expires in 24 hours. The law requires that he be deported if he stays in the country after that time. The prosecutor summarizes the situation to her assistant:

a. It is necessary that the only remaining witness be deported.

b. It is necessary to deport the only remaining witness.

As with need, AcI complements do not pattern with to-infinitives. Rather, like that-clause complements, their modal potential is limited to teleological interpretations.35

Layoffs. A new round of layoffs is taking place at a certain company, with employees being called individually and discreetly into the boss’s office to receive the bad news. Somehow, a secretary hears that her friend Kate is next on the boss’s list. She knows Kate will be devastated when she gets the call. The secretary calls a mutual friend immediately with the news. She says, “You won’t believe what’s happening here, . . .

a. False It is necessary for Kate to go in now.

b. True Kate must go in now.

c. False It is necessary that Kate go in now.”

The layoffs scenario was designed to show that must and necessary are different in terms of their modal potential, and that the conclusion suggested by Palmer’s (2001) discussion of necessary – namely, that the modal can express deontic modality in the configuration in (154a) – is incorrect.

(i) Kate must come in now. ≈ It is necessary for Kate to come in now.
(Palmer, 2001, 7)
4.5 Descriptive generalizations: subject orientation or complement types?

How should the generalizations about “deontic-friendly” and “deontic-averse” configurations of versatile modals be stated? In the languages we looked at in sections 4.3 and 4.4, infinitival complements were found in the “deontic-friendly” configurations, while nominal (DP and CP) complements were found in “deontic-averse” configurations. Thus, descriptively, the ability of versatile necessity modals to express deontic modality seems to correlate with their complement being a verbal infinitival clause. The goal of this section is to try and tease apart two factors that may be relevant for the expression of priority modality, thereby settling on a clearer descriptive generalization of the data.

I start with a review of the syntax of need in English (Section 4.5.1). While need to is a raising predicate (Harves 2008), need with noun phrase complements and complements that are (for)-to-infinitives, is not. In terms of syntactic types, (for)-to-infinitives are analyzed as CPs (Pesetsky 1991, following Bresnan 1972). Thus, there are two factors that must be considered when describing the modal potential of need: the syntactic type of the modal’s complement, and whether or not the modal is thematically related to its subject.

These two factors can be teased apart on syntactic as well as semantic grounds (Section 4.5.2). The first argument comes from subjectless examples of the nominal complement construction of ‘need’ and ‘must’ in Hebrew. The modals are subject to the constraint on modal potential in this configuration despite not being thematically related to a subject. This fact argues against tying the teleological interpretation of versatile modals to the presence of an overt subject. Second, it can be shown that even when a modal is thematically related to an overt subject, it does not depend on the subject as a source for its modal backgrounds, in the sense that the relevant priorities the modal is sensitive to need not be “priorities of the subject.”

Section 4.5.3 concludes with a statement of the constraint on the expression of deontic modality based on the type of complement in construction with a versatile modal.
4.5.1 The syntax of need

What do we know about the syntactic properties of the constructions need appears in? One thing that is clear is that the facts are complex. Standard diagnostics suggest a split in the syntactic properties of constructions with need, depending on the particular type of nonfinite complement the modal combines with. The relevant contrast is between to-infinitives and (for)-to-infinitives, and the key distinction (or a correlate thereof) is whether or not phrases from the nonfinite clause can raise into the matrix clause. With to-infinitives, there is evidence that the matrix subject originates from within the infinitive. In contrast, in construction with (for)-to-infinitives, the matrix clause does not contain phrases of foreign origin. Finally, it is hard to tell these two options apart in the nominal complement construction, and indeed both options have been proposed.

Harves (2008) provides three arguments that need behaves like a raising verb when its complement is a to-infinitive, contrasting in this respect with the desire verb want (p. 213).\footnote{Cormack and Smith (2002, 141) also state that need+to is a raising verb (along with have+to), but they do not provide arguments for this position. Brennan (1993) identifies as raising predicates those semi-modals that appear with expletive subjects and tensed clausal complements (e.g., It has to be true/the case that asbestos is dangerous) in addition to taking verbal complements and having lexical subjects (e.g., Asbestos has to be dangerous). She further assumes that semi-modals that only appear in the latter type of configuration are control predicates (p. 88, footnote 20). This criterion identifies need to as a raising predicate.} First, need with a verbal complement can appear with expletive subjects (155d), patterning with the canonical raising verb seem (155b) and contrasting with want (155f).

(155)
\begin{itemize}
  \item a. An adult seems to be in every car.
  \item b. There seems to be an adult in every car.
  \item c. An adult needs to be in every car.
  \item d. There needs to be an adult in every car.
  \item e. An adult wants to be in every car.
  \item f. *There wants to be an adult in every car.
\end{itemize}

(Harves, 2008, 213)
Second, *need to* behaves like the raising predicate *seem* when the verb in its embedded clause is passivized. In comparing *need* and *want*, only *need* maintains the interpretation it has in (156c) under passivization (156d).

(156)  
  a. The doctor seemed to examine every patient.  
  b. Every patient seemed to be examined by the doctor.  
  c. The doctor needs/wants to examine every patient.  
  d. Every patient needs/wants to be examined by the doctor. (*Meaning shift for want*)  

(Harves, 2008, 213)

A third argument for the raising properties of *need to* comes from semantic considerations of scope. It has been observed that quantificational subjects interact with the quantificational component of a raising predicate in a way that suggests that they originate in a non-surface, lower position in the tree. Parallel interpretations are not found in control constructions. In the case of control, a quantificational subject may only take surface, wide scope with respect to the matrix predicate (May 1977, 1985, and see Wurmbrand 1999 with respect to raising properties of modal auxiliaries). (157) exemplifies this contrast between the raising predicate *likely* and the control verbs *try* and *promise*.

(157)  
  a. Someone from New York is likely to win in the lottery.  
     (*Someone > likely, likely > Someone*)  
  b. Someone from New York tried/promised to win in the lottery.  
     (*Someone > tried/promised, *tried/*promised > Someone*)  

(Wurmbrand, 1999, 606)

---

37 On the “passive = active” test for raising predicates, see e.g., Bresnan (1972, 118ff.), Postal (1974).
Harves (2008) observes that *need* (but not *want*) behaves like a raising predicate on this test.\(^{38}\)

(158) Someone needs/wants to be home when the kids get back.

a. *need/*want > Someone

b. Someone > *need/want

(Harves, 2008, 213)

The judgments in (158) diagnose a difference between *need* and *want*, although it is not clear that what is diagnosed is a *syntax of raising* versus a *syntax of control*. First, one could imagine syntactic operations that would create the crucial scope order (“*need > Someone*”) from an underlying control structure, simply by moving the main predicate above the base position of the subject (Wurmbrand, 1999, 608). In addition, it is not clear that failure to show narrow scope readings of the subject in fact implies control in the case of *want*, *try* and other predicates. Despite the behavior of *try* in (157b), for example, recent research has challenged the traditional classification of this verb as a control predicate, suggesting instead that it is an aspectual head that instantiates a raising structure (Cinque, 2006).\(^{39}\) For evidence from Italian that ‘want’ itself can instantiate a monoclausal (more “raising like”) structure, see Hacquard (2006, 2008).

It is clear, in summary, that despite the conceptual similarity between goals and desires (i.e., what we need to and want to do), *need* and *want* behave differently syntactically when complemented by *to*-infinitives. Only *need* exhibits properties of a raising verb in this configuration.\(^{40}\) A diachronic perspective provides a final bit of supporting evidence

---

\(^{38}\) The infinitive can also contain an eventive predicate (e.g., *Someone needs to cook dinner*, interpreted as expressing a need that someone – but not anyone in particular – make dinner). This seems problematic for the view that eventive, future oriented infinitives correlate with a syntax of control (Stowell 1982, Bošković 1997, Martin 1996). For discussion of the temporal approach to raising versus control, see Wurmbrand (2006), Moulton (2009), and references cited therein.

\(^{39}\) See Grano (2011) for a proposal about what the updated semantics of ‘try’ might look like under these assumptions.

\(^{40}\) This conclusion contradicts certain statements in the literature, for instance:
for this conclusion. Higgins (1990) notes in passing that need may belong to a small class of historically object control verbs in English that have developed into raising verbs. Higgins’s main claim is about the verb happen, and seem, likely, and need are mentioned as potentially belonging to the class as well (ibid., p. 23).

To make things more interesting, need and want do pattern alike in the Accusativus cum Infinitivo (AcI) construction we encountered in Section 4.2. From the work of Bresnan (1972), we know that the superficial similarity between (159a) and (159b) does not entail that their syntactic structures are the same. When we run through Bresnan’s diagnostics (ibid., pp. 153ff.), we find that need patterns with want and not with believe in sentences with AcI complements.41

(159)  
a. We need/want him to be in Amherst.  
b. We believe him to be in Amherst.

First, the clause that is the complement of need can begin with for (seen clearly when this unit is displaced, as in (160b)).42 Second, the subject of the infinitive cannot be promoted to the subject position in the matrix clause when need is passivized. (Note that the ungrammatical sentences in (161) become grammatical if the embedding verb is believe.)

[. . . ] need behaves like a true (necessity) modal when it selects a bare VP complement, having only an uninflected nonfinite form, though it also occurs as a regular control verb taking a to-infinitive complement.
(Stowell, 2004, 621)

41 Bresnan cites Aijmer (1972) as observing many of the same facts.
42 Pesetsky (1991) observes that need, unlike want, seems to allow for outright. In contrast, want is odd with an overt for unless the verb is not in an AcI environment (ibid., p. 49):

(i) Bill wants (for) Sue to leave.
(ii) a. Bill wants very much *(for) Sue to leave.
    b. *(For) Sue to leave is wanted by all of us.
Third, the subject of need’s verbal complement may not be a reflexive pronoun ((162), again contrasting with believe, e.g., The presenters believe themselves to be speaking clearly).

(160)  
a. He needs for his generalization to be robust.  
b. What he needs is for his generalization to be robust.  

(161)  
a. *The generalization is{needed/wanted/desired/preferred . . .} to be robust.  
b. *There is {needed/wanted/desired/preferred . . .} to be a robust generalization.  

(162) *The presenters need themselves to speak clearly.

A fourth difference Bresnan (1972) mentions is the possibility of “Equi-NP deletion” with want type AcI, but not with believe type AcI. Although both need and want take simple to-infinitives as complements (in addition to AcI complements), we have already seen that need behaves like a raising verb in this configuration, unlike want which behaves like a control verb (Harves, 2008).

Bresnan (1972) argues that the properties of the two types of AcI reflect two distinct syntactic structures. Want type verbs, she proposes, take clausal complements that include a complementizer (pronounced as for when overt), whereas believe type verbs take complements which are smaller sentential phrases (“the bare S,” p. 170). Pesetsky (1991) accepts the view that overt for is a complementizer, and further argues that for has a null counterpart, 0 for, in sentences like (159a).43

The mixed behavior of need poses a challenge to the neat classification of modal and attitude verbs as belonging to the “believe type” or the “want type.” Need behaves like want in its AcI configuration, but unlike want when the infinitival it combines with is a subjectless string. A summary of the syntactic facts reviewed here suggests a non-uniform analysis, as in Table 4.4.

43 Ibid., pp. 48ff. and §3.2.10. For a recent analysis of the syntax and semantics of believe type AcI, see Moulton (2009, 182ff.).
I. Abe: needs [IP t_i to leave ] \textit{IP complement with raising}

II. Abe needs [CP for/\emptyset_{for} Sue to leave ] \textit{CP complement}

Table 4.4: Syntactic structures for \textit{need} with nonfinite complements.

Of course, \textit{need} also takes simple nominal complements. A sizable portion of the literature on intensional transitive verbs deals with the syntactic and semantic properties of the need-DP construction, but consensus regarding its analysis has not yet been reached.\(^{44}\)

The more recent studies argue that \textit{need’s} complement is not a full concealed clause in this configuration (contra McCawley 1974 and Larson et al. 1997). The two alternative structures they propose are shown in (163). The first option is that \textit{need} takes a small clause (SC) complement with some covert material (Harley 2004, Schwarz 2006 for a subset of cases, Harves 2008).\(^{45}\) The second option, due to Schwarz (2006), is that the modal takes a simple noun phrase complement in certain cases, and the semantics is burdened with the task of deriving a proposition out of this nominal complement.

(163) a. Abe: needs [SC t_i/PRO_i HAVE/GET the car ].

b. Abe needs [\exists the car ].

A small clause analysis of the need-DP construction, with raising, fits with the verb’s raising behavior when complemented by a \textit{to}-infinitive (case I in Table 4.4). A control syntax and a simple syntax as in (163b), both characterized by \textit{need} being thematically related to a matrix subject, fit with the diagnosis that the modal takes a “closed” clausal complement in the AcI construction (case II in the table).


\(^{45}\) Note that the proposals mentioned differ in details relating to the decomposition of the embedded predicates in the small clause (see Harley 2004) as well as the composition of verbal \textit{need} itself (see Harves 2008, Harves and Kayne 2012). I gloss over these details here.
In principle, the two analyses make different predictions about the ability of the subject to take scope under the modal in sentences like *Someone needs the car*. The analysis in (163b), but not those in (163a), directly predicts that a subject quantifier would obligatorily take scope over the modal. And in fact, (164) shows that surface scope is the only option in the need-DP construction, unlike what we saw for *need to* in (158) above.

(164) Someone needs/wants the car.
   a. *need/*want > Someone
   b. Someone > need/want

What makes it difficult to distinguish between these options, however, is the very assumption that the embedded clause in the analysis in (163a) is a small clause. Williams (1983) observed that subjects of small clauses contrast with subjects of bigger *to*-infinitives in their ability to scope under an embedding predicate. This is shown with respect to *seem* in (165) (an example from Johnson and Tomioka 1998).46 In recent work, Moulton (2012, forthcoming) proposes that these facts be given an analysis that does not rule out raising from within the SC. (See also Lechner 2007 for discussion in the context of *need*.)

(165) a. Someone seems to be angry at the Federal Government.
    (wide scope and embedded scope for *someone*)
   b. Someone seems angry at the Federal Government.
    (only wide scope for *someone*)
    (Johnson and Tomioka, 1998)

I do not have new arguments that will help decide the question of what the need-DP construction looks like syntactically. The strategy I will pursue in describing the variable modal potential of *need* is thus to focus on those cases where the properties of the structure are relatively uncontroversial (i.e., the need-IP and need-CP constructions). Once we have

a handle on what could be behind the restriction in those cases, it should be possible to go back and evaluate the different analyses of the need-DP construction.

4.5.2 Subjectless constructions and the source of priorities

The need-IP construction is the only “deontic-friendly” construction of modals like ‘need.’ It is characterized by the lack of a thematic relation between the modal and the sentence’s syntactic subject. The need-CP construction, in contrast, is characterized by the presence of a thematic relation between the modal and the subject, and it is crucially “deontic-averse.” These facts are consistent with the hypothesis that a thematic relation to a subject is sufficient and necessary for limiting the modal potential of versatile modals that take nominal complements.

In this section, I raise two issues for this hypothesis. The first is syntactic and is based on subjectless sentences in the need-CP (and need-DP) construction. As we saw, these sentences are limited in their modal potential just like corresponding sentences with overt subjects. Thus, in order to maintain the hypothesis that a thematic relation is responsible for the limited modal potential of priority modals, one would have to argue that the relevant thematic relation exists even when there is no overt subject. The second concern is semantic in nature. In reviewing how thematic relations to subjects have traditionally been used to explain constraints on the interpretation of modals, I will argue that “subject orientation” as it is usually understood cannot be the explanation of the teleological interpretation of ‘need’ in the relevant constructions. The point here is simply that the modal’s thematic subject is not always the source for the priorities that teleological need is evaluated with respect to. Thus, in order to maintain the hypothesis, a new way of thinking about subject orientation would have to be filled in. In addition, the new definition would have to distinguish between priorities of various kinds, in order to explain why deontic modality cannot be expressed when the modal is thematically related to a subject (but teleological modality can).
I begin with some background on how the notion of “subject orientation” has been applied in the literature to explain the interpretation of priority modals.

An early intuition in the literature was that root modals are thematically related to their subject, in contrast to epistemics (Hofmann 1976, Ross 1969, Perlmutter 1970, 1971, Jackendoff 1972). This thematic relation was proposed, most notably by Brennan (1993), to be responsible for the range of modality types these modals allow. The idea was that when a modal is thematically related to a subject, it is forced to be “about” the individual denoted by that subject (ibid., p. 181). To capture this idea theoretically, Brennan (1993) proposed that root modals use property-level conversational backgrounds to determine their quantification domains, and that these conversational backgrounds consist of relevant properties of the individual denoted by the modal’s subject. On this analysis, the modal backgrounds of may in (166), for example, consist of two properties of Joan’s: having registered to vote, and living in Racine. The deontic statement is true if and only if Joan votes in Racine’s mayoral elections in at least one accessible world (maximally close to an ideal) in which she has these two relevant properties (Brennan, 1993, 182-185).

(166) Joan lives in Racine and is registered to vote. She may [has the right to] vote in Racine’s mayoral elections.

(Brennan, 1993, 181)

The property-level approach that was developed in Brennan’s work was designed to account for “ought to do” deontics of the kind exemplified by (166). It is insightful, but out-of-the-box it fails to distinguish between deontic and teleological modalities. Both varieties seem to be “about” an individual in the intuitive sense, and both display grammatical properties associated with root modals, and not with epistemics.

47 In the context of historical change, see Bybee et al. (1994) on “agent oriented modality.”
More recently, responding to criticisms of the traditional approach to subject orientation and aiming to provide a unified analysis of epistemic and root modals,\(^{48}\) Hacquard (2006, 2010) has argued that the linguistic factor that is responsible for the expression of root modality is relativity to the event described by the VP embedded under the modal.

The correct generalization seems to be that root modality is centered around the event described by the VP and its participants. In most cases, the main participant is the subject, and hence properties of the subject are highlighted. In other cases, however, the location or properties of other participants of the event are more relevant[.]

(Hacquard, 2010, 92)

This characterization is proposed to apply to “ought to do” deontics, teleological modals, and other root modals. Since we are primarily concerned with the expression of teleological modality (as in the nominal complement constructions of ‘need’), let us focus on this type. In the teleological interpretation of *have to* in (167), for example, the modality would be centered around an event of Mary’s taking a train to Paris last night. The fact that Mary, the subject, had a need to go to Paris is responsible for the teleological flavor of the modality.\(^{49}\) More generally, goal-orientation seems to depend on some participant in the embedded event having the priorities that the modal claim is based on.

(167) (Last night) Mary had to take the train (to go to Paris).

(Hacquard, 2010, 110)

However, when we consider examples like (168), it becomes clear that this interpretation of event relativity is problematic. First, even when the subject is the main participant in the embedded event, it is often not the case that their own preferences line up with the relevant goal. (168a) provides such an example. Individuals other than the subject (here,

\(^{48}\) Challenges for the traditional analysis of root modals as two-place (control) predicates were raised by Bhatt (1998), Hackl (1998), and Wurmbrand (1999).

\(^{49}\) Since the event that anchors the modality may be non-actual, one has to worry how it is identified in the general case. When the modal statement is generic, Hacquard (2010) proposes that the relevant event is “a train-taking by Mary in all generic worlds” (p. 110).
the boss in charge of the construction) are responsible for defining the relevant priorities and preferences. In examples like (168b), it is the speaker whose preferences matter for the teleological claim.\textsuperscript{50} The sentences in (168c)-(168d) are examples with VP participants that clearly do not have any preferences or priorities of their own. A head of lettuce does not have a preference for keeping long after it has been picked; an orchid does not prefer to thrive (in the sense of comparing the possibility of thriving to the possibility of not thriving and preferring the former over the latter). It seems impossible, finally, to access the relevant priorities through the locations of the VP events described. These priorities are not centered around the construction site in (168a), the whereabouts of the worker in (168b), the fridge in (168c), or the location of the plant in (168d).

(168)  
\begin{enumerate}
\item In order to finish the construction ahead of schedule, the poor workers need to work 12-hour shifts.
\item [A worker’s visa is about to expire. The worker wants to leave his current boss.]
\begin{quote}
Boss: I like that worker, but he needs a new visa (in order to stay with us).
\end{quote}
\item The lettuce needs to be dry (in order for it to keep longer).
\item An orchid needs bright light and moderate humidity (in order to survive and thrive).
\end{enumerate}

This discussion highlights the need for a new understanding of subject orientation, if some notion of this kind is to provide an insight into the interpretation of priority modals.

\textsuperscript{50} Portner (2009) makes a similar point. In discussing what he classifies as a teleological counterpart of Feldman’s (1986) “ought to do”/“ought to be” split among deontics, Portner notes a difference in the interpretation of (i) and (ii) below. On one natural interpretation of (ib), the modal implicitly references the subject’s goals, Mary in this case. Example (ia) is different: it is not subject oriented when it admits a paraphrase like, ‘It is necessary, given our goal to get all the students to their destination, that at least two of them take the subway.’ The modal can take “wide scope” with respect to the subject, referencing goals of the discourse participants (you and me, for example), not the goals of any two or more students.

(i)  
\begin{enumerate}
\item At least two of the students should take the subway.
\item Mary should take the subway.
\end{enumerate}

(Portner, 2009, 186)
Whether or not the thematic subject of a versatile modal is part of what semantically (i.e., grammatically) prevents it from receiving deontic interpretations is still to be determined.

This brings us to the second worry about using subject orientation as an explanation for a grammatical split among the priority modalities. Recall that in the Hebrew nominal complement construction, versatile modals are restricted to teleological modality even when they are not related to a subject overtly. The relevant contrast is between examples like (169a), which can receive a deontic interpretation, and (169b), which cannot.

(169)  [Fences, repeated from (137) and (with elaboration) from (151).]

a. ṭariχ/χayav-im livnot kan gader.

need.M.SG/must-M.PL build.Inf here fence
‘A fence needs to be built here.’

b. ṭariχ/χayav-im se-tihye/tibane kan gader.

‘A fence is needed here.’ (only pseudo-deontic possible)

The form of the modal ‘must’ in these examples, χayav-im, has masculine plural agreement which is characteristic of predicates in impersonal subjectless constructions in Hebrew. In contrast, the form of the modal ‘need,’ ṭariχ, exhibits masculine singular agree-

The modal noun χοβα ‘obligation’ shows a similar contrast in modal potential, although the contrast is somewhat more subtle: (170b) with a CP complement seems less open to receiving a deontic interpretation than (170a) is, with an infinitival complement. The important point about this example is that the noun ‘obligation’ is never associated with a thematic subject in the language. It is primarily teleological when its complement is a CP, and it is this type of complement that correlates with an obligatory teleological interpretation of the other necessity modals as well.

(170) [In the fences scenario:]

a. χοβα lehakim kan gader.
   obligation construct.Inf here fence
   ‘A fence needs to be constructed here.’

b. χοβα fe-tukam kan gader.
   obligation that-constructed.Fut.F.SG here fence
   ‘A fence needs to be constructed here.’

51 See also Rosén (1977, 220ff.) on the χγ”m (an acronym in Hebrew for ‘lacking person-number’), and further discussion by Kuzar (1992, 1996).

52 Modals contrast with other predicates with sentential complements in that an expletive can be realized overtly only with the latter. The variant of (169b) with ‘need’ in (i) is ungrammatical, and contrasts with the example with ‘clear’ in (ii). As noted by Berman (1980, 767), the expletive is optional in the construction in (ii).

(i) *ze ṭsarצ fe-tihye kan gader.
   it need that-be.Fut.F.SG here fence

(ii) ze barur fe-tihye kan gader.
   it clear that-be.Fut.F.SG here fence
   ‘It is clear that there will be a fence here.’

53 Expletive ze is homophonous to a demonstrative pronoun and a pronominal copula. See Heller (1999, 2002), Sichel (2008), Greenberg (2008), Danon (2012) for recent work on its copular use.
Berman (1980) notes that subjectless constructions of modals differ “in meaning and pragmatic import” from the constructions in which an overt subject is realized. She characterizes the difference in terms that are reminiscent of the traditional intuition regarding subject orientation of root modals.

In the subject-modal sentences, some modality is attributed to the subjects’ relation to the predications: ‘you must’, ‘Dan can’ do something or be affected in some way by an event. In the predicate-complement constructions, the modality is attributed to an event or situation. What is being asserted is that a certain circumstance is possible (likely, mandatory etc.)
(Berman, 1980, 769)

The problem raised by the sentences in (169) is that they both exemplify the subjectless complement construction of a modal, yet the modal has a different range of possible interpretations in each case. If we were to key the modality in both cases to a situation – perhaps a situation in which the relevant rule or priority is salient or has been expressed – it would be difficult to explain why certain sub-varieties of priority modality are impossible as backgrounds only in (169b). On a theory that derives the content of ordering sources from priorities that are salient in a situation, rules as well as priorities would make their way into an ordering source if salient in context. An analysis along these lines would leave the constraint against deontic modality in the nominal complement constructions unexplained.

4.5.3 Conclusion

The generalization about the modal potential of versatile strong necessity modals will be stated with reference to syntactic categories of the modal’s complement, as follows.

(171) (The puzzle of priority modal flavors)

When a versatile strong necessity modal takes a nominal complement (a DP, a subjunctive CP, a for-CP), it can be interpreted teleologically (based on the primary priorities in the context), but it cannot be interpreted deontically (based on some salient deontic content).

With a verbal complement, the modal can receive both interpretations.
The theoretical basis for the phenomenon described in this chapter has not been attempted here and will be the topic of forthcoming work.
APPENDIX A

A KRATZERIAN CONTEXT DEPENDENT POSSIBLE WORLDS SEMANTICS FOR MODALS

Every journey in the formal semantics of modality starts with possible worlds (Kripke 1959, 1963, Lewis 1973, 1986, Kratzer 1977, 1981, 1991). For our purposes, it will not be necessary to commit to a particular position about the metaphysical reality of possible worlds. As entities in our semantic ontology, possible worlds represent all the possible ways things could be (without committing to whether or not these possibilities actually exist alongside the one world we live in; see e.g., Lewis 1973, §4.1, for the realist’s approach to possibilia). A world contains a complete record of the history, present, and future of the individuals that inhabit it. With the modal realist, we will assume that individuals are world-bound. You and I live in the actual world. When we talk about how things could have been different for some individual, we necessarily make a claim about counterparts of that individual in different possible worlds (Lewis, 1968, 1973, 1986).

Modal expressions are linguistic devices for exploring the space of possibilities. They draw attention to contextually relevant subsets of possible worlds and allow us to say something about what happens in them. Specifically, we assume that modals are quantifiers over possible worlds, lexically specified for force (as universal or existential quantifiers) and relativized to two contextual parameters called conversational backgrounds (Kratzer, 1981, 1991, 2012). Each time a modal is uttered, the conversational backgrounds determine anew which worlds it quantifies over. This setup allows for a uniform semantics for

---

1 See von Fintel and Heim (2009) for an introduction to intensional semantics and modality that includes excellent references, historical notes, and exercises. Portner (2009) is a rigorous introductory text that also offers a comprehensive survey of various theoretical approaches to the topic.
modal expressions. The fact that many modal words boast a wide range of interpretations (epistemic, deontic, teleological, and so forth) is usually explained by their dependency on contextually supplied conversational backgrounds.\(^2\)

According to Kratzer (1981, 1991, 2012), zooming in on the relevant possible worlds is a process that proceeds in two steps. The first step involves a conversational background known as a modal base. Modal bases are functions from worlds to sets of propositions. These propositions may correspond to available pieces of evidence in that world or relevant circumstances that hold in it. (Phrases like In view of such-and-such information . . . can be used to make the nature of the modal base explicit.) The second step involves a conversational background called an ordering source. Ordering sources also map worlds to sets of propositions. These propositions are of a normative nature; they encompass bodies of laws, desires, goals, and other types of ideals. As their name suggests, ordering sources are used for ordering or ranking possible worlds. The relevant subset of possible worlds for any given modal claim is thus identified by first finding all the worlds in which all the propositions in the modal base are true and subsequently ranking these worlds according to how close they come to the ideal determined by the ordering source.

The binary relation \(\ll_{g(w)}\) between possible worlds that is induced by an ordering source \(g(w)\) is defined as in (172). If we take \(g(w)\) to represent some kind of ideal in a world of evaluation \(w\) (how politicians should behave, what we would like to do over the weekend, etc.), we say that one possible world comes at least as close to the ideal as another just in case the set of propositions in \(g(w)\) that are true in the first world is a superset of the set of propositions in \(g(w)\) that are true in the second. If the superset is a proper superset, we say

---

\(^2\) Kratzer’s focus was on modal expressions in German and English. In recent years, researchers have uncovered languages with modals that do not vary in the type of modality they express (van der Auwera et al. 2005, Matthewson et al. 2006). German-type languages also have modal words that are restricted lexically to particular modality types (see Kratzer 1981, §6). See Nauze (2008, §2.7.5) for a typological summary.
that the first world is better than the second according to the ordering source. Let’s write $\triangleleft_{g(w)}$ to indicate the relation “better in terms of an ordering” between possible worlds.\footnote{For all worlds $v$ and $z$ in $W$ and a set of propositions $A$: $v \triangleleft_A z$ if and only if $v \triangleleft_A z$ and it is not the case that $z \triangleleft_A v$. The notation I use here differs from that in Kratzer (1981, 1991), but the content of the definitions is the same.}

(172) For all worlds $v$ and $z$ in $W$ and a set of propositions $A$:

$$v \triangleleft_A z \iff \{ p \in A : p(z) \} \subseteq \{ p \in A : p(v) \}.$$  

(Kratzer, 1981, 298)

Let’s step through an example to illustrate the working of this definition. Suppose we are on a roadtrip and we need to decide where to go next. Our agenda in terms of tourist attractions and general preferences is given in (173). Suppose further that given our current location and the time we would spend in each place, we cannot visit both Amherst and Hartford today.

(173) • Visit the Emily Dickinson Homestead in Amherst.
• Visit the Mark Twain House in Hartford.
• Be positive and in a good mood.

The worlds we are concerned with are those in which relevant facts about our trip hold: we are at our present location, the distance to Amherst and Hartford is what it is, the opening hours of the Homestead and the House are as they are, and so forth. There are eight relevant (classes of) possible worlds to consider. These include worlds in which we happily go to Amherst and visit the Homestead ($w_1$ represents these), worlds in which we get to Amherst in a bad mood but manage to visit the Homestead ($w_2$), and worlds in which we have a cheery positive drive but get to Amherst after the Homestead has closed ($w_3$). There are also worlds in which the worst of the worst happens: we have a miserable drive to Amherst and arrive there after the Homestead has closed ($w_4$). Similar characterizations apply to the worlds in which we go to Hartford (these are represented by worlds $w_5$ through
The preferences in (173) correspond to the following propositions (as sets of possible worlds).

\begin{enumerate}
\item \(p_1 = \{w \in W \mid \text{we visit the Dickinson Homestead in } w\} = \{w_1, w_2\}\).
\item \(p_2 = \{w \in W \mid \text{we visit the Twain House in } w\} = \{w_5, w_6\}\).
\item \(p_3 = \{w \in W \mid \text{we are in a good mood in } w\} = \{w_1, w_3, w_5, w_7\}\).
\end{enumerate}

Now, which ordering source(s) represent the preferences that we have in this scenario? Premise Semantics does not determine an answer to this question. The ordering source could be the set consisting of the three basic propositions in (174), but it could also contain propositions derived by conjoining and disjoining these basic propositions (among other options; see Kratzer 1979, 1981, 1989, Lewis 1981, and also Kratzer 2012 for discussion). To illustrate some of the issues involved in this choice, let’s see what happens if we assume that the ordering source is simply the set of propositions \(\{p_1, p_2, p_3\}\). Call this ordering source \(g_1(w)\).

It follows from (172) that \(w_1 \triangleleft_{g_1(w)} w_2\), since \(w_1\) is a member of all the propositions that \(w_2\) is a member of \((p_1)\), and one more \((p_3)\). Similarly, \(w_1 \triangleleft_{g_1(w)} w_3\). Note that \(w_2\) and \(w_3\) themselves are incomparable according to the ordering. (There is no subset relation between the sets of propositions they are members of, namely \(\{p_2\}\) for \(w_2\) and \(\{p_3\}\) for \(w_3\).)

Formally, the binary relations induced by ordering sources have the properties of a preorder: a transitive and reflexive relation which is not necessarily antisymmetric and not necessarily complete. This means that there may be worlds in the modal base that tie with respect to how close they come to an ideal (e.g., for \(w_3\) and \(w_7\) in our example, \(w_3 \triangleleft_{g_1(w)} w_7\) and \(w_7 \triangleleft_{g_1(w)} w_3\)). There may also be worlds that are not comparable at all. (For instance, \(w_2\) and \(w_3\) are incomparable because neither stands in the \(\triangleleft_{g_1(w)}\) relation to the other.)

Figure A.1 shows part of the ordering induced by \(g_1(w)\) in the roadtrip scenario. An arrow in the graph connects two worlds \(u\) and \(v\) if and only if \(u \triangleleft_{g_1(w)} v\) (but note that arrows
that can be inferred from transitivity have been suppressed). Bidirectional arrows designate ties.

![Diagram with nodes labeled w1 through w8]

Figure A.1: Ordering worlds with a single unstructured ordering source. Depicted is the ordering induced on w1 through w8 by the propositions in (174) (with pairs that can be inferred from transitivity suppressed). Amherst-worlds are colored red, Hartford-worlds are colored blue.

Given the preferences in (173), the sentence in (175) is intuitively false.

(175) In light of our preferences, we have to go to Amherst today.

Kratzer’s definition of necessity relativized to a modal base and an ordering source (what she calls human necessity) correctly predicts the sentence to be false.

(176) A proposition p is a human necessity in a world w with respect to a modal base f and an ordering source g iff

\[
\forall u \in \bigcap f(w). \exists v \in \bigcap f(w). (v \mathrel{\triangleleft_g(w)} u \land (\forall z \in \bigcap f(w). z \mathrel{\triangleleft_g(w)} v \rightarrow z \in p)).
\]

(Kratzer, 1981, 298, notation adapted)

Sentence (175) comes out false according to (176), since there are Hartford-going worlds that are unordered with respect to the Amherst-going worlds. In particular, no Amherst-going world in \(\bigcap f(w)\) is at least as good in terms of the ordering as \(w_5\), so \(w_5\) prevents the condition in (176) from being satisfied.

Now suppose we add a new preference to our list in (173), a preference which makes the Amherst-worlds more attractive than the Hartford-worlds. Perhaps we want to use as little
gas as possible, and Amherst happens to be closer to our current location than Hartford. If we add this preference to $g_1(w)$, the result would be a new ordering source (call it $g_2(w)$) that induces the ordering shown in Figure A.2. With $p_4$ added to the ordering, the tie between $w_4$ and $w_8$ is broken such that $w_4 \lessdot_{g_2(w)} w_8$. If going to Amherst means saving gas, then going there is better than going to Hartford, even if the trip is unfruitful and annoying. Similarly, $w_3 \lessdot_{g_2(w)} w_7$ according to $g_2(w)$.

(177) 
- $p_1 = \{w \in W| \text{ we visit the Dickinson Homestead in } w\} = \{w_1, w_2\}$.
- $p_2 = \{w \in W| \text{ we visit the Twain House in } w\} = \{w_5, w_6\}$.
- $p_3 = \{w \in W| \text{ we are in a good mood in } w\} = \{w_1, w_3, w_5, w_7\}$.
- $p_4 = \{w \in W| \text{ we use little gas in } w\} = \{w_1, w_2, w_3, w_4\}$.

![Figure A.2: Ordering worlds with a single unstructured ordering source (continued). The figure depicts the effect of adding $p_4$ to $g_1(w)$ as an independent proposition.](image)

Note that given the new set of preferences, (175) is intuitively a true statement. However, since $w_5$ is still unordered with respect to any Amherst-going world, the condition for necessity in (176) is not fulfilled.

There are a number of possible responses to this result. The obvious response from the perspective of Premise Semantics would be to blame the problematic result on the particular
choice of ordering sources in the examples. Concretely, instead of using \{p_1, p_2, p_3\} as an
ordering at the outset, we could have also used an ordering source that is more “structured.”
An ordering source like \{(p_1 \cap p_3) \cup (p_2 \cap p_3), p_1 \cup p_2 \cup p_3\}, for example, reflects
the intuition that the preference for being in a good mood is secondary to the preference for
visiting one of the museums. An interesting question would then be how, in general, basic
“unlumped” preferences are structured into a single ordering source to represent cascaded
preferences (see Katz et al. to appear for relevant discussion).

Another idea would be to keep \(p_4\) separate from the original set of preferences, \{\(p_1, p_2, p_3\}\), and to use it to refine the ordering they induce (in particular, allowing comparison
of worlds that are incomparable in the original ordering). This is the intuition behind von
Fintel and Iatridou’s (2008) analysis of weak necessity modals, and have to in (175) would
be an example of a strong necessity modal that makes use of the same machinery. This
example would support Kratzer’s original idea that modals in general may be sensitive to
more than a single ordering source.

Actually, it is a simplification to assume that there is never more than one
ordering source involved in modal reasoning. Suppose I draw conclusions
which involve established facts, the Encyclopedia Britannica, the local newspaper and the gossip I picked up at the corner. And suppose further that the
established facts have priority over the Encyclopedia Britannica, the Encyclo-
pedia Britannica has priority over the local newspaper and the local newspaper
has priority over the gossip I picked up at the corner. How do we reason in
such a case?
I think that the semantics of modals which I have presented so far can be ex-
tended in a straightforward way to handle these cases. The interpretation of
a modal expression would have to depend on a modal base \(f\) and a finite se-
cquence of ordering sources \(g_1, \ldots, g_n\). For any world \(w\), \(g_1(w)\) would induce
an ordering on \(\bigcap f(w)\) in the usual way. \(g_2(w)\) would – if necessary – refine
this ordering in undoing the ‘ties’ left by its predecessor, and so on for every
successive member in the sequence.
Probably, we can’t assume that the different ordering sources form a natural se-
cquence with respect to having priority over each other. There may be ordering
sources which have equal priority. This all sounds as if it were the beginning
of my next paper.
(Kratzer, 1981, 321)
APPENDIX B

TRIPLE RELATIVITY: NEGOTIABLE NORMS, NON-NEGOTIABLE NORMS, AND FACTS

A unified conversational background $M^w$ is a set of sets of propositions. It is initialized with the value of $\{f(w)\}$, for some contextually determined Kratzerian modal base $f$. Primary ordering sources, representing non-negotiable norms, are merged into the unified conversational background via the operation of repeatable compatibility restricted union, which invokes Frank’s (1996) “compatibility restricted union” for conversational backgrounds (ibid. p. 42). The set of accessible worlds forms the domain of quantification for modals like must and can. It is retrieved from the unified conversational background after the $n$-th primary ordering source has been added to it. Secondary ordering sources, representing negotiable norms, may order the set of accessible worlds in the usual way, following Kratzer (1981, 298).

Unified Conversational Background For a Kratzerian modal base $f$, a sequence of primary ordering sources $<g_1, g_2, \ldots, g_n>$, and a world of evaluation $w$, we define the unified conversational background $M^w$:

$$M^w = M^w_n,$$

where:

$$M^w_0 = \{f(w)\},$$

$$M^w_i = M^w_{i-1} + g_i(w), \text{ for } 1 \leq i \leq n.$$

Repeatable Compatibility Restricted Union For a set of sets of propositions $\mathbb{M}$ and a finite set of propositions $A$:

$$\mathbb{M} + A = \{M \cup X : M \in \mathbb{M}, X \subseteq A, M \cup X \text{ is consistent}, \forall Y \subseteq A \text{ if } X \subseteq Y \text{ then } M \cup Y$$
is inconsistent).

(Following Frank 1996)

**Accessible Worlds** The set of worlds accessible from a world $w$ given the unified conversational background $\mathbb{M}^w$:

$$\bigcup\{m : m = \bigcap M \text{ for some } M \in \mathbb{M}^w\}.$$


BATES, DOUGLAS, MARTIN MAECHLER, AND BIN DAI. 2008. lme4: Linear mixed-effects models using S4 classes. URL http://lme4.r-forge.r-project.org/, R package version 0.999375-24.


JAEGGER, T. FLORIAN. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59:434–446.


ROSS, JOHN ROBERT. 1976. To have ‘have’ and to not have ‘have’. In *Linguistic and literary studies in honor of Archibald A. Hill*, ed. Mohammad Ali Jazayery, Edgar C. Polomé, and Werner Winter, volume 1, 263–270. Lisse: Peter de Ridder Press.

RUBINSTEIN, ANAT. 2010. *Needing to have* but not *needing*: The limited modal potential of intensional transitive verbs. Poster presented at the 23rd Annual CUNY Conference on Human Sentence Processing, New York University, March.


———. to appear. Modality, weights, and inconsistent premise sets. In Semantics and Linguistic Theory (SALT) XXII.


