

THE UNIVERSITY OF CHICAGO

PREFERRING TO LEARN: AN ATTITUDINAL APPROACH TO POLAR QUESTIONS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE HUMANITIES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF LINGUISTICS

BY
MICHAEL TABATOWSKI

CHICAGO, ILLINOIS

AUGUST 2022

Copyright © 2022 by Michael Tabatowski
All Rights Reserved

For Mom,

who would have been the only one to read it from cover to cover.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	v
ABSTRACT	viii
1 INTRODUCTION	1
1.1 Terminology used in this dissertation	1
1.2 Chapter 2: Issues in the semantics/pragmatics of polar questions	4
1.3 Chapter 3: An attitudinal analysis of polar questions	7
1.4 Chapter 4: Neg-bias and NPIs in polar questions	12
2 ISSUES IN THE SEMANTICS AND PRAGMATICS OF POLAR QUESTIONS	18
2.1 Asymmetry in polar questions	24
2.2 Felicity and bias in polar questions: evidence is not enough	26
2.3 Polar questions are about their radical proposition	33
2.4 Asymmetry by Manner implicature	38
2.5 Asymmetry by extra structure	43
2.6 Conclusion	52
3 AN EXPRESSIVE-ATTITUDINAL ANALYSIS	54
3.1 Desire in polar questions	57
3.1.1 Modeling polar-question desire	62
3.1.2 Goal orientation in polar questions	66
3.2 The attitudinal analysis of polar questions	69
3.3 Polar questions as expressives	70
3.4 Polar questions in context	73
3.4.1 <i>Yes</i> and <i>No</i> in polar questions	74
3.4.2 Polar question felicity: The Evidence Condition revisited	75
3.5 Conclusion	81
4 NEGATION BIAS AND NEGATIVE POLARITY ITEMS IN POLAR QUESTIONS	83
4.1 Characterizing neg-biased questions	85
4.1.1 Inner and outer negation	87
4.2 Previous analyses of neg-bias	95
4.3 Extending our analysis to neg-biased questions	110
4.3.1 From believing to learning	112
4.4 Against the inner-/outer-negation distinction	120
4.4.1 Neg-biased questions and polarity	120
4.4.2 Negative polarity items in neg-biased questions	123
4.4.3 NPIs in polar question: a domain-widening account	129
4.5 A pragmatic account of inner and outer negation	134
5 CONCLUSION	146
REFERENCES	148

ACKNOWLEDGMENTS

What a relief it is to be writing my dissertation acknowledgements! Near the end of this long process—one that, inevitably, I failed to fully comprehend at the outset—it really does feel good and fitting to thank the people to whom I’m most indebted, and to consider the whole sweep of this stage of my life that’s finally drawing to its end.

Preeminent thanks goes to Itamar Francez, who has been exactly the brilliant chair I needed throughout the dissertation process. I am grateful for his approach to doing linguistics: opinionated, passionate, with a constant and very real sense of stakes. During my time in this field, I have struggled with a kind of apathy to theory: presented with some analytical choice point, I struggle to feel any sense of investment in one option over another. This is, I’m sure, not a problem Itamar could conceive of having. I’m certain I could not have developed the commitments that I now have without him; I don’t know that I could have completed grad school with any other chair. I appreciate too that Itamar’s intellectual life is not defined by linguistics—that he has a similar sense of investment and interest in so many domains of life, which has made him a great mentor and friend over these years. Thank you, Itamar.

Thanks to my committee members Chris Kennedy and Karlos Arregi, who also served on my qualifying paper committees. Chris is a profoundly insightful thinker; an early comment of his so often turns out to be exactly right in a way that you can only properly appreciate later. And Karlos has an incredibly keen and systematic approach to linguistics that I can only call ‘honorable’—never willing to accept an insufficiently motivated claim, but because of a real intellectual rigor, and never because of a lack of charity. I admire both of you greatly. Thank you for everything.

I’m grateful to the other faculty members who have made Linguistics at Chicago such a richly rewarding intellectual environment for me, especially Anastasia Giannakidou and John Goldsmith, but also many, many others. Anastasia served on my first qualifying paper committee and has been a great advisor and instructor both, and her influence is certainly

present in many aspects of this dissertation. John, who never advised me, left a lasting impression for the way that he taught Phonology II in my first year of grad school: with a love for the field and its intellectual history that was nothing less than beautiful. Thank you, all.

At New York University, Chris Barker and Tim Leffel taught and TAed my first undergraduate semantics class ten years ago, where I discovered the charm of the enterprise of formal semantics and immediately fell in love with the field. Chris and Tim's generous support and encouragement when I decided to pursue semantics was critical. I'm particularly glad to still call Tim a friend today. Thank you.

I'm grateful to my cohortmates Dan Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kat Montemurro, Brandon Rhodes, and Amara Sankhagowit for their companionship at the beginning of and throughout our program; and to Laura Stigliano and Yenan Sun for accepting me as an honorary member of their cohort. Thanks, everyone, and thanks to all of the other friends I've made during graduate school. Thank you, too, to all of my friends outside of academia.

How could words express the role of my brothers and sisters Mary, Tommy, Steve, Abby, Sam, and Thea? Without them, exactly as they are, I would not be myself. To Dad: Thank you for your perfectly constant support and love. To Mom: thank you for the gift of life and for everything after. I miss you terribly.

I have often said that grad school was worth it simply because it brought Eszter and me together. Eszter: you are the love of my life, perhaps the only person who really understands me, brilliant and gregarious and hard-working—so many things that complement me and make my life infinitely more beautiful. Thank you for being my wife and my best friend. This dissertation is only one more thing that I could not have accomplished without you.

Lastly, there is the one whose love is presupposed in every good thing I have mentioned so far. All glory and honor and thanks be to God for the splendor of creation, and for giving us human beings the faculties to perceive and be moved by its beauty. As Simone Weil

puts it: “The true definition of science is this: the study of the beauty of the world”¹. And, elsewhere: “The beauty of the world is Christ’s tender smile for us coming through matter”². Entirely independent of its role as an academic discipline, formal linguistics is and always will be a real method for us to contemplate the beauty of the world: it allows us to direct our attention toward some part of the structure of the world that is otherwise hidden to human perception. In the light of a good analysis a tangled mass of data comes to instantiate a clear and perfect pattern, and everything irrelevant to it burns away like tinder. But the beauty present in such an analysis is only a sensible reflection of a suprasensible beauty present in language itself—a beauty that is just one manifestation of God’s superabundant love for us, a love whose most perfect form was realized on the cross.

“From his fullness we have all received, grace upon grace” (John 1:16). Every analysis is a prayer about the graces we have received. This imperfect little one is mine. So I give my profoundest thanks to God, that he has allowed me the last seven years in the contemplation of his beauty, and for every other good thing which comes from him.

Paris

June 2022

1. *The Need for Roots*.

2. *Waiting for God*.

ABSTRACT

The semantics and pragmatics of polar questions remains a surprisingly undertheorized phenomenon. Standard proposition-set accounts (Hamblin, 1976; Groenendijk and Stokhof, 1984), which take question meanings to be identifiable with the (appropriately defined) set of their answers, have been developed and deployed mostly for wh-questions; but when applied to polar questions, these accounts predict massive synonymy, and they therefore struggle to account for the differential acceptability of polar questions across contexts. For instance, imagine that your interlocutor sees you writing with your left hand. They ask:

- (1) a. Are you left-handed?
- b. #Are you right-handed?
- c. Are you not right-handed?
- d. #Are you not left-handed?

On the basis of this and other evidence, a number of authors have recognized the need to distinguish the radical proposition in some way (Krifka, 2001; Romero and Han, 2004; Roelofsen and Farkas, 2015) and identified epistemic and evidential conditions on polar question felicity (Büring and Gunlogson, 2000; Sudo, 2013). But no account yet derives the felicity of these questions in various contexts from their systematically different meanings.

This dissertation has two main parts. First, I present a novel analysis of polar questions on which they express the speaker’s *conditional preference for doxastic update* with the radical proposition. This preference is structured by an ordering source consisting of the speaker’s goals and the necessary conditions to achieve those goals. This combination of what I term an *asymmetric semantics* (systematically assigning different meanings to polar questions $p?$ and $\neg p?$) with an attitudinal denotation reliant on speaker goals provides an account for fine-grained polar question felicity across contexts, as well as the “functional heterogeneity” (Kaufmann, 2012; Condoravdi and Lauer, 2012) of polar questions, which sees them able to perform various speech acts like offering, requesting, conversation-starting,

and more (Bolinger, 1978).

In the second part of the dissertation, I extend this analysis of polar questions to what I term *neg-biased* polar questions, more commonly called ‘high-negation questions’ (Ladd, 1981), such as the following:

- (2) a. Aren’t you right-handed?
- b. Aren’t you not left-handed?

The major puzzle of neg-bias is how the “high negation” of these questions systematically gives rise to an inference of speaker bias without contributing propositional negation of the radical (Goodhue, 2018). As the account developed in the first part of this dissertation treats all polar questions as related to the speaker’s doxastic preferences, it provides a natural basis for analyzing this bias. I argue that negation scopes over the entire attitudinal meaning of the question, expressing the speaker’s lack of a preference for doxastic update, while an independent convention of use for questions requires that the speaker wants to know the radical proposition if true. These two components conspire to guarantee that the speaker has a prior belief in the radical proposition. Here I also discuss negative polarity items in polar questions, which have been a sticking point in the literature on neg-biased questions especially. I argue that a long-assumed “inner/outer” ambiguity in neg-biased questions is an illusion owing to the independent ability of NPIs to appear in polar questions, and in particular the pragmatic effect of choosing an NPI contributing domain-widening (Kadmon and Landman, 1993).

Zooming out to consider broader consequences, this analysis is one on which polar questions serve as a way of belief management, and perhaps bear more semantic similarity to imperatives than to wh-questions. It also implicitly problematizes the concept of polar questions as a cross-linguistically stable phenomenon; languages may—and do—have different strategies for expressing speaker interest of various kinds in the truth of a proposition, or the answer to a question, of which the canonical English polar question is just one.

CHAPTER 1

INTRODUCTION

This dissertation attempts to address a few questions regarding the semantics and discourse functions of polar questions, an issue for which there is a surprising dearth of work. Core aspects of polar questions are consequently still poorly understood. This dissertation attempts to address some outstanding issues regarding the felicity of polar questions and the interpretation of ‘high negation’ by developing a novel semantic analysis of canonical polar questions as expressing a speaker’s **conditional preference to learn** the radical proposition of the polar question. Though perhaps at first unintuitive, the account has a number of benefits for the analysis of polar questions of various types in various contexts. This introduction chapter provides a chapter-by-chapter overview of the argument advanced in the dissertation, starting with a brief summary of the novel terminology deployed throughout.

1.1 Terminology used in this dissertation

In order to make the discussion clearer, I use a few idiosyncratic terms to refer to types of polar questions and relevant propositions. First, I draw a distinction between **canonical polar questions** and **neg-biased polar questions**. Canonical polar questions are just that: those polar questions formed from one radical proposition, of either polarity, which appear with a fronted auxiliary verb and no cliticized negation. For example:

(1) Examples of canonical polar questions

- a. Is Ida in the yard?
- b. Can everyone hear me?
- c. Do Muslims not eat pork?

I refer to polar questions with negation syntactically ‘high’, expressing a characteristically biased interpretation, as *neg-biased* questions. Neg-biased questions can have radicals of

positive or negative polarity:

- (2) a. Aren't you hungry?
b. Aren't you not hungry?

The 'neg-biased' terminology is motivated by the inadequacy of existing terminology like 'high-negation polar question' for the same construction. First, in an important sense, it is the bias we are interested in, and not the high position of negation. Although this dissertation is concerned solely with English, neg-bias has been claimed to arise in other languages, like German and Japanese Büring and Gunlogson (2000); Sudo (2013). But what diagnoses the presence of such a question form in a language is the bias inference, and not the presence of negation in a relatively high position *per se*. And neg-bias interpretations can arguably arise even when there is low negation even in English (what AnderBois 2019 refers to as 'Gladiator' questions). Consequently, the name of the phenomenon should refer to its distinctive meaning, and not to its syntactic properties.

Additionally, the term 'high-negation polar question' requires confusing locutions like 'low-negation high-negation polar questions' for examples like (2-b), where a neg-biased question has a radical of negative polarity. In the terminology of this dissertation, these are 'negative neg-biased questions' or 'neg-biased questions of negative polarity', or something else along those lines, in contrast to 'positive neg-biased questions' like (2-a).

Throughout, I refer to the proposition that a polar question is based on as the *radical proposition*, and the logical negation of that proposition as the *inverse proposition*. These terms were chosen to be independent of polarity, as questions can be of positive or negative polarity. Some examples of questions and their radical/inverse propositions follow:

- (3) a. Are you hungry? (Positive canonical polar Q)
Radical proposition: That you're hungry
Inverse proposition: That you're not hungry

- b. Are you not hungry? (Negative canonical polar Q)
Radical proposition: That you're not hungry
Inverse proposition: That you're hungry
- c. Haven't you seen *Meet Me in St. Louis*? (Positive neg-biased polar Q)
Radical proposition: That you have seen *Meet Me in St. Louis*
Inverse proposition: That you have not seen *Meet Me in St. Louis*
- d. Haven't you not seen *Meet Me in St. Louis*? (Negative neg-biased polar Q)
Radical proposition: That you have not seen *Meet Me in St. Louis*
Inverse proposition: That you have seen *Meet Me in St. Louis*

Using the terms *radical* and *inverse proposition* allows us to refer to the propositions that many theories take polar questions to introduce without using terms like *yes* and *no answer*. Referring to these propositions as the *yes* and *no* answers is undesirable for two reasons. First, it presupposes that *yes* and *no* map one-to-one onto the two propositions in all cases. However, this is not true; in negative polar questions, *yes* and *no* do not determine unique propositions (Kramer and Rawlins, 2009, i.a.). Second, it equates the *response* to a polar question, with *yes* or *no*, to the propositional answers to the question. But these are distinct phenomena; *yes* and *no* are used to respond to a polar question in discourse, but the propositional answers to the question are determined solely by the semantics of the question. In order to keep the discussion maximally clear, then, it is good to have a way to refer to the proposition that serves as the basis of the question, and its negation, that makes no reference to how interlocutors respond to a question in discourse.

With these few terminological points out of the way, we can proceed to a brief summary of the dissertation.

1.2 Chapter 2: Issues in the semantics/pragmatics of polar questions

After this introduction, Chapter 2 of this dissertation sets the stage for the puzzle of polar questions. In addition to specifying a “bare” semantics—say, the set of answers to the question—any analysis of polar question meaning must shed light on how the puzzle of how they fit into discourse generally. To illustrate this puzzle, note that two polar questions may appear to be semantically identical, in the sense that they determine the same set of answers, but perform different roles in discourse. For instance, suppose that at a party, an acquaintance pulls out a pack of cigarettes and, gesturing at it, asks you:

- (4) a. Are you a smoker?
b. Are you a non-smoker?

In this context, (4-a) is a significantly more natural question than (4-b). Most naturally it might, for example, invite you out for a cigarette. In contrast, (4-b) is a bit odd; it’s hard to imagine why your interlocutor would choose to say it. If pressed to imagine a way of making (4-b) felicitous, we might come up with less likely scenarios: that the speaker wants to smoke indoors, perhaps, but is worried that you may be non-smoker and therefore irritated by the cigarette smoke. He then may ask (4-b) in order to help decide whether it’s safe to light up indoors.

Despite the intuitive differential felicity between the questions, though, a (logical) answer to one of them is necessarily an answer to the other. If we conceive of a question as merely asking for the assertion of one of two logically exclusive propositions, then it’s clearly impossible to resolve the first question without resolving the second, or vice versa: if you answer *No* to (4-a), then you would necessarily answer *Yes* to (4-b), and the other way around.

The puzzle, then, is why the questions seem to do different things. Why is (4-a) a good way to make an offer, but (4-b) not? This pattern extends beyond offers and across the

whole range of polar questions: we choose particular forms to perform particular speech acts in particular contexts.

Remarkably, few existing proposals attempt to account for this differential felicity of polar questions across contexts. Those accounts that do (e.g. Büring and Gunlogson, 2000; Sudo, 2013; Gyuris, 2017) generally locate the felicity of polar questions in the evidential properties of the context. For instance, Büring and Gunlogson (2000) formulate the following “Evidence Condition”: a polar question $p?$ is felicitous only when there is no evidence for the truth of the inverse proposition $\neg p$ (i.e., the negation of the radical). This Evidence Condition accounts for cases like the following example, where the addressee’s reluctance to eat is taken to be evidence for *You are sick*:

- (5) [*Addressee is not eating the food on his plate*]
 Are you sick? / #Are you not sick?

As there is evidence for *You are sick*, the question *Are you not sick?* is not licensed: there is evidence for its inverse proposition. But we can note straightaway that the Evidence Condition won’t get us far in the smoking scenario—there’s no evidence regarding the radical proposition *You are a (non-)smoker*. Whatever makes (4-a) a felicitous offer for a smoke break, but (4-b) not, does not obviously have anything to do with evidence.

This is just one example of the failure of an evidential condition to do justice to the intricacies of polar question felicity. And existing accounts of question semantics shed little light on a way forward. These typically analyze polar questions as partitioning logical space, or the common ground, into two cells: one cell denotes the radical proposition, and one denotes its logical inverse. With a semantic representation of this shape, then, matrix polar questions asked in discourse are taken to prompt interlocutors to choose one or the other proposition as a continuation of the discourse. But though this picture is appealing for a number of reasons (not least because it comports with standard analyses of wh-questions, which dominate the literature on questions), it doesn’t tell us what makes a polar question

felicitous or not in a given context.

The reason for this, I will argue, is in part that partition analyses by default are **symmetrical**: they assign the same meaning to a question with a negative radical $\neg p$ and a question with a positive radical p (Krifka 2015, 2017 refers to this property as the question's being 'bipolar'). On a naïve partition semantics, there is no semantic difference whatsoever between the following two questions, which are based on a positive and negative radical proposition, respectively:

- (6) a. Is the cat hungry?
- b. Is the cat not hungry?

But even at the most general pretheoretic level, these questions are clearly different — not only in their felicity conditions, or in what contexts they might most naturally be uttered, but more fundamentally in their *aboutness*. In (6-a), the speaker wants to learn *about* the proposition that the cat is hungry; but in (6-b), the speaker wants to learn about the proposition that the cat is *not* hungry. And this, I will argue, is the ultimate source of the difference between the two questions, and among polar questions generally. Put simply, polar question meanings require an **asymmetrical semantics**: from the representation of a polar question, we should be able to retrieve the radical proposition. And, moreover, the radical must be the proposition that the question is *about* in some concrete sense.

More complex analyses attempt to overcome the asymmetry problem in a few ways: by building extra structure into the question, e.g. in the structured-meaning approach of Krifka (2001); or by more abstract structure not present in the meaning of the question per se, e.g. 'highlighting' (Farkas and Roelofsen, 2017); or, sometimes, by appealing to markedness constraints, e.g. in Romero and Han (2004). But these approaches, even when they successfully distinguish the radical and inverse propositions of a polar question, fail to capture the core *aboutness* of the question; they don't explain why we ask some questions in some contexts, and other questions in other contexts. So although they formally distinguish

between questions with negative and positive polarity, they don't provide an account of differential felicity.

1.3 Chapter 3: An attitudinal analysis of polar questions

In Chapter 3, I propose a solution to the problem of differential polar question felicity by analyzing polar questions as expressing a *conditional preference to learn* the radical proposition. For instance, *Are you a smoker?* is assigned a meaning something like the following paraphrase:

(7) If you're a smoker, I prefer to come to believe so.

This approach has two immediate benefits. The first is that it provides a straightforward solution to the symmetry problem: a question is about its radical proposition in the sense that it expresses a preference to know that proposition, and not some other. The second is that it provides a basis for analyzing the felicity conditions of polar questions more generally, across a range of contexts: as they denote modal propositions, polar questions are subject to predictable felicity conditions, and we can avoid merely stipulating what contexts license polar questions.

This account crucially makes use of a goal-oriented ordering source to account for polar question felicity across contexts. In effect, questions are felicitous when they further a speaker's goal, in a particular sense to be made clear later. Speakers' goals structure their preferences; agents prefer states of affairs that bring them closer to achieving their goals. Polar questions are asked in light of goals, and the speaker's goals within a particular context are thus the primary driver of polar question felicity.

Consider the cigarette scenario again. In a context where my goal is to offer you a cigarette, learning that you are a smoker furthers that goal: I want to know that you're a smoker in order to successfully offer you a cigarette. (Presumably, it's pointless or rude

to offer a cigarette to a non-smoker.) Learning that you are a non-smoker, on the other hand, does not further my goal to offer you a cigarette. Therefore, the question *Are you a non-smoker?* is infelicitous in this context.

Now suppose that my goal is not to offer you a cigarette, but rather to determine whether it's acceptable to smoke indoors near you. I might ask, assuming that non-smokers dislike indoor smoking, *Are you a non-smoker?* In this case, the question is much more acceptable. Crucially, the only thing that has changed is my intention as the asker of the question.

This illustrates a kind of goal-dependence in the felicity of polar questions: if we construe the speaker as having one or another goal, we predict different questions to be felicitous, even if other properties of the context remain the same.

This approach also provides a natural explanation for a family of observations from Bolinger (1978), who notes that canonical, bare polar questions can be felicitously used to perform certain speech acts like offering, requesting, and asking whether an inference is licensed, but disjunctive polarity questions cannot. This is the *functional heterogeneity* of polar questions (a name given to the same phenomenon in imperatives by Condoravdi and Lauer 2012): they can be used to perform a variety of speech acts.

- (8) a. Do you want a coffee? (Offer)
b. #Do you want a coffee or not?
- (9) a. Will you marry me? (Request)
b. #Will you marry me or not?
- (10) a. You're hardly eating. Are you sick? (Inference-licensing)
b. You're hardly eating. #Are you sick or not?
- (11) a. Do you like basketball? (Conversation starter)
b. #Do you like basketball or not?

On the present account, the bare polar questions in the (a) examples denote propositions

expressing the speaker’s preference to learn the radical proposition, should it be true. That preference makes reference to a doxastic modal base and an ordering source reflecting the speaker’s goals. It is in light of those goals in each case (e.g., in (8-a), the speaker’s goal to give his interlocutor an appropriate drink) that the speaker asks his questions, i.e. expresses his preference to learn the relevant proposition.

The disjunctive polarity questions in the (b) examples, on the other hand, genuinely offer the interlocutor the choice between two mutually incompatible propositions. This is what Biezma (2009) refers to as the *cornering effect* of disjunctive polarity questions: they offer the addressee two coequal propositions, forcing a choice between them, which yields a sense of insistence on the part of the speaker, but no bias toward one or the other.

As with disjunctive polarity questions, questions asking about a negative radical proposition, too, are generally infelicitous in these same contexts, because they don’t further the relevant goal of the speaker:

- (12) # Do you not want a coffee? (Offer)
- (13) #Will you not marry me? (Request)¹
- (14) You’re hardly eating. #Are you not sick? (Inference licensing)
- (15) #Do you not like basketball? (Conversation starter)

It is impossible, or at least very difficult, to construe the speaker as preferring to know the relevant proposition given the implicit goal of each of these questions. For instance, if my goal is to offer you a drink you want, it is not plausible that I prefer to learn that you do *not* want a coffee, if you don’t want a coffee; I want to find a drink that you *do* want². Similarly,

1. Note that *Won’t you marry me?* is fine—this is an example of the neg-biased question, covered in Chapter 4.

2. In order for this sentence to successfully perform the speech act of offering a coffee, the following condition would need to hold: in order to offer the interlocutor a coffee, the speaker would want to know that the interlocutor does *not* want a coffee. Trivially, we can make this question a felicitous offer in a context where the interlocutor says something like, *Please give me something that I don’t want*. In that case,

if my goal is to know why you're hardly eating, it can't be that I prefer to know that you're not sick (if you are not sick); that knowledge wouldn't further my goal; I won't be any closer to a reason for your lack of appetite. It is difficult for a partition analysis to explain what goes wrong in these examples. On such an approach, the negated questions proffer the same two cells as their positive counterparts; any answer to them is semantically identical to an answer to the positive question. If the function of a polar question is merely to determine which of the two cells the actual world is in, both the positive and negative question should suffice.

Zooming out from the details of the account, the analysis advanced in this dissertation divides the landscape of questions differently from most, if not all, approaches to polar questions so far. In identifying polar questions as expressing a speaker's attitude toward the radical proposition, we will not posit a particularly close relationship between *wh*-questions and polar questions; in effect, we are eschewing the notion of a semantic "core" to questions shared by *wh*- and polar questions. Polar questions are instead taken to be more like imperatives (at least on some accounts, e.g. that of Kaufmann 2012) in expressing a modal proposition.

Although drawing a distinction between *wh*-questions and polar questions may seem like an undesirable consequence, there are conceptual benefits to such an approach. One is that it provides a new perspective on the range of polar-like questions in and across languages, which is quite varied and robust. For instance, in English, we have at least the following constructions:

- (16) a. Is the cat hungry? (Positive polar question)
b. Is the cat not hungry? (Negative polar question)
c. Is the cat hungry, or (is it) not (hungry)? (Alternative polarity question)
d. The cat is hungry? (Rising declarative)

the speaker might offer her a coffee with (12). Generating a more naturalistic example, however, is difficult.

- e. Isn't the cat hungry? (Neg-biased question)
- f. The cat is hungry, isn't it? (Positive tag question)
- g. The cat isn't hungry, is it? (Negative tag question)

The breadth of this menagerie of constructions has been a source of difficulty in the literature on polar questions. Can they all be subsumed, along with *wh*-questions, under the same analytical framework—for instance, a proposition-set account of question meanings? Most accounts choose a few constructions and attempt to derive properties of them; none has attempted to provide an account of all of them. Theorists committed to a proposition-set approach also have to make a decision about where to distinguish among these types: in the semantics, or in the conventional but non-semantic component of language meaning. (For an overview of this puzzle with respect to some of these question types, see Farkas and Roelofsen 2017.)

Our analysis splits up *wh*-questions and polar questions, conceiving of polar questions as expressing a speaker preference, rather than denoting a semantic object corresponding to an abstract question-meaning (e.g. a set of propositions). This conception of polar question meanings provides a new degree of freedom in assessing the menagerie of polar-like questions above. Canonical polar questions express a conditional preference to learn the radical proposition; but in principle, there are a number of different attitudes that one could bear towards a proposition, or a set of propositions, beyond a conditional preference to know. Consequently, we don't have to assume that all (or any) of these other question types express the same attitude.

This says that $p?$ expresses that, for every doxastically accessible p -world of the speaker's, in light of the speaker's goals, they would prefer that world where they come to believe p to that world minimally altered such that they don't come to believe p . In Chapter 3, a simplification of this account is presented,

This account rejects universality of question meaning: it's not necessarily the case that

all the question forms above share something like a core partition semantics, also shared with wh-questions. Polar questions rather express a particular kind of interest or investment in the truth of a particular unsettled proposition, and strategies for doing so may vary within and across languages.

1.4 Chapter 4: Neg-bias and NPIs in polar questions

In Chapter 4, I extend the analysis to what I term *neg-biased* polar questions (more commonly referred to as something like ‘high-negation questions’): those that appear with high negation cliticized to a fronted auxiliary verb, which come with a characteristic biased interpretation (Ladd, 1981):

- (17)
- a. Aren’t you hungry?
 - b. Wasn’t Finland not in the USSR?
 - c. Didn’t anyone lock the front door?
 - d. Isn’t there somewhere to eat around here?

Neg-biased questions are interesting for a number of reasons, posing significant problems for a theory of polar questions. Most importantly, the overtly high negation present in neg-biased questions does not provide propositional negation of the question’s radical proposition, according to a battery of tests (Goodhue, 2018). That is, there are systematic differences between negative-polarity canonical questions like (18-a) and positive neg-biased questions like (18-b). While (18-a) asks about the negative-polarity proposition *Muslims do not eat pork*, (18-b) expresses a speaker’s belief that Muslims do eat pork:

- (18)
- a. Do Muslims not eat pork? (negative canonical question)
 - b. Don’t Muslims eat pork? (positive neg-biased question)

It is, therefore, the presence of high negation in polar questions that is correlated with the

neg-bias reading, and not the presence of negation *per se*; it must appear cliticized to the fronted auxiliary verb to trigger the neg-bias reading. An ideal account of neg-bias would locate the contribution of negation somewhere in the denotation of a question, but outside of the radical proposition. The presence of negation in this specific location, then, would allow us to derive the bias character of neg-biased questions.

A strategy along these lines has been the standard approach in the literature on neg-bias since Ladd (1981), and underlies the accounts of Romero and Han (2004), Krifka (2015, 2017), AnderBois (2019), and Goodhue (2018). However, this comes with significant challenges, especially for proposition-set views of polar questions. These accounts have no clear place to put negation besides immediately outside the question radical. On a proposition-set account, a polar question has a semantic representation of the form $Q(p)$, where p is the radical proposition and Q is a partition-forming operator. The only two possible places for negation to appear are above p , but below Q , which is the denotation of a negative canonical question; or above Q entirely, which is not possible, as the meaning of a question is supposed to be a set of propositions, and it's not clear what it means to negate a set.

Consequently, proposition-set accounts generally assume the presence of some sort of operator that scopes in between the question-forming operator and the radical proposition. This extra operator in the structure of the question provides a 'high' position for negation to appear that is nevertheless still below the polar-question-forming operator. This is exemplified in Romero and Han's 2004 VERUM account, which posits that in the case of neg-biased questions, an epistemic operator of some kind scopes in between Q and the radical proposition p , and negation appears above that operator. Thus, if we call this operator ' \square ' (following Goodhue 2018), we have the following representation for a neg-biased polar question.

- (19) **Generic partition semantics representation for neg-biased question:**
 $Q(\neg\square(p))$

These accounts, however, have outstanding problems. First, it's *prima facie* unclear how to

derive the bias inference from such a representation, although a pragmatic account has been made at least in Goodhue (2018). Second, introducing a special epistemic operator only for neg-biased questions introduces a host of issues in its own right. For one, it posits an asymmetry between neg-biased polar questions and canonical polar questions: neg-biased questions contain the epistemic operator, but canonical polar questions do not. That is, such an account generally still assigns a representation like the following to a canonical polar question.

- (20) **Generic partition semantics representation for canonical polar question:**
 $Q(p)$

Accounting for neg-bias by means of some silent epistemic operator also gives negation an odd status: negation becomes an indicator of the bias inference, but does not directly trigger it. Instead, it's the presence of the \square operator that triggers the bias inference; and for some reason, negation must appear in the sentence in order for the \square operator to be licensed. And in fact, because a proposition-set analysis assigns the same denotation to representations $Q(\phi)$ and $Q(\neg\phi)$, negation ends up not contributing *anything* to the semantics of a neg-biased question. The difference between a canonical question and a neg-biased question is solely a consequence of the \square operator, whose covert presence is signalled by the overt presence of negation. The attitudinal account developed in Chapter 3 provides an alternative way to approach this issue. On this account, negation can simply scope above the entire question, without the need to posit some extra operator.

In order to make this work out, we amend our denotation of polar questions so that they express a speaker's preference for epistemic update, i.e., for a 'coming-to-believe' event to take place. Consequently, the denotation that we will end up assigning to a polar question $p?$ is the following, where $\mathbf{ctb}^{v,t}(p)(x)$ refers to an event of x 's coming to believe p at world v and time t :

$$(21) \quad \llbracket p? \rrbracket_x^w = \forall w' \in \text{Dox}_x^w[p] [(\text{Sim}^{w'}(v)(\mathbf{ctb}^{v,t}(p)(x)) \\ <_{\text{Goals}^w(x)} (\text{Sim}^{w'}(v)(\neg\mathbf{ctb}^{v,t}(p)(x)))]$$

Whereas a canonical polar question expressed a speaker’s conditional preference to learn, then, a neg-biased polar question expresses a speaker’s ambivalence towards learning; it says that it’s not the case that the speaker prefers to learn the radical proposition p if true. The following neg-biased question, then, has the following paraphrase:

- (22) a. Aren’t you hungry? (*You are hungry* = p)
 b. It’s false that for every doxastically accessible p -world, that world would be better if I updated my doxastic state with p than if I didn’t.

I argue that there are two ways for this to be true: either the speaker never believes the radical p at that world, or the speaker does believe p at that world. The former scenario is ruled out by an independent constraint on speaker goals, which order the speaker’s preference, which says that goals cannot be furthered by speaker ignorance. Consequently, the question entails that the speaker believes p at some doxastically accessible world.

We can safely assume that agents have perfect knowledge of their own beliefs. In other words, if an agent believes something at one doxastically accessible world of theirs, then they must believe it in every doxastically accessible world; it is impossible to be unsure whether you believe something. Consequently, since the neg-biased polar question entails the speaker believes p at some doxastically accessible world, it also entails that the speaker believes p at the actual world.

This, then, is the source of the bias inference; it arises as a result of the negation of a preference to learn. Although this reasoning process might seem baroque, it actually instantiates a fairly simple phenomenon: it says that I might be at least as well off, goal-wise, with doxastic stasis as with doxastic update. The only reasonable way this might be true is if I already believe the proposition in question.

This approach to neg-biased questions avoids a number of the pitfalls of other accounts. For one, it doesn't need to posit a special operator in the representation of neg-biased polar questions. Consequently, the neg-bias meaning is derived simply by negating the meaning of a canonical polar question, which is pleasantly parsimonious. The special meaning of a neg-biased question is caused by the introduction of negation at a sufficiently high position in the semantic representation of the polar question.

Additionally, this account derives neg-bias as a semantic consequence of high negation, rather than positing a pragmatic story that derives neg-bias. This is in turn useful for analyzing the 'inner/outer negation' distinction that has been posited in the literature. Briefly, it is commonly taken to be the case that there are *two* interpretations for a neg-biased question: one that licenses NPIs, and one that doesn't (Ladd, 1981; Büring and Gunlogson, 2000; Romero and Han, 2004; Sudo, 2013; Reese, 2007). In the inner-negation case, the speaker has a prior belief that *p*, which is somehow violated or contravened. In the outer-negation case, the speaker straightforwardly believes that *p*, irrespective of whether that belief has been contravened or not.

- (23) Didn't you eat anything? (Inner-negation)
 (I thought you did, but now I'm not sure. / #I think you did.)
- (24) Didn't you eat something? (Outer-negation)
 (I thought you did, but now I'm not sure. / I think you did.)

Since Ladd's (1981) claim that these instantiate different structures, this phenomenon has been a problem for accounts of neg-bias. Some recent accounts, e.g. AnderBois (2019) and Goodhue (2018), dispute that it represents a genuine structural ambiguity at all.

In this chapter, I further argue that the inner/outer distinction can be analyzed as a scalar implicature. In particular, assuming that certain NPIs like *any* contribute domain-widening (Kadmon and Landman, 1993), a neg-biased question with *any* is semantically

weaker than the neg-biased question with *some*. Consequently, the neg-biased question with *any* gives rise to a scalar implicature: that the speaker believes the same question with *some* is false. Since *any* contributes domain widening, this amounts to saying that the speaker believes that the radical proposition with an unwidened domain is false.

This is an appealing story in part because the question with *any* is strictly more marked than the question with *some*: it is licensed only in the violated-expectation context, whereas the question with *some* has no such restriction. This is natural on a pragmatic account, as it's the choice to use a weaker expression that gives rise to the implicature.

This pragmatic account, I argue, explains a number of troublesome properties of neg-biased questions with *any*. For one, it provides an account of the 'violated expectation' inference: the speaker widens the domain for a particular reason; namely, that they believe the question is false on the standard domain. The most natural reason to do this is if you have encountered evidence which suggests that the only way to make your belief true is to widen the domain of quantification.

Another interesting property this pragmatic account sheds light on is the 'expectation' flavor of the speaker's prior belief in the *any*-question. Existing accounts lack a good account of why the question (23) expresses something more like an expectation than a well-justified belief. But if the function of *any* is to expand the domain, this makes perfect sense. The primary situation in which a speaker might expand a domain of quantification in order to make a proposition true is one in which they have an expectation that, for some sufficiently wide domain, the proposition must be true, but they lack evidence for any individual.

CHAPTER 2

ISSUES IN THE SEMANTICS AND PRAGMATICS OF POLAR QUESTIONS

This chapter presents some outstanding problems in the semantics of polar questions. It starts by identifying the issue of *symmetry* in an analysis of polar questions: an analysis is symmetrical just in case it assigns the same meaning to two questions formed on radicals of opposite polarity. We will then see that an asymmetric analysis is necessary to capture the differential felicity and bias profiles of questions across contexts. The section then surveys a number of approaches attempting to build an asymmetry into the representation of the question by various means.

It is a popular assumption in semantics, going back to Hamblin (1958), that the meaning of a question is equal to its possible answers (henceforth, the *Answerhood Postulate*). Hamblin takes this to be a genuine equivalence in meaning: “[k]nowing what counts as an answer *is equivalent to* knowing the question” (162; emphasis added). (For our purposes, answers are construed as propositions answering the question, not as natural-language expressions with a syntactic form or pragmatic effects.) To sharpen the definition, we can state the Answerhood Postulate as follows:

- (1) **Answerhood Postulate:** The semantic value of a question is fully determined by the set of its possible propositional answers.

Since formal semantics generally takes sentence meanings to be equivalent to truth conditions, the meaning of a question is puzzling in a fundamental way: questions don’t have any immediately obvious truth conditions. The Answerhood Postulate offers a way out of the puzzle of questions semantics, allowing a treatment of question meanings as built in a simple way from truth-conditional meanings, i.e., propositions. To capture the semantic contribution of a question, for which the notion of truth conditions seems inappropriate, we

simply characterize the set of (propositional) answers that the question determines. This is the approach inaugurated formally by Karttunen (1977) and Groenendijk and Stokhof (1984). These approaches to wh-questions assign them a meaning that abstracts over the meaning of the wh-constituent, yielding a set of propositional answers as the denotation of the question. Suppose we're in a universe of discourse with three kids: Gus, Ida, and Evie. In set-builder notation, the wh-question *Which kid did Steve watch?* would be assigned the following meaning, in this case equaling a set of three propositions:

$$\begin{aligned}
 (2) \quad \llbracket \text{Which kid did Steve watch?} \rrbracket &= \{\mathbf{watched}(x)(s) \mid x \in \llbracket \text{kid} \rrbracket\} \\
 &= \{\mathbf{watched}(g)(s), \mathbf{watched}(i)(s), \mathbf{watched}(e)(s)\}
 \end{aligned}$$

A straightforward extension of this approach to polar questions has polar questions denote a set of two propositions, one of positive polarity and one of negative polarity. On such an account, the syntactic constituent whose meaning is abstracted over might be, for instance, a possibly phonologically null Polarity head (Kramer and Rawlins, 2009; Holmberg, 2015) whose semantic domain is a set containing an identity or negation operator:

$$\begin{aligned}
 (3) \quad \llbracket \text{Is Gus hungry?} \rrbracket &= \{Op(\mathbf{hungry}(g)) \mid Op \in \{\lambda p.p, \lambda p.\neg p\}\} \\
 &= \{\mathbf{hungry}(g), \neg\mathbf{hungry}(g)\}
 \end{aligned}$$

Without saying more, accounts along these lines imply massive synonymy across polar questions. For instance, it predicts that polar questions formed from positive and negative radicals are synonymous, because they proffer the same set of two alternatives. The following pairs of questions, then, should all be synonymous:

- (4) a. Is Gus hungry?
 b. Is Gus not hungry?
- (5) a. Did you eat something?

- b. Did you not eat anything?
- (6)
- a. Are his shoes tied?
 - b. Are his shoes untied?

Similarly, but distinctly, the proposition-set account suggests synonymy between questions whose radicals are logically opposite. For instance, each of the following pairs of positive-polarity questions should mean the same, at least insofar as their radicals are taken to be logically complementary:

- (7)
- a. Is your dog male?
 - b. Is your dog female?
- (8)
- a. Is the door open?
 - b. Is the door closed?
- (9)
- a. Are you left-handed?
 - b. Are you right-handed?
- (10) (Context: we're competitors in a race)
- a. Are you going to stop running?
 - b. Are you going to continue running?

It's not immediately clear whether we want to locate the difference between these two questions in the semantics or in some other component; for example, perhaps the difference between these questions is purely pragmatic. At the moment, we won't deliberate on this issue. Here, it suffices to note that a proposition-set analysis predicts synonymy in at least two ways: synonymy of questions differing only in the polarity of their radical proposition, and synonymy of questions whose radicals are logically opposite, but not because of polarity.

The Answerhood Postulate has been called into question repeatedly with reference to polar questions. Krifka (2001) notes that certain polar questions have the same set of

propositional answers as certain alternative questions, but admit different natural-language answers. For example, he observes the following two questions ask about whether the door is open, but only the polar question can be answered with the response particles *Yes* or *No*.

- (11) a. Is the door open? {*The door is open, the door is closed*}
 Yes./No./It's open./It's closed.
- b. Is the door open or closed? {*The door is open, the door is closed*}
 *Yes./*No./It's open./It's closed.

On its face, this poses a puzzle for the Answerhood Postulate. Both questions admit two possible propositional answers: that the door is open, or that it is not, which is synonymous with the door being closed. Since the Answerhood Postulate says that a question's meaning is identical to its propositional answers, these questions should have the same meaning. At the same time, the questions admit different *natural-language* answers.

Krifka (2001) solves this problem by eschewing a proposition-set analysis in favor of *structured meanings*, a family of accounts according to which the meaning of a question is a function that, applied to its answer, yields a proposition. On his analysis, the two questions have different meanings, despite admitting the same propositional answers. Solutions of this form requires jettisoning the Answerhood Postulate: question meanings, though obviously intimately related to their propositional answers, cannot be fully characterized by them. And this is just one illustrative example of a problem that arises across the class of non-wh-questions for the Answerhood Postulate. Stated most generally, the problem is that languages have multiple strategies for asking about a proposition and its logical inverse, and these strategies have different properties. But all of these forms of question have just two propositional answers: a proposition, and its logical inverse. In English, for example, we have at least the following:

- (12) a. Is the cat hungry? (Positive polar question)

- b. Is the cat not hungry? (Negative polar question)
- c. Is the cat hungry, or is it not? (Alternative polarity question)
- d. The cat is hungry? (Rising declarative)
- e. Isn't the cat hungry? (Neg-biased question)
- f. The cat is hungry, isn't it? (Positive tag question)
- g. The cat isn't hungry, is it? (Negative tag question)

The Answerhood Postulate would have it that these questions all have the same semantic value: a set containing the alternative propositions *The cat is hungry* and *The cat is not hungry*. Presumably, then, their differences in discourse effects, conditions of use, and so on are due to something other than their semantic value. The nature of their differences is a matter of debate, as we shall see in the next subsection. For now, it suffices to note that this variety of question forms makes for a remarkable contrast with *wh*-questions. Modulo constructions like *wh the hell* (e.g. in *What the hell did you do?*), there's generally only one way to form a particular *wh*-question. And even if we vary the polarity of a *wh*-question, the two resulting questions cannot really be construed as asking the same thing. For instance, consider the following pair:

- (13) a. What did Ida eat?
 b. What did Ida not eat?

It has been claimed that matrix *wh*-questions by default asks for an exhaustive (*mention-all*) answer (Groenendijk and Stokhof, 1984; Dayal, 2010; van Rooy, 2004), i.e., a complete specification of all their true answers. In that sense, an answer to (13-a) also constitutes an answer to (13-b): the exhaustive answer to both determines a set of worlds consistent with the actual world with respect to what Ida ate. Nevertheless, it's clear what the difference is between these two questions. Briefly put, (13-a) asks for a list of items that Ida ate; (13-b) asks for a list of items that she did not eat. Answering one with the other is infelicitous, as

it is not a congruent answer (in the sense of Groenendijk and Stokhof 1984):

(14) *Context: The relevant items in the pantry are cookies and candy)*

A: What did Ida eat?

B: #She didn't eat the cookies.

A's question asks for a proposition (or series of propositions) of the form *Ida ate x*, but B's answer is of the form *Ida didn't eat x*. Consequently, B's answer is most naturally read as implying ignorance about what Ida *did* eat, instead offering the most informative answer he has knowledge of. This sort of pragmatic meaning gets a straightforward account Gricean account (Grice, 1975)¹. Ultimately, there is little confusion about in what contexts a wh-question of positive polarity might be asked over a wh-question of negative polarity: we ask (13-a) if we want to add the propositions *Ida ate x* to our knowledge state, and we ask (13-b) if we want to add the propositions *Ida did not eat x*. The fact that an exhaustive answer to one entails an exhaustive answer to the other is immaterial.

Returning to positive and negative polar questions, let's now consider a pair differing in polarity:

- (15) a. Did Ida eat the cookies?
b. Did Ida not eat the cookies?

Here, the relationship between the two questions is much tighter: an answer to one is necessarily an answer to the other, as there's no possibility of something like a *mention-some/mention-all* ambiguity. In appearance, the two questions are the same, modulo polarity. And since these two questions necessarily share the same set of propositional answers, the Answerhood Postulate would have it that we cannot semantically distinguish between them. Intuitively, however, there is a difference in meaning between these sentences, and in

1. Alternatively, ruling out ignorance, it might be read as expressing reluctance to address the question directly, or something else along those lines.

fact it parallels the distinction between positive- and negative-polarity wh-questions. Just as *What did Ida eat?* asks for a list of things that Ida ate, and not for a list of things that she didn't eat, (15-a) asks if Ida ate the cookies; it doesn't ask if she didn't eat them. *Mutatis mutandis*, the same holds for (15-b). In other words, the difference between (15-a) and (15-b) is in their *aboutness*, where each is *about* its own radical proposition.

2.1 Asymmetry in polar questions

Standard approaches to the semantics of wh-questions, as we have seen, assign them a meaning that abstracts over the wh-constituent, yielding a set of propositional answers as the denotation of the question. Suppose we're in a universe of discourse with three kids: Gus, Ida, and Evie. In set-builder notation, the wh-question *Which kid did Steve watch?* would be assigned the following meaning, in this case equaling a set of three propositions:

$$(16) \quad \llbracket \text{Which kid did Steve watch?} \rrbracket = \{\mathbf{watched}(x)(s) \mid x \in \llbracket \text{kid} \rrbracket\} \\ = \{\mathbf{watched}(g)(s), \mathbf{watched}(i)(s), \mathbf{watched}(e)(s)\}$$

A straightforward extension of this approach to polar questions then has polar questions denote a set of two propositions, one of positive polarity and one of negative polarity. On such an account, the syntactic constituent abstracted over might be, for instance, a possibly phonologically null Polarity head (Kramer and Rawlins, 2009; Holmberg, 2015) whose semantic domain is a set containing an identity or negation operator:

$$(17) \quad \llbracket \text{Is Gus hungry?} \rrbracket = \{Op(\mathbf{hungry}(g)) \mid Op \in \{\lambda p.p, \lambda p.\neg p\}\} \\ = \{\mathbf{hungry}(g), \neg\mathbf{hungry}(g)\}$$

Assuming that multiple negation is semantically vacuous, such polar-question meanings are **symmetrical** in an important sense². They are symmetrical in that, from the denotation

2. Though see AnderBois (2019) for an analysis that hinges on multiple negation *not* being semantically

of the question, we can no longer retrieve the identity of the question radical. The question might as well have been asked with a negative radical, which yields the same meaning:

$$\begin{aligned}
 (18) \quad \llbracket \text{Is Gus not hungry?} \rrbracket &= \{Op(\neg\mathbf{hungry}(g)) \mid Op \in \{\lambda p.p, \lambda p.\neg p\}\} \\
 &= \{\neg\mathbf{hungry}(g), \neg\neg\mathbf{hungry}(g)\} \\
 &= \{\mathbf{hungry}(g), \neg\mathbf{hungry}(g)\}
 \end{aligned}$$

This is the crux of a symmetrical analysis as we'll define it: it systematically assigns the same meanings to polar questions differing in polarity. In effect, then, it equates positive and negative polar questions. This definition is given below:

(19) **Symmetrical semantics for polar questions**

A semantics for polar questions is **symmetrical** if and only if, for any proposition p , it assigns the same denotation to a polar question with radical p and a polar question with radical $\neg p$.

It's important to note that we conceive of symmetry here as a property of analyses, and, strictly speaking, not a property of semantic objects or of natural languages. An analysis of polar questions is symmetrical (or not) depending on the meanings it assigns to them. We will, though, use the term somewhat loosely in the subsequent discussion, and refer to *denotations* as “symmetrical”. To extend our definition a bit, a denotation (within an analysis) for a polar question with radical p is symmetrical if and only if it is synonymous with a polar-question denotation with a radical $\neg p$. That is, a denotation for a polar question is symmetrical if it is assigned by a symmetrical semantics for polar questions.

Not all symmetrical analyses are proposition-set analyses; the structured meanings account of Krifka (2001), for example, is symmetrical despite not being a proposition-set analysis. What's important isn't the two-cell structure, per se, but the fact that the theory

vacuous for a particular Inquisitive Semantics framework.

doesn't Apart from Giannakidou and Mari (2019), however, I am not aware of any analyses of polar questions that can, at their core, be meaningfully described as asymmetrical: Hamblin (1958); Büring and Gunlogson (2000); Kramer and Rawlins (2009); Farkas and Bruce (2009); Roelofsen and Gool (2010); Krifka (2015); Holmberg (2015); Roelofsen and Farkas (2015); AnderBois (2019); Theiler (2020); Giannakidou and Mari (2019) all give symmetrical accounts³. (Krifka (2015), AnderBois (2019), and Nicolae (2013) additionally give asymmetrical accounts for what they consider to be special cases of polar questions, but retain a symmetrical analysis in the general case.)

However, this section, and the dissertation generally, will argue against the assumption that polar questions should receive a symmetrical analysis. Having eschewed a symmetrical analysis, we will be much better positioned to give a real explanation for what polar questions do in context, which is express a speaker's interest in a particular proposition.

In this section, we will first see evidence against a symmetrical account, coming from the felicity conditions of polar questions. We'll see that these depend crucially on properties of the radical proposition.

2.2 Felicity and bias in polar questions: evidence is not enough

Under what conditions can we ask a polar question? This problem turns out to not be trivial to address, and a full characterization of the felicity conditions of polar questions is still lacking. This problem is exacerbated by longstanding confusion about the difference, if any, between the concepts of polar-question *felicity* and polar-question *bias*; and, additionally, whether bias is a property of contexts, individuals, or linguistic expressions.

In order to perform certain speech acts, we have to ask particular questions, and not others. Büring and Gunlogson (2000) observe that positive polar questions “are not (necessarily) neutral”. In a seminal observation, they note that in the following context, *Is it*

3. Roelofsen and Farkas (2015) and Theiler (2020)'s analyses distinguish positive and negative polar questions by their discourse-oriented side effects, namely the “highlights” they introduce to the discourse, as we will see later. But insofar as the denotation of the question is concerned, they give symmetrical accounts.

sunny? is an infelicitous question, but *Is it raining?* is felicitous:

(20) *Context: Addressee walks into a windowless room, dripping wet, in a raincoat.*

What's the weather like out there? {Is it raining? / #Is it sunny?}

Büring and Gunlogson take this example, and other like it, to indicate that (positive) polar questions are infelicitous when there is evidence for their inverse proposition, i.e. the negation of their radical proposition. They refer to this as the **Evidence Condition** on positive polar questions (adapted from their (20)):

(21) **Evidence Condition on Polar Questions:**

[A polar question $p?$ is felicitous only if] there is no compelling contextual evidence against p (i.e., there is either no evidence bearing on whether p , or there is evidence for p)

In the raincoat scenario, the wet raincoat is contextual evidence for both of the propositions *It is raining* and *It is not sunny*: if the addressee comes in, dripping wet and wearing a raincoat, we are much more likely to conclude that it is raining and that it is not sunny. *Is it raining?* is therefore felicitous; there is no evidence against the radical proposition *It is raining*. In fact, there is evidence for the radical proposition, a situation in which polar questions are felicitous. *Is it sunny?*, though, is out: the raincoat provides evidence for the inverse proposition, *It is not sunny*.

But what is the nature of the Evidence Condition—and what is its relationship to bias in questions? Büring and Gunlogson (2000) don't directly refer to the Evidence Condition as expressing a *bias* of the question. Instead, they evidently conceive of the Evidence Condition as a condition of use, which relates the question to its context of utterance. Polar questions, therefore, are “not neutral” precisely to the extent that they are subject to the Evidence Condition. For example, the question *Is it raining?* can be taken to license the inference,

by the Evidence Condition, that there is no evidence to the effect that it is *not* raining. In that sense, the question itself is inherently not neutral.

Contexts, too, are “not neutral”, insofar as they provide evidence that bears on question felicity. For example, a context in which the addressee is dripping wet is not evidentially neutral, at least with respect to the question *Is it raining?*. (However, the same context *is* evidentially neutral for many other polar questions: namely, those for which a wet raincoat cannot be construed as evidence for the inverse proposition.)

For Büring and Gunlogson, then, the concept of bias doesn’t directly play into the account of polar question felicity, although it lurks under the surface. As they frame it, the speaker doesn’t use a question to express a bias that he or she holds; rather, the speaker chooses the form of their question based on evidential properties of the context. Therefore there is a systematic relationship between properties of contexts and felicitous polar questions, but it has nothing to do with the speaker’s attitude toward the radical proposition⁴. Sudo (2013) builds upon the analysis of Büring and Gunlogson (2000). Conceptually, for Sudo, all questions are “biased” in a particular sense: namely, that from their being asked, we can infer something about the context. In other words, questions always tell us something about the context in which they’re asked. Sudo takes there to be two distinct types of bias that operate independently: **evidential bias** on the one hand, related to what evidence is jointly available to conversational participants in context; and **epistemic bias**, on the other, related to what the asker of the question believes; this will become relevant in the discussion of neg-biased questions in Chapter 4.

The origin of the Evidence Condition is something of a mystery. Why are we apparently forbidden from asking questions in the presence of evidence for the inverse proposition? It is not clear why the felicity of a question should be conditioned on epistemic properties of the

4. It is an interesting question whether the role of the speaker’s bias might come from what they take to be sufficient evidence: is this something that depends primarily on the speaker’s assessment? Büring and Gunlogson (2000) don’t explore this idea, but it may be a way to get from contextual properties to speaker bias.

context; and, indeed, we might naïvely expect the opposite of the Evidence Condition: that questions are felicitous when there is evidence for the inverse proposition. Such a question would be highly informative, in the sense that it would reduce the number of active possible worlds by a great deal (cf. the analyses in van Rooy and Šafářová 2003, van Rooy 2004, which are couched in information theory).

Moreover, there are counterexamples to the Evidence Condition. One class of scenarios that contravenes the Evidence Condition is what we might call obligation-checking. Suppose that we come home and the cat is meowing and pacing around her empty food bowl. It seems likely she has not eaten. I say to you:

(22) Did you feed the cat?

In this context, the cat's behavior constitutes clear evidence that you did not feed the cat. However, somehow, the question is still felicitous, contra the Evidence Condition. Interestingly, in the same context, I can also use the same question with a negative radical proposition:

(23) Did you not feed the cat?

The Evidence Condition, in fact, predicts that this should be impossible. If questions forbid evidence for the inverse proposition, then the only contexts that license both polar questions $p?$ and $\neg p?$ should be evidentially neutral. Any context that provides evidence toward either the radical or the inverse propositions, as this one, should only allow one question form.

These obligation-checking questions are perhaps a more general case of 'suspicion' uses of polar questions, in which a speaker asks a question because of an underlying suspicion or belief, even in the presence of evidence for the inverse proposition. For instance, a bartender might say to someone seated at the bar:

(24) Are you over 21?

Presumably, being seated at a bar is evidence for being over 21. And adding evidence against the addressee's being 21 does not make the question infelicitous, either. If the addressee accidentally reveals a driver's license with a suspiciously recent date of birth, both (24) and its negative form, *Are you not 21?* are licensed.

Additionally, and perhaps more crucially, the Evidence Condition is not sufficient to determine when polar questions are infelicitous. That is, there are a number of scenarios in which polar questions are infelicitous despite meeting the Evidence Condition. This suggests that, at the very least, more needs to be said than simply the Evidence Condition to constrain the felicity of polar questions. After all, if the only requirement is that there be no evidence for the inverse proposition, then most polar questions should be licensed in most contexts. But, depending on the context, polar questions are in fact often infelicitous. This can be seen most clearly with polar questions differing with respect to polarity. For instance, to make an offer, a speaker uses the radical proposition corresponding to the state of affairs required for uptake of the offer, i.e. (typically) a positive radical:

- (25) *[Making an offer]*
- a. Do you want a beer?
 - b. #Do you not want a beer?
 - c. Do you like beer?
 - d. #Do you not like beer?

Similarly, when requesting something, a speaker forms a question built on the radical proposition that they wish to be true, independent of evidential factors. Given the nature of requesting, this, too, is usually—but not always—a positive radical (see (26-e)–(26-f)):

- (26) *[Making a request]*
- a. Do you have sparkling water?
 - b. #Do you not have sparkling water?

- c. Will you marry me?
- d. #Will you not marry me?
- e. Will you not bother me while I'm working?
- f. #Will you bother me while I'm working?

What rules out *Will you bother me while I'm working?* as a request? It cannot possibly be evidential properties of the context, i.e., that the context provides evidence for *You will not bother me while I'm working*; that is clearly not the case here. And note that the same question is perfectly fine on a sort of purely information-seeking reading: one on which, e.g., I want to know if you will bother me so that I can plan accordingly.

What seems to be relevant here is not evidence, but rather the speaker's intentions and goals—and the relation of those goals to the radical proposition of the question. When the speaker's goal is to request that the addressee abstain from bothering them, then only the negated radical is felicitous. But when the speaker's goal is something else—say, planning their day—then both the negative and positive radical are felicitous. Most of the existing literature is concerned with cases like the latter, which are what Bolinger (1978) refers to as 'inference-licensing' questions. The function of such a question is to determine whether the speaker is licensed in making the inference of the radical proposition. For example, the raincoat scenario from Büring and Gunlogson (2000):

[Addressee walks into the room dripping wet, wearing a raincoat]

Is it raining? / #Is it sunny?

Here, in terms of inference-licensing, the speaker asks the question *Is it raining?* as a strategy to determine if they can infer that it is raining from the wet raincoat. In terms of speaker goals, we might say that the speaker's goal is to find an explanation of the wet raincoat; they ask the question *Is it raining?* because they think the radical proposition, if true, would be such an explanation.

But in addition to inference-licensing, however, there are a number of other possible

speech acts that can be performed by polar questions—in the domain of imperatives, this is what Condoravdi and Lauer (2012) refer to as ‘functional heterogeneity’, and Kaufmann (2012) as ‘functional inhomogeneity’. In other words, polar questions do more than just request information. Two speech acts that we have seen polar questions perform already are offers and requests. Bolinger (1978) also highlights ‘conversation-starting’ as a function that polar questions can perform:

- (27) *[Starting a conversation]*
So, do you like baseball?

There also appear to be wish-like uses of polar questions, in which a speaker asks a question solely because they wish for the radical proposition to be true:

- (28) *[Looking out the window at the rain falling heavily]*
a. Will it stop raining soon? / #Will it not stop raining soon?
b. Will it not rain for ten minutes? / #Will it rain for ten minutes?

- (29) *[Before the addressee leaves on a long journey]*
Will you get home safe? / #Will you not get home safe?

How to account for these distinctions in felicity across contexts? The Evidence Condition doesn’t seem to provide a way forward; the infelicitous questions seem to be bad for a reason independent of what evidence is available. Instead, we need some way of referring to speaker goals in licensing polar questions in various contexts. Intuitively, what goes wrong is that the radical proposition bears the wrong relationship to the speaker’s goals. In other words, we need an asymmetrical semantics for polar questions that highlights how a polar question is about its radical proposition, and requires the speaker have a particular goal-oriented relationship to that proposition.

2.3 Polar questions are about their radical proposition

Polar questions are a domain in which the distinction between semantics and pragmatics becomes especially hazy. A large part of this problem is due to the fact that polar questions lack an obvious hook to hang a semantic analysis on. They, like questions of any kind, lack truth conditions, the hallmark explanandum of a semantic analysis; but in the analysis of wh-questions, at least, a well-defined and tractable problem has been identified in the compositional role of the wh-word. Polar questions, in contrast, are a puzzle. As a way out of this quandary, it has been often supposed, therefore, that the semantic core of polar questions relates to their response particles (in English, *yes* and *no*).

This subsection argues against two related assumptions, made to greater or lesser degrees in the literature. The first is that *yes* and *no* have a privileged relationship to polar questions: that by understanding the function of *yes* and *no* with respect to polar questions, we can understand polar question semantics. The second assumption we will dispute is that polar questions need a symmetrical semantics—and, consequently, that polar questions should be thought of as “polar” at all.

Polar questions are indeed stereotypically associated with *yes* and *no*. Analyses making this association in part grow out of the Answerhood Postulate of Hamblin (1958) and the resulting proposition-set analyses of wh-questions we have already seen. These approaches take questions to denote the sets of their answers; but it’s important to be clear on what is meant by *answer*. Strictly speaking, *Yes* and *No* are response particles, not answers. *Yes* and *No* are anaphoric to propositions, and it’s those **propositional answers** that are usually meant by *answer*. This is an especially important point to keep in mind with respect to languages with more than two response particles for questions. In Hungarian and Romanian, for example, there are dedicated “reversing” particles used to respond with a positive proposition to a negative question (Roelofsen and Farkas, 2015). However, even in those languages, a proposition-set approach would assume there to be just two propositional answers.

One simple but important reason *Yes* and *No* are stereotypically associated with polar questions is just that the folk name for polar questions is “*yes-no* questions”. This bit of terminology is unfortunately question-begging: it explicitly equates the question type with its possible (discursive) answers, and thereby forecloses any real possibility of an analysis that deemphasizes the roles of *Yes* and *No*.

It’s important to note that *Yes* and *No* are by no means linguistic universals. In addition to languages with more than simply *Yes* and *No*, some languages have no response particles at all. In Mandarin Chinese, for instance, answers to polar questions repeat the predicate of the question (N.B., POLQ is the polar question-forming particle):

- (30) A: Zhāng Sān è ma
 Zhang San hungry POLQ
 ‘Is Zhang San hungry?’
- B: è / bú è
 hungry / NEG hungry
 ‘Yes.’ / ‘No.’

And importantly, response particles can be used to respond to assertions as well as questions, as a number of authors have noted (Holmberg, 2015; Farkas and Bruce, 2009; Roelofsen and Farkas, 2015). These responses to assertions carry many of the same properties as response-particle responses to questions. For instance, and as we shall see in greater detail later, answer particle responses to polar questions in English differ depending on the polarity of the question. In particular, positive polar questions $p?$ only license *Yes*, p and *No*, $\neg p$ answers (31):

- (31) A: Do Muslims eat pork?
 B: Yes, they do. / No, they don’t.
 B’:#No, they do. / #Yes, they don’t.

Negative polar questions, in contrast, license all four combinations:

- (32) A: Do Muslims not eat pork?
B: Yes, they do. / No, they don't.
B': No, they do. / Yes, they don't.

This has been the topic of intense discussion in the semantics of polar questions. But these patterns extend beyond the domain of polar questions proper, as Farkas and Bruce (2009); Farkas and Roelofsen (2017) observe. Assertions can also be “answered” with response particles; and in those contexts, response particles exhibit the same polarity sensitivity as in response to questions. We see an example here:

- (33) A: Muslims eat pork.
B: Yes, they do. / No, they don't.
B': #No, they do. / #Yes, they don't.
- (34) A: Muslims don't eat pork.
B: Yes, they do. / No, they don't.
B': No, they do. / Yes, they don't.

We see in the response particle responses to positive and negative assertions the same pattern as in questions: positive assertions only license *Yes, p* and *No, not p*. Negative assertions, in contrast, license all four combinations of *Yes/no* with *p/not p*. This would suggest that *Yes* and *No* are not the special domain of polar questions.

These are examples of what Holmberg (2015) calls “rejoinders” to assertions, which he distinguishes from answers per se: they are not answers, because they don't respond to a question. The primary piece of evidence that Holmberg marshals for treating *Yes* and *No* differently in response to questions and assertions is that there are verb-echo languages, for example Finnish, where repeating the verb is a felicitous strategy for answering questions but not for responding to assertions:

- (35) A: Jussi puhuu ranskaa hyvin.
‘Jussi speaks French well.’
B: *Puhuu.
speaks
Intended: ‘Yes.’

This objection amounts to finding an asymmetry between answer particle responses to assertions and questions. However, we don’t need response particles to exhibit exactly the same behavior in response to questions and assertions across languages. We may very well find an independent reason that verb-echo responses to assertions are infelicitous in verb-echo languages. In English, at least, *Yes* and *No* seem to pattern alike across questions and assertions. Note that in imperatives, too, we see a difference from the polar question case; imperatives can be responded to with *no*, but *yes* is significantly less acceptable. Instead, we prefer to signal uptake with *sure*, *OK*, etc.:

- (36) A: Go to your room!
B: No. / ??Yes. / Sure. / OK. / Fine.

The entire range of conventionalized particle responses to utterances includes not just *Yes/No*, but also *Yeah/Nah*, *OK* and *Sure*, and many more. And accounting for all of them is far beyond the scope of an analysis of one utterance- or clause-type, such as questions. The response particles *Yes* and *No*, therefore, don’t appear to be a sufficient subject for the semantics of polar questions. And while any complete theory of polar questions must account for the patterns of response particle responses, that account cannot be the entire story—or even the primary subject of the account. The analysis of *Yes* and *No* as polar question responses should, ideally, make use of the same mechanisms that license response particle responses to assertions. And if the distribution of *Yes* and *No* differs cross-linguistically with respect to questions vs. assertions, this is certainly worth explaining for polar questions,

possibly in terms of their uptake conditions; but it is too strong to think that if and only if we understand polar answer responses do we understand polar questions.

If not the distribution of *Yes* and *No*, then what are we accounting for with a theory of polar questions? I argue that the primary explanandum for a theory of polar questions should be the *felicity conditions* of polar questions. Under what conditions is one particular polar question felicitous? The semantic representation that we assign to polar questions should first and foremost shed light on this issue.

It is natural to ask for someone's hand in marriage with *Will you marry me?* but not *#Will you not marry me?*: this is a consequence of the semantics of polar questions. These two questions mean different things; and when you intend to mean to ask for someone's hand in marriage, only one can do so. Even more generally, if p and q are propositions such that q means the same as $\neg p$, we also want to distinguish the questions $p?$ and $q?$. For example, consider the following two questions:

- (37) a. Is John a bachelor?
 b. Is John married?

It matters which of these questions we choose to utter in which context, even though they both resolve the same issue. As Büring and Gunlogson (2000) note for similar examples, (37-a) is odd in a context where the speaker just noticed that John is wearing a wedding band. (37-b), in contrast, is a perfectly natural question in such a context.

There are two major families of accounts that attempt to build asymmetry somewhere into the meaning of polar questions. First, there are **Manner Implicature Accounts**. These assign polar questions a symmetric semantics, and then attempt to account for the distinction between positive and negative polar questions as resulting from a Manner implicature. When a speaker chooses to utter a negative polar question, a hearer infers a reason for the speaker's use of that marked alternative, which is nevertheless semantically identical to the positive polar question. Existing Manner Implicature accounts (Romero and Han,

2004; Csipak and Zobel, 2014) are more speculative than fully articulated: to my knowledge, there is no spelled-out account getting us from synonymous question meanings with different forms to the observed differences between positive and negative polar questions.

Second, there are **Extra-Structure Accounts**: the structured-meaning account of Krifka (2001) and the highlighting account of Roelofsen and Gool (2010); Roelofsen and Farkas (2015). These accounts distinguish the radical proposition by building something on top of the otherwise symmetrical meaning of the question.

2.4 Asymmetry by Manner implicature

This subsection addresses two accounts, due to Romero and Han (2004) and Csipak and Zobel (2014), that attribute polar question asymmetry to something like a Manner implicature (Grice, 1975). According to these accounts, polar questions built from a positive radical ($p?$) and polar questions built from a negative radical ($\neg p?$) have the same meaning: a partition of the Common Ground into two cells. In order to retrieve the identity of the radical proposition, they make reference to properties of the speaker's **manner of expression**, namely the way that the speaker chose to pronounce the question.

First, Romero and Han (2004) are primarily concerned with accounting for the interpretation of negation-biased questions, as in (38):

(38) Aren't you hungry? (cf: Are you hungry? Are you not hungry?)

These questions, appearing with high negation cliticized to the auxiliary verb, come with a characteristic "biased" interpretation that has been the subject of a great deal of debate. However, as negation bias is the subject of Chapter 3, we'll set that part of their analysis aside for now; an in-depth account of their approach to negation bias can be found in Chapter 3.

For now, then, we'll focus on their characterization of the relationship between the two

“cells” of the partition determined by a canonical (non-negation-biased) polar question. Building on an observation from Bolinger (1978), Romero and Han (2004) note that polar questions built on a positive radical can be felicitously used to make a request (39), or to ask whether an inference is licensed (40), but negative polar questions cannot (their examples 88–89):

- (39) a. Will you (please) help me?
b. #Will you (please) not help me?

- (40) a. What’s the matter? Why aren’t you working? Are you tired?
b. What’s the matter? Why aren’t you working? #Are you not tired?

Romero and Han (2004) speculate that something about being the “pronounced cell” of the polar question is relevant here. Because the speaker chose to pronounce that cell, she must have some special interest in that proposition. Consequently, they identify the choice of which cell to pronounce (i.e., whether to form a question from a positive or negative proposition) as bearing in on the speaker’s “intent” regarding the question. But the exact mechanism underlying this process—as well a clear definition of “pronounced cell” or speaker “intent”—is left unspecified.

Csipak and Zobel (2014) present an ingenious illustration of the asymmetry problem, although they don’t phrase it as such. Suppose that two teams played a baseball game: the Tigers and the White Sox. Suppose furthermore that I know that the game is over, and that one team won, but I don’t know which team it was. A partition analysis for polar questions predicts the following questions all have the same meaning.

- (41) a. Which team won?
b. Did the Tigers win?
c. Did the White Sox win?

(Note that these questions should be strictly synonymous on a partition account, as *which* carries a presupposition that exactly one of the relevant alternatives satisfies the relevant property.) Csipak and Zobel (2014) observe that in German, there is a difference in the licensing of certain discourse particles across question types. In particular (among other particles), the constituent wh-question (41-a) does not license the particle *etwa*, but the polar questions do. This differential felicity is unpredicted if we assign all the sentences in (41) the same semantic value.

However, Csipak and Zobel (2014) draw an odd conclusion, at least from our perspective exploring an attitudinal approach to polar questions. Rather than concluding that this differential felicity is evidence that polar questions and wh-questions have different semantic representations, they presuppose that the representations of these two question types are indeed the same. Instead, they take this differential felicity as evidence that “partition structure is not a decisive factor” in accounting for the licensing of *etwa*. By *partition structure*, presumably they mean the way in which a question partitions logical space. But this presupposes that polar questions have the same partition structure as (certain) wh-questions (in certain contexts).

Csipak and Zobel conclude instead that *etwa* is sensitive to properties of the “Explicitly Identified Answer”, which is the radical proposition. In particular, a question containing *etwa* expresses, according to them (an adaptation of their (29a)):

- (42) $\llbracket \text{etwa} \rrbracket(?, \phi)$: The speaker realized he mistakenly believed the inverse proposition to be more likely than the Explicitly Identified Answer [i.e., the radical proposition]

Since wh-questions don’t have an Explicitly Identified Answer, they don’t license *etwa*. And while this is good as a characterization of the particular sensitivity of polar questions with *etwa* to their radical propositions, it can’t constitute a whole story for polar questions. In particular, what their story lacks is a way of distinguishing the radical proposition in the meaning of the question: an asymmetric representation for the question, in other words.

Interestingly, both Csipak and Zobel (2014) and Romero and Han (2004) identify **pronunciation** (equivalently, for Csipak and Zobel, **explicit identification**) as the operative factor for distinguishing pairs of polar questions differing in polarity. That is, they assume the partition analysis is correct for polar questions: the questions $p?$ and $\neg p?$ really *are* synonymous. Having committed to this relatively strong claim, then, they are forced to distinguish the two questions at the level of discourse, by making reference to *what expression* the speaker chose to explicitly pronounce. In effect, accounts of this kind must rely on something like reasoning about the Gricean maxim of Manner (Grice, 1975) to distinguish (what they take to be) otherwise synonymous questions. Given two utterances with identical semantic content, differences in the manner of their production account for their different properties.

Romero and Han's Manner-invoking account is complicated, though, by an incomplete characterization of the (in)felicity of positive and negative polar questions. Recall from (39) and (40) that they claim positive, but not negative, polar questions can be used to perform requests and ask whether an inference is licensed. Though this is true for the particular contexts they present, it's not true in general. A negative question can be a request for the addressee to abstain from an action, as in (43). Similarly, a negative question can ask whether a negative inference is licensed, as in (44). In those cases, we get exactly the opposite pattern; the questions need negation, or else they're infelicitous in the relevant context.

- (43) a. Will you (please) not bother me while I'm working?
 b. #Will you (please) bother me while I'm working?

- (44) a. What's the matter? Why aren't you eating? Do you not like Thai food?
 b. What's the matter? Why aren't you eating? #Do you like Thai food?

In (43), the speaker is indeed performing the speech act of a request; they're requesting that the addressee refrain from bothering them. Similarly, in (44), the speaker asks whether the

inference *You don't like Thai food* is licensed. It's clear, then, that questions with negative polarity can perform the speech acts discussed by Bolinger (1978), contra Romero and Han (2004).

Data like (43) and (44) are inherently problematic for an account that wants a Manner-implicature explanation of the difference between a positive and negative polar question. Manner implicature accounts generally associate structural complexity or marked forms with abnormal or marked circumstances (Levinson, 2000). For instance, the choice of the complex locution *cause to die* rather than *kill* suggests an indirect or complex event—for instance, *cause the mouse to die* might describe an event of placing the mouse somewhere it can't escape, where it eventually dies. *Kill*, conversely, suggests a comparatively stereotypical event, say, catching it in a trap:

- (45) a. Mary caused the mouse to die.
b. Mary killed the mouse.

In the case of polar questions, though, the structurally more complex form is always going to be the question built from the negative proposition. Consequently, we'd expect questions with negative radicals to have strictly more limited distribution or to be used in more atypical situations. But this isn't what we find—e.g. in (44) or (43). In those examples, the speaker's interest was in a negative proposition, so they asked about a negative proposition. This context-sensitivity is going to be unpredicted by a Manner implicature account, which, making reference to properties of the speaker's manner of expression, should not be sensitive to discourse context.

And in general, this sort of account is a conceptually unappealing way to characterize the difference between these ostensibly synonymous questions; it occludes the compositional relationship between the question and the radical proposition. The reason that the speaker's "intent" (in the words of Romero and Han 2004) varies with the choice of radical proposition isn't that the radical proposition happens to be the answer that's pronounced; it's because

the question is, necessarily, compositionally built up from the radical proposition.

Both of these accounts, when they make reference to something like “the pronounced cell”, are trying to retrieve semantic content that isn’t actually there in the denotation of the question by referring to the manner of expression the speaker chose. They want to make reference to p , the radical proposition; but that proposition is “invisible” at the level of the meaning of the question. Instead, they want to retrieve that semantic content by getting it from properties of the speaker’s choice of words.

But retrieving the proposition from the speaker’s “choice of pronunciation”, rather than using that proposition in the derivation at the point that it’s necessary, is fundamentally non-compositional. That proposition is a semantic object; but the speaker’s choice of manner of expression is a syntactic (or even phonological) object. We can’t assign the two questions the same meaning, and then retrieve the radical proposition from the utterance after it’s been uttered.

It’s important to note that Romero and Han (2004), especially, are not primarily occupied with the question of distinguishing positive and negative polar questions. As a result, their discussion of the pronounced cell shouldn’t be taken as suggesting a complete or well-articulated theory of this distinction. They themselves acknowledge that this distinction might be cashed out in semantic, rather than pragmatic, terms. The argument we make here, though, is that the language of “pronounced cell” cannot be taken literally; at best, it can be a mnemonic for the radical proposition. What we need is an account that builds asymmetry into the meaning of the polar question. One such account is the highlighting mechanism introduced by Roelofsen and Farkas (2015), the topic of the next subsection.

2.5 Asymmetry by extra structure

Roelofsen and Farkas (2015), building off of Roelofsen and Gool (2010), use a semantic operation of “highlighting” couched in Inquisitive Semantics to retrieve the identity of the radical proposition. Starr (2009), earlier, also suggests a highlighting approach, but he

doesn't spell it out, leaving it a suggestion for future work. For both Roelofsen and Gool (2010) and Roelofsen and Farkas (2015), this is useful especially for modeling response particle patterns in particular, and the operation is developed primarily to achieve this end.

Intuitively, highlighting is a way of tracking propositions (*possibilities* in the Inquisitive Semantics framework) that are made salient by an utterance, typically for later anaphora. In the case of polar questions, a positive polar question and a negative polar question highlight a positive and a negative proposition, respectively:

(46) Are you hungry? (highlights *You are hungry*)

(47) Are you not hungry? (highlights *You are not hungry*)

Already we can see from these examples that highlighting as a way of achieving asymmetry represents a significant improvement over the Manner-implicature suggestions of Romero and Han (2004); Csipak and Zobel (2014). Since polar questions necessarily highlight their radical proposition, we won't necessarily predict that negative questions—being the marked variant—correlate with atypical situations.

Highlighting, though, is primarily developed to account for *Yes* and *No* answers to polar questions and assertions, not for the felicity conditions of polar questions. One consequence of this is that it is built to track highlighted negative and positive propositions separately. But it's not clear that for an account of polar question felicity conditions, we need to do this. Roelofsen and Farkas (2015), on the other hand, want to track propositions of negative and positive polarity separately because the meanings they assign to *Yes* and *No* depend on the polarity of the question.

A brief summary of the highlighting approach to polar questions follows. First, as we have noted, Roelofsen and Farkas (2015) are (insofar as highlighting is concerned) primarily interested in accounting for the different uses of *Yes* and *No* in response to polar questions

of different polarity. In order to explain this, we briefly need to introduce their feature-based account of answer particles.

The empirical puzzle is that in responses to positive questions, *Yes* and *No* can only associate with a positive or negative proposition, respectively, as in (48). But in response to a negative question, both can associate with either the positive or negative proposition (Kramer and Rawlins, 2009), as in (49).

- (48) A: Are you hungry?
B: Yes, I am. / No, I'm not.
B': #No, I am. / #Yes, I'm not.

- (49) A: Are you not hungry?
B: Yes, I am. / No, I'm not.
B': No, I am. / Yes, I'm not.

What is it about negative questions that allows every combination of *Yes / No* and *I am / I'm not*, when positive questions don't allow it? Roelofsen and Farkas (2015) give a feature-based framework to answer this question. They propose that four polarity features are relevant: [AGREE], [REVERSE], [+], and [-]. [AGREE] and [REVERSE] are **relative features**: they encode whether the propositional answer agrees in polarity with the question, or whether it is the opposite of the question. [+] and [-] are, on the other hand, **absolute features**: they encode the polarity of the propositional answer without reference to the question: either positive or negative. Response particles are then the spellout of any of the following four polarity feature bundles:

- (50) a. [AGREE, +]
b. [REVERSE, +]
c. [AGREE, -]
d. [REVERSE, -]

The last component of the feature-based approach to *Yes* and *No* is that any bundle containing [AGREE] or [+] can be realized as *Yes*, and any bundle containing [REVERSE] or [-] can be realized as *No*. Consequently, we have the following possible spellouts for each bundle.

- (51) a. [AGREE, +] = *Yes, p*
 b. [REVERSE, +] = *Yes, p* **or** *No, p*
 c. [AGREE, -] = *Yes, not p* **or** *No, not p*
 d. [REVERSE, -] = *No, not p*

We can observe that, in response to a positive polar question, only (a) and (d) are coherent feature bundles: agreeing in polarity with the question necessitates a positive-polarity answer, and reversing the polarity of the question yields a negative-polarity answer. Thus, only *Yes, p* and *No, not p* are possible answers to positive polar questions, as they are the only possible realizations of those two feature bundles.

For negative polar questions, contrarily, only (b) and (c) are coherent feature bundles. We can reverse the polarity of the question, yielding a positive answer; or we can agree with the polarity of the question, yielding a negative answer. But these bundles can both be spelled out as *Yes* or *No*, because of our rules for spellout, which say that anything containing either [AGREE] or [+] can be spelled out as *Yes*, and anything containing either [REVERSE] or [-] can be spelled out as *No*. Consequently, negative polar questions allow *Yes, p*, *No, p*, *Yes, not p*, and *No, not p* as possible answers.

Crucially, in order for this theory to work out, we need to be able to refer to the polarities of the question radical and the propositional answer. In particular, the answer particle needs to be able to retrieve the polarity of the question radical for the **relative features** [REVERSE] and [AGREE], and the polarity of the answer proposition for the absolute features [+] and [-].

Roelofsen and Farkas (2015) achieve this by introducing **highlighting**. Expressions are assumed to systematically add propositional discourse referents to a highlighting structure,

which for a proposition ϕ is an ordered pair $\llbracket\phi\rrbracket^\pm = \langle P, N \rangle$, where P is a set of positive propositions and N is a set of negative propositions that are available for anaphoric relations, like those invoked in the use of response particles. $\llbracket\phi\rrbracket^\pm$ for a proposition ϕ is defined recursively as follows (a subset of Roelofsen and Farkas 2015's (60), glossing over disjunction and irrelevant aspects of Inquisitive Semantics):

- (52)
- | | | |
|----|---|--|
| a. | $\llbracket p \rrbracket^\pm = \langle \{p\}, \emptyset \rangle$ | Positive highlights for p |
| b. | $\llbracket \neg\phi \rrbracket^\pm = \langle \emptyset, \{\overline{\cup\llbracket\phi\rrbracket^\pm}\} \rangle$ | Negative highlights for $\neg\phi$ |
| c. | $\llbracket ?\phi \rrbracket^\pm = \begin{cases} \langle \emptyset, \{\alpha\} \rangle, & \text{if } \llbracket \phi \rrbracket^\pm = \langle \emptyset, \{\alpha\} \rangle \\ \langle \{\cup\llbracket\phi\rrbracket^\pm\}, \emptyset \rangle & \text{otherwise.} \end{cases}$ | Highlights for question $?\phi$ |

In plain language, highlighting is computed recursively, presumably over the course of the derivation. In the base case, an atomic proposition p introduces a positive highlight p . Negation of an expression ϕ introduces a single negative highlight consisting of the complement of ϕ . Questions inherit the highlights of their radical propositions. (We are ignoring disjunction and conjunction here, as it isn't directly relevant to our discussion.)

Finally, a presupposition associated with each feature ensures that feature bundles spelled out as *Yes* or *No* guarantee the correct semantic properties (Roelofsen and Farkas's (68–71)):

- (53)
- | | |
|----|--|
| a. | <p>[+] presupposes that its prejacent expresses a proposition containing a single possibility α, which is highlighted and has positive polarity:</p> $\llbracket \text{prejacent} \rrbracket = \{\alpha\}^\downarrow \text{ and } \llbracket \text{prejacent} \rrbracket^\pm = \langle \{\alpha\}, \emptyset \rangle$ |
| b. | <p>[-] presupposes that its prejacent expresses a proposition containing a single possibility α, which is highlighted and has negative polarity:</p> $\llbracket \text{prejacent} \rrbracket = \{\alpha\}^\downarrow \text{ and } \llbracket \text{prejacent} \rrbracket^\pm = \langle \emptyset, \{\alpha\}^\downarrow \rangle$ |
| c. | <p>[AGREE] presupposes that its prejacent highlights a unique possibility α, and that the context provides a unique most salient antecedent possibility β such that α and β contain precisely the same possible worlds and have the same</p> |

polarity.

- d. [REVERSE] presupposes that its prejacent highlights a unique possibility α , and that the context provides a unique most salient antecedent possibility β such that α is the complement of β and has the opposite polarity.

With all these parts in place, we have an account of polar-question answer patterns that correctly models the English facts, as well as providing a framework that can be extended to languages with other patterns, e.g. Romanian or Hungarian.

For Theiler (2020), highlighting allows us to distinguish between otherwise semantically identical polar and wh-questions, which is for her necessary for an analysis of the discourse particle *denn* in German. To illustrate the contrast, she provides a similar example to Csipak and Zobel (2014)'s example of the two teams (Theiler's (4)):

(54) [A and B know exactly two people called Anna. One of them lives in Munich, the other one in Berlin. This is commonly known among A and B.]

A: Vorhin hat Anna gerufen.

Earlier today, Anna called.

B: Welche Anna meinst du denn?

Which Anna do you DENN mean?

B':#Meinst du denn Anna aus München?

Do you DENN mean Anna from Munich?

As with the White Sox and Tigers example earlier, this context makes it so that there are only two possible answers to the *which* question. Consequently, a partition analysis for a polar question should yield the same semantic representation for the wh-question and the polar question.

Highlighting as a semantic mechanism provides the propositional discourse referents that are tracked in discourse. Highlighting is a formal bit of machinery that tracks the discourse

referents and propositions that we need for various semantic purposes: for Roelofsen and Farkas (2015), explaining patterns of response particles and *so* anaphora; for Theiler (2020), additionally, licensing certain discourse particles by referring to properties of highlighted content. But this mechanism raises a number of questions that aren't fully addressed by any of the treatments of highlighting that currently exist.

It's important to note that this isn't a criticism of highlighting as a formal tool to analyze a particular phenomenon, like polar-answer response particle patterns. Highlighting is akin to the background of structured-meaning denotations for polar questions in Krifka (2001); it retains some information that we need later in the derivation, which may very well be theoretically necessary. But the primary conceptual problem with highlighting as a way of encoding polar-question asymmetry is that it doesn't clearly exist as a phenomenon on its own. Apart from a relatively vague notion that highlighting is the way in which an expression "makes a proposition salient", we don't have a good understanding of what a highlight really is.

To clarify this objection, it's instructive to compare highlighting to a structurally similar but better-understood and better-motivated phenomenon: focus. According to standard theories of focus (e.g. Rooth, 1992), focus alternatives are calculated in a similar recursive way through the course of the derivation of a sentence⁵. Like highlighting for Roelofsen and Farkas (2015), the **focus semantic value** of a sentence is computed at a separate tier from the conventional **ordinary semantic value**: in parallel, but not directly interacting. While the ordinary semantic value of an expression tracks its conventional truth-conditional meaning, the focus semantic value tracks relevant alternatives introduced by the focus structure of the expression. For instance, the following two sentences, with focus on the indirect object and the direct object, respectively, have the same ordinary semantic value, but different focus semantic values.

5. Strictly speaking, Roelofsen and Farkas (2015) don't provide a compositional analysis of the highlights of a sentence; instead, they provide a recursive definition over expressions in the Inquisitive Semantics metalanguage.

- (55) Ida introduced Gus to TOMMY.
- a. Ordinary semantic value: **introduce**(t)(g)(i)
 - b. Focus semantic value: {**introduce**(x)(g)(i) | $x \in E$ }
- (56) Ida introduced GUS to Tommy.
- a. Ordinary semantic value: **introduce**(t)(g)(i)
 - b. Focus semantic value: {**introduce**(t)(x)(i) | $x \in E$ }

The focus semantic value is, intuitively, a set of alternatives of the same type as the expression in question, generated by abstracting over the semantic position of the element in focus. The focus semantic value is then made available to focus-sensitive operators—for instance *only*, among others.

Focus is typically conceived of as part of a strategy for discourse management: focus marks information as discourse-new, tracking what is “backgrounded” and what is not in a conversational context (Zimmermann and Onea, 2011). Focus is marked within linguistic expressions using a variety of strategies across languages. In English, focus is marked primarily with intonational prominence (Selkirk, 1995; Büring, 2006). In languages like Hungarian, there is a dedicated position in the clause (É. Kiss, 2002) for focused items. In Kwa (Niger-Congo) and others, focused items are marked with an overt focus-marker morpheme, e.g. *wè* in Kwa specifically (Aboh, 2007). Although semantic focus is known not to correspond one-to-one with overt syntactic markers of focus (Zimmermann and Onea, 2011), there is nevertheless a close relationship between the two. And focus, analyzed in terms of alternatives, bears a clear relationship to wh-questions, which similarly generate a set of alternative propositions, abstracting over the position of the wh-word.

Focus, then, is a phenomenon with various reflexes in different languages, one for which we need a semantic account. Highlighting, though, is much more mysterious. First, it’s unclear that highlighting is a coherent *phenomenon* at all. The need for highlighting exists only within a particular analysis of a particular phenomenon; namely, an analysis of

questions that doesn't provide some other way of retrieving the identity of the radical of the question. But there is not an obvious candidate for an independent, theory-neutral linguistic phenomenon corresponding to highlighting, in the same way that focus corresponds to the semantic accounts that we propose for it.

Consequently, the limits of what we want to explain with recourse to highlighting are unclear. With focus, there is an intuitive domain of interest: expressions get focus-marked, which systematically contributes to the meaning of the sentences that they occur in. For instance, when I utter (57) vs (58), with focus on different constituents, the structure of my contribution to the discourse is systematically different:

(57) Ida ate the ICE CREAM BARS.

(58) IDA ate the ice cream bars.

In (57), we might take it that I presuppose that Ida ate something, and I assert that it was the ice cream bars. In (58), I presuppose that someone ate the ice cream bars, and I assert that this person was Ida. The information that is backgrounded, and the novel information presented as my own contribution, is clearly tightly correlated with the choice of expression and what is marked as focused.

It is not clear that highlighting has an overt realization in any language. For polar questions in English, highlighting seems not to be so much “marked” as simply a reflection of the fact that polar questions are built from one proposition rather than another. I am not aware of claims that any language overtly marks highlighted content in a uniform or semi-uniform way, as they do focused content. Consequently, we want to attribute to highlighting in the first place: how do we constrain it as a theoretical tool? What we want is a way to characterize *why* a polar question is about its radical proposition. And without a better story about what highlighting is, a highlighting account will struggle to explain the aboutness of a polar question; it can only do so insofar as it equates aboutness with being introduced into

the highlighting structure.

An additional puzzle, due in part to the lack of clarity about the identity of highlighting as a phenomenon, is how we might distinguish operations on highlighted vs. asserted content in the general case. A proposition p introduces a positive highlight p , and $\neg p$ introduces a negative highlight $\neg p$; in many cases, then, highlighted and asserted content are one and the same. We could, therefore, reframe a great deal of our pragmatics or conventions of use in terms of highlighting.

Finally, and most critically, highlighting still only gets us part of the way toward an account of polar question felicity. Specifically, it doesn't tell us how the felicity of a question is related to the context in which it is used. We know that felicity conditions depend on the identity of the question radical; but highlighting only gives us the radical, and doesn't tell us a story about how the choice of radical relates to the context. At the very least, then, we will need some sort of account on top of a highlighting account. This is a problem shared with Krifka (2001)'s structured-meanings account of question meanings: even with an analysis that formally disambiguates polar questions by polarity, we need a story for explaining why *that* question is felicitous in *that* context.

Summing up, then, the issue with highlighting as a general approach to polar-question asymmetry is that we don't have a good understanding of what highlighting consists in, if anything, as a linguistic phenomenon. But even with it, more needs to be said to derive the felicity conditions we observe.

2.6 Conclusion

This section has presented a number of approaches to the semantics of polar questions; but each is largely concerned with the behavior of question-answer pairs, rather than the role of questions within discourse, the various speech acts they can perform, and why conversational participants might choose one question over another. What do we want instead? Ideally, the privileged status of the radical proposition should follow straightforwardly from its position as

the radical of the question. We shouldn't have to stipulate that the radical has a special status because the rules for highlighting guarantee that the radical is highlighted; the aboutness of the question should follow in a natural sense from its representation.

CHAPTER 3

AN EXPRESSIVE-ATTITUDINAL ANALYSIS

We have seen in the previous chapter that a semantics for polar questions ought to be able to distinguish the radical proposition: that is, it must be asymmetrical. Ideally, too, the representation of the question should tell us something about the function of the question in various contexts. The current proposal, then, is simple: polar questions express the speaker's preference to know the radical proposition, if it is true. This meaning is asymmetrical in the sense that the speaker's preference is toward the radical proposition ϕ , which could be of negative or positive polarity; and preferring to learn p is not the same as preferring to learn $\neg p$.

A number of the outstanding problems in polar-question semantics become clearer if we take an attitudinal approach to their semantics. In particular, this sort of approach can inform our answers to longstanding issues regarding the felicity conditions and discourse effects of polar questions. Accounts on which a polar question denotes merely a set of propositions require a lot of theoretical machinery to get to the level of discourse effects or felicity conditions, and there is no satisfactory story on the market yet that relates those to the semantic representation in a principled way.

In recent years especially, a number of approaches have built into the semantic representation of imperatives what might more usually be left to the pragmatic component of language. This is what von Stechow and Iatridou 2017 call *strong theories* of imperatives—in contrast with *minimal theories*, which assign a bare-bones semantic value to imperatives. The analysis in the present paper is in part inspired by strong theories of imperative semantics, in particular those which take them to be fundamentally modal (Lewis, 2001; Portner, 2004; Kaufmann, 2012, 2020; Kaufmann and Kaufmann, 2021, i.a.).

Interrogatives have not been the beneficiary of a slew of recent modal approaches in the same way that imperatives have. We can speculate about reasons for this: one is likely that the field of question semantics is dominated by wh-questions, which is generally more

concerned with more tractable, traditionally semantic phenomena like scope and exhaustivity than with issues like felicity conditions, pragmatic effects, and so on. As a consequence, the usefulness of “strong theory of questions” has not really been considered. However, a good deal of recent work has acknowledged the need to build more structure into the meaning of polar questions in order to account for high-negation bias in particular¹ (Romero and Han, 2004; Repp, 2007; Northrup, 2014; Krifka, 2017; Goodhue, 2018; AnderBois, 2019, i.a.). Consequently, there is a gap for a “strong theory of questions” in the literature.

As a first pass, we might paraphrase the attitudinal approach to polar questions as expressing something paraphrasable as, *I want to know the true proposition from between p and $\neg p$* ; this is similar to the ‘tell me truly’ family of accounts cited by Harrah (2002). In other words, questions would receive a logical paraphrase like the following:

- (1) a. Are you hungry?
- b. Tell me truly that you’re hungry or that you’re not hungry.

But this sort of account doesn’t address the asymmetry problem for polar questions: you can’t retrieve the radical from the representation of the question, and even if we distinguished it, it’s not clear how that would be a question about the radical. A modal approach, however, should allow us to build asymmetry into the semantics for polar questions in a reasonably elegant alternative way: by not including the inverse proposition $\neg p$ in the scope of the modal operator at all, but by using a conditional instead.

As a first, pretheoretic natural-language approximation, then, we can understand the meaning of a question $p?$ to be like the following:

- (2) **Paraphrase of a polar-question meaning $p?$:**
If p , I want to believe p .

1. For our analysis of high-negation bias, see Chapter 4.

In other words, a polar question expresses the speaker's desire to have p in their doxastic state, if true. As it stands, this paraphrase needs to be sharpened before it's practical. In particular, we don't want this meaning for a question to be an assertion; in that sense, it's not like a normal attitude report. For instance, attitude reports, but not polar questions, can be responded to with *That's true/false*, and so on:

- (3) a. A: I want to know if you're hungry.
B: That's not true. You don't care if I'm hungry or not.

- (4)
(4) A: Are you hungry?
B: #That's not true. You don't care if I'm hungry or not.

The polar question doesn't seem to proffer the proposition *I want to know if you're hungry* in the same way as the assertion does: that proposition can't be referred to with anaphoric *that*, for example. And we have the clear intuition that questions cannot be true or false. For this reason, we'll pursue an analysis for polar questions along the lines of that proposed by Chernilovskaya et al. (2012a) for wh-exclamatives. On that account, exclamatives like *What big teeth you have!* do not put forth a proposition to be accepted by other conversational participants, but rather "directly" add the expressive content to the common ground: in the case of exclamatives, that the speaker is surprised by the addressee's degree of tooth-size.

Similarly, polar questions directly add the propositional content that the speaker wants to know so-and-so to the common ground, without it being 'placed on the table' at all. In this way, conversational participants will ultimately be able to refer to the radical proposition by anaphora, but not to the expressive content of the question, which is the speaker's desire to know the radical proposition.

3.1 Desire in polar questions

As for *want*, importantly, this should be understood as something like the performative directive use of *want* in the following examples, taken from Condoravdi and Lauer (2011). In examples like these, the utterance of a *want*-sentence performs speech acts stereotypically associated with imperatives: demanding, requesting, advising, conceding, and so on.

- (5) a. *[Mother to child]*
I want you to clean your room before playing.
- b. *[Doctor to patient]*
I want you to take these pills for a week.
- c. *[Affirming an offer]*
I want you to take the last cigarette.
- d. If it is that important to you, I want you to go.

These examples generally don't admit a purely informational interpretation, on which the speaker is asserting something about his or her desires. As a consequence, they don't allow responses that target their propositional content, like *#That's true/false*, *#You're lying*, and so on. Instead, appropriate responses to these *want*-sentences involve uptake (or not) of the command, advice, etc.: *OK*, *I will*, or *No way*, for example.

Kaufmann (2012) argues that this is the source of the sense of truth-value-lessness that imperatives carry: they do carry truth values in a strict sense, but their function is to command and not to assert, like the *want*-sentences above. This is semantically achieved by building in two presuppositions to the modal semantics of the imperative. First, Kaufmann (2012) has imperatives presuppose a *decisive* modal flavor: conversational participants will make decisions based on what is true according to it. Second, imperative modality carries *epistemic authority* on the part of the speaker: the speaker knows the epistemic-modal state of affairs perfectly. When these two conditions conspire, Kaufmann argues, modals, even of

the non-imperative variety, have an imperative, rather than assertive, interpretation.

In a similar way, we can take the meaning of a polar question to be an expression of a desire that, rather than simply asserting something true or false, performs the function of *asking if*: proffering a possible state of affairs, and requesting to be informed if that state of affairs holds.

Kaufmann (2012) identifies three conditions on directives, including bouletic modals, that she takes to be necessary for performative readings of deontic modals. First, **Authority**: the utterer of the directive must be an epistemic authority. For Kaufmann, this means that it's a joint belief in the discourse that the speaker knows what should be done. Second, an **Epistemic Uncertainty Constraint**: the speaker must not take the truth of the prejacent for granted.

Finally, there is an **Ordering Source Restriction**: briefly, that there's a salient decision problem that carrying out the directive resolves, or else that the speaker is expressing a personal preference. This restriction is meant to account for the inference that speakers want their imperatives to be carried out, and all conversational participants take it to resolve the relevant decision problem. For instance, B's answers in the following discourse are infelicitous:

- (6) A: How do I get to Harlem?
B: Take the A train, although...
...I don't care if you do.
...maybe it doesn't go to Harlem.

Kaufmann's criteria for directives are summarized below:

- (7) Kaufmann (2012)'s criteria for directives
- a. **Authority**: Conversational participants take the speaker to be an epistemic authority on what one must do in the relevant context.

- b. **Epistemic uncertainty:** The speaker takes both the prejacent and its inverse to be possible.
- c. **Ordering source restriction:** Conversational participants take there to be a salient decision problem in the context that is resolved by carrying out the directive, or else it expresses bouletic modality.

For polar questions, the situation will be somewhat different, as we analyze them as expressing desire and not deontic modality *per se*. Therefore, we can see these as something more like heuristics than requirements. But it is worth noting that these conditions generally obtain for the desire expressed by polar questions as well. The **authority condition** is typically trivial: speakers are generally taken to be epistemic authorities on what they themselves want to know, except perhaps in very odd contexts. It is generally not felicitous to respond to a polar question with *Hey, wait a second—you don't have the authority to ask me that!*²

The **epistemic uncertainty condition**, transposed onto polar questions, amounts to the condition that the speaker is ignorant about which of the radical or inverse propositions holds. This is indeed generally the case; utterances like the following, which expressly contradicts the epistemic uncertainty condition, are infelicitous.

(8) I know whether Finland was in the USSR. #Was Finland in the USSR?³

The primary counterexample to epistemic uncertainty—i.e., an example of polar questions used without ignorance—is **quiz questions**: questions used to assess the addressee's knowledge. When a teacher includes a question on a quiz, it is generally not taken to imply

2. For an example where this is the case, we might imagine a context in which a soldier is speaking with the President of the United States about a clandestine mission. There, it might be the case that the soldier is jointly taken not to have epistemic authority about what he should know: the President decides what information is knowable. In that case, some polar questions from the soldier might be infelicitous; the President might respond *You can't ask that*. Other contexts involving the law and privileged information may come to mind as well.

3. Note that this is much better with *Do you think Finland was in the USSR?*

ignorance on his part. In *wh*-questions, for example:

(9) [On a quiz]

When was the French Second Republic founded?

But it's worth noting that quiz questions do not generally come in the form of polar questions—at least in English. If we want to quiz someone's knowledge of a particular proposition, we generally use true/false questions and not polar questions. Witness the oddness of (10-b), compared to (10-a):

(10) [On a quiz]

- a. True or false: The French Second Republic was founded in 1850.
- b. ? Was the French Second Republic founded in 1850?

A question like (10-b) is much more natural in a conversational context than in a quiz context: it really seems to express that the speaker wants to know something. To my mind, then, allowing for polar questions to have non-ignorance implying quiz readings is not a desirable goal for a theory of polar questions.

Kaufmann's final condition, the **ordering source restriction**, is highly relevant for polar questions. This is necessary to get the right sense for *want*: we have been paraphrasing polar questions with the word *want*, but the relevant relation is usually a sort of preference of the speaker **in light of his own goals**: e.g., because learning it is a necessary or sufficient condition to achieving those goals. Someone usually asks a question because they consider learning the truth of the radical proposition relevant to their own actions.

Arguably, though, we can get non-goal-oriented readings, as in 'wish' readings for imperatives (Kaufmann, 2012). Someone might ask a question simply because they want the answer to be *Yes*. Imagine a small child who asks a parent:

(11) Can we have chocolate cake for breakfast every day?

or who wakes up on December 23 and asks:

(12) Can it be Christmas yet?

These sorts of sentences are most felicitous with an existential modal; ?? *Will we have chocolate cake for breakfast every day?* is markedly stranger. But a similar wish-like reading for polar questions, without any overt modality, can be seen in what I'll term **hope readings**—polar questions where the speaker expresses a wish to know the affirmative answer. For instance:

- (13) a. [After cooking a labor-intensive meal]
Is it good?
- b. [Before the addressee goes on a dangerous hike]
Will you get home OK?
- c. [Waking up after a week of constant rain, before looking out the window]
Has the rain stopped yet?

It's hard to imagine that the speaker in these scenarios, especially in the latter, is genuinely asking of her interlocutor which of the two propositions is true. Asking *Will you get home OK?* isn't a way of asking a neutral, information-seeking question, proffering two alternative propositions; rather, it is a way of expressing a desire to know the radical proposition *in particular*. The speaker is not really willing to countenance a *No* answer at all. For an analysis where a polar question means *If p, I want to believe so*, hope readings are not surprising. The reason the speaker would like to know *You'll get home OK* if it's true is that learning it would assuage their worry.

3.1.1 Modeling polar-question desire

How do we model the *want* component of a polar question semantically? We might have it that *x wants that p at w* is true iff *p* is true in every world in $Boul_x^w$, the set of worlds consistent with *x*'s desires at world *w*:

(14) **Preliminary analysis of *want*:**

$$\llbracket x \text{ wants } p \rrbracket^w = \forall w' \in Boul_x^w : p(w) = \text{True}$$

But this simple picture is not sufficient for an analysis of *want*, as Stalnaker (1987); Heim (1992) show. The problem with this approach is that *want* is not upward-entailing. That is, if I want *p*, and *p* entails *q*, it does not follow that I want *q*. But if we take desire to be characterizable by a set of worlds consistent with what one wants, then it should be upward-entailing. The following is one of Heim's primary examples of this puzzle, borrowed from Stalnaker (1987):

- (15) a. I want to get well.
b. I want to have been sick.

Clearly, *I get well* entails *I have been sick*. But (15-a) does not in turn entail (15-b). If (15-a) is true, according to our preliminary analysis, then all my bouletic worlds are ones in which I get well. Since getting well entails having been sick, all those worlds are also worlds in which I have been sick. Therefore, from (15-a) we predict that we can infer (15-b): all my bouletic worlds are ones in which I have been sick, too.

Heim proposes to solve this problem by implementing Stalnaker (1987)'s insight that to want something is to prefer it to be true. Following the approaches of Lewis (2001) and Kratzer (1981), Heim proposes that desire involves agents maintaining an ordering over worlds, according to their preferences. First, agents maintain a **preference ranking over worlds** \leq_x^w :

(16) For any entity $x \in E$, $w, u, v \in W$:

$u \leq_x^w v =_{\text{def}}$ In w , u is at least as preferable as v to x .

That is, a world u is ranked at least as high as a world v just in case x doesn't prefer v over u . From this ranking on worlds, we can define a ranking on propositions: agents prefer a proposition p over a proposition q just in case they prefer every p -world over every q -world.

(17) For any entity $x \in E$, world $w \in W$, propositions $p, q \in \wp(W)$:

$p \leq_x^w q =_{\text{def}}$ for all $w' \in p$, $w'' \in q$, $w' \leq_x^w w''$

Heim (1992) proposes a semantics for *want* that uses this bouletic ordering instead of simply quantifying over worlds in x 's bouletic state. Supposing an actual world w , we first look at all the worlds that x takes to be live possibilities for the actual world—that is, Dox_x^w . We then look at the most similar worlds to w in Dox_x^w that make the proposition p true⁴. x *wants* p is true, then, if all the most similar p -worlds in Dox_x^w are higher ranked than all the most similar $\neg p$ -worlds in Dox_x^w . In other words, all else being equal, x prefers p to be true over $\neg p$.

To simplify the formal representation of the denotation of *wants*, Heim introduces the **similarity selection function** $\text{Sim}_w(p)$, which picks out the worlds most similar to w where p is true:

(18) $\text{Sim}_w(p) =_{\text{def}} \{w' \in W \mid w' \in p \text{ and } w' \text{ resembles } w$
no less than any other world in $p\}$

Consequently, Heim (1992)'s meaning for a sentence containing *want* is as follows⁵:

4. Note that Dox_x^w doesn't need to be realistic, i.e. to contain the world of evaluation w —agents may have false beliefs.

5. Since this analysis is couched in dynamic semantics, it doesn't have the form of a compositional denotation for *want*.

$$(19) \quad \llbracket x \text{ wants } p \rrbracket^w = \forall w' \in \text{Dox}_x^w : \text{Sim}^{w'}(p) \leq_x^w \text{Sim}^{w'}(\neg p)$$

In other words, according to x 's beliefs and preference ranking \leq_x^w , it is better to be in a maximally similar p -world than a maximally similar $\neg p$ -world. This is the operationalization of bouletic modality that we'll apply to polar questions in Section 3.2, which express a **desire to know if**. x 's desire to know p means he would prefer to be in a world where he knows p to one where he doesn't.

We will, however, need a more specific notion than simple preference in the sense of Heim (1992). As we discussed earlier, the characteristic performative bouletic modality is one in which the speaker's preference is like an **effective preference** in the sense of Condoravdi and Lauer (2012), or **decisive** in the sense of Kaufmann (2012): it structures (or can be taken by agents to structure) their choices of behavior, and the speaker is committed to behaving as if he prefers it. In that sense, the notion of preference we want to capture here is very different from 'mere' desire. My desire to be rich, for example, doesn't necessarily structure my choice of behavior; I might want to be rich but have no realistic hope of achieving it. Nor must I commit myself to behaving as if I actually want it; no one will take me to be a liar if I claim to want to be rich, but then spend all my money as soon as I get it.

In contrast, when one asks a question, that expression of a desire to know must structure one's subsequent behavior. It will not do to ask a question and then show no interest in learning the answer—to ask a question and then walk off before it's given, for example. Having asked the question, I must show some interest in learning the answer. In this sense, a question is like any other directive. It is not acceptable to tell a child, *Hold my hand before we cross the street!* and then thrust your hands deep into your pockets. Having uttered the directive, I must show some interest in the named state of affairs coming to pass. This is the core of what it means to be a **effective preference** for Condoravdi and Lauer (2012): effective preferences determine our choice of action.

How does this work for polar questions? The intuition I want to plumb here is that

polar questions always have a reason to be asked. It's always pertinent to ask the speaker, *Why do you want to know?*; the reason for asking is always lurking behind the question. In effect, a question expresses a speaker's desire to know something *in light of her goals*. The reason the speaker wants to know what they're asking about is that having the relevant knowledge orders him toward the achievement of some goal. Asking the question "furthers" the goal—that is, brings it closer to reality.

This goal orientation of polar questions is especially clear when they're used to perform the speech act of offering. If someone asks you *Do you like Beaujolais?*, bottle in hand, it's clear that his goal is to serve you a drink. In light of that goal, his question makes perfect sense: he wants to know if you like Beaujolais because learning it will allow him to serve you a drink that you like.

But compare that to a situation in which your interlocutor comes into your office in the middle of the workday and asks, apropos of nothing, *Do you like Beaujolais?* Here, we instinctively try to find a goal, even if it's far-fetched, in order to make the question felicitous again. Is he offering wine, though it's the middle of the workday? Is he considering purchasing wine, but wants input from an expert? Is he collecting the names of coworkers who enjoy Beaujolais for a scavenger hunt? We want a reason; it is surprisingly difficult, if it is possible at all, to construe the question as being asked "just because"—that is, without the motivation of some particular goal. This leads us to the Question-Goal Hypothesis:

(20) **Question-Goal Hypothesis:**

Polar questions are always asked in the furtherance of a goal. (*In the furtherance of a goal* meaning: the speaker takes it to be necessary or sufficient to know the relevant information in order to achieve the relevant goal.)

The Question-Goal Hypothesis is a claim about natural-language polar questions. And it is a natural one if we take polar questions to be primarily expressions of an attitude toward their radical proposition. To ask the question is to express an interest in a particular piece

of knowledge, and not merely to carve up logical space into cells. That interest is part of the meaning of the question, and it proceeds from our preexisting goals, conversational or otherwise.

The goal-oriented nature of a polar question persists even if we try very hard to cancel the inference. The speaker in (21), far from successfully canceling the inference that his question has a motive, mostly succeeds in sounding like a bad liar fishing for some information.

(21) Do you like Beaujolais? I'm not asking for any particular reason.

This isn't to say, of course, that we can't ask questions out of a general sort of curiosity or desire to know things. Consider someone who asks something like the following.

(22) I like knowing trivia about people. Do you like Beaujolais?

But then here the speaker's goal is to know a piece of trivia about you, and it's clear how the question furthers that goal: learning that you like Beaujolais is a sufficient condition to achieve that goal.

3.1.2 Goal orientation in polar questions

How do we model the goal orientation of polar question modality? As a first pass, we could take the set of x 's goals at w and time t , call it $\text{Goals}^{w,t}(x)$, and order worlds by how many of those goals they make true, à la Kratzer (1981). The definition for this **Goal set** is given below:

(23) For any x, w, t :

$$\text{Goals}^{w,t}(x) = \{p \in \wp(W) \mid x \text{ has it as a goal that } p \text{ in } w \text{ at } t\}$$

A set of propositions (a *conversational background*) induces an ordering on worlds according to how many of the propositions it contains are made true at each world. Worlds that make

maximally many true are ranked highest. Here, then, the worlds where x achieves the most possible goals are highest-ranked. For a simple example involving questions, suppose that x 's goal is to serve his interlocutor y a drink that she likes. Ignoring other, irrelevant goals of x , and suppressing temporality, we have the following goal set:

$$(24) \quad \text{Goal}^{w,t}(x) = \{x \text{ serves } y \text{ a drink she likes}\}$$

In order to advance this goal, x asks his interlocutor y the following question:

$$(25) \quad [\text{Speaker goal: give the interlocutor a drink he likes}]$$

Do you like Beaujolais?

According to the usual ordering induced by a conversational background (Kratzer, 1981), worlds are higher-ranked just in case they make more propositions in the background true. Remember that we want polar questions to mean that the speaker wants to know p , i.e. preferentially orders worlds where x knows p over worlds where x doesn't. But ordering worlds based on (24) doesn't guarantee that x -knows- p worlds will be ranked above x -doesn't-know- p worlds.

Suppose that x has two drinks he can serve y : Beaujolais, or tonic water. For all x knows, y likes only one, both, or neither. x 's ultimate goal is to provide y with a suitable drink, but he doesn't care which one it is; it's up to y 's preferences. All x 's top goal worlds, then, are worlds where he serves y a drink; in some of them that drink is Beaujolais, and in others it is tonic water.

Therefore, in some of the top-ranked worlds, y likes both drinks, but x only gives y tonic water. Consequently, it's not true that in *every* world in which x achieves his goals and y likes Beaujolais, x knows it. In worlds where y likes both and x ends up serving y tonic water, x never knows anything about Beaujolais.

This issue is even starker with conversation-starter questions, like the following:

(26) So, do you play any sports?

Now here, the speaker’s goal is merely to start an engaging conversation, and it has nothing to do with sports *per se*. The question might have been chosen as a guess at a topic that would lead to an engaging conversation; but the speaker could very easily have asked about a million other topics. If we look simply at those worlds in which the speaker succeeds in starting an engaging conversation—even those in which the addressee *does* play sports—there will be worlds where the speaker never knows that the addressee plays sports; the conversation goes in another direction.

So, going back to the act of offering with *Do you like Beaujolais?*, what is it in virtue of that the speaker’s preference ranking puts worlds where he knows *You like Beaujolais* over worlds where he doesn’t? Our ranking can’t simply order goal-worlds above non-goal-worlds; this isn’t fine-grained enough. Intuitively, we want worlds to be ordered by \leq_x^w such that a higher-ranked world is one which is ‘closer’ to being a goal world than worlds ranked below it, even if it isn’t a goal world *per se*. Among non-goal-worlds, then, there should still be a ranking according to which some are better than others.

We’ll say that a world w is **goal-preferable for** x over a world v if and only if w is closer to achieving x ’s goals than v . But what exactly does it mean to be goal-preferable? We’ll operationalize this as **achieving more subgoals**: a world is closer to being a goal world for x if it achieves more subgoals for x .

Subgoals are those future states of affairs that x takes to be necessary to achieve his ultimate goal g . They are those states of affairs that get him closer to achieving g , even if he doesn’t achieve it. Each goal in an agent’s goal set determines a set of **subgoals**, defined as follows:

(27) For any $x \in E, w \in W, t \in T, g \in \text{Goal}^{w,t}(x)$:

$$\text{SUBGOAL}_x^{w,t}(g) = \{p \in \wp(W) : \text{for some } t' \text{ temporally following } t, x \text{ takes } p(t') \text{ to be a necessary condition for } g \text{ in } w\}$$

In the Beaujolais example, the speaker’s goal is to serve his interlocutor a suitable drink. What are possible subgoals for this goal? Here, in addition to others, he must first know of a drink that he has that his interlocutor likes it. Therefore, one (but certainly not the only) subgoal of *Serve y a drink she likes* is *Learn a drink that y likes that x has*. It’s in virtue of this proposition in the set of all of x ’s subgoals that x knows that y likes Beaujolais is preferable to x doesn’t know that y likes Beaujolais.

The ordering over worlds in Dox_x^w , therefore, is:

- (28) **Ordering on worlds** $\leq_{\text{Goals}^{w,t}(x)}$ for any speaker x , time t , worlds u, v, w :
 $u \leq_{\text{Goals}^{w,t}(x)} v$ if and only if u makes at least as many propositions in $\{p : \exists g [g \in \text{Goals}_x^{w,t} \ \& \ p \in \text{SUBGOAL}_x^{w,t}(g)]\}$ true as v does.

As Heim (1992) does, we can extend this definition to propositions: a proposition p goal-outranks a proposition q just in case every p -world outranks every q -world:

- (29) $p \leq_{\text{Goals}^{w,t}(x)} q =_{\text{def}} \forall u \in p, \forall v \in q : u \leq_{\text{Goals}^{w,t}(x)} v$

Bouletic modality for polar questions, then, is to be construed as preference based on a ranking of worlds according to how close they get to being a goal world for the agent, based on the agent’s assessment of what is necessary to achieve the goal. Satisfying more necessary conditions for a goal makes a world higher-ranked.

3.2 The attitudinal analysis of polar questions

Now we are in a position to fully articulate out our analysis of a polar question p ? as expressing an agent’s *desire to know if p*. To preview our analysis, the meaning of the polar question *Is Ida hungry?* uttered by a speaker x in world w will carry the following semantic content:

(30) $\llbracket \text{Is Ida hungry?} \rrbracket_x^w =$ For every world w' consistent with what x takes to be the case and where Ida is hungry, a minimally different world w'' where x knows *Ida is hungry* is goal-preferable to a minimally different world w'' where x does not know *Ida is hungry*.

Polar questions are built on top of question radicals, which are propositions, i.e. sets of worlds. The meaning of any question $p?$ is defined as follows: it is that proposition that is true just in case the speaker goal-prefers learning p over not:

(31) $\llbracket p? \rrbracket_x^w = \forall w' \in \text{Dox}_x^w[p][(\text{Sim}^{w'}(v)(p \subseteq \text{Dox}_x^v)) \leq_{\text{Goals}^w(x)} (\text{Sim}^{w'}(v)(p \not\subseteq \text{Dox}_x^v))]$

In other words, if we take the set of x 's doxastically accessible worlds and order them by how many (sub)-goals of x 's that they make true, then every p -world where x knows p is equally good as, if not better than, the closest world where x does not know p .

3.3 Polar questions as expressives

Polar questions clearly do not function as assertions. So if their meaning is indeed propositional, as I have argued here, then it must be something that behaves differently in discourse from an assertion, which presents a proposition as a contribution to the common ground. Polar questions are thus, on my account, *expressive*: they express the speaker's mental state, rather than proffering a proposition.

I follow Chernilovskaya et al. (2012a) and Castroviejo Miró (2009) in assuming that the difference between expressive and propositional content is the 'directness' with which it is added to the Common Ground. Chernilovskaya et al. (2012a) consider the expressive component of wh-exclamatives like *How big your teeth are!*, noting five properties of the expressive component (that the speaker is surprised at the size of the addressee's teeth) of such utterances:

- (32) a. *Expressive content cannot be coherently denied by the speaker of the exclamative.*
 #No, you aren't surprised by that.
- b. *It cannot be explicitly accepted in the way the asserted content can.*
 # Yes, you are surprised by that.
- c. *It cannot be explicitly confirmed in the way the asserted content can.*
 # That's right.
- d. *It cannot be challenged by the audience, not even as a lie.*
 # You're lying, you're not surprised.
- e. *It can be commented on by questioning beliefs that are necessary for the attitude expressed.*
 What do you mean? My teeth aren't that big.

Polar questions conform to these properties as well. It's not possible to target the expressive *want to know p* component of the meaning of a polar question by responding e.g. *That's true/false, Oh, I see*, or so on. However, one can comment on the expressive component by questioning the necessary beliefs for holding that preference. For instance, if I believe that you really don't care one way or the other what the answer is:

- (33) A: How are you feeling today?
 B: What do you mean? You don't care how I'm feeling.

How can contributions to the Common Ground be made directly, without input by other discourse participants? In general, contributions to the Common Ground must be agreed upon by all conversational participants, since the Common Ground contains information that is jointly accepted. Identifying the contribution of exclamatives to the Common Ground as 'automatic', then, would seem to require forcing conversational participants to believe the content of the exclamative.

Chernilovskaya et al. (2012b)'s solution is that the content of exclamatives gets directly

entered into the Common Ground in virtue of their being an expression of the utterer’s mental state. Compare this to Stalnaker’s 1987 example of a goat walking into the room: the fact that a goat is present is immediately added to the Common Ground, in virtue of it being perceptually evident to everyone. An exclamative is similar. By saying *What big teeth you have!*, the speaker’s surprise is entered into the Common Ground by means of a convention of use, which says that exclamatives are only used when their propositional content is true; Chernilovskaya et al. (2012b)’s “Lewis-convention on exclamatives” (adapted):

- (34) **Lewis-convention on exclamatives:** A speaker only utters an exclamative with descriptive content p in a world w if the speaker experienced a mental event directed toward p before the utterance time.

This convention is something that holds in a speech community: it is accepted among all participants in the discourse that it holds. Then it follow from this convention, plus the fact that the speaker uttered an exclamative, that the speaker had the requisite mental event (namely, in our case, being surprised at the addressee’s tooth-size).

If we take polar questions to carry propositional content, then a similar convention holds for them: a convention of use requires that polar questions only be uttered when the speaker actually bears the mental state (of preferring-to-know) that is expressed by the question. That expressive content is added to the Common Ground ‘directly’ upon utterance. So we have the following convention of use on polar questions:

- (35) **Lewis-convention on polar questions:** A speaker only utters a polar question with expressive content p in a world w if p is true at w at the time of utterance.

For example, if I utter *Is Gus hungry?*, then the following line of reasoning gets us to the direct contribution of the expressive content to the Common Ground:

- (36) 1. Conversational participants are following the Lewis-convention on polar questions.

2. The speaker uttered a polar question with expressive content p .
3. Therefore p must be true at the time of the speaker's utterance.

Thus it is really accurate to say that the polar question *expresses* a desire to know, rather than to say e.g. that it asks a question or proffers options to interlocutors. Speakers ask questions in order to express that they wish to know the radical proposition, if it is true.

3.4 Polar questions in context

Now we may turn to the consequences of the expressive-attitudinal approach to polar questions. To start with, we predict that it's not possible to felicitously ask a question without wanting to know the answer. This, as the core semantic contribution of a polar question, means that the speaker really does prefer to know p if p is true. And this seems right; the following utterance, for example, is clearly infelicitous.

(37) #If you're hungry, I don't want to know. Are you hungry?

This pair of sentences is much more felicitous if their order is flipped. But in that case, the resulting text is most naturally read as the speaker taking back their stated interest in learning the relevant proposition. There is the flavor of a retraction—the speaker is, in the second sentence, retracting their question, not merely asserting an additional fact:

(38) Are you hungry? (Never mind,) I don't want to know.

Indeed, something like *Never mind, I don't want to know*, or *I take that back—I don't want to know*, is the most natural way to retract a polar question in English.

3.4.1 Yes and No in polar questions

How do the response particles *Yes* and *No* work with respect to polar questions on this semantics? I assume, following Krifka (2015), that *Yes* and *No* are anaphoric to propositions introduced compositionally by the polar question, just as they are to propositions introduced by other means, such as assertions. *Yes* is anaphoric to a proposition and denotes that proposition itself; *No*, on the other hand, denotes the negation of the proposition it is anaphoric to.

Polar questions can introduce propositional discourse referents in maximally two places. In a positive polar question, one discourse referent is introduced, corresponding to the positive question radical. (For simplicity of presentation, we're representing the semantic contribution of the polar question as PolQ.) Here, the possible propositional discourse referent is marked with an underline:

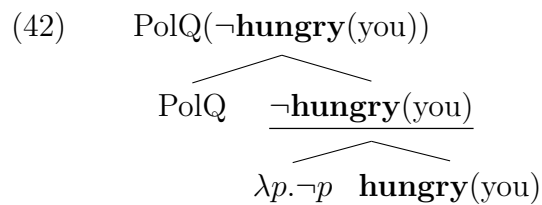
(39) Are you hungry?

(40) PolQ(**hungry**(you))
 └──┬──
 PolQ hungry(you)

The response particle *Yes* picks up the only available discourse referent, **hungry**(you), introduced immediately below PolQ, and refers to it.

In the case of negative questions, though, there are two possible discourse referents: one containing negation, and one not. That is, both constituents denoting propositions introduce a discourse referent, both marked with an underline below.

(41) Are you not hungry?



This is the core of Krifka’s analysis of the response particles *Yes* and *No*. With a negative question, because there are two possible propositional anaphors and two possible response particles, there is a total of four possible answers:

- (43) a. Yes, I am. (*Yes + positive-polarity anaphor*)
 b. Yes, I’m not. (*Yes + negative-polarity anaphor*)
 c. No, I am. (*No + positive-polarity anaphor*)
 d. No, I’m not. (*No + negative-polarity anaphor*)

Combining this analysis of response particles with the expressive-attitudinal analysis of polar questions, we have an approach that deemphasizes the role of polar response particles quite a bit. Polar answer particles are anaphoric to the radical proposition of the question, but in a way that doesn’t directly relate to the function of the question at all. That is, the asker’s question expresses the speaker’s attitude toward the radical proposition, and then the answers *Yes* or *No* either affirm or deny the radical.

3.4.2 Polar question felicity: *The Evidence Condition revisited*

The *want to know* meaning for polar questions explains why they, but not other types of questions, may be used in Bolinger contexts: namely, inference-licensing uses, offers, requests, and conversation starters. In particular, each of these contexts provides an obvious goal to generate a set of subgoals to use as an ordering source. In inference-licensing, the speaker’s goal is to know an explanation for something; in offers, to provide the addressee with something; in requests, to obtain something; and in conversation starters, to have a rewarding conversation.

Many, if not all, cases of polar questions in the literature are what Bolinger would likely refer to as inference-licensing uses. For instance, the (simplified) raincoat example from Büring and Gunlogson (2000):

- (44) [Addressee walks into the room dripping wet, wearing a raincoat]
Speaker: Is it raining?

Here, the speaker's ultimate goal is to know the reason that the addressee is dripping wet. In effect, the speaker is trying to determine if they can infer from the wet raincoat that it's raining. Now, learning that it's raining is not, per se, a subgoal for the goal *know why Addressee is dripping wet*: there are many possible reasons that someone might be wet, so it's not the case that the speaker believes that learning it's raining is a necessary condition for learning why the addressee is dripping wet. In fact, even if it *is* raining, the rain might not be the reason that the addressee is dripping wet. (It's logically possible that, for example, a bucket of water fell on them from a door.) But what *is* a subgoal of *know why the addressee is dripping wet* is the following proposition: that the speaker knows of *something* that can account for the addressee's being wet. In order to know why the addressee is dripping wet, the speaker must know that there is some actual state of affairs that might account for it. And this is the subgoal that's satisfied by asking *Is it raining?*

This reasoning present in the raincoat scenario characterizes a class of question-asking scenarios that amount to venturing a hypothesis: the speaker asks a polar question in which the radical, if true, would explain some observed phenomenon. These are what Bolinger (1978) refers to as *inference-licensing* uses of polar questions:

- (45) a. [Addressee isn't eating]
Are you sick?
b. [Speaker hears barking behind a door]
Do you have a dog?

This line of reasoning explains why *Is it raining?* is felicitous in the raincoat scenario, but *Is it sunny?* is not. Learning that it's raining would further the speaker's goal of learning why the addressee is dripping wet; but learning that it is *sunny*, in contrast, does not. It doesn't get the speaker any closer to achieving his goal, as learning that it's sunny isn't entailed by any subgoal of *Learn why addressee is dripping wet*. In fact, it's difficult to conceive of *any* goal in this context that is furthered by learning if it's sunny. This is why the question seems infelicitous.

Recall the Evidence Condition from Büring and Gunlogson (2000):

(46) **Evidence Condition on Polar Questions:**

[A polar question $p?$ is felicitous only if] there is no compelling contextual evidence against p (i.e., there is either no evidence bearing on whether p , or there is evidence for p)

The Evidence Condition rules out scenarios in which a question is asked in the presence of evidence for the inverse proposition. I suggest that the Evidence Condition follows from natural assumptions about Bolinger's inference-licensing polar questions and the goal-oriented preference semantics for questions. In particular, what goes wrong in these scenarios is that the radical proposition of the question ($p?$) cannot possibly be an explanation for the observed evidence E , as E suggests $\neg p$. In the raincoat scenario, for example, learning that it is sunny will not satisfy the speaker's goal-oriented desire to know why the addressee is dripping wet. Consequently, learning p cannot be a subgoal of the speaker's goal to know why E is the case; and so the speaker cannot ask $p?$ in order to further that goal.

One prediction of this account is that questions should, in fact, be acceptable in the presence of evidence for the inverse proposition, so long as the speaker's goal is right. For example, consider the following scenario:

(47) [Speaker and addressee get home and see that the cat is meowing and pacing in front

of its empty food bowl.]

Did you feed the cat?

This is a problematic example for the Evidence Condition: the cat's behavior provides clear evidence for the inverse proposition *You didn't feed the cat*. Why, then, is the question nevertheless felicitous in this context, contra the Evidence Condition? First, note that this question is best in a context where the addressee is expected to feed the cat; it's much less felicitous if the addressee has no such obligation:

- (48) [Speaker and addressee get home and see that the cat is meowing and pacing in front of its empty food bowl. Both know that it's the speaker's job to feed the cat.]
#Did you feed the cat?

There are a number of different goals that might motivate the speaker to ask the addressee *Did you feed the cat?*. Perhaps the speaker's goal is not to determine why the cat is behaving in this way, but rather to determine whether the addressee did his chores. In that case, knowing that the addressee fed the cat is entailed by this goal.

There are a number of similar scenarios contravening the Evidence Condition, involving what we might call 'violated expectations': the speaker expected that something ought to have happened, and seeing evidence that it didn't, asks about whether the obligation to do the thing in question was fulfilled. For example, consider the following scenarios:

- (49) a. [A bill reading 'ELECTRIC BILL OVERDUE' is in the mailbox]
Did you pay the electric bill?
- b. [A teacher is collecting homework one-by-one from students. The teacher's addressee doesn't hand the teacher anything]
Did you do your homework?
- c. [The speaker's home is untidy, and the speaker employs a cleaner]

Did the cleaner come today?

In these situations, the speaker observes evidence that the radical proposition doesn't hold, but asks the question for some independent reason: they hope it to be true, perhaps. Similarly, scenarios in which the speaker is hopeful or suspicious that the radical proposition is true, even despite evidence for the inverse proposition, contravene the Evidence Condition:

- (50) a. *[Speaker notices addressee is wearing a gold ring on their ring finger]*
Are you single?
- b. *[Speaker and addressee are seated at a bar having a drink]*
Are you over 21?

In these scenarios, the speaker's question is best read in a hopeful tone and a suspicious tone, respectively. That is, in (50-a), the speaker is hopeful that the addressee is single—and so perhaps the ring is fake, for example. And in (50-b), the speaker is suspicious that the addressee is under 21, even though their presence at the bar would generally suggest that they are indeed over 21.

Crucially, none of these counterexamples to the Evidence Condition inference-licensing questions in the sense of Bolinger (1978): although the context is such that there is evidence for the inverse proposition of the question, the question is not intended to determine whether the speaker is licensed in inferring the radical proposition from the available evidence. That is, it's the speaker's evidence-independent goal—to find a romantic partner, for example—that motivates the question, and not the evidential properties of the context at all.

Requests and offers, too, don't care about the evidential properties of the discourse context. For instance, even if the context suggests that you'll say *No*, I might offer you a drink just to be polite.

- (51) a. *[Addressee spent a long time finishing his last beer]*

Do you want another beer?

Again, here we seem to be able to ‘look past’ the evidence in the context, because the speaker’s goal is not directly related to that evidence at all. The attitudinal analysis, which uses speaker goals as an ordering source for the preference expressed by polar questions, is able to account for this neatly. What really matters is the speaker’s goals, and not the evidential properties of the contexts. In contexts with particular evidential properties, of course, speakers are more or less likely to have certain goals: if there is some salient thing in the discourse, it’s more likely that questions will be about that thing. However, ultimately, it’s the speaker’s goal—and the hearer’s ability to infer a goal—that matters.

I argue, therefore, that polar question (in)felicity is primarily about the likelihood of the speaker having a goal that is furthered by the particular question. When questions are infelicitous, it is because it’s impossible (or very difficult) to construe a goal for the speaker to be furthering with the question. For example, the wet raincoat scenario of Büring and Gunlogson (2000):

(52) *[The addressee walks into the room in a dripping wet raincoat.]*

Is it raining? / #Is it sunny?

Here, *Is it sunny?* is odd because there is no obvious goal the hearer can assign to the speaker that is furthered by that question. In this context, the only really salient goal one can ascribe to the speaker is determining a reason for the wet raincoat. But this question gets much better if we simply add some reason for the speaker to hope that it is sunny:

(53) *[The addressee walks into the room in a dripping wet raincoat. The speaker has been hoping all morning that the weather will be nice enough for him to go suntan later.]*

Is it raining? / Is it sunny?

In this elaborated context, perhaps the speaker can be faulted for not inferring from the wet

raincoat that the weather is basically resolved. However, his overriding interest in the sun shining seems enough to make the question felicitous again, without changing the evidential properties of the context at all. One way to see this clearly is simply to preface the speaker's question with something like *I really hope it's sunny*. This change to the discourse context has nothing to do with evidence, and everything to do with speaker goals.

In other words, then, what is relevant to judgements of polar question felicity is **goal construal**: given a polar question and a context, can we assign a reasonable goal to the speaker that is furthered by the question? Polar questions are infelicitous when we cannot imagine a reason that the speaker would ask the question, and not because of properties of the context per se.

3.5 Conclusion

This chapter has presented an analysis of polar questions that treats them as expressions of the speaker's goal-preference to know the radical proposition if true. For each speaker goal, we determine a set of subgoals, which are those propositions that the speaker takes to be necessary to achieve their goal. The meaning of a polar question is that for every possible world in the speakers doxastic state, they would achieve more subgoals if they knew the radical proposition at that world than if they didn't. We therefore assign the following expressive content to a polar question p :

$$(54) \quad \llbracket p? \rrbracket_x^w = \forall w' \in \text{Dox}_x^w[p][(\text{Sim}^{w'}(v)(p \subseteq \text{Dox}_x^v)) \leq_{\text{Goals}^w(x)} (\text{Sim}^{w'}(v)(p \not\subseteq \text{Dox}_x^v))]$$

This denotation, combined with the Lewis-condition on polar questions, means that speakers will ask polar questions only when that expressive content is true. The content, in turn, will be added automatically to the Common Ground, rather than presented as a contribution to the discourse like assertions are.

The goal-orientation of polar questions means that polar question felicity is intimately

tied to the speaker's overall communicative and action-oriented goals, and less tied to evidential properties of the context than has been previously claimed. This helps account for the functional heterogeneity of polar questions: the fact that they can be used to perform a number of speech acts beyond simply requesting information from an interlocutor. Additionally, the response particles *Yes* and *No* get a diminished role on this account; they may be used to respond to polar questions, but they do not play a particular role in the semantics or conditions of use of the question per se.

CHAPTER 4

NEGATION BIAS AND NEGATIVE POLARITY ITEMS IN POLAR QUESTIONS

Since Ladd (1981) inaugurated it as a research question, the issue of neg(ation)-biased polar questions has vexed researchers. Syntactically, these are polar questions that, in English, appear with preposed negation, cliticized to an auxiliary verb. With respect to their meaning, these questions uniformly carry an inference, of some kind, to the effect that the speaker expects or believes the radical proposition to be true. The radical proposition may be of either positive polarity, as in (1), or negative, as in (2).

- (1) Wasn't Lithuania in the USSR?

Bias inference: The speaker believes that Lithuania was in the USSR.

- (2) Wasn't Finland not in the USSR?

Bias inference: The speaker believes that Finland was not in the USSR.

The primary puzzle in the analysis of neg-biased questions is determining the semantic role of the negative morpheme *-n't*. While the presence of this morpheme makes these questions appear to be overtly negated, they don't seem to be semantically negative: that is, for example, the above questions don't ask about the propositions *Lithuania wasn't in the USSR* and *Finland wasn't not in the USSR*, respectively; see 4.4 for concrete arguments against the claim that neg-biased polar questions necessarily express a question with a negative radical. Instead, the contribution of high negation in these questions seems to be the bias inference: that the speaker believes the radical proposition to be true.

There are two families of analyses of neg-biased questions. The first, generally earlier, family takes them—or a subset of them—to be genuinely semantically negative, i.e. to be questions about a negative-polarity radical proposition. This **negative radical approach** is the core of the argument advanced by Ladd (1981), van Rooy and Šafářová (2003), and, in a

fashion, AnderBois (2019), among others. The second family of accounts takes the negation in neg-biased questions to scope over a silent operator, typically some kind of epistemic or discourse operator, outside of the question radical. Exemplars of this **wide-scope negation approach** are Romero and Han (2004), Krifka (2015, 2017), and Goodhue (2018). This chapter advances an argument in the latter camp: we will locate the contribution of negation in neg-biased polar questions outside of the question radical, taking the widest scope possible. This will necessitate extended our analysis of polar questions from expressing a preference to know to instead expressing a **preference to learn**.

A second issue regarding neg-biased polar questions comes from the claim, due to Ladd (1981), that neg-biased questions have two readings: an *inner negation* and an *outer negation* reading, which ostensibly differ with respect to where negation appears in the logical form of the question. According to Ladd, the two readings are diagnosable with polarity items: neg-biased questions containing negative polarity items obligatorily receive inner-negation readings, and those containing positive polarity items obligatorily receive outer-negation readings.

With respect to their meanings, briefly, inner-negation neg-biased questions carry something like a **violated-expectation** inference: the speaker thought p , but now doubts it for some reason, and so asks a neg-biased question about p . That is, the speaker’s expectation that p was violated. Outer-negation neg-biased questions, in contrast, carry no such violated-expectation inference; they express a more simple bias on the part of the speaker toward the radical proposition p . An example of each follows (underlining highlights the NPI/PPI forcing purported inner/outer readings, not prosodic emphasis):

- (3) Haven’t you eaten anything today? (“**Inner-negation**” neg-biased Q)
Violated-expectation bias inference: Speaker expected that you had eaten something today, but something called that expectation into question.
- (4) Haven’t you eaten something today? (“**Outer-negation**” neg-biased Q)

Bias inference: Speaker believes that you've eaten something today.

Inner-negation questions add a wrinkle to the landscape of neg-biased questions: they seem to express a different kind of bias. However, some recent work has disputed the inner-/outer-negation distinction in the first place (e.g. Goodhue, 2018; AnderBois, 2019). This chapter contributes a new argument to this effect, positing no special distinction between inner- and outer-negation neg-biased questions. I will argue that the inner-/outer-negation distinction is in fact an illusion, one which arises from the presence of NPIs in the polar question, which are independently licensed in polar questions and have nothing to do with negation.

The structure of this chapter is as follows. In Section 4.1, we characterize the empirical phenomenon of neg-bias. In Section 4.2, we critically assess existing approaches to neg-bias. Section 4.3 extends the analysis developed in Chapter 3 to high-negation questions. Section 4.4 presents an argument against the inner-/outer-negation distinction. Finally, Section 4.4.3 expands the scope from NPIs in neg-biased questions to NPIs in polar questions generally.

4.1 Characterizing neg-biased questions

To start with, we may note that neg-biased questions can appear with radical propositions of both positive and negative polarity. Consider the following two examples of neg-biased questions:

- (5) a. Don't Muslims eat beef?
b. Don't Muslims not eat pork?

In (5-a), the speaker asks about the proposition that Muslims eat beef, and the question, with high negation, expresses some kind of bias toward the truth of that proposition. In (5-b), the speaker asks about a negative-polarity proposition, and the attendant bias is toward that same negative-polarity proposition.

We will see more evidence later, from Goodhue (2018), that the negative morpheme -

n't in a neg-biased question does not contribute genuine propositional negation. For now, it suffices to observe that in (5-b), there is no intuitive sense of “double negation”; the question really seems to contain just one propositional negation, contributed by *not* preceding the VP *eat pork*. This distinguishes the neg-biased questions above with the following low-negation questions that would be their ostensible synonyms:

- (6) a. Do Muslims not eat beef?
b. Do Muslims not not eat pork?

Clearly, this pair of questions is not (necessarily) synonymous with their neg-biased counterparts. And in particular, it's difficult to imagine why anyone would ever utter the question (6-b): the choice of double negation renders the question confusing.

It is worth noting that in limited circumstances, questions with low negation can receive a neg-biased interpretation. There are what AnderBois (2019) dubs “Gladiator low-negation questions”; Romero and Han (2004), too, characterize them as “archaic”. The following are examples, adapted from AnderBois (2019):

- (7) a. Are we not men?
b. Are you not amused?
c. If you prick us, do we not bleed?

Each of these questions expresses the same bias as a high-negation neg-biased question: *Aren't we men?*, and so on. But as is now traditional in the literature, we will bracket these “Gladiator” low-negation neg-biased questions, and leave them to some future work to sort out. The primary issue with these questions is that their synchronic status is not well-established: is this really a productive construction in contemporary English, or is it an archaic construction that remains salient because of its presence in literature?

However, it should not be impossible to extend a scopal analysis of neg-bias to these

Gladiator neg-biased questions. If what is responsible for neg-bias is the presence of negation in some particularly high position in the derivation of the sentence, above an operator that takes scope over the radical proposition of the question, then Gladiator neg-biased questions might receive an analysis that allows preverbal negation to scope out of the radical proposition entirely.

An interesting issue that Gladiator neg-biased questions raise is the historical development of neg-bias in English, and its relationship at various stages to the cross-linguistic realizations of neg-bias. Obviously, not all languages allow negation to appear in the syntactically high position characteristic of (non-Gladiator) neg-biased questions; but various languages have been claimed to allow neg-biased questions, principally Germanic. There does not yet exist an articulation or a theoretical account of the varieties of constructions across languages that have been claimed to allow neg-biased interpretations.

One complication in the characterization of negation bias comes from the introduction of the **inner-/outer-negation** distinction, which we describe now.

4.1.1 *Inner and outer negation*

Ladd's (1981) foundational work on neg-biased questions posited two distinct readings for a neg-biased question, termed the *inner-negation* and *outer-negation* readings. The terminology comes from a purported variability in the location of negation. In inner-negation neg-biased questions, Ladd's intuition is that negation is inside the radical proposition; the question is, in some sense, about a negative radical. In outer-negation neg-biased questions, contrarily, Ladd infers that negation is located somewhere outside the radical. Consider the following neg-biased question uttered in two different contexts, adapted from Ladd¹.

1. Some of Ladd's original examples used the NPI *either*, which has turned out to be problematic, as not all English speakers accept *either* in polar questions of any type (Sailor, 2013). As Goodhue (2018) observes, too, the meaning of *either* is not especially well understood, which makes it less than ideal in critical examples. As Ladd's argument relates to negative polarity items, and not to *either* in particular, we use examples with *any* instead, which does not have these problems.

(8) **Outer-negation neg-biased question:**

[A group including a vegetarian is debating where to go out to eat. The speaker has heard that there's a vegetarian restaurant nearby.]

Isn't there a vegetarian restaurant around here?

(9) **Inner-negation neg-biased question:**

[The same group is debating where to go out to eat. They check a smartphone app listing restaurants in the area, but don't see anything vegetarian.]

Isn't there a vegetarian restaurant around here?

For Ladd, these two cases exemplify two distinct readings for the neg-biased question. In the first, outer-negation reading, the speaker has a prior belief that there's a vegetarian restaurant nearby: i.e., the speaker believes that p . He then asks *isn't p?* in order to, as Ladd puts it, “double-check” whether p is true. Negation is *outer* in the sense that what is double-checked is the proposition p , and negation is somehow located outside of that proposition.

In the second case, the speaker still believes that there's a vegetarian restaurant around, but this belief isn't necessarily informed by hearsay or experience; instead it might have the character of an expectation or generic belief, such as *Neighborhoods like this one have vegetarian restaurants*. When faced with evidence against this belief in the form of a failed web search, the speaker asks the neg-biased question in order to “double-check” the proposition $\neg p$: that there not a vegetarian restaurant around here. This is the *inner-negation* reading of the neg-biased question, because the proposition being double-checked is of negative polarity.

This typifies an interesting two-step process of licensing characteristic of inner negation, but not outer negation. In the first step, the speaker has some prior belief. In the second step, the speaker's belief is called into question by some utterance or piece of evidence in the discourse context. The inner-negation reading is licensed by this mismatch between the

speaker's belief and the encountered evidence. Interestingly, and unlike the outer-negation question, the question double-checks the truth not of the speaker's prior belief, but of its negation: the proposition suggested by the encountered evidence. For instance, in our example above, what is double-checked is the proposition *There is no vegetarian restaurant around*, even though the bias, intuitively, is toward the positive-polarity proposition *There is a vegetarian restaurant around*; this is what was his prior belief.

Why not make the more parsimonious assumption that these two readings are merely two different ways of satisfying the same meaning? Ladd's reason for positing a genuine ambiguity comes from the role of negative polarity items. In particular, the presence of a negative polarity item forces an inner-negation reading. If we take the contexts from before and replace *a vegetarian restaurant* with the NPI-containing DP *any vegetarian restaurants*, we see that an outer-negation reading is unavailable, making the question infelicitous, while the inner-negation reading is natural.

(10) **Outer negation + NPI = infelicitous**

[A group including a vegetarian is debating where to go out to eat. The speaker has heard that there's a vegetarian restaurant nearby.]

#Aren't there any vegetarian restaurants around here?

(11) **Inner negation + NPI = felicitous**

[The same group is debating where to go out to eat. They check a smartphone app listing restaurants in the area, but don't see anything vegetarian.]

Aren't there any vegetarian restaurants around here?

Ladd's conclusion is that there is an ambiguity in neg-biased questions related to the location of negation. In outer-negation questions, negation is outside the radical, and therefore cannot license NPIs. In inner-negation questions, the presence of negation inside the question radical licenses NPIs. We will argue against making the inner/outer distinction at all in Section 4.4.

But it is an important distinction to keep in mind when assessing accounts of neg-biased questions, many of which take the distinction for granted and spend a great deal of effort attempting to derive it.

It is good to highlight here that inner- and outer-negation readings are presumed to be two distinct, generally available interpretations of neg-biased questions. That is, it's not the case that an inner-negation reading requires an NPI. Rather, an NPI forces an inner-negation reading, such that a neg-biased question *without* an NPI is ambiguous between an inner- and outer-negation interpretation.

Büring and Gunlogson (2000) are the first to attempt to break down the bias inference of neg-biased questions. Along with their Evidence Condition on (canonical) polar questions, discussed in detail in Chapter 2, they present Evidence Conditions on neg-biased questions: one for inner-negation and one for outer-negation questions:

(12) **Evidence Conditions on neg-biased questions** *Isn't p?*

For outer-negation questions: There is no compelling contextual evidence for p (i.e., there is either no evidence with respect to p , or there is evidence against p).

For inner-negation questions: There is compelling contextual evidence against p .

Büring and Gunlogson present some scenarios to bring out the intuition behind these evidential conditions, but they are not uncontroversial. They take there to be three kinds of contexts: those with evidence for the radical proposition p , those with evidence against it—presumably, the same thing as evidence for the inverse proposition $\neg p$ —, and those with no evidence for p or against p .

First, let's consider contexts with evidence for the radical proposition p . They claim that neither kind of neg-biased question is felicitous with evidence for p , as evidenced by the following example (where *some* forces an outer-negation reading, and *any* an inner-negation reading):

(13) *[A group is debating where to go out to eat. The speaker has been to a vegetarian restaurant nearby before.]*

#Isn't there some/any vegetarian restaurant around here?

In the outer-negation case (with *some*), however, this does not accord with my own judgment, at least in the context of this piece of evidence². It does seem fairly clear that a speaker may utter *Isn't there some vegetarian restaurant around here?* even if she has been to the restaurant before, so long as she is not entirely certain, or otherwise willing to accept that she may be wrong: hedging for politeness, for example. In other words, if there is a condition, the condition doesn't really seem to be on evidence, but on the speaker's epistemic state.

Next, let's consider the second type of context, in which there is evidence against the radical proposition p , i.e., there is evidence for the proposition $\neg p$. Inner negation neg-biased questions, Büring and Gunlogson claim, are licensed only in these contexts. For example, where the evidence against p comes in the form of an interlocutor claiming something which suggests p is false:

(14) *[A group is debating where to go out to eat. One member says that there are no health food restaurants around.]*

Aren't there any vegetarian restaurants around here?

In other words, inner-negation questions seem to want contextual evidence that suggests $\neg p$. However, interestingly, this evidence can't be too strong. Above, for example, the evidence against there being vegetarian restaurants comes in the form of an interlocutor saying that there are no health food restaurants—which suggests, but does not entail, that there are no vegetarian restaurants. A discourse in which the interlocutor claims outright that $\neg p$ is too strong; this yields infelicity:

2. It is, additionally, unclear whether we really want to treat a speaker's prior experience as a kind of 'evidence'.

(15) [A group is debating where to go out to eat. One member says, “There are no vegetarian restaurants around.”]

#Aren’t there any vegetarian restaurants around?

Outer-negation questions, too, are felicitous in the presence of evidence against the radical proposition p . For instance:

(16) [A group is debating where to go out to eat. One member says that there are no health food restaurants around.]

Aren’t there some vegetarian restaurants around here?

The picture painted by Buring and Gunlogson (2000), then, is summarized in the following table, where different evidential conditions are across the x axis, and question types are on the y axis.

	For p	Against p	Neutral wrt p
(17) Outer-negation question	#	✓	✓
Inner-negation question	#	✓	#

Sudo (2013), recognizing the limitations of the contextual-evidence-only characterization presented by Buring and Gunlogson (2000), suggests a two-dimensional classification. According to his account, questions come with an **evidential bias** condition, but also an **epistemic bias** condition.

For Sudo, a context can provide either negative or positive evidence with respect to a proposition p , or it can be neutral. Negative evidence makes p less likely, positive evidence makes p more likely, and a neutral context has no effect. Based on this classification of evidence types, Sudo identifies five possible evidential conditions a question can carry:

(18) **Possible evidence conditions** on polar questions from Sudo (2013)

- a. [+positive]: The question is felicitous only with evidence for the radical proposition p .

- b. [**−positive**]: The question is felicitous only in the absence of evidence for the radical proposition p .
- c. [**+negative**]: The question is felicitous only with evidence against the radical proposition p .
- d. [**−negative**]: The question is felicitous only in the absence of evidence against the radical proposition p .
- e. [**neutral**]: The question is felicitous only if there is no evidence for or against the radical proposition p .

On top of this five-way classification of evidential conditions, Sudo (2013) identifies **three types of epistemic bias condition**: positive, negative, and none. As opposed to the evidential conditions, which impose restrictions on contexts in which a question is felicitous, epistemic conditions care only about the speaker’s epistemic state. Positive epistemic bias requires the speaker believe p , or possibly believed p before entering the current context; Sudo is noncommittal on this point. Negative epistemic bias requires the speaker have believed $\neg p$. And “no” epistemic bias imposes no particular restriction.

Clearly, this is a powerful classification scheme: with five possible evidential conditions and three possible epistemic conditions, we predict at least 15 possible bias configurations, of which only a few are instantiated by actual question types. In particular, in English, Sudo identifies the following patterns:

		Evidential bias	Epistemic bias
(19)	Outer negation	−positive	positive
	Inner negation	+negative	positive
	Canonical polar question	−negative	none

This classification of question bias types can be seen as an expansion on the one offered by Büring and Gunlogson (2000). Whereas Büring and Gunlogson only allowed questions to impose the evidential conditions *positive* and *negative*, this scheme allows a much more fine-grained description of the particular bias conditions imposed by any question.

Note that the [–negative] evidential bias of canonical polar questions, in the last row of the table, is an endorsement of Büring and Gunlogson’s (2000) Evidence Condition on polar questions, which says that (canonical) polar questions are felicitous only if there is no evidence for the inverse proposition.

Note, too, that two things are shared by the two types of neg-biased questions. The first is positive epistemic bias, i.e. that the speaker believes the radical proposition p ; and the second is a lack of contextual evidence for the radical proposition p . In inner-negation questions, this is strengthened to a requirement of evidence for the inverse proposition $\neg p$.

We will argue later, in Section 4.4.3, that the distinctive pattern for inner-negation questions is due to the presence of the NPI, which contributes a domain-widening effect (Kadmon and Landman, 1993). For both Büring and Gunlogson (2000) and Sudo (2013), inner-negation questions are distinguished by their apparent requirement to be in a context that provides evidence *against* the radical proposition, rather than one merely lacking evidence *for* the radical proposition. This seems to be a difference of degree rather than kind: inner-negation questions like contexts where we’re very much disinclined to infer p , and outer-negation questions like contexts where we’re only somewhat disinclined to infer p . For us, this difference will come from the use of an NPI; expressing a domain-widened question, it will implicate that the speaker doesn’t believe the radical proposition with an unwidened domain. This state of affairs, consequently, is best instantiated by a context with evidence “against” p (i.e., suggesting but not directly entailing $\neg p$). This approach will allow us to treat inner- and outer-negation questions as one and the same, modulo the presence of the (independently licensed) NPI.

Admitting for now, though, the possibility that inner- and outer-negation questions really represent two different types of constructions, the following picture emerges from Büring and Gunlogson (2000) and Sudo (2013). Neg-biased questions carry with them an epistemic bias on the part of the speaker, and (to a greater or lesser extent) are preferable in contexts without evidence for the radical proposition p . Zooming in, both claim that outer-negation

neg-biased questions want a context that does not provide evidence for the radical proposition p , which we dispute; neg-biased polar questions with PPIs seem OK, so long as the speaker is willing to accept the possibility that they may be wrong. Inner-negation neg-biased questions want a context that provides evidence for $\neg p$.

Neither Büring and Gunlogson (2000) or Sudo (2013) present a way for getting from the denotation of a question to these conditions. Instead, both seem to perceive the conditions as something like conditions of use or presuppositions: they are conventional requirements on asking the question, not inferences derived from the meaning of the question. Even accepting that, however, we have no real idea where these conditions come from: why do questions care about evidential or epistemic bias? Why should there be an (at least) 15-way typology of bias in questions?

As these questions remain unanswered, it is best to think of these bias typologies as primarily descriptive: they present generalizations about the contexts in which polar questions of various kinds are licensed, but they are probably not underlying requirements on the use of those polar questions. Instead, we want them to follow from more general conditions on asking questions, and from the meanings of polar questions, including neg-biased questions in particular.

In the next section, we will move on to discuss existing accounts that do attempt to derive neg-bias from the semantics and the conventions of use regarding polar questions, remembering the generalizations made by Büring and Gunlogson (2000) and Sudo (2013).

4.2 Previous analyses of neg-bias

Ladd's (1981) seminal paper does not provide an articulated analysis of high negation in polar questions. The first major account to do so that we will consider is the Verum account of Romero and Han (2004), which analyzes the inner/outer distinction in terms of a scope interaction with a VERUM operator (Höhle, 2018). The denotation of VERUM, a silent morpheme, is a propositional operator Romero and Han take to be synonymous with

emphatic *really*; it contributes that the speaker takes the proposition it operates on to be in the Common Ground, in every epistemically accessible world (their 7). Here, $\text{Epi}_x(w)$ is x 's epistemic state at w , a set of worlds; and $\text{Conv}_x(w)$ is the set of worlds where all of x 's conversational goals are fulfilled.

$$(20) \quad \llbracket \text{VERUM}_i \rrbracket = \llbracket \text{really}_i \rrbracket^{g/xi} \\ = \lambda p_{\langle s,t \rangle} \lambda w. \forall w' \in \text{Epi}_x(w) [\forall w'' \in \text{Conv}_x(w') [p \in \text{CG}_{w''}]]$$

Semantically, it takes a proposition p and returns a new proposition saying that for the parametrized agent x , every world epistemically accessible for x is such that, in order for x to achieve his conversational goals, p must be in the Common Ground. They abbreviate this meaning as “FOR-SURE-CG $_x$ ”, as it intuitively expresses that x takes p to be “for sure in the Common Ground”.

In order to derive inner- and outer-negation readings for polar questions, VERUM interacts with both negation and the polar-question-forming operator Q . Q , which takes a proposition p and returns the set $\{p, \neg p\}$, yielding a question meaning, obligatory takes widest scope.

This leaves two possible scope orderings for VERUM and negation, which Romero and Han take to correspond to the inner- and outer-negation interpretations of the question, respectively. These scope orderings are given below, with an example neg-biased question followed by an intuitive paraphrase in the Romero and Han (2004) framework.

$$(21) \quad Q \gg \text{VERUM} \gg \text{NEG} \gg p \quad (\text{Inner negation})$$

Aren't there any vegetarian restaurants around?

Paraphrase: ‘Is it for sure that there are no vegetarian restaurants around?’

$$(22) \quad Q \gg \text{NEG} \gg \text{VERUM} \gg p \quad (\text{Outer negation})$$

Aren't there some vegetarian restaurants around?

Paraphrase: ‘Is it (false that it's) for sure that there are some vegetarian restaurants around?’

The outer-negation question includes the contribution of negation in parentheses because the Q operator gives rise to a symmetrical semantics for polar questions. As it takes a proposition and returns the set containing it and its negation, a theory depending this Q operator to form question meanings cannot distinguish between questions with positive and negative polarity. Consequently, the negation ‘washes out’, in a sense, on Romero and Han’s account of outer negation: the high negation doesn’t actually contribute anything to the semantics of the question *per se*.

There are two primary issues with the Verum account. The first is that it proposes two interpretations for neg-biased questions, one of which is based on a negative question radical; viz. the outer-negation interpretation. An outer-negation question is supposed to have a meaning like follows:

$$(23) \quad \llbracket \text{Didn't you eat something?} \rrbracket = Q(\neg\text{FOR-SURE-CG}_x(\text{you ate something}))$$

The proposition under the Q operator, $\neg\text{FOR-SURE-CG}_x(\text{you ate something})$, is of negative polarity. Consequently, we should expect this question to behave like any other negative-polarity question radical. In particular, it should give rise to the four-way response-particle pattern that we observe for negative polar questions:

- | | | |
|------|---------------------|---------------------------|
| (24) | Are you not coming? | (Negative polar question) |
| | a. Yes, I am. | (Yes, p) |
| | b. No, I am. | (No, p) |
| | c. Yes, I'm not. | (Yes, $\neg p$) |
| | d. No, I'm not. | (No, $\neg p$) |

But the purported inner-negation neg-biased question does not pattern like this. Instead, it patterns like a positive polar question:

$$(25) \quad \text{Aren't there some vegetarian restaurants around?} \quad (\text{Inner-negation question})$$

- | | | |
|----|----------------------|-------------------|
| a. | Yes, there are. | (Yes, p) |
| b. | *No, there are. | (*No, p) |
| c. | *Yes, there are not. | (*Yes, $\neg p$) |
| d. | No, there are not. | (No, $\neg p$) |

This evidence suggests that the Verum analysis on which a scopal ambiguity gives rise to the inner/outer readings of neg-biased questions is not on the right track; if it were, then the outer-negation question should behave like a negative question.

But what if we assume that there is no genuine inner/outer distinction, and that all neg-biased questions arise because of the scope ordering $\text{VERUM} \gg \text{NEG} \gg p$? On such an “adapted Verum” analysis, negation doesn’t have the option of scoping over VERUM; instead, it’s the presence of VERUM that gives rise to the neg-biased interpretation.

An adapted Verum account would be odd, because we can see that high negation (i.e., *-n’t* cliticized to the auxiliary verb) is what conditions neg-bias. Specifically, if a polar question is formed with high negation, then it obligatorily has neg-bias; and (modulo the Gladiator readings discussed earlier), if the polar question does *not* contain high negation, then it can’t be neg-biased. But an adapted Verum account, which says that neg-bias comes from VERUM scoping over negation, would necessarily locate the source of neg-bias in the role of VERUM, and not the role of negation. Consequently, on the adapted Verum account, it’s strange that high negation should be required to trigger bias.

Why is VERUM not independently available? In fact, it is, according to Romero and Han (2004), but apparently only overtly—either with something like *really*, or else with intonational prominence (though see Goodhue 2018 for arguments against treating intonational prominence as expressing VERUM). But neither *really* nor intonational prominence are necessary in neg-biased questions. There is, then, an odd state of affairs whereby negation is necessary to trigger the neg-biased interpretation, but doesn’t contribute the bias itself. Instead, the somewhat mysterious VERUM operator intervenes between the question-forming

morpheme and negation to yield the intended meaning, but only in neg-biased questions.

Conceptually, an analysis along the following lines would be ideal. A positive polar question $p?$ and the corresponding neg-biased question *Isn't p ?* differ only in one way: that the latter has negation somewhere outside the question radical, in a position associated with the syntactically high negation. The adapted Verum account gets close to this, but it fails in that it has to posit the presence of a silent VERUM operator in neg-biased questions, rather than in all questions. Instead, it would be preferable if, for example, in neg-biased questions negation simply scoped over the question-forming morpheme. This is the shape of analysis that we will pursue in Section 4.3 with the theory of polar questions developed in Chapter 3.

van Rooy and Šafářová (2003) present an intriguing account of neg-biased questions that is couched in Bayesian decision theory. On their account, questions are licensed according to the speaker's utility value assigned to the answer. In particular, positive polar questions are licensed when the speaker's utility value for the (positive) radical proposition p is higher than for the question's inverse proposition, $\neg p$. Neg-biased questions, on the other hand, are licensed when the speaker's utility value for $\neg p$ is greater than for that of p . Finally, disjunctive polar questions (e.g. *Is he coming, or is he not?*) are licensed when these utilities are approximately equal.

Without getting into the formal details, which are irrelevant to our purposes, van Rooy and Šafářová (2003) identify two independent factors that drive utility. The first is informativity: more informative answers have higher utilities. Information-theoretically, an answer is more informative the less likely it is, as learning an unlikely proposition removes many worlds from the Common Ground, greatly reducing uncertainty about which world we're in. Questions based on unlikely radicals, therefore, generally have higher utility than questions in the same context based on likely radicals. On the one hand, this seems right; it is generally not good to ask questions about things that are extremely likely, as it feels unnecessary.

(26) Is the sun going to rise tomorrow?

On the other hand, this seems to conflict with the Evidence Condition on polar questions from Büring and Gunlogson (2000), which says that polar questions are infelicitous in the presence of evidence for the inverse proposition. In this framework, questions are expected to be generally better the more evidence there is against their radical proposition, at least insofar as we assume that evidence against the radical proposition lowers the likelihood of the radical being true. But this doesn't seem to accord with the Evidence Condition.

The second relevant factor for utility value that they identify is speaker goals, or speaker desires: a proposition p can have high utility (for a speaker) if it makes a speaker's goal proposition g more likely. For this, they give several examples. Consider a medical examination, e.g. one whose goal is to assess a child's wellbeing. The doctor might ask the parent, *Is your child apathetic?*, but not because learning this is very informative—perhaps it's quite likely in context, which makes it relatively uninformative. Instead, the doctor's goal in this context is something like having a diagnosis for the child, or determining if the child suffers from some particular illness. It is in light of this goal that the question has high utility.

There are a number of issues to raise with the van Rooy and Šafářová (2003) account, but also a number of great insights. The first issue is that, in conditioning felicity on the utility of the negative proposition, it suggests that neg-biased polar questions are simply the negation of positive polar questions. The evidence from answer particles alone suggests that this is not correct; neg-biased questions like *Aren't you hungry?* differ from genuinely negative (canonical) polar questions like *Are you not hungry?*

Recall that van Rooy and Šafářová (2003) predict polar question felicity when the radical is of higher utility than the inverse proposition. But it also says that neg-biased questions, like *Isn't p ?* are felicitous when $\neg p$ has greater utility! Consequently, their account cannot distinguish at all between the felicity conditions of genuinely negative polar questions (*Are you not hungry?*) and neg-biased polar questions (*Aren't you hungry?*); both depend on the

utility of $\neg p$. This is symptomatic of a kind of confusion that has plagued the literature on neg-biased questions: a confusion over what the radical proposition of a neg-biased question is, and whether it's the same as a question with negative polarity. This is a question that we will return to in Section 4.4.

The goal-based side of this account, though, bears a number of similarities to the analysis presented in Chapter 3 of this dissertation, which finds polar questions licensed when they contribute to achieving a speaker's goals—though this is cashed out on our account as a semantic notion, where goals provide the preferential ordering source for a speaker's doxastic state. Our account certainly shares with van Rooy and Šafářová (2003) a commitment to the role of speaker goals in licensing polar questions: it is impossible to accurately describe polar question felicity without reference to the speaker's goals. However, ultimately, the conflation of neg-biased and negative-polarity questions poses serious difficulties for van Rooy and Šafářová (2003), and it is not immediately clear how one might minimally modify their account to avoid this confusion.

Krifka (2015, 2017) presents a novel account of polar questions, including neg-biased polar questions, couched in a commitment space framework based on Cohen and Krifka (2014)³. On this account, a **commitment state**, the most basic characterization of the state of a discourse at a particular point in time, is modeled as a set of propositions. Commitment states are updated by illocutionary acts. Assertion of the proposition ϕ , for example, simply adds the proposition ϕ to the commitment state. A **commitment space** C is defined as a set of commitment states that contains its intersection, which is nonempty: $\cap C \neq \emptyset$, and $\cap C \in C$.

A commitment space C represents both the discourse state and the licit discourse moves at a point in a conversation. Rather than merely updating discourse states, then, the pri-

3. The two accounts are substantially the same, as far as the empirical coverage and structure of the analysis is concerned. Krifka (2015) presents an account on which questions express a request for commitment, whereas Krifka (2017) sees them as a request for assertion. However, as far as I can tell, nothing critical hangs in the balance for polar questions.

mary function of a speech act in conversation is to move between discourse *spaces*. To take the simple example of assertion, update with the assertion of the proposition ϕ does more than simply add ϕ to the commitment state; it constrains future licit discourse moves. For instance, it disallows the assertion of $\neg\phi$. Update of a commitment space, then, proceeds according to the following definition (Krifka’s 2015 example (3)):

(27) Update of a commitment space C with an illocutionary act \mathfrak{A} :

$$C + \mathfrak{A} = \{c \in C \mid \cap C + \mathfrak{A} \subseteq c\}$$

Krifka refers to $\cap C$ as the *root* of C : as the intersection of all the commitment states in c , themselves sets of propositions, it represents the set of propositions that every participant has committed to up to the current moment in the discourse. Therefore the definition for commitment space update via an illocutionary act says: update with \mathfrak{A} retains only the commitment states c in C that contain at least the propositions in $\cap C + \mathfrak{A}$. If a conversation only contained instances of assertion, then the conversation proceeds by successively adding information to get a new conversational root, and any future update must be consistent with the root at that moment.

However, we want to do more than simply add assertions. The last ingredient is a means of keeping track of the history of commitment state updates. This is necessary because Krifka’s account needs the ability to reject or otherwise cancel the contribution of other speech acts; and in order to do so, we need to be able to look back at the discourse before that contribution. For instance, I might respond to an assertion of yours with *No*, essentially rejecting the uptake of that speech act and “moving the clock back” on the commitment space. To keep track of past developments in the discourse, we make a list of discourse spaces, where speech acts act on the top element of the list, adding a new discourse space to the top each time. Rejection, then, amounts to retrieving the penultimate state and adding it to the top of the list. Krifka refers to this list as a **commitment state development**.

Functionally, there are two ways to update the discourse on this account. On the one

hand, we can add a proposition to every state in the commitment space, which adds it to the root, the intersection of all the commitment states in a space. Alternatively, we can keep the root the same, but add information to other states in the commitment space. This amounts to giving (or taking away) possible future developments in the discourse, as each of these states represents a possible future root.

Krifka (2015, 2017) proposes that there are two kinds of questions, differing by how they constrain future developments of the discourse as outlined above; he further proposes that proffering different discourse developments accounts for (some of) the observed bias across polar question types. For example, the *rising declaratives* of Gunlogson (2002), which have declarative syntax but rising intonation associated with questions, have been claimed to exhibit a bias toward the radical proposition:

(28) Gus won the race?

Questions, including rising declaratives, do not add a proposition to the root of the commitment space. Instead, they constrain the development of the discourse. Krifka claims that rising declaratives are ‘monopolar’, admitting only a discourse continuation in which the addressee commits to the radical proposition (here, *Gus won the race*). In effect, asking the rising declarative above puts the addressee in a position where their only licit move, besides rejecting the previous move, is to commit to *Gus won the race*.

For Krifka (2015), canonical polar questions, on the other hand, are ambiguous between a monopolar (‘biased’) and bipolar (unbiased) interpretation:

(29) Did Gus win the race?

Allegedly, this question has a monopolar interpretation like the rising declarative, on which it asks the speaker to commit to *Gus won the race*; but more standardly, it has a bipolar interpretation, on which it invites the speaker to commit to either *Gus won the race* or the

logical inverse, *Gus didn't win the race*. But it's not entirely clear how we are meant to distinguish between the putative monopolar and bipolar readings of canonical polar questions. The only examples Krifka gives of monopolar readings for neg-biased questions are questions of negative polarity, which, he claims, (can) have a biased interpretation, toward the negated proposition, i.e. the question radical:

(30) Did I not win the race?

This is assigned a meaning the same as the rising declarative *I didn't win the race?*, i.e., it creates a discourse context in which the addressee must either commit to the negative proposition or reject the discourse move. But clearly these constructions are different. For one, a rising declarative can be used to respond to an assertion of the radical proposition, but a negative question cannot:

- (31) a. A: You didn't win the race.
 B: I didn't win the race?
- b. A: You didn't win the race.
 B: #Did I not win the race?

This is behavior that puts negative polar questions in line with other polar questions, but is incongruent with rising declaratives. And from a purely conceptual standpoint, it's odd that polar questions—or perhaps only a subset of them—should have an ambiguity that makes them potentially synonymous with rising declaratives. It would be preferable, of course, that all canonical polar questions have the same interpretation, which is distinct from rising declaratives, which clearly pattern differently.

However, it's this purported ambiguity in canonical polar questions that Krifka exploits to account for neg-biased questions, which he analyzes as expressing a request for the *denegation* of the speech act of asking a monopolar question, i.e., a disavowal of the speech act on the

part of the addressee. The high negation of neg-bias actually appears in the specifier of the *commitment phrase*; the commitment phrase's complement is the TP whose denotation is the radical proposition.

A canonical polar question (on its monopolar interpretation) presents the addressee only with discourse continuations in which they commit to the radical proposition. A neg-biased question, on the contrary, presents the addressee only with discourse continuations in which they do *not* commit to the radical proposition. This, then, is Krifka's source for the bias of neg-biased polar questions.

There are a number of issues to raise with the commitment-space account of neg-bias. The first is the purported ambiguity in polar questions, on which they can alternatively express a request for commitment to the radical proposition, or a request for commitment to either one of the radical or the inverse propositions. On the one hand, this presents a (partial) solution to the asymmetry problem discussed in Chapter 2; the monopolar interpretation is asymmetrical. But on the other hand, the only real reason to posit this ambiguity is in order to have a source construction for the neg-biased question (Krifka 2017 refers to it as an 'assumption').

A better state of affairs might be to jettison the 'bipolar' interpretation entirely; this would yield an analysis similar to the one in this dissertation, although couched in a speech act framework and not a modal/attitudinal one. On such an account, a polar question simply presents to the interlocutor discourse futures in which the interlocutor commits to the radical proposition. This would, of course, require a different analysis for rising declaratives, such that they don't systematically receive the same interpretation as canonical polar questions.

A higher-level issue with the Krifka-style discourse-space approach to neg-bias has to do with the discourse roles and speech acts performable with polar questions. Krifka's accounts basically build in the interlocutor to the meaning of a question; they require an interlocutor who can commit to a proposition, and their only real function is to impose a requirement on that interlocutor. But polar questions can be used in the absence of an interlocutor. For

instance, there are wish uses of polar questions that don't require an interlocutor:

(32) Will it stop raining?

(33) Does the mockery never stop?

Additionally, second-person-directed wish-type polar questions can be directed toward inanimate objects. Here, it doesn't really seem plausible to interpret the question as imposing discourse restrictions on the addressee:

(34) [*Said to a computer*]
Will you stop freezing?!

In this case, it's not clear how we could 'coerce' the question into being directed towards the computer's future discourse moves. The expressive-attitudinal analysis of this dissertation fares better with all of these wish-type uses, as on that approach, each of these questions simply expresses the speaker's conditional preference to know the radical proposition, according to their goals. In other words, since the addressee—and in particular, the addressee's discourse moves and commitments—are not built into the semantics of the question, these don't require any special story.

Regardless of the status of canonical polar questions, however, Krifka's commitment-space approach to neg-bias questions itself also raises some questions. In particular, how can we get from a request for the denegation of a commitment to bias toward that proposition on the part of the speaker? Krifka's explanation of this is not particularly specific, e.g. it does not provide a pragmatic account that gets us from the commitment-space semantics for neg-biased questions to the inference of speaker bias. Krifka (2017) does provide two hints. First, in the "outer-negation" case, he suggests that the bias inference is the result of a speaker's goals:

(35) *[A group is trying to decide where to eat. The speaker has been to a vegetarian restaurant nearby.]*

Isn't there a vegetarian restaurant around here?

Here, Krifka (2017) refers to a rhetorical strategy on the part of the speaker, namely 'to appeal to the addressee to exclude certain options in order to find a solution, here the best restaurant choice' (394). But this can't be right: the speaker's request in a neg-biased question is for the addressee to refrain from committing to ϕ , and the addressee's refraining from commitment can hardly be taken to inform the group's choice of restaurant, except very obliquely. A similar argument is made for the "inner-negation" neg-biased question: that the neg-biased question is used to rule out options. However, this doesn't get us to the bias inference.

We can imagine a pragmatic story of neg-bias on the commitment-space account according to which the request for the addressee not to commit to the radical proposition p expresses the speaker's confidence in p : because the speaker is so sure of p , they request that the addressee simply not commit to it. But this account, too, would be odd, because neg-biased questions are extremely naturally answered in the affirmative:

(36) A: Isn't there a vegetarian restaurant around here?

B: Yes, there is.

For this reason alone, the "request not to commit" story seems to be insufficient; neg-biased questions intuitively request an assertion of the radical proposition just as much as canonical questions do. At the very least, a great deal more needs to be said for Krifka's account to really deliver an explanation for neg-bias.

The last approach to neg-bias that we will consider is that of Goodhue (2018, 2022), which is inspired by Krifka (2017). Goodhue attempts to provide a genuine pragmatic story for why a Krifka-like account would result in the bias inference characteristic of neg-biased

questions. He takes this to follow from the “unbalanced” partitioning induced by a high-negation question, providing a fairly complex pragmatic story that attempts to derive the bias as an implicature.

In particular, Goodhue assumes that high-negation polar questions involve some modal operator \Box with universal force, possibly of “conversational modality” as in Krifka (2015), but possibly of epistemic modality. Negation scopes over this operator, identified with Krifka’s Assert or Commit operator, as it does in Krifka’s account; however, instead of the Request operator of Krifka (2017), the outermost question-forming operator simply returns a set containing its prejacent and the negation of its prejacent. This gives rise to the following partitioning of logical space for a neg-biased question *Isn’t p?*:

(37) Goodhue’s denotation for a neg-biased question *Isn’t p?*:

$$\{\neg\Box p, \Box p\}$$

For a canonical polar question, on the other hand, the two options presented to the addressee are that they commit to p , or to $\neg p$:

(38) Goodhue’s denotation for a canonical polar question *p?*:

$$\{\Box p, \Box\neg p\}$$

As in Krifka’s account, this builds the assertion operator into the meaning of the question directly. On Goodhue’s account, neg-bias comes from the speaker’s choice to ask the question with the ‘unbalanced’ partition of *Isn’t p?*, rather than a canonical polar question. This is a sort of Quantity implicature—both answers to (38) entail an answer to (37), but not vice-versa. That is, the $\Box p$ cell of the canonical polar question entails $\Box p$ in the neg-biased question; and $\Box\neg p$ in the canonical polar question entails $\neg\Box p$, which is strictly weaker. Therefore, the canonical polar question is ‘stronger’ than the neg-biased question.

This is what Goodhue identifies as the source of the bias inference: since the speaker

chose to ask the weaker question, they must have had a reason. This reason, in turn, is that they are not ignorant of whether p or $\neg p$: if they were, they would have asked the canonical polar question.

This is a story about why there is a bias, but so far it does not tell us why the bias is for the radical proposition p . Finally, Goodhue derives the bias for p (as opposed to $\neg p$) from an argument about speaker-addressee coordination: the speaker wants to know whether the addressee shares their bias. If the speaker were biased for $\neg p$, then the $\neg \square p$ cell of the neg-biased question would not resolve that issue. But if the speaker is biased for p , then it does.

As a development of Krifka's account, Goodhue's runs into a few of the same problems. One is that it is inherently discourse-oriented: it relies on the speaker reasoning about the addressee and the addressee's actions. But, as with canonical polar questions, neg-biased questions can be asked even in the absence of an animate addressee:

(39) *[Said while flipping past TV channels alone]*

Isn't there anything good on?

The account of neg-bias that we will develop, based on the account of polar questions developed in Chapter 3, eschews a partition semantics entirely, but is in broad strokes somewhat similar to this account. In particular, we will also have negation take wide scope in neg-biased questions, which allows us to derive the bias inference.

However, because the attitudinal approach of this dissertation doesn't use a partition semantics, negation can straightforwardly take widest scope. This will have the happy consequence of allowing us to derive the bias inference straightforwardly, without a baroque pragmatic story. Instead, the speaker's bias will follow from the effect of negating a conditional preference to learn.

4.3 Extending our analysis to neg-biased questions

Existing accounts of neg-bias run into a few high-level problems, which we will briefly recap here.

First, a number of accounts characterize the bias of neg-biased questions in terms of evidence and speaker belief, but don't relate it to the semantics of the question (Büring and Gunlogson, 2000; Sudo, 2013; Gyuris, 2017). As these accounts don't provide a semantic story, and rather categorize question types by bias characteristics, they can only really be taken to provide a description of the felicity conditions of neg-biased questions, and not to provide an account of them *per se*.

Second, within the domain of semantic accounts, a prevailing issue is cohesion within an analysis of polar questions. The Verum account of Romero and Han (2004), for example, doesn't clearly fit into an analysis of canonical polar questions; the Verum operator is recruited to play a special role in neg-biased questions, but doesn't appear in canonical polar questions. For Krifka (2015, 2017), too, neg-biased questions are built from a special, and not particularly motivated, case of polar questions—the monopolar polar questions, which present the addressee with a single discourse continuation, as opposed to the more canonical bipolar polar questions.

Third, semantic accounts struggle to explain how we get from semantically high negation (outside the question radical, above some operator embedding the radical proposition) to the inference of speaker bias. In part, this is due to confusion about what bias consists in: is bias a property of contexts, of speakers, or of the utterance itself? Ever since Ladd (1981), it has been common to presuppose that bias is located in the context, and to provide an explanation of neg-bias by presenting a story for why a speaker might choose that question in such a 'biased context'. This is the shape of the argument that Krifka (2015, 2017) sketches for deriving the bias inference. Such an explanation, however, can not be sufficient unless it explains why we use neg-biased questions *only* in biased contexts, and never in neutral contexts.

Finally, semantic accounts based on a proposition-set partition semantics do not provide a satisfying account of the special role of negation in neg-biased questions. Before the introduction of the partition-forming operator Q , we have a semantic object with negative polarity, introduced by the high negation. But the partition-forming operator Q effaces the role of negation: once it's applied, our denotation for the question can no longer tell us whether the argument of Q was of negative or positive polarity. What derives the bias on these accounts, then, is not the presence of negation, but the partition structure involving some kind of operator (e.g. VERUM or Goodhue 2018's \square operator). The high negation of neg-bias, then, does not directly yield the bias inference. Instead, it serves as a signal that the question denotes a partition whose elements contain that operator, which is the ultimate source of the bias.

An ideal account of neg-biased polar questions should have them share the semantics of canonical polar questions, with the difference between the two being that neg-biased polar questions locate the semantic role of negation in a designated position outside of the question radical. In turn, the presence of negation should straightforwardly give rise to the bias inference, and not merely serve as a signal to an 'unbalanced partition' (Goodhue, 2018), but bear no explicit relationship to negation.

This chapter presents an account that avoids these issues. First and foremost, our account of neg-biased polar questions will be identical to our account of canonical polar questions, modulo the introduction of external negation. This account, additionally, does not require the introduction of a covert operator like VERUM that must co-occur with high negation to yield the bias reading. Instead, the polar question merely represents a particular kind of preference on the part of the speaker, and the neg-biased question, by negating that, represents the speaker's lack of such a preference. Finally, this account straightforwardly related neg-bias to negation. Neg-bias is derived as a semantic inference from the speaker's expression of a lack of preference for epistemic update.

The shape of the present account is as follows. We retain the primary insight of Chapter 3,

namely that polar questions express a speaker’s doxastically oriented conditional preference. However, in order to account for neg-biased questions, the speaker’s preference must regard *doxastic update*, and not merely knowledge. This preference for update will allow us to capture the behavior of neg-biased questions while retaining the insights of the attitudinal analysis of polar questions. We will, in turn, reinterpret the ‘preference to know’ component as separate convention of use (Condoravdi and Lauer, 2012) of polar questions.

With that in place, the neg-biased question express that a speaker does not prefer to have their doxastic state updated with respect to the radical proposition—that is, the speaker could do without doxastic update. However, as a separate convention of use, ‘preference to know’ component remains in effect. The upshot is that the speaker expresses that they do *not* prefer to update their knowledge state regarding p if p , but they independently prefer to *know* p if p . These two conditions conspire to guarantee that in neg-biased questions, the speaker has a prior belief that p .

4.3.1 From believing to learning

Unfortunately, it will not do to simply negate our existing denotation for polar questions. As a reminder, this is the denotation we developed for a polar question $p?$ asked by a speaker x in a world w :

$$(40) \quad \llbracket p? \rrbracket_x^w = \forall w' \in \text{Dox}_x^w[p][(\text{Sim}^{w'}(v)(p \subseteq \text{Dox}_x^v)) \leq_{\text{Goals}^{w(x)}} (\text{Sim}^{w'}(v)(p \not\subseteq \text{Dox}_x^v))]$$

In other words, a polar question expresses that, for every world in a speaker’s doxastic state where p , the speaker would achieve more (sub)goals if the world were minimally different such that they believed p . If we simply stick a negation sign in front of this denotation, then we would have (by the equivalence of $\neg\forall x[Px]$ and $\exists x[\neg Px]$):

(41) First-pass attempt at neg-bias:

$$\llbracket \text{Isn't } p? \rrbracket_x^w = \exists w' \in \text{Dox}_x^w[p][\neg(\text{Sim}^{w'}(v)(p \subseteq \text{Dox}_x^v)) \leq_{\text{Goals}^{w(x)}} (\text{Sim}^{w'}(v)(p \not\subseteq \text{Dox}_x^v)]$$

$Dox_x^v))]$

The above attempt at an analysis negates (at some p -world) that x 's knowing p is preferable to x 's not knowing p . This, in turn, amounts to saying that x 's ignorance of p is preferable to x 's knowing p at that world. In effect, we flip the order of the two propositions:

(42) First-pass attempt at neg-bias:

$$\llbracket \text{Isn't } p? \rrbracket_x^w = \exists w' \in \text{Dox}_x^w[p] [\llbracket (\text{Sim}^{w'}(v)(p \not\subseteq \text{Dox}_x^v)) <_{\text{Goals}^w(x)} (\text{Sim}^{w'}(v)(p \subseteq \text{Dox}_x^v)) \rrbracket]$$

In plainer words, this says that for some world in the speaker's doxastic state where p , that world would be better if the speaker were ignorant of p than if the speaker knew that p . But clearly this is not the meaning of a high-negation question: it does not express that the speaker might prefer to remain ignorant of p if p !

What goes wrong with this first-pass analysis? Intuitively, we're negating the wrong thing: we end up with an analysis that says the speaker doesn't mind (i.e., does not have a preference against) *not* believing p if it's true. But it's not the speaker's belief of p that we want to negate; crucially, we want the speaker to believe p in the end! So what is the element of the meaning of the question that is targeted by negation in neg-biased questions? I propose that, instead, what is negated in a high negation question is not the speaker's belief, but the speaker's **doxastic update with p** .

In other words, what a neg-biased question ought to express is the following meaning:

(43) There is some p -world in x 's doxastic state where as many (sub)goals are accomplished in the minimally different world where the speaker does not **update his doxastic state with p** as in the minimally different world where the speaker does update his doxastic state with p .

And, consequently, the meaning of a canonical polar question will have to be changed to

the following, where it expresses a preference for doxastic update, and not a preference for a static belief:

- (44) For every p -world in x 's doxastic state, more (sub)goals would be accomplished if world were minimally different such that the speaker **updated his doxastic state with p** .

For canonical questions, this change will not result in any obvious differences the preference to believe from Chapter 3. But in order to cash this out formally, we need a concrete notion of doxastic update with respect to a proposition.

For ease of explication, we'll describe doxastic update as an event of 'coming to believe', an event which causes the agent to be in a state of belief. Coming-to-believe could probably also be characterized in a dynamic epistemic logic (Van Benthem, 2011) or a similar framework. Here we use an adapted version of neo-Davidsonian event semantics (Maienborn, 2019; of Philosophy & Linguistics Terence Parsons and Parsons, 1990):

- (45) **Definition:** Come-to-believe (ctb) event:

$$\mathbf{ctb}^{w,t}(p)(x) =_{\text{def}} \exists e[\text{COME-TO-BELIEVE}^{w,t}(e) \wedge \text{PATIENT}^{w,t}(x)(e) \wedge \\ \exists s[\text{CAUSE}^{w,t}(s)(e) \wedge \text{BELIEVE}^w(p)(x)(s)]]$$

where $\text{BELIEVE}^w(p)(x)(s)$ iff for every time t in the trace (temporal extent) of state s , $p \subseteq \text{Dox}_x^{w,t}$.

This definition says that x 's coming-to-believe p is an event which causes a state of belief that p in x . This is the essential object of the speaker's preference when asking a polar question. Again, the event semantics representation of this meaning is not necessarily crucial. What is crucial is that the object of the speaker's preference be *updated* with the radical proposition, rather than merely a state of knowing the radical proposition; how that update is expressed should not necessarily be relevant.

For a canonical polar question, then, the speaker goal-prefers coming to believe p over not: every epistemically accessible world is such that, if we minimally change it so that there is an event of the speaker coming to believe p , it achieves more speaker (sub)goals than if we minimally change it so that the speaker does not come to believe p . That denotation is given below:

$$(46) \quad \llbracket p? \rrbracket_x^w = \forall w' \in \text{Dox}_x^w[p] [(\text{Sim}^{w'}(v)(\mathbf{ctb}^{v,t}(p)(x)) \\ <_{\text{Goals}^w(x)} (\text{Sim}^{w'}(v)(\neg\mathbf{ctb}^{v,t}(p)(x)))]$$

This formulation bears deep similarities to Krifka’s (2017) analysis of (monopolar) polar questions as the combination of a REQUEST and an ASSERT operator, essentially requesting (by limiting future legal discourse moves) that an interlocutor assert the radical proposition. Here, Krifka’s ‘request’ is analogous to the speaker’s goal-oriented preference, and Krifka’s (addressee’s) assertion is analogous to the speaker’s doxastic update. Whereas Krifka’s question-asker requests an assertion that p , our question-asker prefers to learn that p . But this analysis of polar questions is different in a number of important respects. For example, it is couched solely in a speaker’s doxastic state, and does not require an interlocutor. As such, it is expressive, and not discourse-oriented at all; the primary effect of a polar question is to express something akin to *Boy, I’d love to learn that p!*. This will make it easier to extend to the neg-biased questions, as we will see in a moment.

One thing to keep in mind is that an agent is introspective with respect to his beliefs: he always knows the contents of his own current doxastic set. And this fact plus the denotation for a polar question above actually allows us to derive the generalization that asking a polar question requires the speaker not believe p . In effect, the question expresses a preference for a coming-to-believe event, but there is no such doxastically possible event if the speaker already believes p .

We can show this informally by a reductio argument: suppose that a speaker believed p and nevertheless asked $p?$. The question expresses the speaker’s preference to come to

believe p over not coming to believe p : that for every doxastically possible world, the closest world to that where there is a come-to-believe- p event achieves more (sub)goals than the closest world where there is no such event. In order for this to be the case, there must be similar worlds where the speaker does not currently believe p : the come-to-believe- p event causes belief in p , and events can't cause states that start in the past. Additionally, because the similarity function Sim ranges over worlds in the speaker's doxastic set, those similar worlds where the speaker does not currently believe p must be in the speaker's doxastic set. But we stipulated that the speaker believes p , and from the speaker's doxastic introspection, in every doxastically accessible world, they believe p . Therefore, there are some worlds in which they both believe p and don't believe p : a contradiction.

Interestingly, the same argument does not, strictly speaking, hold if the speaker believes that $\neg p$ —if we allow that speakers can remove propositions from their doxastic set. If the speaker believes $\neg p$ now, then there may very well be doxastically possible worlds where the speaker nonetheless comes to believe p later; they simply, at some point, disavow their previous belief that p . In other words, this account predicts an asymmetry between the acceptability of polar questions depending on the speaker's prior belief: if he believes p , then a question is infelicitous; but if he believes $\neg p$, then the question is felicitous so long as the speaker admits the possibility that he revises his belief later.

And in fact, this doesn't seem like a bad prediction. If I do really believe that the Tigers will win the game, then it's odd for me to ask *Will the Tigers win the game?*—I already believe it, so why am I asking? But if I believe the Tigers will *not* win the game, the question seems better; it expresses my preference to come to believe that proposition if it's true, even if that requires revising my current belief that the Tigers will lose.

Now we can move on to neg-biased questions. Negating the denotation for a polar question should straightforwardly give us the meaning of a neg-biased question. Since the meaning of a question says that every doxastically accessible world w' satisfies $[\text{Sim}^{w'}(v)(p) <_{\text{Goals}^{w}(x)} \text{Sim}^{w'}(v)(\neg p)]$, this amounts to saying that, for some doxastically accessible

world, the ordering is reversed (and the ranking relation induced by the set of (sub)goals is weakened to ‘at least as highly ranked as’). In other words, we basically change the universal quantifier to the existential, and flip the order of the preference:

$$(47) \quad \llbracket \text{Isn't } p? \rrbracket_x^w = \exists w' \in \text{Dox}_x^w[p] [(\text{Sim}^{w'}(v)(\neg \mathbf{ctb}^{v,t}(p)(x)) \\ \leq_{\text{Goals}^w(x)} (\text{Sim}^{w'}(v)(\mathbf{ctb}^{v,t}(p)(x)))]$$

This denotation says that, for at least some doxastically accessible p -world, the speaker achieves more (sub)goals if that world is minimally changed so that there is *no* come-to-believe- p event than if there is one.

There are two ways that there might *not* be a come-to-believe- p event: either x never believes that p , remaining ignorant (or believing $\neg p$); or x already believed that p , and so there is no come-to-believe event after the utterance time. Let’s call these the *Never Believed* and *Always Believed* scenarios, and consider them in turn.

In the *Never Believed* scenario, there is some doxastically accessible p -world where x never believed that p . Consequently, then, x must not believe p in the actual world, by introspection: if it’s consistent with x ’s beliefs that he doesn’t believe p , then he simply doesn’t believe p in the actual world. The meaning of the neg-biased question, then, would be that in some p -world, x achieves more goals by remaining ignorant of p than by learning p .

In the *Always Believed* scenario, there is some doxastically accessible world where x already believes that p , and so never comes to believe it. The neg-biased question semantics, then, says that x achieves at least as many (sub)goals in this world as in the closest doxastically accessible world where x comes to believe that p . As in *Never Believed*, it therefore must be the case that x believes p in the actual world: if it’s consistent with x ’s beliefs that x believes p , then x simply believes p . But that means that there cannot be a closest doxastically accessible world where x comes to believe that p : x already believes p , so there are no epistemically possible worlds where he does not! Consequently, a neg-biased question

is trivially true so long as x believes p , as the lower-ranked proposition is necessarily false.

Therefore, we have two predicted possible scenarios: that the neg-biased polar question expresses the speaker's lack of a preference for doxastic update because he believes p , or because he does not believe p :

(48) **Predicted readings of neg-biased polar questions**

Never Believed: The speaker doesn't believe p , and even if p is true, might achieve as many (sub)goals without coming to believe p as he would if he did come to believe p

Always Believed: The speaker believes that p , and if p is true, he would achieve as many goals without coming to believe p as he would if he did come to believe p (which is impossible)

Why, then, is it that neg-biased polar questions seem to only have the Always Believed interpretation? I propose that this has to do with a secondary requirement: that ignorance cannot further goals. This can be cashed out a number of ways; I am non-committal on the best implementation. For now, we can consider it a generalization on what is a possible (sub)goal. In particular, I formulate the following generalization:

(49) **Anti-ignorance condition on (sub)goals:** Speakers never take it to be necessary to remain ignorant in order to achieve some goal.

This anti-ignorance condition says that a speaker's goal structure will never include remaining ignorant about the truth of a proposition. We don't want to add *Learn true propositions* to the goal set *tout court*, as that would license all polar questions: if speakers have a general goal to learn true propositions, then learning any proposition will further it, and we lose the ability to rule out any polar questions. But there are other ways that one could build essentially this same effect into the system. For instance, we might take it to be a condition of

use of polar questions that the speaker wants to know the radical proposition if true, which is a restatement of our original semantics for polar questions. However, this seems unprincipled, as it would have to be at some other level of representation that avoids being negated by the high negation of neg-biased questions and would mostly recapitulate the semantic content of the question. It seems better to rule out the Never Believed interpretation of neg-biased questions by some general principle about speaker goals.

As we noted before, the Always Believed interpretation of neg-biased questions is trivially true so long as the speaker believes the radical p . That is, it doesn't actually depend on the speaker's preferences, since there are no doxastically possible worlds where the speaker comes to believe p . However, the question still communicates that the speaker believes the radical proposition p , so its contribution is not entirely trivial. Moreover, this prediction actually seems plausible. Compared to canonical polar questions, neg-biased questions seem much less context-sensitive. At the very least, the role of bias seems to make the speaker's goal structure irrelevant for determining felicity of neg-biased questions; the question really seems to only depend on the speaker's beliefs.

For example, whereas (50-a) and (50-b) can perform different roles in different contexts, the felicity of their corresponding neg-biased questions (51-a) and (51-b) really just seems to depend upon the speaker's belief state, and nothing else.

- (50) a. Are you hungry?
b. Are you not hungry?
- (51) a. Aren't you hungry?
b. Aren't you not hungry?

The resulting picture, then, is that neg-biased questions straightforwardly negate the meaning of a polar question. This means that the speaker has some doxastic world where they either currently believe p or do not believe p . A ban on ignorance as a speaker goal rules

out the latter scenario, leaving the only possible interpretation that the speaker currently believes p . And since the speaker has a doxastically accessible world where he believes p —that is, according to what he believes, he *might* believe p —he must believe p at the actual world.

This, then, derives the neg-bias inference, that the speaker believes the radical proposition, as a semantic consequence of the representation of the neg-biased question. The fact that we derive this semantically, rather than pragmatically, will be useful as we move forward to consider canonical and neg-biased questions with NPIs, and the strength relationships among them.

4.4 Against the inner-/outer-negation distinction

With an analysis of both canonical polar questions and neg-biased questions in place, we can flesh out an argument against the inner-/outer-negation distinction that has been made in the literature. This section will argue against making this distinction, instead preferring to locate the differences in acceptability between the two purported question types in the presence or absence of a negative polarity item, which has been traditionally taken to force inner-negation readings.

4.4.1 *Neg-biased questions and polarity*

Since Ladd (1981), it has been common to take the licenser of NPIs in inner-negation neg-biased questions to be high negation itself, which is present in the question radical. However, such an argument only works if neg-biased polar questions can truly be of negative polarity, i.e. have negation in the question radical.

Goodhue (2018) provides a bevy of tests that shows definitively that this is not the case: negation is not present in the neg-biased question radical. We have already seen that, in response to a neg-biased question, we observe the response-particle pattern characteristic of

positive-polarity questions:

- (52) Aren't there any vegetarian restaurants around?
- a. Yes, there are.
 - b. No, there aren't.
 - c. *Yes, there aren't.
 - d. *No, there are.

This suggests that, so far as polarity is concerned, neg-biased questions are of positive polarity. Goodhue also offers the following two tests. Essentially, these tests show that neg-biased questions pattern like positive polar questions, and not like negative polar questions, for a range of polarity-sensitive phenomena.

First, he notes that projective content, e.g. the presupposition of *again*, may include negation in negative-polarity canonical polar questions, but not in neg-biased questions. For instance, in the following questions containing *again*, the presupposition that the proposition has happened before contains negation only in the canonical polar question with negative polarity (Goodhue's (31)).

- (53) a. Did Danielle come to class again?
(*presupposes*: Danielle has come to class before.)
- b. Did Danielle not come to class again?
(*presupposes*: Danielle has, at least once, not come to class before.)
- c. Didn't Danielle come to class again?
(*presupposes*: Danielle has come to class before.)

The same pattern obtains with other presupposition triggers, e.g. *too*.

Second, Goodhue notes that only under negation can *until*-adverbials modify clauses with verbs denoting punctual events, such as *discover a hole in the wall* (De Swart, 1996).

When the clause is of positive polarity, the sentence is ungrammatical:

- (54) a. *The girls discovered the hole in the wall until midnight.
b. The girls didn't discover the hole in the wall until midnight.

In polar questions, the same pattern obtains. Crucially, neg-biased polar questions do not admit *until*-adverbials⁴:

- (55) a. *Did the girls discover the hole in the wall until midnight?
b. Did the girls not discover the hole in the wall until midnight?
c. *Didn't the girls discover the hole in the wall until midnight?

If neg-biased questions are ambiguous between an inner- and outer-negation interpretation that tracks whether negation appears in the question radical, then the neg-biased question (55-c) should simply force an inner-negation interpretation, rather than being wholly unavailable. The natural conclusion, then, is that neg-biased questions simply never contain negation in the question radical *per se*.

Crucially, this conclusion does not mean that (at least some of) the constructions that have been referred to as 'inner-negation' and 'outer-negation' are not genuinely different from each other. Rather, we simply shouldn't conclude that the presence or absence of negation in the question radical is what gives rise to the difference between them. It may be that some other property of the questions is what gives rise to this perceived distinction.

I argue that this property is simply the presence of a negative polarity item. The presence of a negative polarity item has long been taken to be diagnostic of an inner-negation reading; but what if the inner-negation reading consists solely of the neg-biased question containing an NPI? In order to investigate this issue, we first need to look at NPIs in neg-biased questions, and the particular examples that have been of interest in the literature so far.

4. Unless the lower clause is explicitly negated, e.g. in *Didn't the girls not discover the hole in the wall until midnight?*

4.4.2 *Negative polarity items in neg-biased questions*

There is reason to doubt that Ladd (1981), when identifying inner-negation questions, actually identifies a unique phenomenon. This section presents three reasons to doubt it. First, a number of the examples he points to are problematic for most speakers of English—namely, those with *either*. Second, one of the examples he presents, involving *even*, has been shown in later work (e.g. Guerzoni, 2003) to be independent of negation—and present in canonical polar questions as well as neg-biased questions. Finally, examples he invokes involving strong NPIs, e.g. *lift a finger*, are known to only have rhetorical uses, which accounts for the lack of ambiguity that he identifies for them. In sum, the NPIs that form the basis of analyses of inner-/outer-negation neg-bias are problematic in a number of respects, and many of them give rise to phenomena that we now know not to be related to neg-bias.

In general, polar questions license NPIs. As a result, we don't need negation to be semantically present in the radical proposition to license NPIs at all. And because polar questions generally license NPIs, it's not surprising that neg-biased polar questions do as well, even setting aside the role of high negation in those questions. Consequently, this section concludes, it is more parsimonious to identify the properties associated with 'inner-negation' interpretation with the presence of negative polarity items, independently licensed in polar questions, rather than with an alternation in the position of negation.

A number of Ladd's (1981) original examples used the NPI *either*, which has attracted some later critique. Sailor (2013), for example, experimentally confirms that relatively few English speakers admit the grammaticality of Ladd's original example with *either*:

(56) Isn't Jane coming either?

Indeed, my own dialect of English doesn't allow *either* in this context, and I have not found a speaker who accepts it—informal polling finds a general sense of confusion at the meaning of such questions. Ladd further claims that in his dialect of English, (56) is grammatical,

but the corresponding canonical polar question with *either*, below, is ungrammatical, a distinction which he attributes to the licensing role of negation:

(57) *Is Jane coming either?

If one has such a dialect of English, it's reasonable to think that the NPI might be licensed by the high negation of neg-biased questions. However, the much more common state of affairs seems to be that English speakers simply reject both of these. Consequently, *either* is not a very useful test case as a negative polarity item in questions. Instead, NPIs like *anything* are less clearly problematic:

(58) Did Ida eat anything?

(59) Didn't Ida eat anything?

Another example that Ladd points to regards *even*. *Even* is now known, from work taking place after the publication of Ladd (1981), to be ambiguous between an “easyP” and a “hardP” interpretation in some contexts, e.g. polar questions and the antecedents of conditionals (Wilkinson, 1996; Rooth, 1985; Guerzoni, 2003). In positive sentences, *even* is unambiguous, presupposing that its prejacent proposition is low on a scale of likelihood relative to other focus-alternative propositions. (In the following examples, SMALL CAPS marks focus.)

(60) Ida even ate the NATTO.

This sentence asserts that Ida ate the natto, and presupposes that it was unlikely for her to have eaten natto, as opposed to other relevant edible things. Guerzoni (2003) refers to this as the ‘hardP’ presupposition. (The original examples regarded solving problems on an exam.) In negative-polarity questions, the opposite presupposition obtains:

(61) Ida didn't even eat the NATTO.

This sentence asserts that Ida didn't eat the natto, but presupposes that it was a very *likely* thing for her have eaten. This is what Guerzoni (2003) calls the 'easyP' presupposition.

Crucially, when *even* appears in polar questions and the antecedent of conditionals, we observe a systematic ambiguity between easyP and hardP readings. The following examples provide contexts that bring out the easyP and hardP readings of a polar question and a conditional antecedent containing *even*, respectively:

(62) a. Did Ida even eat the NATTO?

easyP context: Ida loves natto and eats it whenever possible. However, yesterday she wasn't eating much.

hardP context: Ida hates natto and never eats it, but yesterday she was a very adventurous eater.

b. If Ida even ate the NATTO, then we'll take her out for ice cream.

easyP context: We're looking for an excuse to take Ida out for ice cream, and she loves to eat natto.

hardP context: We want to reward Ida for trying food she really doesn't enjoy, and she hates natto.

Ladd's example involving *even* picks up on exactly this ambiguity. Here is his purported outer-negation example; in it, *Didn't he even vote for Reagan?* receives a hardP interpretation, presupposing that voting for Reagan is a very unlikely thing.

(63) [*A and B are former left-wing activists discussing the recent activities of a colleague.*]

A: Did you hear John's decided to go to business school?

B: Yeah. I can't believe how much he's changed these days—**didn't he even vote for Reagan?**

A: That's what somebody told me.

The next example is Ladd's purported inner-negation case involving *even*. In this one, *Didn't he even vote for Reagan?* receives an easyP interpretation, presupposing that it is a very likely thing for John to have done.

(64) *[A and B are staunch Republicans.]*

A: What's John been up to these days? I haven't seen him at the Club in ages.

B: Haven't you heard? He says he's disillusioned with two-party politics—he's joined Common Cause, gave a lot of money to the Citizens' Party...

A: **Didn't he even vote for Reagan?**

B: Not as far as I know.

For Ladd, these examples is taken to contain outer negation and inner negation, respectively. In the first, the speaker's prior belief is that John voted for Reagan: the question “double-checks” the positive-polarity proposition p , the content of the speaker's prior belief, rather than the negative-polarity proposition $\neg p$. In the second, the speaker's prior belief is also that John voted for Reagan; however, confronted with evidence that John is moving away from the Republican party, the speaker's neg-biased question “double-checks” the negative proposition $\neg p$, which he infers from his interlocutor.

But these readings seem to map on perfectly to the hardP/easyP interpretations of polar questions. We can get the same distinction with canonical polar questions, showing that the ambiguity is independent of the presence of neg-bias:

(65) *[A and B are former left-wing activists discussing the recent activities of a colleague.]*

A: Wow, John has sure changed lately. He's going to business school!

B: I'll say! Did he even vote for Reagan? **(hardP reading)**

This hardP reading for the polar question is most naturally read as the speaker asking

whether he is licensed to infer that John voted for Reagan. This results in an ‘expectation’ inference with respect to the radical proposition that is similar to the corresponding outer-negation hardP example from Ladd. Similarly, the easyP interpretation of the canonical polar question gives rise to an ‘expectation against *p*’ inference like the neg-biased question:

(66) *[A and B are staunch Republicans.]*

A: Wow, John has sure changed lately. He donated to the Citizens’ Party!

B: I’ll say! Did he even vote for Reagan? (easyP reading)

This easyP reading is most naturally read as the speaker’s being unsure that John met the minimal standard for being a Republican. In sum, the pattern seems to persist even in the absence of neg-bias, which casts doubt on its status as evidence for the inner/outer-negation ambiguity.

The semantic landscape of *even* is remarkably complex crosslinguistically (Giannakidou, 2007), and the varieties of approach to *even* are too varied to get into here, as they are not directly relevant. For now, it suffices to recognize that *even* in polar questions gives rise to a systematic ambiguity that mirrors the one cited by Ladd as evidence for the inner-/outer-negation distinction. Consequently, its presence in neg-biased questions cannot be evidence for that distinction.

One final example that Ladd considers involves the minimizer NPI *lift a finger*. Minimizers, like *lift a finger*, *a red cent*, or *give a shit*, are often taken to denote minimal values on some scale (Ladusaw, 1979; Fauconnier, 1975; Giannakidou, 2011; Israel, 2011). It has been noted that, in polar questions, these only give rise to ‘rhetorical’ readings, which primarily function to express the speaker’s high degree of belief in the *no* answer (Guerzoni, 2003; Abels, 2003):

- (67) a. I was cleaning the apartment all day. Did you lift a finger to help?
b. Did he ever spend a red cent on me?

- c. Does Paul give a shit about climate change?

Although it is possible to respond to these questions with *Yes* and *No*, indicating that they do not play purely rhetorical roles, these questions are not most naturally read as sincerely seeking information from an interlocutor.

Ladd (1981), looking at minimizers in neg-biased polar questions, claims that they obligatorily receive inner-negation readings. He claims this arises from the fact that *lift a finger* ‘has no positive counterpart’ (166) that gives a neg-biased question containing *lift a finger* only an ‘inner’ reading, which ‘double-checks’ the negative-polarity proposition:

- (68) Aren’t you going to lift a finger to help?

For Ladd, this example shows that only the reading “double-checking” the negative proposition is licensed—the inner-negation reading. But there is a more parsimonious explanation: as with the positive polar questions, this question seems to function primarily to express the speaker’s belief that the addressee is very unlikely to help. Like with the examples involving *even*, this seems to be one particular instantiation of a phenomenon that occurs across polar questions: minimizers give rise to rhetorical interpretations, regardless of whether in a neg-biased or a canonical polar question.

Much of the literature on these phenomena regarding NPIs in polar questions wasn’t in place when Ladd (1981) presented his seminal findings, so it obviously is not a fault of his that he did not consider it. But a great deal of work on neg-biased polar questions has taken place since then, and, to my knowledge, no one has observed yet that a great deal of the original claims regarding NPIs in neg-biased questions are based on phenomena that are not at all unique to neg-bias, but rather shared across polar questions.

The picture that emerges, then, is actually much more attractive than the one painted by Ladd (1981), where negation interferes in a subclass of neg-biased questions in order to derive properties of neg-biased questions containing NPIs. Neg-bias does not, apparently, greatly

alter polarity properties of the question at all. Instead, as we have seen, *even* and minimizer NPIs contribute the same thing to canonical and neg-biased polar questions alike. Coupled with the evidence from Goodhue (2018) showing that neg-biased questions do not contain negation in their question radicals, it seems that neg-biased questions are genuinely much more similar to canonical polar questions than has been supposed. This is not a surprise on the analysis presented in this dissertation, which treats them as the negation of standard polar questions, essentially turning a universally quantified modal into an existentially quantified one.

Our goal, then, should not be to show that neg-biased questions are systematically structurally different when inner-negation readings obtain. Instead, we should attempt to show how the presence of NPIs of various types alters the felicity conditions and bias profile of polar questions, and try to find a unified account that does not make explicit reference to neg-bias at all. On such an account, *anything* would contribute the same thing to both of the following questions, and the issue of neg-bias is basically orthogonal.

- (69) a. Did you see anything/something?
 b. Didn't you see anything/something?

This is the approach presented in the next section, which takes the contribution of NPIs in polar questions to express domain widening in the sense of Kadmon and Landman (1993), in both neg-biased and canonical polar questions.

4.4.3 *NPIs in polar question: a domain-widening account*

What meaning does a negative polarity item contribute to a polar question, or in general? Most work on NPIs in polar questions that does not explicitly deal with neg-bias attempts to explain the *licensing* of NPIs (Guerzoni, 2003; Nicolae, 2013), asking the question: in virtue of what semantic properties of polar questions are NPIs licensed? But our question here is, instead: what does the choice to use an NPI contribute to the meaning of the polar

question?

To address this issue, I assume Kadmon and Landman's (1993) proposal that NPI *any* contributes domain-widening. On this account and accounts following it, NPI *any* allows for a larger domain of quantification across some relevant dimension than corresponding non-NPIs like *some* and *a*. For instance, the question (70-a) supposes a domain of books consisting of e.g. stereotypically reading material; but (70-b) widens this domain, allowing for books that are possibly less stereotypical.

- (70) a. John didn't borrow a book from me.
b. John didn't borrow any book from me.

Such an analysis allows us to make sense of how these assertions seem to have different 'flavors', but retain the standard assumption that *any* is an existential quantifier in the same way that *a* is.

For Kadmon and Landman (1993), and accounts following them like Krifka (1995), a primary goal is to explain NPI licensing: why are NPIs licensed in only some contexts? Their answer is that NPIs are licensed when domain-widening yields a semantically stronger expression. Under negation, for instance, widening the domain 'reduces the tolerance for exceptions', in the words of Kadmon and Landman (1993): the sentence is more likely to be true with *any* than with *a*. The opposite is true for positive-polarity sentences, whence the licensing of NPIs under negation.

This argument for NPI licensing has been criticized by a number of authors, who claim that it does not capture the semantic behavior of the full range of NPIs (Giannakidou, 2007). However, that domain-widening is a function of NPIs has remained a useful and widely supported claim. For our purposes, we do not need to assume that NPIs are licensed by domain-widening strength accounts, but only that domain-widening is one of the things that NPIs do.

Looking at polar questions, the difference in flavor introduced by *any* seems even more

pronounced than in assertions:

- (71) a. Did John borrow a book from you?
b. Did John borrow any book from you?

The question (71-b) really seems to ask for an answer that considers a domain consisting of every possible book, no matter how short, bad, etc. it may have been.

According to the account of polar questions advanced in this dissertation, a canonical polar question expresses the speaker's preference to come to believe the radical proposition if true. For instance, the following question says: *if John borrowed a book from you, I prefer to know so.*

- (72) Did John borrow a book from you?

What does the choice to widen the domain of quantification contribute to the meaning of the question? As Van Rooy (2003) observes, domain-widening makes a *Yes* answer more likely: by expanding what counts as a book, I increase the likelihood that something John borrowed counted as a book. For van Rooy, who presents an information-theoretic account similar to that presented for neg-biased questions in van Rooy and Šafářová (2003), this leads to the claim that the function of NPIs in polar questions to balance the partition structure induced by a question, making the *Yes* answer more likely when the speaker otherwise would expect a *No* answer. Van Rooy and Šafářová generalizes that speakers choose questions with high entropy, i.e. (in the polar question case) those that make the *Yes* and *No* answers equally likely.

On the conditional-preference account, things look a little different. By widening the domain of quantification, the speaker expresses that their goals are such that even learning that John borrowed the most atypical or irrelevant book would further them:

- (73) Did John borrow any book from you?

When I utter this question, I say that my preference ranking is such that my goals are furthered (i.e., more (sub)goals are achieved) if I learn that John borrowed any book from you, regardless of its status. This amounts to the ‘reduced tolerance for exceptions’ that Kadmon and Landman (1993) cite. However, this doesn’t mean that the speaker is otherwise biased toward a *No* answer to the same question without an NPI, as Van Rooy (2003) suggests. Instead, we can provide a pragmatic account.

By widening the domain of quantification, the speaker expresses their interest in more possible ways to satisfy the question radical. Consequently, the question containing an NPI is semantically stronger than the question without an NPI. And this comes without any need to define a notion of strength specific to questions; the NPI-containing polar question asymmetrically entails the question without the NPI. For instance, if I prefer to learn if John borrowed a book from set E , then I prefer to learn if he borrowed a book from a subset of E . But the inverse does not hold:

(74) Every world in which John borrowed a book from E is such that, if I learned that John borrowed a book from E , I would achieve more goals than if I didn’t.

Entails: Every world in which John borrowed a book from $D \subseteq E$ is such that, if I learned that John borrowed a book from D , I would achieve more goals than if I didn’t.

(75) Every world in which John borrowed a book from D is such that, if I learned that John borrowed a book from D , I would achieve more goals than if I didn’t.

Does NOT entail: Every world in which John borrowed a book from $E : D \subseteq E$ is such that, if I learned that John borrowed a book from E , I would achieve more goals than if I didn’t.

The relative strength of polar questions containing NPIs gives rise to the possibility of deriving a scalar implicature from a polar question *not* containing an NPI that the speaker

had the option of using: a state of affairs that, to my knowledge, has not been addressed in the literature. Because the question containing an NPI is stronger, when a speaker chooses not to use an NPI, they implicate that the stronger question is false.

In particular, Gricean reasoning according to the maxims of Quantity and Quality derives the following scalar implicature for a polar question not containing an NPI: if the speaker believed the stronger question (containing the NPI) were true, they would have said it. But they didn't say it, so they must believe it is false. Moreover, because polar questions are about goal-oriented doxastic preferences, and we have perfect epistemic access to our own preferences, from the speaker's belief that it is false follows that the stronger question is simply false. For a clear example of such a quantity implicature, consider the following scenario and possible alternative questions:

- (76) *[The speaker is a TSA agent, speaking to a traveler who is about to go through a metal detector.]*
- a. ?Are you wearing metal objects?
 - b. ?Are you wearing some metal objects?
 - c. ?Are you wearing a metal object?
 - d. Are you wearing any metal objects?

Each of the polar questions not containing an NPI is odd in this scenario. But why is this? I argue that they implicate something obviously false. In particular, by choosing not to use an NPI, the speaker implicates that there is some part of an extended domain of metal objects for which they do not care whether the addressee is wearing something from it. But a TSA agent is naturally concerned with whether you are wearing any metal at all. Consequently, the speaker's choice to use a question without an NPI is odd.

The issue of the pragmatic contribution of polarity items to polar questions is complex and surely goes beyond the scalar inference identified in the TSA agent scenario. Indeed, across the domain of polar questions, it often seems that questions with NPIs are less marked,

and otherwise identical questions with PPIs are more marked. In general, a question with an NPI feels more neutral, whereas the question with an equivalent PPI seems to come with extra pragmatic inferences. For instance, the following question with *anyone* is quite natural, but is odd with *someone* or *a person*:

- (77) a. Has anyone told you that you're beautiful?
b. ?Has someone told you that you're beautiful?

Whereas (77-a) is a neutral way to inquire whether the addressee has ever experienced being called beautiful, (77-b) seems reserved for marked situations. For example, a speaker might ask it if their goal is to determine whether to infer *Someone told you that you're beautiful*. This doesn't follow from the scalar-implicature account sketched above. It may be due to a specificity inference from the PPI *some* (Enç, 1991; Farkas, 2002). This is an interesting avenue for future research, but as it doesn't bear directly on the issue of inner-/outer-negation, we must bracket it for now.

4.5 A pragmatic account of inner and outer negation

Remember that, according to Ladd (1981), inner and outer negation are diagnosed primarily by the presence of negative polarity items. Neg-biased questions containing NPIs are taken to instantiate inner negation; these inner-negation questions “double-check” the inverse proposition. Outer-negation questions, in contrast, “double-check” the radical proposition. For Ladd, it's because of this semantic difference that inner-negation questions express a sort of contravened belief: I believed p , but something makes me now think $\neg p$. So now I'm asking whether $\neg p$ is really the case after all. Consider the following two examples:

- (78) Haven't you eaten anything? (Inner negation)
(I thought you did, but now I'm not sure. / #I think you did.)

- (79) Haven't you eaten something? (Outer negation)
(I thought you did, but now I'm not sure. / I think you did.)

With a domain-widening account of NPIs in questions, we can do away with the need to posit two distinct semantic structures for these questions (apart from the contribution of the NPI), and explain how this distinction arises pragmatically.

Remember that, in canonical polar questions, a question containing *any* was stronger than a question with *some*, leading to a scalar inference when *some* is used. Because of the introduction of high negation in neg-biased questions, though, the strength relation between the two questions is flipped: a neg-biased question containing a PPI *some* is stronger than a neg-biased question containing an NPI *any*. Consequently, it's the choice to use a *negative* polarity item that gives rise to a scalar implicature in neg-biased questions, as the *any*-question is the weaker alternative. For example, consider the following two neg-biased questions:

- (80) a. Didn't you eat anything?
b. Didn't you eat something?

Here, the *any*-containing neg-biased question is weaker, as it expresses the speaker's belief that the addressee ate something in a widened domain of quantification. The *some*-containing question is stronger.

As before, let's refer to the domain quantified over by a normal quantifier D , and the widened domain quantified over by *anything* E . The *something*-question says: for some doxastically accessible world where you ate something from the smaller domain D , I would achieve at least as many goals if I didn't learn so as if I did learn so. The *anything*-question, in turn, widens the domain to some E that is a superset of D . If there is such a doxastically accessible world where you ate something from D , the smaller domain, then there is necessarily a doxastically accessible world where you ate something from E , a superset of

D. (Namely, the same world.) The same does not follow in reverse, of course: if there's a world where you ate something from the wider domain, it's not necessarily true that there's a world where you ate something from the narrower domain. The question with *something*, then, entails the question with *anything*, but not vice versa.

What does the speaker implicate when they use the weaker neg-biased question containing *anything*? By widening the domain of quantification, they pragmatically communicate that they are unsure that the stronger expression, with the narrower domain of quantification, is true. Moreover, as in the case of canonical polar questions, both expressions are about the speaker's doxastic state, which the speaker necessarily has perfect knowledge of. Consequently, the question straightforwardly implicates that the stronger expression is false. Thus, we have the following semantic and pragmatic content for a neg-biased question containing an NPI:

(81) Didn't you eat anything?

entails: I believe you ate something from a widened domain of quantification *E*.

implicates: I don't believe that you ate something from the standard domain of quantification *D*.

For the stronger question, with *something*, to be false, it must be that the speaker does not believe that you have eaten something in the standard domain. However, as the speaker uttered the neg-biased question, they *do* believe that you have eaten something in the domain of *anything*, the widened domain. This scalar implicature turns out to explain a number of properties of neg-biased questions containing NPIs that have been attributed to a scope distinction à la Ladd (1981).

First, the 'double-checking' intuition of Ladd (1981) receives a natural explanation. According Ladd, in a neg-biased question containing the NPI, the speaker "double-checks" the negated proposition; but in neg-biased question without an NPI, the speaker "double-checks" the positive proposition. This is odd in part because double-checking isn't a particularly clear

notion: questions don't usually double-check, after all. Presumably, for Ladd, (part of) what neg-bias consists in *is* the double-checking operation, and the inner/outer distinction determines what the polarity of the double-checked proposition is. A scopal approach along these lines (in which negation has two scope positions, one inside the radical and one outside) will necessitate some other silent operator that corresponds to the “double-checking” operation, e.g. the VERUM operator of Romero and Han (2004).

With the pragmatic account, we can do away with a notion of double-checking propositions of different polarity. In the NPI-containing neg-biased question, what is happening isn't double-checking, but implicature. Neg-biased questions with *any* systematically implicate that the speaker believes the proposition is false on the standard domain. But the speaker also expresses with the *any*-question that they believe the proposition is true on a widened domain. It is this conflict—that the speaker thinks p is true for the wide domain, but false for the narrow one—that has been analyzed as ‘double-checking’ $\neg p$ against the prior belief p . The speaker's prior belief p in fact was always true with respect to the wider domain, and what Ladd identifies as new evidence suggesting $\neg p$ (for the narrow domain) is in fact an implicature. This has been interpreted as a condition of use of the question, rather than a pragmatic inference.

Second, the pragmatic account also gives insight into the character of the prior belief in inner-negation neg-biased questions, which have a distinctive ‘guess’ or ‘expectation’ flavor. For instance, consider the following scenario:

- (82) *[The addressee is flipping listlessly through TV channels without choosing anything.]*
Don't you like anything?

In the traditional terms from Ladd (1981), the speaker has a prior belief that the addressee likes something; but confronted with evidence against this, he chooses to ask the *anything*-question, ‘double-checking’ the negative proposition that he has inferred from the context. However, this traditional account doesn't explain the strong ‘expectation’ flavor of

the speaker's prior belief. In this context, the speaker's belief is almost certainly something like a generalization or expectation—rather than a well-founded belief that, for some particular thing, the addressee likes it. For instance, the following are all plausible reasons for the speaker to believe that the addressee likes something if he utters (82):

- (83) Reasons for speaker's prior belief that the addressee likes something, given an utterance of (82)
- a. The speaker knows that the addressee almost always finds something he likes
 - b. The speaker believes people, in general, like some things that play on TV
 - c. The speaker believes that, after enough browsing, everyone's standards decline enough to find that they like something

Though they are all different, these are all beliefs that might lead one to suppose that, for any individual person flipping through channels, they will eventually like something. However, they do not support the speaker's belief that the addressee *in particular* likes something. This is the expectation inference: the speaker expects the radical proposition to be true, in virtue of something like a generic belief. In many cases, this inference has a strongly modal flavor: one feels that the neg-biased question with *anything* suggests the speaker's belief that the addressee 'should', 'ought to', or 'must' satisfy the radical proposition.

On the other hand, this question is extremely odd if the basis for the speaker's belief is something concrete. For instance, if the speaker believes that the addressee likes something because the addressee said so, but the speaker simply forgot what that thing is, he would not ask *Don't you like anything?*, but rather:

- (84) Don't you like something?

This expectation inference arises robustly across NPI-containing neg-biased questions. Consider the following examples, where the two continuations following the neg-biased question

offer possible justifications for the speaker's prior belief. Systematically, the neg-biased question containing an NPI is infelicitous if the speaker has concrete evidence for their belief, rather than something like an expectation.

- (85) a. [*The speaker is informed that there was a cockroach in her office.*]
Didn't anyone kill it?
(I think someone should have. / # Someone told me John did.)
- b. Haven't you seen *Chungking Express* yet?
(I think you should have. / # Someone told me you had.)
- c. Won't anybody lend me a dollar?
(I think someone should. / # Someone told me they will.)
- d. Haven't you ever been to the Art Institute?
(I think you should have. / # Someone told me you had.)

In contrast, given a neg-biased questions *without* an NPI, the speaker's belief can optionally be licensed by this sort of expectation, but it can also be concretely justified:

- (86) a. [*The speaker is informed that there was a cockroach in her office.*]
Didn't someone kill it?
(I think someone should have. / Someone told me John did.)
- b. Haven't you seen *Chungking Express* already?
(I think you should have. / Someone told me you had.)
- c. Won't somebody lend me a dollar?
(I think someone should. / Someone told me they will.)
- d. Haven't you once been to the Art Institute?
(I think you should have. / Someone told me you had.)

Why should this expectation flavor should arise so strongly for neg-biased questions with

NPIs, but not neg-biased questions *without* NPIs? This is a mystery on the traditional account of inner-negation, which says that inner negation is used when the speaker's prior belief is contravened by the context. But why is it only one kind of prior belief: namely, expectation?

On the pragmatic account, the source of the expectation inference gets a natural story. Speakers choose to utter neg-biased questions with domain-widening NPIs when they believe that the radical proposition is false with respect to a narrow domain, but have some reason to believe it is true with respect to a wider domain. So by choosing to use the neg-biased question with a domain-widening NPI, I pragmatically express my belief that domain-widening turns the proposition from false to true. But under what circumstances does simply widening a domain of quantification make a proposition go from false to true? This makes the most sense exactly when I have some prior belief or expectation that the proposition *should* be true for some entity in a sufficiently wide domain, but I lack evidence to have the belief for any particular entity.

In effect, on this view domain-widening in neg-biased questions comes to be a kind of 'rescue' operation: I want to salvage my belief in *You like something*, so I widen the domain by using *anything*. This communicates something like: *OK, so my belief is false for the narrow domain. But if I widen the domain, then it's true.* This is why having specific evidence makes for a bad prior belief in 'inner negation' questions: if I had specific evidence, I wouldn't be widening the domain in order to retain my belief; I would make reference to the belief implied by that specific evidence instead.

Now, in contrast, the neg-biased question without an NPI expresses the speaker's belief that the radical proposition holds for some entity in the unwidened domain. This belief might be justified by an expectation, or it might be justified by some particular evidence; there's no reason it should be one over the other. Consequently, the non-NPI-containing ('outer-negation') neg-biased question is fine with either type of belief.

Finally, recall the bias profiles identified by Sudo (2013) for various types of polar ques-

tions (*Isn't*) p ?. In particular, recall that inner-negation questions were taken to require evidence against p , whereas outer-negation questions simply require there be no evidence *for* p . Sudo's claims are reproduced in the following chart:

	Evidential bias	Epistemic bias
(87) Outer negation (no NPI)	–positive	positive
Inner negation (NPI)	+negative	positive
Canonical polar question	–negative	none

This pattern, wherein inner-negation (i.e., NPI-containing) neg-biased questions impose a stronger restriction, can now be explained by the same pragmatic strengthening of the inner-negation question. As we've seen, the NPI-containing neg-biased question implicates that the speaker does not believe the narrow-domain proposition and believes that widening the domain rescues that belief. Consequently, widening the domain is unnecessary unless there is some reason for the speaker to doubt that the radical proposition p is true for the narrow domain.

That there be a reason for the speaker's use of the weaker alternative is, I claim, the same thing as Sudo's (2013) requirement that there be evidence against the radical proposition p . It is in light of some particular reason that the speaker believes the radical is false for the narrower domain, and therefore uses an expression referring to a wider domain. That 'evidence', too, does not actually need to be mutually available in the context. For instance, looking around the house for thumbtacks, I might go into the room where my wife is working and ask:

(88) Aren't there any thumbtacks in this house?

For Sudo, *evidence against* p has to be construed in such a way that it exists in this context. But in this discourse context, there *isn't* jointly available evidence to the effect that there are no thumbtacks in the house. That evidence was available to exactly one participant, and before the time of the discourse. The pragmatic story doesn't impose such a requirement, merely implicating that the speaker has come to believe the radical proposition is false on

an insufficiently widened domain.

Finally, this account provides an alternative way of looking at an argument expressed by Reese (2007); Asher and Reese (2007) regarding speech acts. On their account, a neg-biased question have a complex speech act type: assertion • question. To support this claim, Reese (2007) notes that neg-biased polar questions pass many of the tests for illocutionary force proposed by Sadock (1971, 1974). These consist of discourse markers that target assertions but are infelicitous with questions. For instance, *after all* is infelicitous with canonical polar questions, but licensed with assertions and neg-biased questions:

- (89) You're going to do great at the swim meet.
- a. ... After all, you broke the state record in the 200-m butterfly.
 - b. #... After all, did you break the state record in the 200-m butterfly?
 - c. ... After all, didn't you break the state record in the 200-m butterfly?

Another is *yet*, which can elaborate on assertions but not canonical polar questions:

- (90) a. A: Susan travels by plane once a week.
B: Yet she never springs for business class.
- b. A: Does Susan travel by plane once a week?
B: #Yet she never springs for business class.
- c. A: Doesn't Susan travel by plane once a month?
B: Yet she never springs for business class.

Crucially, Reese (2007) notes that inner- and outer-negation neg-biased questions behave differently with respect to these tests: only 'outer-negation' readings are felicitous. That is, the presence of an NPI causes these neg-biased questions to fail the tests again:

- (91) You're going to do great at the swim meet.

- a. ... After all, didn't you beat some of your opponents last time?
 - b. #... After all, didn't you beat any of your opponents last time?
- (92)
- a. A: Doesn't Susan have some upcoming flights?
B: Yet she won't spring for business class.
 - b. A: Doesn't Susan have any upcoming flights?
B: #Yet she won't spring for business class.

That some neg-biased questions apparently behave like questions, but others like assertions, is what leads to the claim that neg-biased questions instantiate a kind of hybrid speech act. The present analysis of polar questions, though, gives an interesting alternative way to consider this data, which is very briefly sketched here.

A polar question $p?$ is odd on the attitudinal analysis in part because the radical is the restrictor to universal quantification over belief worlds. Essentially, a polar question $p?$ has a (simplified) semantic shape like the following:

(93) **Shape of a canonical polar question:**

$$\forall w [p(w)] [q(w)]$$

This is, presumably, part of what gives polar questions their unique properties: anything explicitly added to the polar question meaning becomes part of the restrictor, and there is no real way to get into the nuclear scope of the question meaning at all. Therefore the polar-question environment is non-veridical (Giannakidou, 1998, et seq.) and downward-entailing, which provides a natural story for why NPIs are licensed there. But since it is in the restrictor of a universal quantifier, from the radical proposition p being used in a question, we cannot conclude anything about whether p is true, or even possibly true. This is why polar questions, at first glance, don't seem to contribute anything at all.

Introducing negation outside of the universal quantifier changes the picture, as it creates the entailment that there is a p -world. By negating (93), we get the shape of a neg-biased

question:

(94) **Shape of a neg-biased polar question:**

$$\neg\forall w [p(w)] [q(w)]$$

This entails that there is an accessible world where p is the case: it is equivalent to the proposition that there is a possible world where p and not q .

(95) **Shape of a neg-biased polar question:**

$$\exists w [p(w) \wedge \neg q(w)]$$

This meaning is, intuitively, more ‘assertion-like’: it says something about the status of p . Perhaps the reason that neg-biased questions without NPIs license assertion-targeting particles like *yet* and *after all* has to do with this equivalence. Without getting into theories of NPI licensing, it is possible that NPIs somehow require—at some level of representation—the negated-universal logical form in (94) in order to license the NPI; but the assertion-targeting particles, in contrast, require a reanalysis of the negated polar question as having the logical form in (95), towards a more assertion-like character. In other words, the two are incompatible because of divergent licensing conditions.

This sketch of an argument, obviously, does not explain the facts observed by Reese (2007); Asher and Reese (2007); however, it may present a way forward that avoids the need to posit things like complex speech acts to account for neg bias. It is also worth noting that here, again, it is the NPI-containing (‘inner negation’) neg-biased questions whose distribution is restricted. This is not surprising on a pragmatic account, according to which NPI-containing neg-biased questions are weaker than their non-NPI containing alternatives, and are therefore marked. Accordingly, another possible explanation is simply that *yet* and *after all* need meanings that are sufficiently strong, and a neg-biased polar question containing an NPI is too weak to license them. Since an NPI-containing neg-biased

question ultimately expresses a speaker's belief that something in an extended domain makes the radical proposition true, but not in the standard domain of quantification, perhaps it is simply too weak to express the level of commitment or semantic strength necessary to license assertion-targeting particles.

In other words, an approach to neg-bias that allows it to negate the meaning of a question *in toto* might shed light on why neg-biased questions seem to be less question-like than other polar questions: that the radical's status as the restrictor of a universal quantifier over doxastic worlds—essentially, an if-clause—plays a role. In a canonical polar question, we cannot infer anything about the status of the radical proposition; but when the entire expression is negated, we can: namely, that the speaker believes the radical to be true.

CHAPTER 5

CONCLUSION

This dissertation has presented an attitudinal account of polar questions, an account according to which their function is to express a mental state of the speaker: namely, their goal-oriented preference to learn the radical proposition, if true. Although this account requires jettisoning the partition semantics typically assumed in the literature on questions, it comes with a number of benefits that make it attractive.

After the introduction in Chapter 1, Chapter 2 identifies some outstanding issues in the semantics of canonical polar questions. Because existing accounts attempt to retain the core properties of analyses that exist primarily for wh-questions, most predict widespread synonymy across polar questions. In order to account for the differential felicity conditions and response particle patterns of otherwise ‘synonymous’ questions, a number of approaches have emerged. These attempt to address what I refer to as the problem of polar question ‘symmetry’: an account is symmetrical in case the meaning of a question $p?$ and a question $\neg p?$ are synonymous. (Crucially, not all symmetrical accounts are partition-based accounts.) Accounts that attempt to build an asymmetry into polar questions include highlighting and structured-meaning approaches. However, these still struggle to explain how we get from a polar question’s semantic value to its felicity in a particular context. Existing descriptions of the felicity conditions of polar questions, too, fail to explain how the generalizations the authors identify relate to the semantics of the question itself.

In Chapter 3, I present a positive proposal that attempts to address this issue. I propose that polar questions denote a proposition that makes reference to the speaker’s mental state: a question $p?$ says that, for every doxastically accessible world where p , the speaker would achieve more of his goals and subgoals if he believed p than if he didn’t. In essence, this is an implementation of the intuition that a question $p?$ expresses a speaker’s *desire to know if p*. Building this structure into the semantics of the question allows us to refer to the set of the speaker’s discursive and action-motivating goals as an ordering source for a

modal preference operator, which explains why polar question felicity is so dependent on properties of the context—and especially the speaker’s intentions in that particular context. A question’s sense of truth-valuelessness comes from the expressive nature of its meaning: rather than presenting the proposition as something to be added to the Common Ground, the question ‘automatically’ adds its content to the Common Ground. This happens because, as the question expresses a speaker preference, it functions as an *expression* of that preference: an action that speakers are entitled to do in virtue of having the relevant preference.

In Chapter 4, I show how this approach can be extended to neg-biased polar questions. In particular, by changing the speaker’s preference from *believing* to *coming to believe*, we can analyze neg-biased questions as straightforwardly negating the denotation of the polar question. This allows us to derive the doxastic bias of the speaker as a semantic inference of the polar question. That semantic account, in turn, feeds a pragmatic account of what has traditionally been taken to be an ambiguity in neg-biased questions: the ambiguity between an inner-negation and an outer-negation interpretation. I argue that what has been taken to be a characteristic inner-negation interpretation is really a consequence of the choice to use a negative polarity item in the neg-biased question. Assuming that negative polarity items like *any* contribute a widened domain of quantification, I argue that the speaker’s choice to ask a question using a widened domain gives rise to scalar implicature: namely, the implicature that the speaker believes the radical of the question with the narrower domain to be false. This provides a useful pragmatic account of the inner/outer negation distinction, according to which the marked nature of neg-biased questions containing *any* is because of their semantically weak nature, rather than a structural difference between inner- and outer-negation questions per se. Ultimately, this dissertation provides an alternative account of polar questions which allows for a unified approach to canonical and neg-biased questions, deriving their properties in a straightforward way without the need to stipulate extra semantic operators in neg-biased questions—an account which feeds a pragmatic account of what has traditionally been taken to be a semantic phenomenon.

REFERENCES

- Abels, Klaus. 2003. Who gives a damn about minimizers in questions? In *Semantics and Linguistic Theory*, volume 13, 1–18.
- Aboh, Enoch Oladé. 2007. Leftward focus versus rightward focus: The kwa-bantu conspiracy. In *SOAS Working Papers in Linguistics*, volume 15, 81–104.
- AnderBois, Scott. 2019. Negation, alternatives, and negative polar questions in american english. In *Questions in discourse*, 118–171. Brill.
- Asher, Nicholas, and Brian Reese. 2007. Intonation and discourse: Biased questions. *Interdisciplinary Studies on Information Structure* 8:1–38.
- Biezma, Maria. 2009. Alternative vs polar questions: the cornering effect. *Semantics and Linguistic Theory* 19:37. URL <https://doi.org/10.3765/salt.v19i0.2519>.
- Bolinger, Dwight. 1978. Yes—no questions are not alternative questions. In *Questions*, 87–105. Springer.
- Büring, Daniel. 2006. Focus projection and default prominence. In *The architecture of focus*, 321–346. De Gruyter Mouton. URL <https://doi.org/10.1515/9783110922011.321>.
- Büring, Daniel, and Christine Gunlogson. 2000. Aren't positive and negative polar questions the same? ucsc/ucla. Ms., 01 2000.
- Castroviejo Miró, Elena. 2009. Deconstructing exclamations. *Catalan Journal of Linguistics* 7:41–90. URL <https://raco.cat/index.php/CatalanJournal/article/view/128410>.
- Chernilovskaya, Anna, Cleo Condoravdi, and Sven Lauer. 2012a. On the discourse effects of wh-exclamatives. In *Proceedings of the 30th West Coast Conference on Formal Linguistics*, 109–119. Citeseer.
- Chernilovskaya, Anna, Cleo Condoravdi, and Sven Lauer. 2012b. On the discourse effects of wh-exclamatives (handout).
- Cohen, Ariel, and Manfred Krifka. 2014. Superlative quantifiers and meta-speech acts. *Linguistics and Philosophy* 37:41–90.
- Condoravdi, Cleo, and Sven Lauer. 2011. Performative verbs and performative acts. In *Proceedings of Sinn und Bedeutung*, volume 15, 149–164.
- Condoravdi, Cleo, and Sven Lauer. 2012. Imperatives: Meaning and illocutionary force. In *Empirical Issues in Syntax and Semantics*, ed. Christopher Piñón, volume 9, 37–58.
- Csipak, Eva, and Sarah Zobel. 2014. A condition on the distribution of discourse particles across types of questions. In *Proceedings of NELS*, volume 44, 83–94.
- Dayal, Veneeta. 2010. *Locality in WH quantification*. Studies in Linguistics and Philosophy. Dordrecht, Netherlands: Springer.

- De Swart, Henriette. 1996. Meaning and use of *not... until*. *Journal of Semantics* 13:221–263.
- Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 1–25.
- Farkas, D. F., and K. B. Bruce. 2009. On reacting to assertions and polar questions. *Journal of Semantics* 27:81–118. URL <https://doi.org/10.1093/jos/ffp010>.
- Farkas, Donka. 2002. Specificity distinctions. *Journal of Semantics* 19:213–243.
- Farkas, Donka F., and Floris Roelofsen. 2017. Division of labor in the interpretation of declaratives and interrogatives. *Journal of Semantics* .
- Fauconnier, Gilles. 1975. Pragmatic scales and logical structure. *Linguistic Inquiry* 6:353–375.
- von Fintel, Kai, and Sabine Iatridou. 2017. *A modest proposal for the meaning of imperatives*. Oxford University Press.
- Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*, volume 23. John Benjamins Publishing.
- Giannakidou, Anastasia. 2007. The landscape of even. *Natural Language & Linguistic Theory* 25:39–81.
- Giannakidou, Anastasia. 2011. Positive polarity items and negative polarity items: variation, licensing, and compositionality. In *Semantics: An international handbook of natural language meaning*, ed. Claudia Maienborn, Klaus von Stechow, and Paul Portner, 1660–1712. Mouton de Gruyter.
- Giannakidou, Anastasia, and Alda Mari. 2019. Modalization and bias in questions. University of Chicago and Institut Jean Nicod, Ms.
- Goodhue, Daniel. 2018. On asking and answering biased polar questions. Doctoral Dissertation, McGill.
- Goodhue, Daniel. 2022. Aren't there multiple causes of bias in polar questions? URL <https://lingbuzz.net/lingbuzz/004741>, lingbuzz/004741, may 2022.
- Grice, H. P. 1975. Logic and conversation. In *Syntax and semantics: Vol. 3: Speech acts*, ed. Peter Cole and Jerry L. Morgan, 41–58. New York: Academic Press. URL <http://www.ucl.ac.uk/lis/studypacks/Grice-Logic.pdf>.
- Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. In *Varieties of formal semantics: Proceedings of the fourth amsterdam colloquium*, ed. Fred Landman and Frank Veltman, 143–170. Foris.
- Guerzoni, Elena. 2003. Why even ask? Doctoral Dissertation, MIT.
- Gunlogson, Christine. 2002. Declarative questions. *Semantics and Linguistic Theory* 12:124. URL <https://doi.org/10.3765/salt.v12i0.2860>.

- Gyuris, Beáta. 2017. New perspectives on bias in polar questions. *International Review of Pragmatics* 9:1–50.
- Hamblin, Charles L. 1976. Questions in Montague English. In *Montague grammar*, 247–259. Elsevier.
- Hamblin, C.L. 1958. Questions. *Australasian Journal of Philosophy* 36:159–168. URL <https://doi.org/10.1080/00048405885200211>.
- Harrah, David. 2002. The logic of questions. In *Handbook of philosophical logic*, 1–60. Springer.
- Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.
- Höhle, Tilman N. 2018. Über verum-fokus im Deutschen. *Beiträge zur deutschen Grammatik* 381.
- Holmberg, Anders. 2015. *The syntax of yes and no*. Oxford University Press.
- Israel, Michael. 2011. *The Grammar of Polarity: Pragmatics, Sensitivity, and the Logic of Scales*. Cambridge, England: Cambridge University Press.
- Kadmon, Nirit, and Fred Landman. 1993. Any. *Linguistics and Philosophy* 16:353–422.
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44. URL <https://doi.org/10.1007/bf00351935>.
- Kaufmann, Magdalena. 2012. *Interpreting imperatives*. Springer Netherlands. URL <https://doi.org/10.1007/978-94-007-2269-9>.
- Kaufmann, Magdalena. 2020. On the performance of modal objects. *Theoretical Linguistics* 46:253–266. URL <https://doi.org/10.1515/tl-2020-0014>.
- Kaufmann, Magdalena, and Stefan Kaufmann. 2021. Iffy endorsements. *Journal of Semantics* 38:639–665. URL <https://doi.org/10.1093/jos/ffab017>.
- Kramer, Grace, and Kyle Rawlins. 2009. Polarity particles: an ellipsis account. In *Proceedings of NELS 39*.
- Kratzer, Angelika. 1981. The notional category of modality. In *Words, worlds, and contexts*. DE GRUYTER. URL <https://doi.org/10.1515/9783110842524-004>.
- Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 209–257.
- Krifka, Manfred. 2001. For a structured meaning account of questions and answers. In *Audiatur vox sapientiae*, 287–319. Akademie Verlag. URL <https://doi.org/10.1515/9783050080116.287>.

- Krifka, Manfred. 2015. Bias in commitment space semantics: Declarative questions, negated questions, and question tags. *Semantics and Linguistic Theory* 25:328.
- Krifka, Manfred. 2017. Negated polarity questions as denegations of assertions. In *Contrastiveness in information structure, alternatives and scalar implicatures*, 359–398. Springer International Publishing. URL https://doi.org/10.1007/978-3-319-10106-4_18.
- Ladd, D Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In *Papers from the... Regional Meeting. Chicago Ling. Soc. Chicago, Ill*, 17, 164–171.
- Ladusaw, William Allen. 1979. *Polarity Sensitivity as Inherent Scope Relations*. The University of Texas at Austin.
- Levinson, Stephen. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, Mass: MIT Press.
- Lewis, David. 2001. *Counterfactuals*. Malden, Mass: Blackwell Publishers.
- Maienborn, Claudia. 2019. 8. event semantics. In *Semantics - theories*, 232–266. De Gruyter. URL <https://doi.org/10.1515/9783110589245-008>.
- Nicolae, Andreea Cristina. 2013. Any questions? polarity as a window into the structure of questions. Doctoral Dissertation, Harvard University.
- Northrup, Oliver. 2014. Grounds for commitment. Doctoral Dissertation, UCSC.
- of Philosophy & Linguistics Terence Parsons, Professor, and Terence Parsons. 1990. *Events in the semantics of english*. MIT Press.
- Portner, Paul. 2004. The semantics of imperatives within a theory of clause types. *Semantics and Linguistic Theory* 14:235.
- Reese, Brian Jon. 2007. Bias in questions. Doctoral Dissertation, The University of Texas at Austin.
- Repp, Sophie. 2007. $\neg(A\&B)$. Gapping, negation and speech act operators. *Research on Language and Computation* 4:397–423.
- Roelofsen, Floris, and Donka F Farkas. 2015. Polarity particle responses as a window onto the interpretation of questions and assertions. *Language* 91:359–414.
- Roelofsen, Floris, and Sam van Gool. 2010. Disjunctive questions, intonation, and highlighting. In *Logic, language and meaning*, 384–394. Springer.
- Romero, Maribel, and Chung-hye Han. 2004. On negative yes/no questions. *Linguistics and Philosophy* 27:609–658.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116. URL <http://www.jstor.org/stable/23748778>.

- Rooth, Mats Edward. 1985. Association with focus (montague grammar, semantics, only, even). Doctoral Dissertation, University of Massachusetts Amherst.
- van Rooy, Robert. 2004. Utility of mention-some questions. *Research On Language And Computation* 2:401–416.
- van Rooy, Robert, and Marie Šafářová. 2003. On polar questions. In *Semantics and Linguistic Theory*, volume 13, 292–309.
- Sadock, Jerrold. 1971. Queclaratives. In *Papers from the Seventh Regional Meeting of the Chicago Linguistic Society*, 223–232.
- Sadock, Jerrold. 1974. *Towards a linguistic theory of speech acts*. Academic Press.
- Sailor, Craig. 2013. Questionable negation. Talk at the 87th Annual Meeting of the Linguistic Society of America.
- Selkirk, Elisabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. *The Handbook of Phonological Theory* 1:550–569.
- Stalnaker, Robert. 1987. *Inquiry*. Cambridge, Mass: MIT Press.
- Starr, William. 2009. Conditionals and questions. *Manuscript, Rutgers University* .
- Sudo, Yasutada. 2013. Biased polar questions in english and japanese. *Current Research in the Semantics/Pragmatics Interface* 28:275–295.
- Theiler, Nadine. 2020. *Denn* as a highlighting-sensitive particle. *Linguistics and Philosophy* 44:323–362. URL <https://doi.org/10.1007/s10988-019-09290-7>.
- Van Benthem, Johan. 2011. *Logical dynamics of information and interaction*. Cambridge University Press.
- Van Rooy, Robert. 2003. Negative polarity items in questions: Strength as relevance. *Journal of Semantics* 20:239–273.
- Van Rooy, Robert, and Marie Šafářová. 2003. On polar questions. In *Semantics and Linguistic Theory*, volume 13, 292–309.
- Wilkinson, Karina. 1996. The scope of even. *Natural Language Semantics* 4:193–215.
- Zimmermann, Malte, and Edgar Onea. 2011. Focus marking and focus interpretation. *Lingua* 121:1651–1670. URL <https://doi.org/10.1016/j.lingua.2011.06.002>.
- É. Kiss, Katalin. 2002. *The syntax of hungarian*. Cambridge University Press.