UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ANAPHORA AND NEGATION

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Lisa Hofmann

March 2022

The Dissertation of Lisa Hofmann is approved:

Professor Adrian Brasoveanu, Chair

Assistant Professor Jess H.K. Law

Assistant Professor Amanda Rysling

Peter Biehl
Vice Provost and Dean of Graduate Studies

Contents

List of Figures

Abstract

Anaphora and Negation

Lisa Hofmann

This dissertation provides a commitment-based account of discourse-polarity and anaphoric polarity-sensitivity by appealing to veridicality distinctions.

It investigates the interaction of anaphora and negation, addressing the overarching question of how human language users combine information from multiple sources into an integrated representation of language meaning. Because the interpretation of anaphora systematically depends on the discourse context, studying its interaction with the local utterance context addresses this overarching question.

The work provides arguments and evidence for theoretically understanding the constraints that negation places on anaphora in terms of speaker commitments and veridicality. Anaphoric accessibility is sensitive to negation because it is sensitive to the veridicality of the embedding context(s) of an anaphor and its antecedent.

On these terms, a formal semantic account of the interaction of anaphora and negation is developed. It is an intensional dynamic account based on Compositional DRT and flat-update systems, where negation and other non-veridical operators are treated as externally dynamic. I argue that we need to conceptualize discourse referents as an intensional and epistemic representation to capture the relevant interactions.

This provides a unified account of several related issues that have received disparate analyses in the previous literature: anaphora to indefinites under negation (double-negation, bathroom-disjunctions, discourse subordination, and cross-speaker cases), the interaction of propositional anaphora and negation, polarity-sensitive negativity-tags, and the question what counts as a negative sentence/utterance for the purposes of anaphora in discourse.

Acknowledgments

I would never have been able to write this dissertation without support from mentors, teachers, colleagues, friends, and family. An acknowledgment of the people who supported my journey towards completing my dissertation is one of the hardest parts that I have had to write. That's because I have no filter, so talking about my personal feelings in a published document is kind of scary to me. But I suppose a long dissertation can also have a long list of acknowledgments. Therefore, I believe that the first thanks should probably go out, on your behalf, as a reader, to my dissertation committee, who encouraged me to make my dissertation more skimmable (by including more headings, highlights, cross-references, etc.).

I am incredibly grateful to the professors who have provided me with mentorship and support throughout the years.

Especially, I want to thank **Adrian Brasoveanu**, who was a fantastic advisor for me. He struck the balance of being a respectful colleague and an excellent mentor at the same time. He was motivating without exercising pressure and allowed me to take responsibility for my own work while being there for me whenever I needed it. I could not begin to do justice to all the lessons he imparted to me by attempting to summarize the most important ones. So will just say that there is no linguist from who I learned more. The attention to formal detail and simultaneous understanding of the big picture that is apparent in his research and teaching will always be something I aspire to.

I don't know what my life would be like if I hadn't met **Daniel Altshuler**, my undergraduate mentor at HHU Düsseldorf, who encouraged and supported my getting a Ph.D. Daniel is the most engaging teacher I have ever met, who has, again and again, inspired me to see the importance of our work, in- and outside of the classroom. He taught me many things, about how to be a better writer, about the significance of the kinds of questions we ask, and about navigating some of the social aspects of academia. His way of making interpersonal connections, which gives rise to inter-disciplinary work, and his way of providing access to his students is a great role model for me.

This work has also greatly benefited from interactions with my dissertation committee:

The conversations I have had with **Jess Law** about the analysis in my dissertation were one of the most fun things I was able to do during my graduate education. Discussing my ideas with a fellow formal semanticist who would take them seriously was exciting and empowering. Jess constantly pushed me to present my ideas more effectively. My dissertation would be a lot harder to read if it weren't for her. There certainly would be fewer diagrams.

My other committee member, **Amanda Rysling**, is a fantastic teacher, and her seminar made me a better linguist. As someone who is not a phoneticist, the main thing that struck me about her work is her way of approaching a problem. She taught me how important it is to make even basic assumptions explicit and to continuously evaluate whether you can justify them. And that often, the questions underlying basic assumptions are the more important ones to ask, especially in an interdisciplinary setting.

My dissertation also would not be what it is now without the mentorship and support I received from other UC Santa Cruz faculty:

I am immensely thankful to **Jim McCloskey** for teaching us syntax in the morning calm (/mudslide) and for introducing me to the topic of anaphora and polarity. He

chaired my first qualifying paper on polarity ellipsis in *'Why not'* questions. His way of constructing an argument and his kind nature have made a lasting impression on me.

Thank you to **Donka Farkas** for being a committee member for my qualifying exam, for providing valuable feedback on my work, and for some extensive discussions. Donka taught me many things about being a professional linguist and about how to engage with existing literature on a phenomenon.

Thank you to **Jorge Hankamer** for being supportive of me as a fellow researcher of anaphora and being on the committee for my second qualifying paper. As a discussion partner, collaborator, or reading group participant, you can always count on Jorge to bring up an interesting piece of data that will lead to discussion, rethinking your hypothesis, or a new experimental design.

Ivy Sichel was never a member of one of my committees, but I am grateful for having had a few different opportunities to interact with her and for her presence in the department. Working with Ivy on Negative Fragments was a tremendous experience for me, where I learned a lot about how to build an argument. I am also grateful that I could be a TA for her Sociolinguistics class. The nuanced way in which she would address social issues in the classroom is something that I hope to be able to emulate one day.

Pranav Anand was a member of the committee for my first qualifying paper. I am grateful for the opportunity to learn from him throughout the years, and I appreciate his perspective. The questions he raises always feel deeply interesting and important.

I also owe a debt of gratitude to **Rajesh Bhatt**, who gave me the chance to visit the linguistics department of UMass Amherst for a year before starting graduate school. His institutional support was instrumental in my becoming a linguist and allowed me to feel like I could be a part of an international research community. I thank him, as well as **Angelika Kratzer**, **Vincent Homer**, and **Kyle Johnson**, for welcoming me into their

department and into the world of academia. There exists a possible world in which I went to graduate school at UMass, and at some point, it was extremely similar to the actual world. Choosing between various excellent options for graduate school was one of the most difficult decisions I ever had to make. No choice would have prevented me from regretting giving up on the others. Looking back, I am beyond grateful for the opportunity to interact with as many different ideas and excellent linguists as I did.

Thank you to **Albert Ortmann**, my first academic mentor at HHU Düsseldorf, who hired me as a research assistant. He taught me a lot about definiteness, morphology, typological variation, and how to do academic research. Also, thanks to **Peter Sutton** for providing very insightful feedback on my BA thesis and to **Sebastian Löbner** for his support.

Many of the people mentioned above also enriched my academic life by teaching excellent classes. Thanks are due to Adrian Brasoveanu, as well as Rajesh Bhatt, for providing me with tools for formal semantics; Pranav Anand for influencing the way I ask questions; and Vincent Homer for introducing me to some key phenomena in natural language semantics. One of the semantics classes I thought back to the most was **Seth Cable**'s introductory class on intensional semantics. Had it not been for his captivating, thought-provoking, and excellently structured course, I would not have been able to write the kind of dissertation I have now, years later.

I also want to say thank you to my undergraduate semantics professors: **Elizabeth Coppock**, for a great Intro to Pragmatics; **Kata Balogh** for introducing me to dynamic semantics; **Hana Filip**, for an exciting class on how certain aspects of nominal and verbal lexical semantics interact compositionally; and (again) Daniel Altshuler, for two fantastic semantics classes.

Besides Amanda Rysling and Adrian Brasoveanu, I have learned a lot about psycholinguistics and experimental methodology from **Matt Wagers**. I am thankful that Matt's classes introduced me to the subject area and enabled me to ask questions I would not even have conceived of before.

I absolutely loved learning about Syntax from Jim McCloskey, who taught me about the expression of polarity and the beauty of minimalism; Kyle Johnson, who leads students deep into the forest of a million thought-provoking example sentences; and Rajesh Bhatt in a pleasantly challenging class on agreement. Thank you also to **Maziar Toosarvandani** for teaching the formal machinery with unparalleled clarity, excellent in-class discussions of assigned readings, which I will strive to emulate when teaching, and an incredible field methods seminar.

Finally, Phonologists like **Junko Ito** and **Gaja Jarosz** broadened my horizon by introducing me to OT, a conception of generative grammar that was wildly different from the formalisms I was familiar with up to that point.

Many thanks to **James Funk, Ashley Hardisty, Logan Roberts, Gwyn Vandevere, and Maria Zimmer** for keeping the UCSC linguistics department running and for your help with administrative matters throughout the years.

Thank you to the UCSC Humanities Division for a Global Outreach Fellowship, which allowed me to join UCSC as an international student at a time when there was no regular funding for international students, and for a Summer Dissertation Fellowship in 2021, which allowed me to focus completing this work.

I am incredibly grateful for my fellow graduate students (in linguistics and other fields) that I was lucky enough to interact with along the way:

Andrew Angeles for always being a friendly face to talk to; **Vishal Arvindam** for being someone that I could always speak to really well, for being a great collaborator,

and for fun nights out; **Jérémie Beauchamp**, for making me feel less alone among the American students, and for numerous delightful evenings of hanging out; **Netta Ben-Meir** for welcoming me to California; **Juan Carlos Dávila** for being a wonderful friend, for building community, and for some of the best nights with music and drinks; **Karen Duek** for being sensitive and insightful, for lovely days of enjoying the sun, and for helping me connect to people when I first moved to Santa Cruz; **Jack Duff** for doing his thing, being a great travel companion and karaoke-jockey, and for holding it down; **Ben Eischens** for brightening our workdays with never-ending pun fun; **Andrew Hedding** for being my friend from the very beginning, for long nights doing homework together, for sharing hikes, meals, and music; **Morwenna Hoeks** for being someone I could always relate to, and for making our time at the 2019 ESSLLI in Riga so great; **Jyoti Iyer** for a lot of great memories, for being a sincere and serious person, who is also funny, and for good conversations; **Nick Kalivoda** for many lovely afternoons of discussing linguistics or politics in the sun, for all the fun facts and interesting conversations; **Max Kaplan** for being great to hang out with and inviting us over to his place; **Margaret Kroll** for being a role model for me from the very beginning; **Ana McTaggart** for teaching me about new forms of friendship and relating to others, for giving me rides and hanging out in the sun; **Susanna Melkonian-Altshuler** for countless conversations about our shared experiences, and for being my friend, no matter where in the world we are; **Jason Ostrove** for being a generous and enthusiastic collaborator, for teaching me about fieldwork, and illuminating discussions; **Jed Pizarro-Guevara** for making this community a better place, for doing things like taking all the first years out to eat, or making everyone go to karaoke (It would not have been the same without you); **Radhika Prasad** for sharing meals with me and for ameliorating feelings of alienation by experiencing them together; **Stephanie Rich** for being really funny, a great Riga travel companion, and for just being able to share some of this time with you; **Tom Roberts** for

being someone I could talk to about semantics, for a good vibe in our shared office, and for being one of the people hosting things at LingHaus; **Anissa Zaitsu** for discussing our twin research projects during our second year; **Giorgia Zorzi** for being dinosaurs together for halloween and fun times in Northampton; **Eric Zyman** for his big nerd energy, for always being up for a discussion, and for being an excellent protest companion.

Thank you, **Patricia Naumann**, **Dana Römling**, and **Katharina Sternke**, the Kitty Hooligans, for making HHU Düssledorf a more lovely place and for bringing queer-feminism to our department.

Thank you to the **organizers of the UCSC wildcat strike** for fighting to make higher education more accessible and standing up for all graduate students and UCSC community members.

Thanks to other **UCSC ling students** who shared (parts of) this journey with me, for making it special: Lalitha Balachandran, Jenny Bellik, Richard Bibbs, Dan Brodkin, Dhyana Buckley, Yaqing Cao, Maya Wax Cavallaro, Mansi Desai, Hitomi Hirayama, Maho Morimoto, Chris Romano, Deniz Rudin, Kelsey Sasaki, Nick Van Handel, Jake Vincent, Lydia Werthen.

Thank you to the other members of the **student community at UMass** who made my time there an amazing experience. These include the 2015 first-year cohort, who welcomed me in their midst: Carolyn Anderson, Alex Göbel, Christopher Hammerly, Jaieun Kim, Brandon Prickett, Michael Wilson, Rong Yin; as well as: Sakshi Bhatia, Thuy Bui, Devika Dutt, Rodica Ivan, Nick LaCara, Jon Ander Mendia, Deniz Ozyildiz, Ekaterina Vostrikova.

Thank you to other **linguists** who I have had the pleasure of interacting with, for discussing our work, for the lessons you taught, or for making our world feel like a community: Chris Barker, Ryan Bennett, Diti Bhadra, Simon Charlow, Sandy Chung, Haoze

Li, Jesse Harris, Kaden Holladay, Laura Kalin, Magdalena Kaufmann, Elin McCready, Grant McGuire, Dean McHugh, Armin Mester, Morgan Moyer, Jonathan Pesetsky, Judith Tonhauser, Alexander Williams, Zhouye Zhao.

Even though most of them might never see it, I also thankfully acknowledge the friends I made outside of academia, who supported me along the way or simply made my life better. I cannot begin to express how thankful I am for each of these beautiful friendships. Thank you to: **Chimene Agblevon** for being one of the most funny and caring people on the planet; **Mizgin Ekmen** for everything, for being you, for your energy, and including me in your vision of what a community can be; **Engels Garcia** for allowing me to share so many aspects of your life, for being patient with me, and the many things I learned from you; **Lenny Guibert** above all for keeping me fit; **Mia Kaven** for being a great neighbor and an insightful collaborator; **Markus Kresin** for making me feel safe, for hours of playing video games and many long-distance calls; **Laura Ludwig** for always telling it like it is; **Lorenalia Mans** for always being down to hang out, and making me feel a lot less lonely so many times; **Lena Maurer** for being dependable, sweet, and a great discussion partner; **Johannes Newzella** for being such a warm, calm, and nurturing presence, **Marie Thöne** for being part of my journey for so many years—I will always look back fondly on our shared time in school. Thank you to **everyone who is part of YAYA/AMALOPA** simply for existing. You give me hope. I am also grateful to Nils Blömeling, Valentin Barbosa, Evelyn Salguero, Olga Siemens, and many others for making my life sweeter.

Finally, I want to thank my family. My brother **Leon Hofmann,** and my grandparents **Roswitha Koczy-Meier** and **Hans-Hubert Meier** are my rock. Always being able to return to them means the world to me.

None of it all would have been possible without **Petra Koczy**, my very first teacher and biggest supporter. My mother is an educator and values education as the highest good—a value she imparted to me. From the very beginning, she gave me easy access to it and allowed me to find joy in it by encouraging me without pressure. Her unconditional love helps me feel safe in this world.

[—German translation of last paragraph:—]

Schließlich möchte ich meiner Familie danken. Mein Bruder **Leon Hofmann,** und meine Großeltern **Roswitha Koczy-Meier** und **Hans-Hubert Meier** sind mein Fels in der Brandung. Immer wieder zu ihnen zurück kommen zu können bedeutet mir alles.

Nichts von allem wäre mir möglich gewesen ohne **Petra Koczy**, meine erste Lehrerin und größte Unterstützerin. Meine Mutter ist Lehererin und hat mir Bildung als höchstes Gut vermittelt. Von Anfang an hat sie mir den Zugang dazu leicht gemacht, und es mir ermöglicht, Spaß darin zu finden, indem sie mich ohne Druck ermutigt hat. Ihre bedingungslose Liebe hilft mir, mich in dieser Welt sicher zu fühlen.

Chapter 1

Introduction

1.1 Research Question and Empirical Scope

The overarching question addressed in this work is how human users of natural language utilize combined information from multiple sources to understand the meaning of what is being communicated via natural language utterances. The way in which this question is posed and addressed here builds on the basic assumptions that understanding the meaning of an utterance is to arrive at a mental representation of the information encoded by that utterance; and that this representation integrates the conventional information associated with the literal content with contextual information from the situation in which it is communicated, how it is communicated, and what has been communicated before.

The semantic interpretation of anaphora is systematically dependent on the discourse context, i.e., contextual information about what has already been communicated in the discourse. That is why the study of anaphora provides a suitable way of addressing this overarching question. In particular, this work focuses on the interpretation of (semantic[1]) anaphora in interaction with negation and other non-veridical operators[2]. In order to better understand what kinds of representations we need to assume to account for anaphora, this work investigates constraints on how anaphora are interpreted dynamically: What kind of information needs to be available in the discourse for a language user to know which interpretations for anaphora are possible, and which are not. This work contributes to an understanding of this question by applying methodology from formal semantics: It uses tools from mathematical logic to provide a model that can predict intuitive native speaker judgments about the acceptability or truth of natural language utterances in particular (kinds of) situations.

The empirical focus of this work is the interaction of anaphora and negation. Studying the interaction allows us to make advancements in our understanding of both phenomena that would not be possible by examining them individually.

On the one hand, it provides a window into the discourse effect of negation. The ways in which the semantics of negation affects the flow of information in discourse has been studied by investigating its interaction with presuppositions (e.g. Karttunen, 1973; Heim, 1992, et seq.).[3] The effect of negation on anaphoric dependencies provides a unique window into how negation affects how information from earlier parts of the discourse may be available later on.

[1]A large body of work going back to Hankamer and Sag's (1976) distinction of deep vs. surface anaphora suggests that some anaphora operate on syntactic structure (VP-ellipsis being a prime example). This work focuses on anaphora that can be characterized as semantic dependencies.

[2]A (unary) non-veridical operator is one that doesn't entail its prejacent. Formally:
A function $f : A \mapsto B$, where A, B are boolean algebras, is veridical (cf. Zwarts, 1995), iff the following holds:

- $f(p) \rightarrow p$,

and non-veridical otherwise. (See also discussion in Giannakidou, 1998).

[3]Further research into the discourse effect of negation investigates how different ways of expressing negation related to information-structural distinctions effects (e.g. Schwenter, 2002; Amaral, 2007, et seq.), or rhetorical structure (e.g. Clark, 1976; Wason, 1965; Horn, 1989; Moxey, Sanford, and Dawydiak, 2001; Amaral, 2010).

On the other hand, the investigation of anaphora to content introduced under negation and other non-veridical operators allows us to understand better how we talk about hypothetical and counterfactual content, which comes with specific constraints and characteristic ways of being interpreted. Originally, classic dynamic semantic analyses of anaphora (Karttunen, 1976; Kamp, 1981; Heim, 1982, 1983; Groenendijk and Stokhof, 1991; Kamp and Reyle, 1993, et seq.) have focused on cases of anaphora to content that interlocutors consider factual. Anaphora to hypothetical—modalized—content have been analyzed in the literature on modal subordination (Roberts, 1987, 1989; Frank, 1996; Frank and Kamp, 1997; Geurts, 1999; Stone, 1999; Brasoveanu, 2006, 2010a).

The interaction of anaphora and negation has been addressed in the literature in a few ways: **(i)** in dynamic analyses of anaphora to indefinites under negation (Krahmer and Muskens, 1995; Hofmann, 2019a; Gotham, 2019; Elliott, 2020; Lewis, 2021; Mandelkern, 2022), **(ii)** noting so-called negativity-tags—complex anaphoric expressions which are characterized by a requirement that their propositional antecedent to be introduced by a negative antecedent sentence (Klima, 1964; Moxey et al., 2001; Brasoveanu, Farkas, and Roelofsen, 2013; Brasoveanu, De Clercq, Farkas, and Roelofsen, 2014; Hofmann, 2019b), and **(iii)** in literature on polarity particles (like English *yes, no*), which are also characterized as anaphoric to propositional content and senstive to the polarity of the antecedent sentence (see e.g. Pope, 1972; Ginzburg and Sag, 2000; Farkas and Bruce, 2010; Cooper and Ginzburg, 2011; Krifka, 2013; Roelofsen and Farkas, 2015).

1.2 Summary of the Main Problems and Contributions

The present work provides a careful investigation of the various ways in which negation and anaphora interact and develops a generalization that appeals to the notion of speaker commitment (Stalnaker, 1978; Gunlogson, 2004) and veridicality (Zwarts, 1995; Giannakidou, 1998). Based on this generalization, the theoretical claim is put forward that anaphora is sensitive to negation because negated content is stored in the discourse as anti-veridical content. Rather than anaphora being sensitive to a sentential notion of polarity, it is sensitive to the status of the introduced content in relation to the speaker's commitments. This theoretical claim of how anaphora and negation interact is implemented in a formal intensional dynamic semantic framework by conceptualizing discourse referents as a representation that is intensional (Heim, 1983; Stone, 1999; Stone and Hardt, 1999; Brasoveanu, 2006, 2010a) and, in addition, understood epistemically, in relation to speaker commitments. The veridicality-based dynamic semantic account provides a unified analysis of anaphora to indefinites under negation, the interaction of propositional anaphora and negation, polarity-sensitive negativity-tags, and of what counts as a negative sentence/utterance for the purposes of anaphora in discourse.

More specifically, the following empirical issues and main points are addressed here, each of which is introduced in some more detail below:

Empirical issues:

1. Anaphora to Indefinites under Negation

2. The Polarity-Sensitivity of Negativity-Tags

3. Discourse-Negativity

Empirical Contributions:

1. A General Picture – Anaphora and Veridicality

2. Contextual Introduction of Negativity

Theoretical Contributions:

1. Intensional drefs in a flat-update system – Arguments and Implementation

2. Epistemic Drefs

3. Discourse-Negativity – Anti-Veridical Propositional Drefs

Empirical Issue #1: Anaphora to Indefinites under Negation. Anaphora constrains negation. A basic example of this is the contrast in (1), due to Karttunen (1976). While an indefinite in a positive sentence (1-a) can provide the interpretation for the subsequent anaphoric pronoun *it*, the same is not the case for an unspecific indefinite in the scope of negation (1-b).

(1) a. *Mary has [a car]*$^{\upsilon_1}$. *It*$_{\upsilon_1}$ *is parked outside.*

b. *Mary doesn't have [a car]*$^{\upsilon_2}$. *#It*$_{\upsilon_2}$ *is parked outside.*

The question of how can this contrast be explained is addressed in classic dynamic semantic theories by assuming that the content in the scope of negation (and other non-veridical operators) is dynamically inert (Kamp (1981); Heim (1982, 1983); Groenendijk and Stokhof (1991); Kamp and Reyle (1993)). This means that content in the scope of negation can affect the interpretation of the discourse only in a very limited and localized way, and is disregarded for the interpretation of a discourse as a whole. As a result, anaphora to negated content (1-b) are ruled out.

However, anaphora to negated content is possible; it is only more constrained than anaphora to non-negated content. Counterexamples to the classic dynamic semantic view have been around for as long as the generalization itself (Karttunen (1976)). They include cases where indefinites under negation can provide an antecedent for a pronoun when the antecedent is embedded under double negation (Karttunen (1976); Krahmer and Muskens (1995)); cases where the antecedent is in the first disjunct of a disjunction and the anaphor in the second one (these ‘bathroom-disjunctions’ are first noted by Roberts (1989), where they are attributed to Barbara Partee, see also Krahmer and Muskens (1995)); cases of modal subordination (see Roberts (1989); examples of modal subordination with negative antecedents are also noted in Frank and Kamp (1997), Geurts (1999)); or cases where the utterance containing the anaphor disagrees with the utterance containing the antecedent.

(2) a. Double negation:

It’s not true that there isn’t [a bathroom]$^{\upsilon_1}$ *in this house. It*$_{\upsilon_1}$ *is just in a weird place.*

b. Disjunction:

Either there isn’t [a bathroom]$^{\upsilon_2}$ *in this house, or it*$_{\upsilon_2}$ *is in a weird place.*

c. Modal subordination:

There isn’t [a bathroom]$^{\upsilon_3}$ *in this house. It*$_{\upsilon_3}$ *would be easier to find.*

d. Disagreement:

A: *There isn’t [a bathroom]*$^{\upsilon_4}$ *in this house.*

B: *(What are you talking about?) It*$_{\upsilon_4}$ *is just in a weird place.*

The question of how the interpretation of anaphora to negated content can be captured while still accounting for the constraints governing the contrast (1) is a classic problem in dynamic semantics.

Empirical Issue #2: The Polarity-Sensitivity of Negativity-Tags. We also see polarity constraining the interpretation and availability of anaphora in the case of certain expressions that are anaphoric to propositional content, notably polar response particles (Pope, 1972; Ginzburg and Sag, 2000; Kramer and Rawlins, 2009; Farkas and Bruce, 2010; Holmberg, 2013; Krifka, 2013; Roelofsen and Farkas, 2015, among others), polar additive tags (Klima, 1964), and information-seeking uses of why-not questions (Hofmann, 2018; Anand, Hardt, and McCloskey, 2021).

(3) Polar response particles [Roelofsen and Farkas 2015: (6-7)]

a. *Peter passed the test.*

agreement: *{Yes / * no}, he did.*

disagreement: *{*Yes / no}, he didn't.*

b. *Peter didn't pass the test.*

agreement: *{Yes / no}, he didn't.*

disagreement: *{Yes / no}, he did.*

(4) Polar additive tags:

a. *Peter passed the test.*

*{So / * neither} did Sue.*

b. *Peter didn't pass the test.*

{ So / neither} did Sue.*

(5) Information-seeking *'Why not?'*:

a. *Peter passed the test.*

Why (not)?*

b. *Peter didn't pass the test.*

Why (not)?

(3) shows that the pragmatic function of English polar response particles is sensitive to the polarity of the antecedent: In the positive context (3-a), *yes* is used to agree, and *no* to disagree. In the negative context (3-b), either particle can be used for both agreeing and disagreeing (Pope, 1972). (4) shows a binary contrast between polar additive tags with *so* and *neither*: the former can only be used with positive antecedents, and the latter only with negative ones. Because *neither*-tags require negative antecedents, Klima (1964) called them negativity-tags. (5) shows that the information-seeking use of *'Why not?'* could also be considered a negativity-tag: It requires a negtive antecedent, in contrast to a *why*-sluice without negation.

Empirical Issue #3: Discourse-Negativity Although negativity-tags show a clear-cut contrast when comparing negative and non-negative sentences (as in (3) through (5), above), the class of utterances that license negativity-tags has been shown to be both syntactically and semantically diverse (Klima, 1964; Brasoveanu et al., 2013, 2014). The class of utterances that pattern as negative in discourse include utterances of sentences with negative markers (like English *not*), but also ones with semantically weaker anti-additive quantifiers (*never, nobody, nothing*), negative downward-entailing quantifiers (*rarely, few* + *NP*), or negative proximatives (*hardly/barely (any* + *NP)*), in various syntactic positions (such as sentence adverbial or argument positions).

Questions Raised. To summarize, anaphora to negated content is possible, while there are constraints associated with the interaction of anaphora and negation. Some anaphora resist negative antecedents (1-b), while others require them (4)+ (5); yet others are interpreted differently based on the antecedent polarity (3). The complex ways in which anaphora interact with negation give rise to the following questions:

1. How can we characterize 'negative content' (for the purposes of anaphora) and the discourse-effect of negation?

2. How can the interpretation of anaphora to negated content be best explained in dynamic semantics?

Investigating these questions, the work makes the following main empirical and theoretical points.

Empirical Contribution #1: A General Picture. Anaphora patterns systematically along the lines of veridicality-distinctions and reference to non-veridical antecedents is not exceptional. This is the case for both individual and propositional anaphora, which pattern in parallel ways. Roughly speaking, in cases where both discourse segments are required to be compatible (such as subsequent assertions by the same speaker), there is a weak veridicality-matching requirement[4]:

A veridical anaphor (one interpreted in a veridical context) needs a veridical antecedent (introduced in a veridical context as well); non-veridical antecedents can be picked up by anaphora that are also non-veridical (and therefore do not require a veridical referent). Negation is not exceptional wrt the constraints it places on anaphora, but instead, it fits into the broader landscape of non-veridical operators.

Theoretical Contribution #1: Intensional drefs in a flat-update system – Arguments and Implementation. Suppose we accept the argument that the discourse-behavior of negation fits into a larger picture of patterns along the lines of veridicality-distinctions. In that case, we will also want to give a unified account for double negation, bathroom-disjunctions, modal subordination with negative antecedents, inter-speaker disagreement, and, more broadly, anaphora to antecedents embedded under negation or other non-veridical operators. We would want a unified account to explain the circumstances under which anaphora have a suitable antecedent, for anaphora in veridical contexts and anaphora in non-/anti-veridical contexts.

This work presents arguments for a flat-update dynamic semantics with intensional drefs. Flat-update in this context means that every expression with the potential to add discourse referents will do so on a global level—no matter the embedding context. Negation is treated as an externally dynamic operator (and more generally, there are no externally static operators in a flat-update system). [5] The constraints on anaphora that govern the basic contrast in (1) (repeated here) are captured in a different way.

[4]Thanks to an anonymous reviewer for this suitable description

[5]See Stone (1999); Stone and Hardt (1999); Brasoveanu (2006, 2010a,b); Elliott (2020); Lewis (2021) for dynamic frameworks which can be characterized as *flat-update* in this sense.

(1) a. *Mary has [a car]*$^{\upsilon_1}$. *It*$_{\upsilon_1}$ *is parked outside.*

b. *Mary doesn't have [a car]*$^{\upsilon_2}$. *#It*$_{\upsilon_2}$ *is parked outside.*

To understand this difference, we distinguish between different kinds of information stored in the discourse. We can distinguish a referential context—storing the salient drefs—and an intensional context—storing the set of possible worlds under consideration. Classic dynamic theories capture the contrast in (1) by appealing to externally static operators, which may introduce a local referential context in relation to which the embedded information is interpreted before it becomes inaccessible. Following analyses of modal subordination from Stone (1999); Stone and Hardt (1999); Brasoveanu (2006, 2010a), the analysis presented here captures (1) by interpreting discourse referents intensionally and in relation to their local intensional context instead. While we have only a global referential context, the discourse effect of embedding contexts is captured intensionally by interpreting drefs relative to different local and global intensional context sets of possible worlds.

Following Brasoveanu's (2006, 2010a) implementation of Muskens's (1996) Compositional DRT, I develop a formal semantic framework that is suited to implement the analysis: Intensional Compositional DRT.

The classic dynamic systems, which do not allow for the global introduction of new drefs in the scope of negation, cannot account for anaphora to negated content. Here, we group analytical approaches which could be applied to address anaphora to negated content into three categories: **(i)** The flat-update dynamic systems described above; **(ii)** analyses under which drefs are introduced conditionally in case the local embedded proposition is true (Krahmer and Muskens, 1995; Gotham, 2019; Mandelkern, 2022); **(iii)** classic dynamic systems in combination with a mechanism of accommodation (e.g. Roberts, 1987, 1989; Nouwen, 2003; Lewis, 2021). Considering the various cases of anaphora to negated content, we compare these kinds of accounts and provide arguments for a flat-update dynamic semantics with intensional and epistemic dref representations.

Theoretical Contribution #2: Pragmatic Drefs. This work explores how the interaction between negation and anaphora can be captured in a flat-update semantics, the consequences for representations of drefs, and the kind of information needed to be stored in discourse. It argues that the systematic interactions between anaphora and veridicality can be explained by appealing to general principles of discourse consistency, as soon as we make the following assumptions:

- **Intensional drefs in a flat update semantics** (as suggested above, following Stone (1999); Brasoveanu (2006))

  All operators are externally dynamic, while drefs store intensional information about the possible worlds in which they are located. E.g., indefinites introduce drefs relative to their local intensional (embedding) context, and pronouns are interpreted relative to their local intensional context.

- **Epistemic status in relation to speaker commitments** (developed here)

  The discourse needs to store the information about the relation between intensional drefs and speaker commitments in order to evaluate constraints on anaphora.

This work shows that these representational assumptions allow us to explain the veridicality-based generalizations on the distribution of possible anaphora in terms of the more general principle that a discourse be consistent. This is interpeted as an argument that the interpretation of anaphora relies on drefs storing some pragmatic information about their epistemic status in relation to the discourse commitments of the interlocutors.

Empirical Contribution #2: Contextual Introduction of Negativity. The negativity-tags in (3) trough (5) have often been characterized as sensitive to a notion of negativity.

Literature addressing the question of what counts as a discourse-negative context for the purposes of anaphora has always acknowledged the syntactic and semantic diversity among these contexts, but I also show that negativity can be introduced contextually.

The generalization that is developed here is the following: As long as a proposition is introduced into the discourse, the inference that the speaker is committed to its falsity may be contributed by a contextually introduced inference. This is enough to make a context negative for the purpose of anaphora. To show this, we will consider data from contexts which Kroll (2019) identified as contexts licensing polarity reversal in sluicing (such as neg-raising, disjunction), as well as contexts in which strong NPIs may be licensed without an overt operator (such as sarcasm, see Horn (2016)).

I also present experimental evidence showing that negativity-tags (in particular information-seeking or factive uses of elliptical *'why not'*-questions) are (to varying extent) licensed following utterances with various kinds of anti-veridical embeddings (anti-veridical attitudes, cases of neg-raising that are uncontroversially derived pragmatically, as well as clauses with overt negation, and more stereotypical cases of neg-raising with *think/believe*).

Theoretical Contribution #3: Discourse-Negativity. The evidence showing that negativity may be introduced contextually or by anti-veridical attitudes suggests that a notion of discourse-negativity that is relevant for anaphora could not be characterized by appealing to a notion of sentential negation. Instead of a syntactic or semantic representation that makes reference to sentences, I suggest that the relevant generalization characterizes discourse-negative utterances as ones that introduce an anti-veridical discourse referent. The result is a commitment-based account of anaphoric polarity-sensitivity and discourse-negativity.

1.3 Structure of the Dissertation

To make these points, the work proceeds in the following way:

Chapter 2: The Dynamics of Negation provides some background on the topics of anaphora, discourse referents and dynamic semantics, anaphora to indefinites under negation, and anaphoric accessibility in classic dynamic semantic frameworks. It goes on to present arguments for a treatment of negation as an externally dynamic operator, and more broadly for a flat-update dynamic semantics that stores intensional and epistemic information about discourse referents. While different factors constrain anaphoric co-reference, this chapter provides arguments that the kinds of constraints that classic dynamic frameworks (Kamp (1981); Heim (1982, 1983); Groenendijk and Stokhof (1991); Kamp and Reyle (1993)) capture intra-sententially by evoking a distinction between externally dynamic vs. externally static operators can be generalized to cross-sentential anaphora by appealing to generalizations in terms of veridicality and a speaker's discourse commitments.

Chapter 3: Intensional CDRT develops a suitable formal framework to capture these generalizations. This framework should have drefs to handle anaphora, intensionality to capture veridicality distinctions, and a representation of the speaker's commitments to represent information about their epistemic state. Chapter 3 introduces such a framework—Intensional CDRT. The chapter introduces basic definitions (basic types, representations of discourse updates, truth and acceptability) and illustrates how basic declarative utterances as well as anaphoric expressions are interpreted.

Chapter 4: Propositional Operators and Propositional Drefs adds propositional operators to intensional CDRT. A (unary) propositional operator introduces a propositional dref corresponding to its prejacent and lexicalizes a relationship between this embedded dref and the embedding context. The information stored in the discourse keeps track of the relationship between propositional drefs and the interlocutors' commitments. The newly introduced dref also acts as the local intensional context for the material in its scope. This chapter also discusses some limitations associated with the system: The truth-conditions we are able to derive for non-specific indefinites in the scope of negation are too weak. Basically, the system can only derive specific[6] indefinites. An argument is provided that a solution of this issue requires that individual drefs are relativized to the possible worlds that their referent exists in (this will be implemented in **Chapter 7**).

Chapter 5: Propositional Anaphora and Veridicality shows how propositional anaphora pattern along the lines of veridicality and how the account in **Chapter 4** can provide an analysis of these patterns. The anaphoric dependency requires that antecedent and anaphor be compatible with an interpretation with the same epistemic status. The speaker is committed to the truth/falsity of the one exactly when they are also committed to the truth/falsity of the other. Since antecedent and anaphor refer to the same proposition, their epistemic status must be the same. While the semantic properties of the embedding context do not have to be precisely the same, there is a weak requirement that the veridicality of the embedding contexts matches. The analysis in ICDRT shows that this property of weak veridicality-matching falls out from discourse-consistency, as long as we keep track of relations between propositional drefs and their relationship to the speaker's discourse commitments.

[6] i.e. scopally specific in the sense of Farkas (1994, 2002)

Chapter 6: Discourse-Negativity and Anaphoric Polarity-Sensitivity provides a pragmatic characterization of anaphoric polarity-sensitivity and discourse-negativity within the framework developed here. This is supported by judgment data suggesting that discourse-negativity may be introduced contextually, and experimental evidence showing that anti-veridical attitudes and pragmatically derived cases of neg-raising introduce discourse negativity (at least to some extent). The chapter goes on to show how the polarity-sensitivity of information seeking/factive uses of elliptical *'why not'*-questions can be analyzed in the system with only minimal extensions.

Chapter 7: Individual Anaphora shows how individual anaphora pattern along the lines of veridicality and formalizes a way of keeping track of veridicality for individual drefs. This is done by introducing individual drefs relative to their local intensional context (following Stone (1999); Stone and Hardt (1999)). Since our system already keeps track of the relationship between salient sets of worlds and the speaker's discourse commitments, relativizing individual drefs in this way will enable us to store the information whether the speaker is committed to the existence of a referent, the opposite, or neither.

I show how this formalization can account for the interaction of negation and individual anaphora, in particular, how it can address the classic problem of anaphora to unspecific indefinites in the scope of negation. The chapter also provides a comparison to other accounts to the same issue, and discusses some implications for related issues.

Chapter 8 concludes the work by summarizing the main points and pointing out some open questions and possible avenues for further research into the interaction of anaphora and negation.

Chapter 2

The Dynamics of Negation

2.1 Introduction

This chapter constitutes a first step towards a formal analysis of the interaction of anaphora and negation. The empirical issues outlined in the introduction (**Chapter 1**) paint the picture that anaphora to negated content is constrained, but possible, and sometimes even required. One of main questions that arises is how the relevant constraints can be captured while permitting counterexamples in an explanatory way. Given that anaphora to negated content (and non-veridical content more broadly) is ruled out in standard dynamic semantic frameworks, this is a classic problem in dynamic semantics.

The current chapter presents relevant background assumptions and arguments for the theoretical stance taken in this work: That the sensitivity of anaphora to negation and polarity is best understood pragmatically, in relation to the speaker's commitments. The chapter does this by carving out generalizations and analytical intuitions on which the formal account presented in upcoming chapters rests. It proceeds in the following way:

Section 2.2 provides background on the topics of anaphora, discourse referents and dynamic semantics, as well as anaphora to indefinites under negation. The reader who is familiar with these topics may want to skip this background section (or read selectively).

Section 2.3 discusses two kinds of counterexamples to the overall generalization that anaphora to content in the scope of non-veridical operators is impossible: Either the antecedent is actually in a veridical embedding context (like in the case of antecedents under double negation), or the anaphor is in a non-veridical context itself (like in the case of modal subordination). The section illustrates why these kinds of cases are problematic for the classic account, and how they fit within a general pattern of along the lines of veridicality.

This section more specifically develops a generalization about anaphoric accessibility, by building on Karttunen's (1976) generalizations about discourse referents, and ideas of how anaphora interact with their local contexts from Stone (1999); Brasoveanu (2006); Kroll (2019).

(1) **Generalization about anaphoric accessibility**

a. An (individual-denoting) expression is a possible antecedent for an individual anaphor just in case the referent of the antecedent exists in the local context of the anaphor.

b. An (proposition-denoting) expression is a possible antecedent for a propositional anaphor just in case the proposition expressed by the antecedent is true in the local context of the anaphor.

(1) highlights the parallels between individual and propositional anaphora, and suggests that anaphoric accessibility—the question of whether or not an anaphoric relation is possible, given a potential antecedent and anaphor—is constrained by the veridicality of

the embedding context. This predicts that the question whether the speaker is committed to the local context (i.e. its veridicality) directly affects the availability of an anaphoric relation—consistently with what we observe.

Section 2.4 provides arguments for treating negation as externally dynamic in an intensional flat-update semantics with epistemic drefs. This approach is compared to accommodation-based accounts and ones that analyze anaphora to negated content by allowing negated content to introduce negated content under certain conditions.

Section 2.5 concludes the chapter by summarizing its main points.

2.2 Background on Anaphora and Discourse Referents

2.2.1 Anaphora and Co-reference

The phenomenon of ANAPHORA is generally understood as a dependency between two linguistic expressions. A prototypical example is given in (2).

(2) *Mary has [**a car**]$^{\upsilon_1}$. **It**$_{\upsilon_1}$ is parked outside.*

In (2), the anaphoric pronoun *it* depends on its ANTECEDENT *a car* for its (semantic) interpretation. In particular, the relationship between the anaphor and its antecedent in (2) is one of CO-REFERENCE: Both denote the same object, Mary's car, and the anaphor inherits this denotation from the antecedent. This relationship is indicated by annotation with the same variable name (υ_1).

Giving an exact theory-neutral characterization of anaphora has proven difficult. The Stanford Encyclopedia of Philosophy entry on anaphora (King and Lewis (2018)) discusses two common characterizations of anaphora as a phenomenon whereby:

(3) a. *'the interpretation of an occurrence of one expression depends on the interpretation of an occurrence of another'*

or

b. *'an occurrence of an expression has its referent supplied by an occurrence of some other expression.'*

King and Lewis (2018) discuss that the characterization in **(a.)**, in terms of dependence on the linguistic context, is too weak, while **(b.)**, formulated in terms of co-reference, is too strong. We can see why when making explicit what is meant by the terms REFERENT and REFERENCE. In a traditional sense, an expression is referential if it denotes a specific individual object. In the compositional formal semantic perspective adopted in this work, DENOTATION can be understood as a semantic value in a Fregean/Tarskian sense. In these terms, we say that an expression is referential if it denotes a specific object, and the referent of an expression is the specific object denoted by the expression (for an overview of theories of reference, see, e.g., Yli-Vakkuri and McGilvray (2010)).

Considering the two possible characterizations of anaphora in (3), **(a.)** is liberal enough to include cases of lexical disambiguation, which we would not want to characterize as anaphoric (*'John is down by the* ***bank*** *of the river.'*, King and Lewis (2018): (1)). **(b.)**, on the other hand, fails to include bound anaphora (4-a) and donkey anaphora (4-b), which are traditionally characterized as anaphora but are not normally thought of as referential because they do not denote a specific object in the domain of reference.

(4) a. Bound anaphora:

Every actress$^{\upsilon_2}$ *believes* ***she***$_{\upsilon_2}$ *deserves more roles.*

b. Donkey anaphora:

Every farmer who owns ***a donkey***$^{\upsilon_3}$ *feeds* ***it***$_{\upsilon_3}$.

The pronouns *she* and *it* in (4) do not refer to a particular actress or a particular donkey. Instead, the denotation of *she* varies based on the value of the quantifier *every actress*. (4-a) could be paraphrased as: 'Each actress (in the domain of quantification) believes that **that actress** deserves more roles'. Because of the syntactic configuration here, which allows for the pronoun to co-vary with a quantifier that commands it, this kind of anaphora is commonly analyzed as a syntactically bound variable. This relationship is indicated by co-indexation with the same variable name here as well.

Similarly, the denotation of *it* varies based on the values of *every farmer* and *a donkey*. (4-b) could be paraphrased as: Each farmer who owns a donkey, feeds **that donkey**. Donkey anaphora of this kind involve co-variation of a pronoun with a quantifier, without command or syntactic binding.

2.2.2 Discourse Referents and Dynamic Semantics

The use of DISCOURSE REFERENTS (drefs) in dynamic semantics (Karttunen (1976); Kamp (1981); Heim (1982); Groenendijk and Stokhof (1991)) presents a generalized notion of reference, which allows for a unification of these three varieties of anaphoric dependencies. Dynamic semantic theories assume that the meaning of natural language utterances is interpreted by sequentially updating and accessing information in a representation of the discourse. A discourse referent is a mapping from a variable name to a referent object and can be stored and accessed dynamically. These theories treat anaphoric pronouns as DISCOURSE VARIABLES, which receive their denotation from a previously established discourse referent. In the basic case (2), the dynamic interpretation of the indefinite *a car* introduces a discourse referent into a representation of the current state of the discourse: The variable υ_1 now maps to an object in the domain, namely Mary's car. This enables a subsequent pronoun to be interpreted with the same variable name, which is now defined in the discourse as mapping to Mary's car.

In the bound anaphor case (4-a), the quantificational antecedent also introduces a discourse referent into the representation (υ_2). Three features of discourse referents in dynamic semantics are essential for the interpretation of bound anaphora: **(i)** the introduction of new mappings is non-deterministic; **(ii)** we can quantify over possible mappings; and **(iii)** mappings from discourse variables to referent objects may be dependent on being in the scope of certain semantic operators, such as quantifiers. The quantificational antecedent *every actress* introduces the discourse referent υ_2, which enables a subsequent pronoun to be interpreted with the same variable name and thus the same mapping(s). Due to the non-deterministic nature of introducing discourse referents **(i)**, υ_2 may map to any actress in the domain of reference. Although the overall interpretation of the quantifier is non-referential (it does not map to a single specific object), each of the non-deterministic quantificational instances, viewed in isolation, are referential in a localized sense (for each possible mapping, υ_2 maps to exactly one object in the domain). The semantics of the quantifier is formalized by quantifying over these possible mappings **(ii)**: Since (4-a) involves universal quantification, the truth conditions require that every single one of the possible mappings satisfying the restrictor of the quantifier (υ_2 is an actress) also satisfy its nuclear scope (υ_2 believes that υ_2 deserves more roles). Because the (singular) 'referentiality' of the quantifier (in the sense discussed here) is tied to specific quantificational instances, a 'co-referential' (singular) pronoun can only be used in the scope of the quantifier, which is evaluated in relation to each of the quantificational instances **(iii)**. As a result, donkey-anaphora in a subsequent clause that is in the scope of the quantifier is acceptable (4-b), but those in a subsequent clause outside of it are not (5-b).

(5) a. *Every farmer who owns* ***a donkey***$^{\upsilon_3}$ *feeds* ***it***$_{\upsilon_3}$.

b. #***It***$_{\upsilon_3}$ *likes hay.*

We can describe this behavior as by characterizing the quantificational operator contributed by *every* as INTERNALLY DYNAMIC and EXTERNALLY STATIC (in the sense of Groenendijk and Stokhof (1991)). It is internally dynamic because, in the scope of the quantifier, a discourse referent is introduced, which contributes a mapping for each of the non-deterministic quantificational instances, and can be retrieved anaphorically within the scope of the quantifier. It is externally static because the same discourse referent is not available outside of the scope of the quantifier, which accounts for the descriptive characterization as non-referential.

Donkey anaphora like (4-b) work like bound pronouns in dynamic semantics: Although there is no syntactic configuration of command between the antecedent *a donkey* and the anaphor *it*, both are in the scope of the quantifier *every farmer* and co-vary with its value. Similarly to (4-a), the antecedent introduces a dref that has a specific mapping for each quantificational instance[1], which allows the dref to provide the denotation for an anaphoric pronoun. Here, both the antecedent and anaphor are in the scope of a quantificational operator, which can be characterized as internally dynamic and externally static.

Discourse referents and their interaction with quantifiers allow for a generalized notion of (co-)reference. Although the anaphor and its antecedent in (4) do not involve 'reference' in a traditional sense, i.e., in the sense that these expressions denote a specific object. We can say that they exhibit a kind of relative reference: They map to a specific object *relative to* each of the possible quantificational instances (see the treatment of quantification in Kamp, 1981; Heim, 1982; Groenendijk and Stokhof, 1991).

[1] I am glossing over strong donkey readings, under which an indefinite antecedent may be interpreted as a universal quantifier as well. This difference is not substantial for the point made here.

The present work is concerned with anaphora considered non-referential for different reasons: Anaphora to non-specific indefinites in the scope of non-veridical operators, such as negation. An initial example is given in (6):

(6) *Mary doesn't have [**a car**]$^{\upsilon_4}$ or a place to park **it**$_{\upsilon_4}$.*

The anaphoric pronoun *it* is neither bound nor a donkey anaphor. Its value does not co-vary with the value of a quantifier. Although both indefinites in (4) and (6) are non-referential they are quite different.

Language encodes (non-)referentiality as (non-)specificity. Although Farkas (1994, 2002), who distinguishes between different notions of (non-)specificity would characterize both (4) and (6) as scopally and epistemically non-specific, Farkas (1997) shows that these kinds of contexts exhibit different kinds of (scopal) non-specificity which are linguistically distinct. In particular, Farkas demonstrates that morphologically marked dependent indefinites in Hungarian and Romanian are only interpreted in the context of universal quantifiers, quantificational adverbs, iterative aspect, or distributive plurals.

Brasoveanu and Farkas (2011) note that these are precisely the contexts that can be analyzed in terms of quantification over variable mappings ranging over possible values, i.e., environments that introduce discourse-plurality in the sense of Brasoveanu (2008). In contrast, strong negative polarity items (NPIs) are indefinites morphologically marked to be interpreted in the scope of negation (see overview in Farkas and Brasoveanu (2020)).

In the sense that it lacks co-variation, the configuration of (6) is more similar to that of the referential anaphor in (2). Nevertheless, the indefinite antecedent *a car* and the anaphor in utterances like (6) are considered non-specific and non-referential.

Although there is a sense in which we might say that the indefinite *a car* and the anaphoric pronoun *it* make reference to Mary's car, or rather the idea of Mary's car, this car is not an actual object. To the contrary: An utterance of the first sentence in (6) asserts that such a car does not actually exist. The reason for this is that the antecedent *a car* is in the scope of negation: It is an unspecific indefinite and non-referential in this sense. Nevertheless, we can use an anaphoric pronoun to 'refer' back to it.

However, like with unspecific indefinites in the scope of quantifiers, the mappings for variables introduced in the scope of negation are dependent on being in the scope of the scope-taking operator (in this case, negation). As a result, a subsequent anaphor that is in the scope of the quantifier is acceptable (as in (6), shown above), but one outside of it is (usually) not (7).

(7) a. *Mary doesn't have [**a car**]*$^{\upsilon_5}$.

b. *#It*$_{\upsilon_5}$ *is parked outside.*

In that sense, this second kind of unspecific indefinite behaves similar to the first one. To employ a generalized notion of (discourse) reference in an explanatory theory of anaphora, we may consider discourse referents being interpreted in relation to propositional operators like the negation in (7): They map to a specific object *relative to* a hypothetical situation which is contrary to what is asserted to be true, and therefore relative to the negated content (Stone, 1999; Stone and Hardt, 1999; Brasoveanu, 2006, 2010a).

2.2.3 Accessibility

ACCESSIBILITY is the question of whether or not an anaphoric relation is possible, given a potential antecedent and an anaphoric expression. Dynamic semantic theories have traditionally treated the question of accessibility as a theoretical notion, which needs to be explicitly addressed in a theory of anaphora (Kamp, 1981; Heim, 1982; Groenendijk and Stokhof, 1991; Kamp and Reyle, 1993). This classic account of accessibility constraints is based on the assumption that negated content (as well as content in the scope of certain kinds of operators) is dynamically inert and could not influence the interpretation of the surrounding context. These assumptions aim to capture the basic fact that, while indefinite noun phrases can generally be antecedents for anaphora, non-specific indefinites in the scope of negation often cannot. This generalization (due to Karttunen, 1976), is illustrated again in (8).

(8) a. *Mary has [a car]$^{\upsilon_1}$. It$_{\upsilon_1}$ is parked outside.*

b. *Mary doesn't have [a car]$^{\upsilon_2}$. #It$_{\upsilon_2}$ is parked outside.*

For example, Dynamic Predicate Logic (Groenendijk and Stokhof, 1991) lexically distinguished between externally static and externally dynamic operators. Negation is characterized as externally static, i.e., assuming that the negated content cannot provide new antecedents for subsequent anaphora. Similarly, the same kind of distinction in Discourse Representation Theory (Kamp, 1981; Kamp and Reyle, 1993) is made by suggesting that negation (and other assumed dynamically inert operators) introduce subordinate parts of the discourse representation which are temporary and inaccessible on the global level.

Counterexamples to the generalization illustrated by the contrast in (8) suggest that accessibility should not be treated as a theoretical notion that is hard-coded into the lexical semantics of propositional operators.

In the following chapters, I argue that we need intensional and epistemic representations of discourse referents to capture the contrast in (8), as well as the counterexamples. The formal analysis in the upcoming chapters shows that no separate theory of accessibility is needed when we assume the appropriate kinds of representations for drefs. Instead, the constraints on anaphoric relations fall out from the general requirement that a discourse be non-contradictory (in particular, this is shown in **Chapters 5&7**, for propositional and individual anaphora, respectively). Therefore, the term ‘accessibility’ will be used purely descriptively here rather than as a theoretical notion.

2.3 Towards Analyzing Anaphora to Negated Content

To work towards an understanding of anaphora to negated content, this section presents some central generalizations about how anaphoric accessibility is influenced by negation. **Section** 2.3.1 gives an overview some related cases that are problematic to the classic dynamic semantic theories of anaphora.

Based on generalizations and data already discussed by Karttunen (1976), the following two subsections discuss why certain cases are problematic to the classic dynamic semantic theories of anaphora, and what kind of generalization would be a suitable basis for a theoretical analysis. To this effect, the section proposes that anaphoric accessibility is influenced by speaker commitment to the referent of an anaphor and its (potential) antecedent.

This is established empirically both on the antecedent side, and the side of the anaphor. In particular, **Section** 2.3.2 shows that accessibility is affected by the veridicality of the embedding context of the (potential) antecedent, and how this generalization can explain double negation and related cases. **Section** 2.3.3 shows that accessibility is affected by the veridicality of the embedding context of the anaphor, and how this part of the generalization can explain cases of modal subordination, local satisfaction, and accommodation of anaphoric requirements.

Discussing these two kinds of empricial problems for the classic account posed by the respective cases, the section shows that anaphora is systematically influenced by the status of both anaphor and antecedent in relation to the speaker’s commitments, which in turn is influenced by the veridicality of their local embedding context. This general pattern provides an argument for treating double negation, modal subordination, and the other cases of anaphora to negated content considered here in a unified way—by appealing to speaker commitments and veridicality.

2.3.1 Accessing the Inaccessible

The traditional dynamic semantic conception of negation as externally static, assuming that the flow of information from the negated content to parts of the discourse outside of the scope of negation is completely interrupted, captures the basic contrast noted above.

Here, we are considering cases where this overall generalization does not apply. The counterexamples with individual anaphora introduced in **Chapter 1** are repeated here:

(9) a. Double negation:

It's not true that there isn't [a bathroom]$^{\upsilon_1}$ *in this house. It*$_{\upsilon_1}$ *is just in a weird place.*

b. Disjunction:

Either there isn't [a bathroom]$^{\upsilon_2}$ *in this house, or it*$_{\upsilon_2}$ *is in a weird place.*

c. Modal subordination:

There isn't [a bathroom]$^{\upsilon_3}$ *in this house. It*$_{\upsilon_3}$ *would be easier to find.*

d. Disagreement:

A: *There isn't [a bathroom]*$^{\upsilon_4}$ *in this house.*

B: *(What are you talking about?) It*$_{\upsilon_4}$ *is just in a weird place.*

According to the standard dynamic perspective, these cases would be ruled out. Broadly, there are two kinds of exceptions here: Double-negation cases, where the two instances of negation cancel each other out, just like they would in a static framework, and discourse subordination cases.

Discourse subordination is not particular to negation but part of a broader phenomenon. Similar kinds of exceptions to the basic pattern can be found in classic cases of modal subordination with modal antecedents (Roberts, 1989).

(10) *[A wolf]$^{\upsilon}$ might walk in.* (Roberts, 1989, 11)

a. *# It$_{\upsilon}$ is gray.*

b. *It$_{\upsilon}$ would eat you first.*

An indefinite DP under a non-veridical modal operator cannot provide an antecedent for the individual anaphor in (10-a), in a non-modalized indicative context. If the anaphor itself is embedded under the modal *would*, however, this is possible (10-b).

Anaphora to antecedents under universal quantifiers represent another related case. Contrary to the generalization introduced in the above **Section 2.2.2**, these kinds of anaphora are possible under certain circumstances, like in the cases of quantificational subordination (Karttunen, 1976), telescoping (Roberts, 1987), or plural anaphora.

(11) a. Quantificational Subordination (Karttunen, 1976)

Every time Bill comes here, he picks up [a book]$^{\upsilon_1}$ and wants to borrow it$_{\upsilon_1}$. I never let him take it$_{\upsilon_1}$.

b. Telescoping: (adapted from Roberts, 1987)

[Every actress]$_{\upsilon_2}$ who won an award walked to the stage. She gave an acceptance speech and thanked her colleagues.

c. Plural Anaphora: (adapted from Roberts, 1987)

[Each candidate]$^{\upsilon_3}$ for the space mission meets all our requirements. They$_{\upsilon_3}$ have a Ph.D. in Astrophysics and extensive prior flight experience.

For the interaction of negation and anaphora, we are concerned with the interaction between anaphora and their intensional contexts and setting aside the cases where anaphora interact with individual quantification in (11). However, some of the points made here for discourse subordination with negative antecedents could also be applied to other kinds of discourse subordination. See Brasoveanu (2006) for a unified account of quantificational and modal subordination and further discussion of the parallels.

Problems for the Classic View. As noted above, cases of discourse subordination are a challenge for classic dynamic semantic theories because they involve anaphora with antecedents from parts of the discourse that are otherwise thought of as inaccessible or opaque and generally not able to introduce new drefs to the overall discourse. Double negation cases and bathroom sentences present a similar kind of challenge because indefinites under negation may be picked up anaphorically here as well.

2.3.2 Double Negation and Speaker Commitment

Krahmer and Muskens (1995), addressing anaphora to indefinites under negation, frame the problem by making reference to the fact that certain equivalences from static logic do not hold in dynamic semantics. Particularly, because negation is externally static, there is no double negation elimination in dynamic semantics. Based on this, (12-b) is not predicted to be dynamically equivalent to (12-a), or license anaphora to material in the scope of negation (12-c).

(12) a. *John brought an umbrella.*

b. Krahmer and Muskens (1995):

It's not true that John didn't bring an umbrella.

c. *It was purple and stood in the hallway.*

Veridical Contexts under Negation

In his seminal paper on discourse referents, Karttunen (1976) introduces the similar examples in (13).

(13) a. *John forgot not to bring [an umbrella]*$^{\upsilon_1}$*, although we had no room for it*$_{\upsilon_1}$.

[adapted from Karttunen (1976): 14b]

b. *Bill didn't realize that he had [a dime]*$^{\upsilon_2}$*. It*$_{\upsilon_2}$ *was in his pocket*

[Karttunen (1976): 16]

(13-a) is in a sense analogous to the double negation case: The indefinite *an umbrella* is in the scope of negation, but also an additional non-veridical operator. As a result, it can be the antecedent to a subsequent anaphor. The ameliorating effect of a second negation can also be provided by the anti-veridical implicative verb *forget*. (13-b) is not exactly analogous to double negation, but here the anaphoric relation is enabled by the fact that the factive attitude *realize* takes narrow scope wrt negation. As a result, a speaker of (13-b) is committed that Bill had a dime, and the indefinite *a dime* can be picked up anaphorically.

In the discussion of these examples, Karttunen, therefore, suggests that it is not the presence of negation itself that modulates the possibility of an anaphoric relation, but the proposition introduced by the sentence containing the indefinite, and whether the speaker is committed to its truth or falsity. He writes:

(14) Karttunen's generalization for non-specific indefinites:

A non-specific indefinite [...] in an affirmative sentence (single sentence or a complement) establishes a discourse referent just in case the proposition represented by the sentence is asserted, implied or presupposed by the speaker to be true. A non-specific indefinite in a negative sentence establishes a referent only if the proposition is implied to be false.

— Karttunen (1976) (p. 13)

This generalization characterizes the cases in (13): (13-b) presupposes that *Bill has a dime*, and (13-a) entails that John actually brought an umbrella, and therefore that it's false that *John didn't bring an umbrella.* According to (14), that makes an anaphoric relation possible here.

Assuming that the negated propositional content is represented as its own local proposition in the scope of negation, we can summarize Karttunen's generalization by stating that an indefinite introduces a discourse referent, only if its most local proposition is *'asserted, implied, or presupposed by the speaker to be true.'* For our purposes, consider a proposition the most local proposition to a DP, iff it is the proposition explicitly expressed by the smallest constituent s.t. the meaning of the DP contributes to the semantic value of the proposition. We can also subsume the different ways in which a speaker might present themselves as believing in the truth of a proposition (asserting, implying, or presupposing) using the notion of speaker commitment (Gunlogson (2004)). With these updates, we reformulate Karttunen's generalization:

(15) Updated generalization for non-specific indefinites (preliminary)

A non-specific indefinite establishes a discourse referent just in case the speaker is committed to the truth of the proposition most local to the indefinite.

(15) holds that a only non-specific indefinites in veridical contexts license subsequent anaphora. The generalization extends straightforwardly to double negation cases more generally:

(9-a) Double negation:

It's not true that there isn't [*a bathroom*]$^{\upsilon_1}$ *in this house.*

It$_{\upsilon_1}$ *is just in a weird place.*

(9-a) entails that *there is a bathroom in this house,* making the negated content true on the level of speaker commitments. As a result, the speaker will also be committed to the existence of a referent for the indefinite.

Entities and their Local Intensional Context

Note that there is an intimate connection between a speaker's commitment to the existence of a referent for an indefinite and their commitment to the truth of the most local proposition to which the indefinite contributes. This connection can be illustrated with regards to the examples above: In (13-a), *an umbrella* refers to an existing umbrella, but it could only be the umbrella that John brought. The speaker, committing to the existence of the umbrella, also has to commit to it being true that *John brought an umbrella*. In (13-b), *a dime* refers to an existing dime, but it could only be the dime that Bill had. Similarly, committing to the existence of this dime, the speaker also commits to it being true that *Bill had a dime*. Instead of talking about the truth of the proposition local to a non-specific indefinite, we could equivalently also talk about the existence of its referent.

(16) Updated generalization for individual anaphora (preliminary)

An individual-denoting expression establishes a discourse referent just in case the speaker is committed to the existence of its referent.

That also allows us to talk about individual-denoting expressions more generally. In particular, we do not need to exclude specific indefinites from this generalization. While the empirical focus of this work lies on non-specific indefinites, a generalization that includes specific indefinites might be more explanatorily useful.

Indefinites that are in a similar syntactic configuration to our non-specific indefinites under negation, like *a paper* in (17-a), can be interpreted specifically as well. Under this kind of reading, they take wide scope over negation. (17-a) means that there is a specific paper (from the list of required readings for the class) that Mary did not read. Since the existential quantification associated with the indefinite is not interpreted in the scope of negation, the speaker will be committed to the existence of its referent. As we might expect, this also allows for subsequent anaphora (17-b).

(17) Context: Mary is enrolled in a class with a list of required readings.

a. *Mary didn't read [a paper]*$^{\upsilon}$.

b. *It*$_{\upsilon}$ *was too dense.*

Propositional Anaphora

Propositional anaphora provide further evidence that negated content can provide antecedents in discourse. In fact, negative sentences systematically make available two propositional antecedents that can be picked up by anaphora (Stone and Hardt (1999); Krifka (2013); Snider (2017)). One corresponding to the whole sentence, illustrated in (18-a), and another corresponding to the abstract negated content (18-b).

(18) [Mary didn't [~~Mary~~ sleep]$^{\phi_1}$]$^{\phi_2}$

a. That$_{\phi_2}$ is very surprising. (*that* $\approx$ that Mary didn't sleep)

b. Even though she claimed that$_{\phi_1}$. (*that* $\approx$ that Mary slept)

The contrast in (18) suggests that constraints on the interpretation of propositional anaphora are similar to the constraints on the interpretation of anaphora to individual entities. In the case of propositional anaphora, the epistemic status of the proposition introducing the discourse referent governs whether it is an accessible antecedent. If the propositional anaphor *that* appears in a veridical context (18-a), its antecedent can only be the proposition *Mary didn't sleep*, i.e., the one to which the speaker is committed. If the propositional anaphor *that* appears in an anti-veridical context (18-b), its antecedent will be the proposition *Mary slept*, i.e., the one that the speaker is committed to being false. This antecedent is not available in (18-a), but only if the propositional pronoun *that* is interpreted in the anti-veridical intensional context. This allows for the antecedent proposition *Mary slept* to be interpreted as true wrt the local intensional context of the anaphor, i.e., the set of worlds corresponding to Mary's claims. In this way, cases like (18-b) are analogous to the modal subordination cases above. The fact that negative sentences systematically make available these two kinds of propositional discourse referents is further evidence against a view of negation as externally static.

If we limit our consideration to propositional anaphora in veridical contexts, for now, we can construct examples that parallel the above ones that Karttunen gave for individual anaphora. The above cases, in which there are propositions in the scope of negation that are in a veridical context due to presupposition projection or double negation, also license propositional anaphora:

(19) a. *Bill didn't realize [that he had a dime]*$^{\phi_1}$.

It$_{\phi_1}$ *was a surprise.* (*it* ≈ that Bill had a dime)

b. *It's not true that there isn't [a bathroom in this house]*$^{\phi_2}$.

That$_{\phi_2}$ *was to be expected.*

(*that* ≈ that there is a bathroom in this house)

The above cases—without negation but with non-veridical intensional propositional embeddings—which do not license individual anaphora to elements in their scope, exhibit the same limitation for propositional anaphora.

We can state a generalization for propositional anaphora that parallels that for individual ones:

(20) Generalization for propositional anaphora (preliminary)

A proposition-denoting expression establishes a discourse referent just in case the speaker is committed to the truth of the expressed proposition.

Explanatory Generalizations

As suggested above, negation is not the only operator that classic dynamic theories treat as externally static. Likewise, disjunction and conditionals—other non-veridical operators—are treated as dynamically inert in the same way. This choice is based on the observation non-veridical propositional operators, in simple cases, preclude subsequent pronouns in indicative contexts.

It also fits with the broader generalization that the question of whether the speaker is committed to an embedded proposition constrains anaphora. Since non-veridical propositional operators do not entail their prejacent, overall, this seems to be the right generalization.

It is further supported by the fact that this successfully rules out some other cases where the speaker is not committed to the truth of the local propositional context of potential anaphoric antecedents, like the cases of indefinites under modals and intensional verbs given in (21).

(21) a. *Bill can make [a kite]$^{\upsilon_1}$.#It$_{\upsilon_1}$ has a long string.*

[after Karttunen (1976): 10b]

b. *John wants to catch [a fish]*$^{\upsilon_2}$. *#Do you see it*$_{\upsilon_2}$ *from here?*

[after Karttunen (1976): 11a]

c. *John failed to find [an answer]*$^{\upsilon_3}$*.#It*$_{\upsilon_3}$ *was wrong.*

[Karttunen (1976): 14b]

d. *I doubt that Mary has [a car]*$^{\upsilon_4}$*.#Bill has seen it*$_{\upsilon_4}$.

[Karttunen (1976): 17]

None of the antecedent sentences in (21) commit the speaker to the truth of the most local proposition of the indefinite DP or the existence of a referent, so anaphoric reference fails. Again, we can construct parallel examples for propositional anaphora:

(22) a. *I doubt that Mary has a car. Bill knows that.*

(*that* $\approx$ that I doubt that Mary has a car)

(*that* $\not\approx$ that Mary has a car)

b. *It's a lie that Bill had a dime. That was a surprise.*

(*that* $\approx$ that it's a lie that Bill had a dime)

(*that* $\not\approx$ that Bill had a dime)

The first sentences in (22), involve non-veridical propositional embeddings. The propositional anaphora in the second sentences are in veridical contexts and can only refer to a previous veridical proposition expressed by the matrix sentence. They cannot refer to the embedded non-veridical propositions.

Classic dynamic semantic theories treat these patterns as lexical idiosyncrasies associated with the semantics of propositional operators. The generalization developed here suggests that the truth-functional properties of a propositional operator offer an explanation of its contribution to the anaphoric potential of an utterance. We might provide a more explanatory theory of anaphoric accessibility by relating it to propositional operators' independently motivated relational meaning component in an intensional semantics.

Conclusion: Commitment and Veridicality Constrains Anaphora

Crucially, Karttunen suggests that anaphora are constrained by the veridicality of their embedding context: Whether or not a potential antecedent is available does not only depend on idiosyncratic properties of embedding operators. More generally, it needs to take into account the epistemic status of its local context—we need to know whether it corresponds to a proposition that is considered true or false.

> *In order to decide whether or not a non-specific indefinite [...] is to be associated with a referent, a text interpreting device must be able to assign a truth value to the proposition represented by the sentence in which the [indefinite] appears.*
>
> — Karttunen (1976) (p. 14)

I adopt from Karttunen's view that the possibility of anaphoric relation is conditioned on the truth of the local proposition (here, conceptualized as the truth relative to the speaker's commitments). I depart from Karttunen by arguing against the hypothesis that the contrasts in question can be explained by asking whether or not a dref is introduced. I argue that anaphoric accessibility is not (only) determined by the properties of the antecedent but crucially constrained by the interpretation of the anaphor. The generalizations stated in this subsection hold that a potential antecedent introduces a dref, just in case that dref is veridical. The main argument for taking into account the properties of anaphora as well when thinking about accessibility comes from anaphora to non-veridical antecedents. This argument is presented in the following subsection.

2.3.3 A Lookahead Problem

In classic dynamic semantics, the distinction is that veridical operators allow the material in their scope to introduce new drefs, while non-veridical operators do not. The previous subsection noted that this generalization is not used to provide an explanation of the dynamic behavior of the operators. However, these frameworks do account for the fact that the properties of non-veridical operators constrain anaphora—by explicitly specifying their dynamic behavior on the level of the idiosyncratic lexical semantics of propositional operators.

Besides the question of explanatory adequacy discussed above, the classic way of constraining accessibility causes a problem of descriptive adequacy. This kind of problem, familiar from the literature on modal subordination, is tied to the fact that the embedding context of the antecedent ultimately determines anaphoric accessibility in these accounts. However, the empirical picture suggests that the embedding context of the anaphor should be taken into account as well. Therefore, accounts where anaphoric accessibility is fully determined by the anaphoric potential of the utterance containing the antecedent, run into a lookahead-problem.

This subsection illustrates and discusses this problem in relation to anaphora to negated content and suggests how taking into account a generalization based on commitment and veridicality could address the problem. This is done by discussing evidence and arguments for the following position:

The constraints that preclude anaphoric reference to negated content do make reference to the kind of embedding context in which a (potential) antecedent appears. However, in the dynamically incremental interpretation of a discourse, the constraints are enforced as part of the interpretation of the anaphor. For individual anaphora, in particular, they can be derived from the assumption that a pronoun presupposes the existence of its referent entity in its local context (Stone and Hardt (1999); Brasoveanu (2006)). For propositional anaphora, the interpretation of a pronoun requires that its referent proposition is true in its local context (see Kroll (2019) for a similar argument for the licensing of sluicing). The claim that the constraints governing anaphoric accessibility are placed on the interpretation of anaphora, rather than antecedents, is supported by data that show that the constraints can be accommodated or satisfied in the local context of the anaphor, as is the case with modal subordination.

Accommodation

The above generalization (16) states that individual anaphora requires that the speaker is committed to the existence of a referent. Similarly, (20) states that propositional anaphora requires that the speaker is committed to the truth of a referent. This requirement may be accommodated in cases where there is an overt antecedent, and the antecedent and the anaphor are interpreted relative to distinct epistemic representations.

Accommodation with Individual Anaphora. This can happen under third-person attitude ascriptions:

(23) *Bill doubts that Mary has [a car]$^{\upsilon_1}$. I have seen it$_{\upsilon_1}$.* [Karttunen (1976): 18]

In (23), the sentence containing the antecedent is about the epistemic state of a third person (*Bill*) and does not say anything about the epistemic status of the proposition *Mary has a car*, in relation to the speaker. Only when the speaker utters *I have seen it* are they committed to the proposition that *Mary has a car,* by way of accommodation. A similar case arises with inter-speaker disagreement:

(2-d)Disagreement:

A: *There isn't* [*a bathroom*]$^{\upsilon_4}$ *in this house.*

B: *(What are you talking about?) It*$_{\upsilon_4}$ *is right over there.*

In (2-d) the antecedent sentence has nothing to say about the epistemic status of the referent of *a bathroom* wrt speaker B, in this case because it is an assertion by speaker A. However, speaker B commits to the truth of the proposition *there is a bathroom in this house*. They refer to the bathroom as if they believe in its existence—using a pronoun in an indicative context. This suggests that a DP can be an accessible antecedent for an anaphor, as long as the speaker (using the anaphor) commits to the existence of its referent. That is due to the requirement that the pronoun has an existing referent. Another consequence is that the speaker is committed to the truth of the most local proposition to the antecedent. (In this particular case, that proposition is an existential statement, and therefore the same proposition *there is a bathroom in this house*.)

Accommodation with Propositional Anaphora. Again, we can construct parallel examples with propositional anaphora:

(24) a. *Bill doubts that [Mary has a car]*$^{\phi_1}$*, but I know it*$_{\phi_1}$ *for sure.*

b. A: *There isn't [a bathroom in this house]*$^{\phi_2}$.

B: *(What are you talking about?) I'm certain of it*$_{\phi_2}$.

Similarly to (23), the first clause in (24-a) does not commit the speaker to the truth of the proposition that *Mary has a car.* Still, it can be an antecedent for *that*. The requirement that this antecedent is true in the context of the pronoun can be accommodated. We see a similar parallel for inter-speaker disagreement in (24-b). Although the assertion of speaker A does not commit B to anything, B can reference the proposition *there is a bathroom in this house* as if they were committed to its truth. Again, this looks like an instance of accommodation.

Upshot. The requirement that the speaker is committed to the existence/truth of the antecedent can be accommodated when the anaphoric pronoun is used. This suggests that the requirement does not, in fact, contribute to the interpretation of the antecedent but rather to the interpretation of the anaphor.

Local Satisfaction

More evidence for the view that the relevant restrictions are invoked in the interpretation of anaphora rather than antecedents is provided by cases where anaphora are evaluated in non-veridical contexts. In these cases, even indefinites in non-factual embeddings are possible antecedents. I suggest that this is the case because the requirement on the existence/truth of the referent is not enforced on a global level but instead satisfied in the local context of the anaphor.

Local Satisfaction in Conditionals. If a conditional antecedent introduces a potential anaphoric antecedent, the latter can be picked up in the conditional consequent without the speaker being committed to its existence/truth.

(25) Conditionals:

If [there is [a bathroom]$^{\upsilon}$ in this house]$^{\phi}$,

a. *it*$_{\upsilon}$ *is hard to find.*

b. *that*$_{\phi}$ *is not surprising.*

Modal Subordination with Individual Anaphora. Another case is constituted by modal subordination (Roberts (1989)), where the modal *would* introduces a hypothetical set of worlds and supplies it as a local context of evaluation for its prejacent proposition, which includes the pronoun *it* (see also Stone (1999); Brasoveanu (2010a)).

(9-c) Modal subordination:

There isn't [a bathroom]$^{\upsilon_3}$ in this house. It$_{\upsilon_3}$ would be easier to find.

The sentence containing the anaphor is interpreted so that in the worlds, where *there is a bathroom in this house*, that bathroom would be easier to find. The antecedent is in a non-veridical context. As a result, the speaker is not committed to the truth of the proposition that there is a bathroom in this house or the existence of said bathroom. Nevertheless, the anaphor, which itself is in a non-veridical context, can anaphorically pick up the hypothetical bathroom. Because the pronoun is interpreted relative to a local non-veridical context, it can pick up a non-veridical discourse referent.

The condition that the antecedent's referent exists is not satisfied in the worlds taken to be true by the speaker. However, it is satisfied locally, in the intensional context of the pronoun. A similar situation arises if both the antecedent and the pronoun are embedded under third-person ascribed speech reports:

(26) *Bill says he saw a lion on the street. He claims it had escaped from the Zoo.*

[after Karttunen (1976): 19]

In (26), the pronoun *it* is interpreted in relation to the set of worlds compatible with Bill's claims. The indefinite *a lion* was introduced in a proposition ascribed to Bill's epistemic state and can therefore be an antecedent for *it* although the speaker of the utterance is not committed to its existence. We saw that an indefinite in a non-factual context could not normally be the antecedent for an anaphor because the existence-requirement of the anaphor would not be satisfied. However, an anaphoric relation can be established if the pronoun is itself interpreted wrt a proposition to which the speaker is not committed.

Parallel Patterns with Propositional Anaphora. For these kinds of cases in which the veridicality-based conditions on anaphora are evaluated locally, we can also construct parallel examples for propositional anaphora:

(27) a. *There isn't [a bathroom in this house]*$^{\phi_1}$*, although that*$_{\phi_1}$ *would be ideal.*

(that ≈ if there were a bathroom in the house)

b. *Bill says [he saw a lion on the street]*$^{\phi_2}$*, but I don't believe that*$_{\phi_2}$.

(that ≈ that Bill saw a lion on the street)

Like in the individual anaphora cases, the antecedents in (27) are non-veridical: In (a.), this is the prejacent of negation (the proposition that there is a bathroom in this house), and in (b.) the complement of *says* (that Bill saw a lion on the street). In both cases, a propositional anaphor refers to these propositions without committing the speaker to their truth. Instead, both anaphoric pronouns are embedded under non-veridical operators, and the requirement that the antecedent proposition is true is evaluated wrt a local non-veridical proposition.

Upshot. The examples in this subsection suggest that the veridicality-requirement on anaphora can be satisfied in their local context. This shows that the local intensional context of the anaphor constrains anaphoric resolution in addition to the local intensional context of the antecedent. An account of the constraints on anaphoric co-reference needs to consider the anaphor's local context. We update our parallel generalizations for individual and propositional anaphora as in (28):

(28) **Generalization about anaphoric accessibility (Final version)**

a. An (individual-denoting) expression is a possible antecedent for an individual anaphor just in case the referent of the antecedent exists in the local context of the anaphor.

b. A (proposition-denoting) expression is a possible antecedent for a propositional anaphor just in case the proposition expressed by the antecedent is true in the local context of the anaphor.

The idea that anaphoric accessibility is determined by which propositions are true in the local context of the anaphor cannot be reconciled with the idea that a given expression can sometimes introduce a dref and sometimes does not—depending on the operators taking scope over it. The constraints on anaphoric co-reference should be captured not as a condition on introducing discourse referents but as a condition on the circumstances under which expression associated with a dref can be an antecedent for a given pronoun.

Conclusion: The Interpretation of the Anaphor Constrains Anaphora

If the constraints on anaphora are enforced as part of the interpretation of the pronoun, it could not be fully characterized by a theory that attempts to predict when a discourse referent is introduced (globally) or not. Instead, any potential antecedent needs to introduce a dref and provide a way of accessing the information to which the conditions on using an anaphor make reference. This point is supported by the fact that the speaker not being committed to the existence of a referent does not make it inaccessible for anaphora in all cases because the requirement that the referent of the antecedent exists can be accommodated or satisfied locally.

2.4 Theoretical Choices

Some previous analyses of the cases that are problematic for the classic dynamic semantic theories treat them as exceptions. Roberts (1989) gives an analysis of modal subordination based on anaphoric modals allowing access to otherwise inaccessible parts of the discourse, which is particular to the semantics of anaphoric modals like *would*. Krahmer and Muskens (1995) give an analysis of the double negation and disjunction cases based on the conventional semantics associated with negation and disjunction operators. Here, indefinites under negation may license subsequent anaphora in some cases: They introduce a dref only under certain constellations of embedding operators (most notably under double negation).

These accounts maintain a view under which negation, or non-veridical operators more broadly, are treated as externally static operators, and cases of anaphoric reference to negated content are exceptional. This is in contrast to views of dynamic semantics that involve flat discourse update, under which all discourse referents are introduced globally but store information about the semantics of their embedding context (Stone (1999); Stone and Hardt (1999); Brasoveanu (2006, 2010a,b)). Here, I argue that the way that anaphoric dependencies interact with negation provides arguments in favor of a flat-update dynamic semantics, under which negation is treated as externally dynamic.

Further, constraints on anaphoric dependencies can be tied to semantic and pragmatic properties of discourse referents. Specifically, the constraints can be captured in terms of speaker commitments: They rely on information about the anaphoric pronouns and their antecedents and the epistemic status of both, relative to the speaker. That is, they rely on the information about whether the pronoun and antecedent refer to something that is part of a speaker's publically assumed reality, something that the speaker considers to be contrary to fact, or something that they have no such commitments about.

2.4.1 Comparing Three Kinds of Approaches

Comparing different possible approaches to anaphora to negated content, we consider analyses of this phenomenon and related phenomena (like modal subordination or complement anaphora) put forward in the literature. The authors introducing the analyses for these related phenomena did not necessarily intend for their work to apply to anaphora to negated content. Here, we compare the various approaches as analyses of anaphora to negated content, and the presented arguments are to be understood as comments on their merit for this particular empirical phenomenon.

Possible (Kinds of) Analyses of Anaphora to Negated Content

The possible approaches to anaphora to negated content compared here can roughly be grouped into three categories:

1. **Accommodation/Bridging/Reanalysis:**

   Following a classic conception under which indefinites in the scope of non-veridical operators never introduce a dref on a global level, these accounts allow for anaphora in certain contexts to access antecedents from otherwise inaccessible parts of the discourse. Accounts of this kind may involve some inferential mechanism that is likened to accommodation or bridging as well as some additional constraints on this mechanism (e.g. Roberts (1989); Nouwen (2003); Lewis (2021)) or retrieve the inaccessible part of the previous context anaphorically (Geurts (1999); Frank and Kamp (1997)), and they are typically developed to account for modal subordination, although the accounts in Nouwen (2003); Lewis (2021) are counterexamples to that.

2. **Conditional Variable Update:**

   Instead of following a classic conception under which indefinites in the scope of non-veridical operators never introduce a dref on a global level, these accounts devise a mechanism that allows an indefinite under negation to introduce a dref only under certain conditions. Accounts of the type are typically developed to account for double-negation (9-a) and bathroom-disjunction (9-b) cases (e.g. Krahmer and Muskens (1995); Gotham (2019); Mandelkern (2022)).

3. **Flat Update Dynamic Semantics with Intensional Drefs:**

   All operators are externally dynamic, but discourse referents store information about their embedding context (Stone (1999); Stone and Hardt (1999); Brasoveanu (2006, 2010a,b)). The constraints that classic dynamic theories capture by appealing to eternally static operators are captured by making reference to relations between discourse referents. These accounts were also developed to account for modal subordination, with the goal of capturing generalizations across different domains.

The individual accounts mentioned here will be discussed and contrasted in some more detail in **Chapter 7**, where a formal account of individual anaphora to indefinites under negation is introduced. In this chapter, we are interested in comparing these classes of accounts on a broader level.

Discussion and Comparison

In the above **Section 2.3**, we have discussed two problems that classic dynamic semantic analyses run into:

1. **Lookahead Problem:**

   In many cases, the availability of an anaphoric relation is not only influenced by the properties of the embedding context of the antecedent but also the embedding context of the anaphor. In a dynamically incremental interpretation of the discourse, the information about upcoming utterances is not available when interpreting a potential antecedent. Therefore, accounts that condition the introduction of drefs on properties of the embedding context run into a lookahead-problem.

2. **Explanatory Adequacy:**

   Negation falls into a larger class of non-/anti-veridical propositional operators. Deriving the referential accessibility of material in its scope by its intensional properties would be more explanatory than hard-coding it into the lexical semantics of operators.

Conditional variable update accounts based on properties of the antecedent (Krahmer and Muskens (1995), and to some extent, Gotham (2019), and Mandelkern (2022), see discussion in **Section 7.5**) also run into the lookahead-problem: Even if these accounts capture the fact that a doubly negative context allows for anaphora to its scope and possibly capture generalizations about veridicality, they still condition the introduction of a dref conditionally on the properties of the antecedent. As such, they also run into the lookahead problem and could not account for cases of discourse subordination.

Accommodation-accounts and accounts which devise mechanisms to make otherwise inaccessible parts of the discourse accessible under certain conditions (Roberts, 1989; Geurts, 1999; Frank and Kamp, 1997; Nouwen, 2003; Lewis, 2021) also run into the problems of generality and explanatory adequacy:

Often, the property of certain operators making parts of the discourse inaccessible is still hard-coded into the lexical semantics of propositional operators. Doing this does not allow for an explanatory account of how the intensional and epistemic properties of the local context and veridicality of embedding operators affect the anaphoric potential of the antecedent. As such, many accommodation analyses could not account for double negation cases (ones that do, like Lewis's (2021) dynamic pragmatic account an addition make reference to some notion of speaker commitment).

In fact, we might imagine an account where the effect of externally static operators could be derived from their intensional properties, that is, any content that the speaker is not committed to cannot contribute new discourse referents. Then, an accommodation-account would need to assume that accommodation is possible, only if it wouldn't cause a contradiction (which, of course, would be a very reasonable and well-motivated constraint).

However, additional constraints would be needed, or at least the question of what can be accommodated would have to be carefully addressed. In particular, accommodation of an antecedent that hasn't been introduced into the discourse generally doesn't (usually[2]) seem to be possible, as suggested by the famous contrast in (29).

(29) Roberts (1987), attributed to Barbara Partee:

a. *I dropped ten marbles and found all but one of them. It's probably under the couch.*

[2]See also the discussion in Nouwen (2003), who draws a parallel between complement set anaphora to negative quantifiers and occasional uses of pronominal bridging. In bridging, however, while the anaphoric referent is not introduced explicitly, it will be in a certain relation to its antecedent.

b. *I dropped ten marbles and found nine of them. # It's probably under the couch.*

While a pronoun (*it*) can be used to refer to the missing marble when it has been mentioned by an explicit DP (*one of them*) (29-a), the same is not possible when the existence of such marble is an implicit inference of the previous discourse (29-b).

Propositional anaphora can be used to illustrate the same point: The propositional anaphor (*that*), used in a counterfactual context in (30-a), may pick up the proposition ϕ, which has been introduced explicitly under negation. However, in (30-b), where the same proposition is an implicit inference, the same is not possible for the pronoun (*that*).

(30) a. *Peter didn't [~~Peter~~ pass the test]*$^{\phi}$. *That*$_{\phi}$ *would have been a surprise.*

b. *# Peter failed the test. That would have been a surprise.*

Now, if an accommodation-account were to assume that anaphora to non-veridical content do require an anaphoric link, but the inferences about their existence which enable this anaphoric link may be accommodated, the difference between an accommodation-basd account and the flat-update account presented here are not actually that big any more.

In an accommodation-world, inaccessible parts of the discourse can be accessed under certain conditions, and in a flat-update-world, no part of the discourse is inaccessible, but any discourse referent has some conditions under which it can be accessed. It seems to me like this difference could not be addressed based on acceptability or truth-value judgment data alone.

As far as I am aware, there is no definitive empirical evidence against the entire class of accommodation-based accounts to anaphora to negated/non-veridical content. For a strongly convincing empirical argument for flat update or accommodation as a suitable model of the parts of linguistic competence that have to do with the interaction of anaphora and negation (or more broadly: veridicality), we may want to look for other psychological, behavioral, or neurophysiological evidence for cognitive representations or processes[3].

The arguments against accommodation-based accounts in the literature which I am aware of are either made **(i)** against specific implementations rather than the whole class of accounts, or **(ii)** on a conceptual or theoretical level. Our reasons for choosing a flat-update account over an accomodation-based one are based on arguments that intensional flat-update accounts are more explanatory than accommodation-based accounts in two ways: First, accommodation-accounts may not be well suited to address the regularity and generality in which negation and veridicality interact, and second, that an flat-update account presented here allows for a more parsimonious explanation of anaphoric accessibility in terms of two very basic constraints on discourse interpretation.

Arguments that accommodation-based analyses may not account for the regularity of anaphora in discourse subordination (and the constraints on it) are noted by Stone (1999), who suggests that the idea of exceptional accommodation as a last resort, as

[3]See Dwivedi, Phillips, Laguë-Beauvais, and Baum (2006); Dwivedi, Drury, Molnar, Phillips, Baum, and Steinhauer (2010) for some initial research on the neurophysiological response to the anomaly caused by veridical anaphora to non-veridical antecedents.

originally intended by Lewis (1979), does not capture the regular nature of anaphora in modal subordination.

To support this, Stone notes another aspect of why these accommodation accounts for modal subordination may not be general enough: The anaphoric parallels between modality and tense, illustrated in the following contrast:

(31) Stone (1999):

a. *John ate a cheesesteak. It {was / # is} very greasy.*

b. *John might be eating a cheesesteak. It {would be / # is} very greasy.*

Stone suggests that nobody would argue that past tense is an externally static operator. Nevertheless, he argues that the constraints in the two domains—a present tense pronoun being incompatible with a past tense antecedent, and an indicative pronoun being incompatible with a modal antecedent—should be analyzed analogously (for more arguments for this independently motivated claim see Stone (1997)).

Further, the intensional flat-update account presented here allows us to reduce the notion of anaphoric accessibility to the independtly motivated general mechanims that **(i)** any DP referent exists in its own local context (/any propositional anaphoric referent is true in its local context), and **(ii)** that interpretable discourses contain consistent, non-contradictory, information. As soon as we assuming an appropriate intensional dynamic representation, these two basic principles can explain why normally, non-veridical operators preclude anaphoric reference, and also allow us to understand the exceptions to this generalization.

Accommodation accounts on the other hand require an explicit theoretical operationalization of (non-)accessibility, and of an accommodation mechanism (and that is in addition to assuming the independently motivated principles **(i+ii)** above).

Therefore, I take intensional flat-update accounts to be more general and more explanatory. In the absence of empirical evidence to decide between them, this is a reason to prefer a flat-update account.

Conclusion: Flat-update Dynamic Semantics with Intensional Drefs

Based on the results about veridicality generalizations and the lookahead-problem, we conclude that the contextual representation of discourse referents has to store the information about their intensional content, i.e., the set of worlds associated with their most local proposition and its epistemic status, i.e., whether the speaker is committed to its truth, falsity, or neither. In other words: a discourse referent needs to store the information about the set of worlds in which they exist and the relation between that set and the set of discourse commitments of the speaker.

2.4.2 Epistemic Representations of Drefs

Flat-Update Dynamic System with Epistemic Drefs

As noted above, this work adopts conceptions of discourse referents as representations which can store information about their embedding context in a flat-update dynamic system (following Stone (1999); Brasoveanu (2006)). In particular, discourse referents are understood as intensional representations which map discourse variables to referents relative to the possible worlds of their local intensional context. More specifically, I adapt the accounts in Stone (1999); Brasoveanu (2006) which have been developed to account for modal subordination (/discourse subordination more broadly in the case of Brasoveanu (2006)), and extend them to new empirical cases of anaphora to negated content.

Above, I contributed additional arguments for a flat-update dynamic semantics (**Section 2.4.1**). Theoretically, I also contribute the point that anaphora to negated content shows that discourse referents need to be understood pragmatically: In addition to the intensional representation of drefs suggested by Stone and Brasoveanu, the discourse needs to store epistemic information about them and keep track of their status in relation to the speaker's commitments. In summary, I suggest that the interactions between negation and anaphora support the view that the anaphoric interpretation of discourse variables relies on information about:

1. **Dynamic Semantics and Discourse Referents:** (Karttunen, 1976; Kamp, 1981; Heim, 1982, 1983; Groenendijk and Stokhof, 1991; Kamp and Reyle, 1993)

   Dynamically updated dref mappings from discourse variables to referents

2. **Intensional Drefs:** (Stone, 1999; Stone and Hardt, 1999; Brasoveanu, 2006, 2010a,b, see also the last chapter of Dekker (1993))

   The relationship between these mappings and the sets of worlds in which the referent exists; and

3. **Epistemic Drefs:** (Developed here)

   The attitude of the speaker towards these sets of worlds.

The generalization developed in **Section 2.3** is based on an argument which suggest that in order to understand the constraints placed on anaphoric relations, we need to consider the semantics of intensional operators, the local propositional contexts they introduce, and the relation of these contexts to the commitments to a speaker. We need to know whether the speaker is committed to the propositions corresponding to these local contexts.

The relativization to a speaker is crucial for distinguishing between veridically and non-(/anti-)veridically introduced propositions, since the speaker will be committed to the former kind, but not the latter. Further, especially the inter-speaker disagreement cases show that show that the epistemic status of a dref should be evaluated in relation to the commitments of the (current) speaker.

The notion of veridicality, as introduced in the literature (see e.g. Zwarts, 1995; Giannakidou, 1998) and referenced above, is a property of propositional operators. Here, we will formally introduce this notion, but also extend the notion beyond its established use. We do this to characterize the epistemic status of propositions and entities in relation to the commitments of a speaker, as well as their representations in discourse. Accordingly, we speak of veridical individuals/propositions as ones that the speaker is committed to being true/existing. We speak of a veridical dref as one that refers to a veridical individual or proposition.

Veridicality as an Epistemic Notion

Formal Definition. We formalize the notion of epistemic status of embedded propositions as the relationship between local intensional contexts and the commitments of a speaker. In natural language, this information is reference by the notion of veridicality.

VERIDICALITY characterizes the relationship between (embedded) propositions and the commitments of a speaker. An embedding context that brings about speaker commitment to its prejacent is a veridical embedding. Here, we formally introduce this notion and extend it to discourse referents.

Formalizing the embedding context as a propositional operator, veridicality can be defined as follows (Zwarts (1995); Giannakidou (1998)):

(32) Definition 1: [Giannakidou (1998): 20]

Let Op be a monadic propositional operator. Then:

a. Op is *veridical* iff $\mathrm{Op}(p) \rightarrow p$ is logically valid. Otherwise Op is non-veridical.

b. Op is *anti-veridical* iff $\mathrm{Op}(p) \rightarrow \neg p$ is logically valid.

(33) Definition 2: [After Giannakidou (1998): 21]

Let Op be a dyadic propositional operator. Then:

a. Op is *veridical* wrt p [q] iff $\mathrm{Op}(p, q) \rightarrow p$ [$\mathrm{Op}(p, q) \rightarrow q$] is logically valid. Otherwise Op is non-veridical wrt p [q].

b. Op is *anti-veridical* wrt p [q] iff $\mathrm{Op}(p, q) \rightarrow \neg p$ [$\mathrm{Op}(p, q) \rightarrow \neg q$] is logically valid.

A propositional operator is veridical with respect to its argument(s), iff its use entails the argument(s). Based on this, we can characterize embeddings under *I know*, *Sue believes*, and *It's not the case* in the following way:

(34) a. Veridical embedding:

[*I know* [*that Mary has a car*]]. ($\rightarrow$ *Mary has a car*)

b. Non-veridical embedding:

[*Sue believes* [*that Mary has a car*]]. ($\not\rightarrow$ *Mary has a car*)

c. Anti-veridical embedding:

[*It's not the case* [*that Mary has a car*]]. ($\rightarrow \neg$*Mary has a car*)

Epistemic Veridicality Relative to the Speaker. The notion of truth or veridicality as referenced in natural language is relative to an epistemic agent's belief state or discourse commitments. This has been argued for mood (Farkas (1992)), negative polarity items (Giannakidou (1998)), and the same is the case for the interpretation of anaphora. In Karttunen's words:

> *[The text-interpreting device] must be sensitive to the semantic properties of verbs that take sentential complements, distinguish between assertion, implication, and presupposition, and finally, it must distinguish what exists for the speaker from what exists only for somebody else.*
>
> — Karttunen (1976) (p. 14)

The relevant distinction for our purposes is about whether the speaker is or is not committed to an embedded proposition being true or false. This notion goes beyond the above definition of veridicality, which is part of the conventional semantics associated with propositional operators.

de Marneffe, Manning, and Potts (2012) show that pragmatic factors influence the assessment of speaker commitment to an embedded proposition by language users, in addition to lexical information (see also related work on contextual influences on facticity assessment by Tonhauser (2016); Tonhauser, Beaver, and Degen (2018); Degen and Tonhauser (2021)). Assumptions about the epistemic authority or trustworthiness of the source of a speech report, prosodic cues, or prior beliefs can affect the assessment of speaker commitment. For example, an utterance of (35) may lead to the inference that Mary does have a car or to the inference that she doesn't, depending on whether a comprehender considers Mary a trustworthy source or a notorious liar.

(35) *Mary said that she has a car.*

A third-person speech report like (35) is a semantically non-veridical embedding (the proposition expressed by *Mary said that she has a car* does not ential that Mary has a car). However, the assessment of speaker commitment to the embedded proposition can be influenced by contextual factors as well.

Therefore, we will use the term 'veridicality'slightly differently from previous uses, which describe a property of propositional operators as defined above. Here, we use it to characterize the epistemic status of embedded propositions, which is influenced by lexical veridicality as defined above, but also by contextual factors of the respective utterance.

Therefore, we speak of the veridicality of the use of an operator in an utterance, and of the intensional context of the embedded proposition:

(36) Definition 3:

Let Op be a monadic propositional operator, and U be an utterance by a speaker S of the proposiion $\mathrm{Op}(\mathrm{p})$. Then:

a. Op introduces a *veridical* embedding intensional context in U iff

•S uses U to conversationally present themself as believing $\mathrm{Op}(\mathrm{p})$, and

•S uses U to conversationally present themself as believing p.

Otherwise Op introduces a non-veridical embedding in U.

b. Op introduces an *anti-veridical* embedding intensional context in U iff

•S uses U to conversationally present themself as believing $\mathrm{Op}(\mathrm{p})$, and

•S uses U to conversationally present themself as believing $\neg\mathrm{p}$.

According to (36), a veridical embedding context is one s.t. by committing themself to the full proposition, the speaker also commits themself to an embedded proposition.

Applying the Notion to Drefs. We want to use that notion to describe the epistemic status of discourse referents, propositional and individual ones. We can say that a dref inherits its veridicality from the context that it was introduced in. Just in case a speaker is committed to the local context of the introducing expression, they are also committed to that dref being veridical (i.e. existent/true in their commitment set). This is related to Stone's (1999) notion of a hypothetical scenario dref.

Based on these considerations about veridicality, I am adopting the following terminology to talk about the epistemic status of propositional discourse referents (relative to a speaker):

(37) Veridicality for drefs:

a. Veridical drefs:

(i) A propositional dref is veridical wrt a speaker S iff it refers to a proposition that S is committed to (and non-veridical otherwise).

(ii) An individual dref is veridical wrt S iff the S is committed to the existence of its referent (and non-veridical otherwise).

b. Anti-veridical drefs:

(i) A propositional dref is anti-veridical wrt a speaker S iff it refers to a proposition that S is committed to being false.

(ii) An individual dref is veridical wrt S iff the S is committed to the non-existence of its referent.

2.5 Conclusion

Anaphora to content introduced under operators which have been characterized as externally static are possible. Here, we focus on anaphora to indefinites in the scope of negation, while understanding them in the broader context of non-veridical propositional operators. Anaphora to negated content may happen: **(i)** When the antecedent is actually in a veridical context (like e.g. in double-negation cases), or **(ii)** when the anaphor is also in a non-veridical context (like e.g. in modal subordination).

The systematic nature of how anaphora and veridicality interact suggests that negation is just a piece of a bigger picture, and that the counterexamples to the original generalization should receive a unified account that treats them as regular cases of anaphora. Here, these points are interpreted as arguments for a flat-update dynamic semantics with intensional and epistemic representations of discourse referents.

Developing this kind of account will allow us to extend analytical insights from Stone's (1999) and Brasoveanu (2006) accounts of modal subordination to anaphora to negated content, give a unified account of phenomena which have received disparate analyses in the literature: Such as anaphora to unspecific indefinites under negation in the cases of double-negation and bathroom sentences, casses of modal subordination with negative antecedents, as well as negativity-tags.

Chapter 3

Intensional Compositional DRT

3.1 Introduction

The formal semantic analyses in this dissertation are framed in Intensional Compositional DRT, an intensionalized version of the dynamic Logic introduced as the Logic of Change in Muskens (1996). I am closely following Brasoveanu's (2010a) implementation of this system, except that I will be contextually manipulating singleton information states, i.e., singleton variable assignments, as opposed to the plural information states (formalized as sets of variable assignments) employed by Brasoveanu. This chapter introduces a basic version of this framework, which lays the grounds necessary to account for the constraints on anaphoric reference laid out in the previous chapters.

The framework presented here has the following properties: It is dynamic, including a representation of discourse referents to which subsequent anaphora may refer. It is intensional, providing a representation of the local propositional context provided by sentential operators. It integrates the pragmatic information of an epistemic representation of the speaker's discourse commitments into an integral meaning representation associated with an utterance.

As argued in **Chapter 2**, these properties are necessary components of a semantic account of anaphora and constraints on anaphoric relations. Additionally, the system introduced here is compositional, allowing us to localize the meaning contribution of single expressions and understand how it contributes to a semantic representation of entire utterances.

This chapter lays out how the system can be used as a framework for interpreting natural language utterances and illustrates this with some simple examples involving propositional drefs for asserted propositions and some simple cases of individual anaphora. This system will be the basis for the analyses of negation and anaphora.

A Dynamic Intensional Semantics

Muskens's (1996) Compositional DRT (CDRT) characterizes natural language meaning both truth-conditionally and dynamically. Natural language meaning is characterized in terms of relations between propositions and relations between utterances in discourse. The contribution of an utterance is formalized dynamically, in terms of its context change potential (CCP, Heim (1982, 1983)), i.e., by characterizing the information that it may add into a representation of the discourse. The meaning of an utterance dynamically updates the current state of the discourse context. Utterances are interpreted with respect to an input context and yield an output context after making the changes required by the update.

The discourse state is conceived of as a referential context that keeps track of representations of discourse referents (Kamp (1981); Heim (1982, 1983); Groenendijk and Stokhof (1991)). This referential context is modeled as a variable assignment, which maps discourse variables to their static referent in a model. A discourse referent is a mapping from a specific discourse variable name to its referent. The interpretation of referential expressions may update the referential context to include new discourse referents (drefs) that allow for anaphora in subsequent discourse (Karttunen (1976)).

In addition to updates to the referential context, an utterance updates an epistemic context representing the interlocutors' beliefs. This epistemic context is modeled as a set of worlds that will be narrowed down by intersection with the sets of worlds associated with propositions uttered in discourse, like the common-ground-based context set from Stalnaker (1978, 2002). Following Gunlogson (2004), I will assume that there is not one epistemic context for the whole conversation but several epistemic contexts relative to the interlocutors. For each interlocutor, this will be the set of worlds compatible with their public discourse commitments.

While dynamic frameworks have been employed fruitfully using drefs for objects of multiple types, including times, events, and predicates, I will use discourse referents for individuals and sets of worlds. Using drefs for sets of worlds, we can represent the epistemic context as a part of the referential context using a special dref for the interlocutors' epistemic state.

In addition to the introduction of new drefs, an update can impose conditions on these referents (Kamp (1981); Heim (1983); Groenendijk and Stokhof (1991); Kamp and Reyle (1993)) relative to a model. Consider (1), an example representation of an update in standard DRT (Kamp and Reyle (1993)):

(1) *Mary slept.*

$$\begin{array}{|c|}\hline x \\ \hline x = \mathsf{Mary}_e \\ \mathsf{sleep}(x) \\ \hline \end{array}$$

The top row of the DRS representing the update in (1) introduces a new dref x: It updates the variable assignment s.t. x refers to an entity in the model. The variable update is *random*. This will be explained formally below. Conceptually, this means that based on only the variable update (without additional conditions on the interpretation of x), x could refer to any entity in the model.

The conditions restrict the random update. The first condition $x = \mathsf{Mary}_e$ restricts the reference. As a result, the interpretation of (1) can only be True if x refers to the entity Mary_e. This condition is contributed by the semantics of the proper name. The second condition $\mathsf{sleep}(x)$ imposes a further condition on x. Here, the verb contributes the predicate, and the DP *Mary* provides its argument.

Truth-conditions in this dynamic framework are contingent on a successful update: (1) is true relative to an initial variable assignment and a model, iff a discourse referent can be added to the variable assignment which maps x to an entity in the model, s.t. x is $Mary_e$ and x sleeps.

The framework presented here is based on classic DRT. So this informal characterization of discourse referents, DRS conditions, and truth-conditional import can be applied to the system presented in this chapter. However, classic DRT does not involve intensionality. To see how DRT is extended to involve intensionality and propositional anaphora, I will proceed to present the formal system and its application for natural language semantics in the following way:

Section 3.2 illustrates the basic assumptions about types, formulas, discourse updates, and the model-theoretic interpretation.

Section 3.3 lays out some assumptions about how this system is applied for the semantic interpretation of simple cases of natural language assertions.

Section 3.4 illustrates the treatment of discourse variables and anaphora.

Section 3.5 discusses in some more detail how the pragmatics of assertion is captured implicitly as part of non-deterministic update, as well as some assumptions about the mechanism.

Section 3.6 illustrates how the system will be applied compositionally.

Section 3.7 concludes the chapter by summarizing its main points.

A complete list of formal definitions for this system can be found in the formal appendix to this chapter (**Appendix A**).

3.2 The Formal System

To allow for a model-theoretic interpretation of dynamic updates, the dynamic DRT representations are taken to be abbreviations of static logic formulas in Gallin's (1975) Ty2 logic. Therefore, the used system of intensional CDRT is a version of Muskens's (1996) logic of change, which emulates dynamics in a static system by construing dynamic meanings, i.e., CCPs, as relations over discourse states. These static relations can be understood as a model of the dynamic discourse update in dynamic semantic systems such as DRT. Viewed as relations, dynamic updates can be evaluated in a static logic.

3.2.1 Types and Objects

The system uses four basic types: t (truth-values), e (entities), w (possible worlds), and s (variable assignments). The basic dynamic objects, i.e. the ones that are manipulated in context are variable assignments.

A CDRT model includes a model frame and an interpretation function. The frame includes the domains for atomic objects D_t, D_e, D_w, D_s, corresponding to the set of truth-values, entities, possible worlds, and variable assignments in the model universe.

The domain of entities includes all entities in the model universe, as well as a dummy element $\#_e$. This dummy element will be used to model cases where a discourse variable does not refer to any entity in the domain. Similarly, the domain of truth values includes a dummy element $\#_t$ in addition to the standard truth-values $\{0, 1\}$.

$\#_e$ corresponds to the 'indeterminate' value in a trivalent logics, s.t. a formula receives an indeterminate truth-value $\#_t$ if any of its constituents evaluate to $\#_e$. This emulates partial functions in a system that uses only total functions (van den Berg (1996); Brasoveanu (2006, 2010a)) but still captures that the use of a pronoun results in unacceptability if used in a context where a referent cannot be determined.

A CDRT update can be understood as a binary relation over discourse states, where discourse states are variable assignments. An update D is a relation of type $s(st)$ between input state i_s and output state j_s (subscripts on terms indicate their type). Natural language (NL) utterances, modeled as updates, will be interpreted in relation to an input context and yield an output context after updating new variable assignments and imposing truth-conditional restrictions on the output context.

Following Brasoveanu's (2010a) implementation of CDRT, I will move away from the classic conception of variables as atoms and variable assignment as functions that map variables to entities. Instead, discourse variables are construed as type-lifted variables functioning as projection functions over variable assignments in the sense of (Landman, 1986, see also Brasoveanu (2006)). Variable assignments are basic objects in the semantic ontology (type s), and a discourse variable takes an assignment as its argument to return an individual.

Individual drefs are functions from assignments to individuals. An individual dref υ is a function of type se from assignments i_s to individuals x_e. Accordingly, the individual $\upsilon_{se}(i_s)$ is the individual that the assignment i assigns to υ. Throughout, I will use $\upsilon_1, \upsilon_2, \upsilon, \upsilon', \ldots$ to denote drefs for individuals.

Among individual drefs, we can distinguish between discourse constants and discourse variables (Muskens (1996) calls these specific vs. unspecific drefs). Discourse constants always refer to the same individual, regardless of the assignment, such as the dref associated with a proper name ($Mary_{se} = \lambda i_s.mary_e$). Discourse variables, on the other hand, are drefs whose referential value differs by assignment. These need to be introduced into the discourse to be assigned a referent by the assignments in the discourse. If no discourse referent for an individual variable dref name has been introduced yet, the assignments in the discourse will map it to $\#_e$ ('nothing').

Propositional drefs are functions from variable assignments to sets of possible worlds. They allow us to make static propositional content available for reference in discourse as part of the object language (Brasoveanu (2006, 2010a)). A propositional dref ϕ is a function of type $s(wt)$ from assignments i_s to sets of worlds (wt). The proposition $\phi_{s(wt)}(i_s)$ is the set of worlds that i assigns to ϕ. I will use use $\phi_1, \phi_2, \phi, \phi', \ldots$ to denote drefs for propositions.

Propositional discourse variables need to be introduced into the discourse before the assignments in the discourse can map them to a set of worlds. If no discourse referent for a propositional variable dref name has been introduced yet, the assignments in the discourse will map it to $\{\langle w, \#_t \rangle \mid w \in D_w\}$, a relation of type (w, t), which maps every world to the indeterminate truth value $\#_t$. This object can be interpreted as a set of worlds, s.t. for any possible world, it cannot be determined whether it belongs to the set or not. This is a way of formalizing the assumption that propositional discourse variables contain no information in a discourse before being introduced as a dref.

In some cases, we may want to consider discourse referents of either type. In these cases, I will use $\delta_1, \delta_2, \delta, \delta'$ to denote drefs of either type (i.e. type $s\sigma$, where $\sigma \in \{(e), (wt)\}$).

3.2.2 DRS Representation of Updates

CDRT updates can be represented as DRSs. A DRS is a boxed representation that abbreviates an update of type $s(st)$. There are two basic kinds of updates in CDRT, which can be represented in DRSs: *Variable updates* and updates including *conditions*.

Variable update

Variable updates model the introduction of new drefs and thereby capture the anaphoric potential of an utterance. A CDRT update that involves only a variable update is represented by the DRS-box in (2-a). (2-b) is a notational variant of this CDRT box-representation in inline CDRT.

(2) a. $\begin{array}{|c|}\hline \delta \\ \hline \\ \hline \end{array}$

b. $[\delta]$

To increase readability, I will provide both notational variants for updates corresponding to utterances throughout and use inline CDRT in definitions, abbreviations, and NL translations.

Like all updates, the DRS in (2) abbreviates a binary relation over assignments. Variable updates with a dref δ require that the input and output assignments i_s and j_s satisfy the relation $i[\delta]j$:

(3) Variable update:

$[\delta] := \lambda i_s.\lambda j_s.i[\delta]j$

The variable update $i[\delta]j$ is a random assignment of values to a variable δ, by determining an output assignment j that is just like the input assignment i, but in addition to all the mappings specified in i, it also assigns some value to δ. Formally, $i[\delta]j$ holds of two assignments i_s and j_s iff i and j differ at most wrt the values assigned to the variable δ (see e.g. Groenendijk and Stokhof (1991); Muskens (1996); Brasoveanu (2006)). This interpretation of $i[\delta]j$ is defined formally as (4).

(4) Variable update:

If $\delta \in \mathbf{Term}_{s\sigma}$, for $\sigma \in \{(e), (wt)\}$, $i, j \in \mathbf{Term}_s$:

$[\![i[\delta]j]\!]^{M,g} = 1$ iff

$\bullet [\![\forall\delta'_{s\sigma}[\mathbf{dVar}(\delta') \wedge \delta' \neq \delta](\delta'(i) = \delta'(j))]\!]^{M,g} = 1$ and

$\bullet [\![\forall\delta'_{s\tau}[\mathbf{dVar}(\delta')](\delta'(i) = \delta'(j))]\!]^{M,g} = 1$, for all $\tau \neq \sigma, \tau \in \{(e), (wt)\}$;

0 otherwise.

A full definition of the assumed interpretation function can be found in (4-d) in **Appendix A**. Variable update is random and non-deterministic. I will illustrate this here with an example for individual drefs and propositional drefs.

Individual Drefs: For a model M_1, where $D_e = \{mary, sue, max, \#\}$, consider an assignment i_s. Let's say i only specifies one discourse referent: the mapping $\upsilon_1(i) = mary_e$. Any other individual discourse variable will be mapped to $\#_e$ wrt i and any propositional discourse variable will be mapped to $\{\langle w, \#_t\rangle \mid w \in D_w\}$. In other words: i doesn't contain any information regarding any drefs other than υ_1.

a.)

	i
υ_1	$mary_e$
υ_2	$\#$
υ_3	$\#$
$\dots$	$\dots$
ϕ_1	$\{\langle w, \#_t\rangle \mid w \in D_w\}$
ϕ_2	$\{\langle w, \#_t\rangle \mid w \in D_w\}$
$\dots$	$\dots$

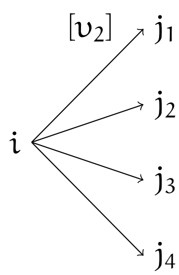
b.)

	i
υ_1	$mary_e$

Table 3.1: An overview of the mappings specified in i

An overview of the assignments for i is given in **Table 3.1 (a)**. It shows that i maps the dref υ_1 to the individual $mary_e$, and that no other discourse variable contains any information wrt i. Hereafter, whenever I give a table to illustrate the mappings associated with an assignment, I will omit the rows showing that a dref is associated with no information. Accordingly, assume that any individual discourse variable not listed maps to $\#_e$ wrt that assignment, and any propositional discourse variable to $\{\langle w, \#_t\rangle \mid w \in D_w\}$. Following this convention, the table can be simplified as in **Table 3.1 (b)**.

Given model M_1 and input assignment i, the relation $i[\upsilon_2]j$ is compatible with multiple values for j. Due to the random nature of variable update, the output assignment allows for υ_2 to refer to any of the entities in D_e. All of these possible mappings correspond to options for possible output values j. The non-deterministic random update, along with the possible output values, is illustrated in **Figure 3.1**.



	υ_1	υ_2
i	Mary_e	#
j_1	Mary_e	Mary_e
j_2	Mary_e	Sue_e
j_3	Mary_e	Max_e
j_4	Mary_e	#

Figure 3.1: Graph illustrating the update $i[\upsilon_2]j$, and overview of the mappings in i, j

In addition to the mapping $\upsilon_1 \mapsto \mathsf{Mary}_e$, which was already specified in the input i, the possible outputs include a new mapping for the dref υ_2. Variable update can introduce new drefs by allowing for new mappings. However, this update is random and doesn't impose any restrictions on the mapping beyond its type. Therefore $\upsilon_2(j)$ could refer to any entity $e \in D_e$, including the dummy entity #.

Random variable update is equivalent to existential quantification. We can see that this is the case, since there is a possible output for each of the elements in the domain. As we will see, we can consider an update D to be true given an input, if the update gives rise to at least one possible output that satisfies the update based on the input. In combination of these assumptions, random variable assignment amounts to existential quantification.

DRS-conditions will introduce further constraints on the possible mappings. This will be illustrated in the following subsection. Before we get there, let's take a look at a variable update with a propositional dref.

Propositional Drefs: Consider a model M_2, where $D_w = \{w_1, w_2\}$. Based on this and our domain of truth values, we can construct our domain of sets of worlds D_{wt}. Since we are using a trivalent logic, the domain of truth values is $D_t = \{1, 0, \#\}$ for any ICDRT model. As a result, D_{wt} contains more objects than just the power set of D_w. While it will include all possible subsets of D_w ($\varnothing, \{w_1\}, \{w_2\}, \{w_1, w_2\}$), which can be equivalently represented as $\{\langle w_1, 0\rangle, \langle w_2, 0\rangle\}, \{\langle w_1, 1\rangle, \langle w_2, 0\rangle\}, \{\langle w_1, 0\rangle, \langle w_2, 1\rangle\}$, $\{\langle w_1, 1\rangle, \langle w_2, 1\rangle\}$). It will also include all possible mappings $D_w \mapsto \{\#_t\}$. The domain of possible propositions D_{wt} with $D_w = \{w_1, w_2\}$ is shown in (5).

(5) $$\begin{aligned} D_{wt} = D_t^{D_w} = \{&\{\langle w_1, 0\rangle, \langle w_2, 0\rangle\},\\ &\{\langle w_1, 0\rangle, \langle w_2, 1\rangle\},\\ &\{\langle w_1, 0\rangle, \langle w_2, \#\rangle\},\\ &\{\langle w_1, 1\rangle, \langle w_2, 0\rangle\},\\ &\{\langle w_1, 1\rangle, \langle w_2, 1\rangle\},\\ &\{\langle w_1, 1\rangle, \langle w_2, \#\rangle\},\\ &\{\langle w_1, \#\rangle, \langle w_2, 0\rangle\},\\ &\{\langle w_1, \#\rangle, \langle w_2, 1\rangle\},\\ &\{\langle w_1, \#\rangle, \langle w_2, \#\rangle\}\} \end{aligned}$$

Consider an assignment i_s. Assume i only specifies one discourse referent: the mapping

$\phi_1(i) = \{w_1\}$.[1] Any other propositional discourse variable will be mapped to $\{\langle w, \#_t\rangle \mid w \in D_w\}$ wrt i and any individual discourse variable will be mapped to $\#_e$. In other words: i doesn't contain any information regarding any drefs other than ϕ_1. An overview of the assignments for i is given in **Table 3.2**.

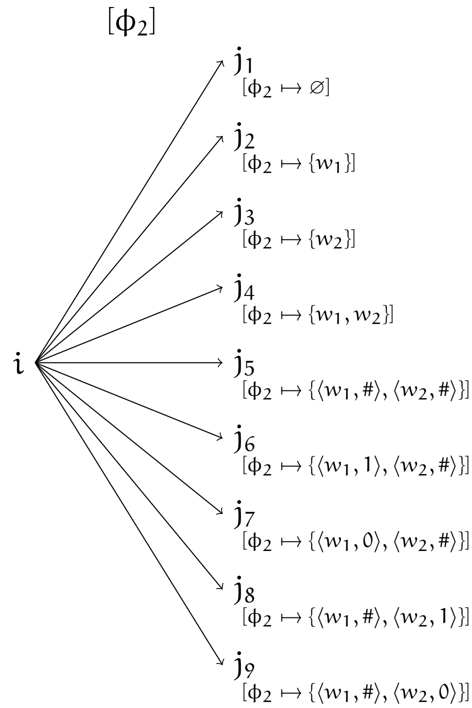
	i
ϕ_1	$\{w_1\}$

Table 3.2: The mappings for i

Given model M_2 and input assignment i, the relation $i[\phi_2]j$ is compatible with multiple values for j. Due to the random nature of variable update, the output assignment allows for ϕ_2 to refer to any of the objects in D_{wt}. All of these possible mappings correspond to options for possible output values of j. The non-deterministic random update, along with the possible output values, is illustrated in **Figure 3.2**.

Like with individual drefs, the variable update introduces a discourse referent by allowing for new mappings. While all of the given mappings are possible due to the randomness of variable update, not all of them are useful for the interpretation of a NL discourse. However, the random update doesn't impose any restrictions on the mapping beyond its type. Any further constraints on dref mappings will be introduced by DRS conditions, which restrict the possible mappings by imposing conditions on the output state.

[1]This is a simplified representation of the object $\{\langle w_1, 1\rangle, \langle w_2, 0\rangle\}$.



	ϕ_1	ϕ_2
i	$\{w_1\}$	$\{\langle w_1, \#\rangle, \langle w_2, \#\rangle\}$
j_1	$\{w_1\}$	$\varnothing$
j_2	$\{w_1\}$	$\{w_1\}$
j_3	$\{w_1\}$	$\{w_2\}$
j_4	$\{w_1\}$	$\{w_1, w_2\}$
j_5	$\{w_1\}$	$\{\langle w_1, \#\rangle, \langle w_2, \#\rangle\}$
j_6	$\{w_1\}$	$\{\langle w_1, 1\rangle, \langle w_2, \#\rangle\}$
j_7	$\{w_1\}$	$\{\langle w_1, 0\rangle, \langle w_2, \#\rangle\}$
j_8	$\{w_1\}$	$\{\langle w_1, \#\rangle, \langle w_2, 1\rangle\}$
j_9	$\{w_1\}$	$\{\langle w_1, \#\rangle, \langle w_2, 0\rangle\}$

Figure 3.2: Graph illustrating the update $i[\phi_2]j$, and overview of the mappings in i, j

DRS Conditions

DRS conditions are conditions on the output state. Accordingly, their type is st. Because the system emulates a dynamic system within a static framework, dynamic terms abbreviate static relations. On the dynamic side, a DRS condition is the result of combining a dynamic relation with all its arguments. On the static side, a DRS-condition abbreviates a property of assignments. The abbreviation is exemplified in (6):

(6) $Sleep_{\phi}\{v\} :=$

$\lambda i_s.\forall w[w \in \phi(i)](sleep_{e(wt)}(v(i))(w)),$

for $v \in \mathbf{Term}_{s(we)}$, $\phi \in \mathbf{Term}_{s(wt)}$

The dynamic relation $Sleep$ requires two arguments: an individual dref for the sleeper and a propositional dref for the intensional context. The latter is indicated as a subscript on the predicate. (6) states that the dynamic term $\mathtt{Sleep}_{\phi}\{\upsilon\}$ is interpreted as a condition on an assignment i, by assuring that the corresponding static predicate holds of the entity referred to by υ at i in all the worlds in the propositional referent of ϕ at i.

This mechanism of point-wise checking whether the relation holds for each world in the local intensional context was introduced by Stone (1999), who suggests that lexical meanings encapsulate quantification over possible worlds. The interpretation of lexical relations like, e.g., the predicate introduced by the verb *sleep* is interpreted like this. We call this ENCAPSULATED QUANTIFICATION because interpreting DRS-conditions in this way involves static quantification over singleton possible worlds. That is separate from the random variable update, which the dynamic semantic framework uses to dynamically quantify over individuals or sets of possible worlds (i.e., propositions).

DRS conditions can be included in updates. An update that includes only the condition (6) could be represented as in (7), shown here in box-notation and inline-notation.

(7) a.

$\mathtt{Sleep}_{\phi}\{\upsilon\}$

b. $[\mathtt{Sleep}_{\phi}\{\upsilon\}]$

(7) abbreviates a binary relation over assignments as defined in (8).

(8) $[\mathtt{Sleep}_{\phi}\{\upsilon\}] := \lambda i_s.\lambda j_s.i = j \wedge \mathtt{Sleep}_{\phi}\{\upsilon\}(j)$

An update with a condition doesn't add new mappings to the output state, but restricts the possible output values.

To illustrate how updates including conditions constrain the reference of individual and propositional drefs, let us consider a concrete example of interpreting the update in relation to a model and a discourse state:

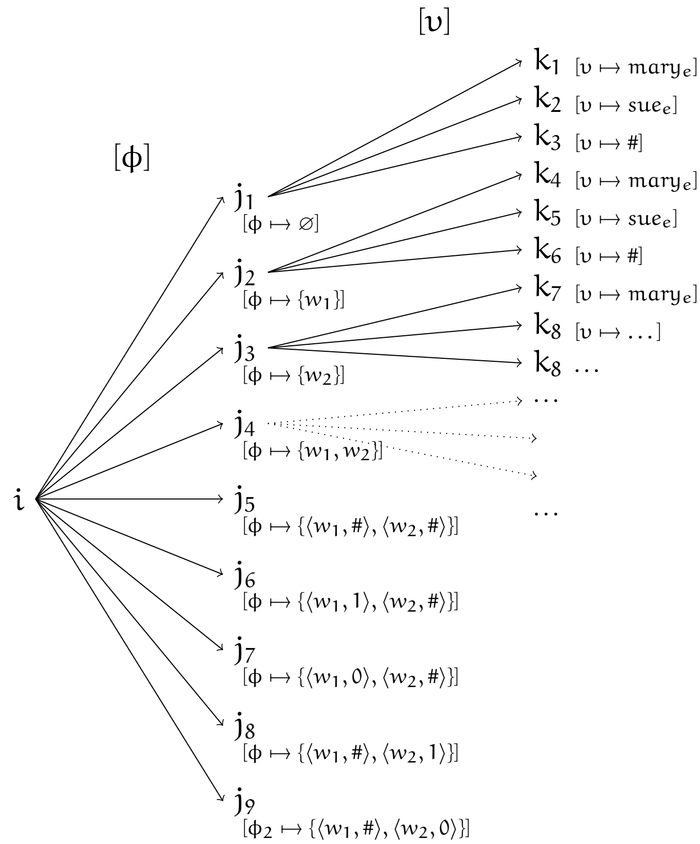


Figure 3.3: Graph illustrating update of i with $[\phi]$, then $[\upsilon]$

For a model M_3, assume $D_e = \{Mary_e, Sue_e, \#\}$, $D_w = \{w_1, w_2\}$, as well as an idealized initial discourse state i, which contains no information wrt any discourse variable names. Before we illustrate how conditions restrict dref mappings, let us assume that i is updated with a propositional dref ϕ and an individual dref υ. The possible output assignments should involve all possible combinations for mapping these variable names to objects of the corresponding type. This is illustrated in **Figure 3.3**.

If a condition makes reference to these discourse variables, the update can restrict their possible values. The DRS (7) abbreviates the following update:

(9) $[\mathtt{Sleep}_{\phi}\{\upsilon\}] =$

$$\lambda i_s.\lambda j_s.i = j \wedge \forall w[w \in \phi(j)](\mathtt{sleep}_{e(wt)}(\upsilon(j))(w))$$

Because the of the statement $i = j$, this update works as a test. It doesn't make any changes to the possible values of the discourse state. However, filters possible output values by imposing the condition that the specified relation holds of the assigned referents of the discourse variables. As a result, any assignment that doesn't make the condition true is not a possible output any more. (9) states that a possible output j is s.t. $\upsilon(j)$ sleeps in each $w \in \phi(j)$.

For concreteness, consider the interpretation of $\mathtt{sleep}_{e(wt)}$ in M_3:

(10) The semantic value of *sleep* in M_3:

$$[\![\mathtt{sleep}]\!]^{M_3} = \begin{bmatrix} \mathtt{mary}_e \mapsto \begin{bmatrix} w_1 \mapsto 1 \\ w_2 \mapsto 1 \end{bmatrix} \\ \mathtt{sue}_e \mapsto \begin{bmatrix} w_1 \mapsto 1 \\ w_2 \mapsto 0 \end{bmatrix} \\ \#_e \mapsto \begin{bmatrix} w_1 \mapsto \\ \#_t \\ w_2 \mapsto \\ \#_t \end{bmatrix} \end{bmatrix}$$

(10) shows that in M_3, $\mathtt{mary}_e$ sleeps in all worlds and $\mathtt{sue}_e$ sleeps only in w_1. If the first argument of $\mathtt{sleep}_{e(wt)}$ is $\#_e$, it will return $\#_t$ for any world. This reflects the intuition that we cannot decide the truth of a proposition if the semantic value of its constituents is cannot be determined.

Based on this, the possible output values in **Figure 3.3** are filtered by an update with (9). Only those assignments k that assign values to ϕ and υ in a way s.t. $\upsilon(\mathsf{k})$ sleeps in all worlds in $\phi(\mathsf{k})$ are possible outputs. Since sue_e sleeps in w_1 and mary_e sleeps in w_1 and w_2, any assignment that maps υ to sue_e and ϕ to $\{w_1\}$ or any assignment that maps υ to mary_e and ϕ to any subset of $\{w_1, w_2\}$ will satisfy this condition. For any assignment k that maps ϕ to the empty set, the statement $\forall w[w \in \phi(\mathsf{k})](\mathsf{sleep}_{e(wt)}(\upsilon(\mathsf{k}))(w))$ is vacuously satisfied. For any assignment k that maps υ to sue_e and ϕ to a set of worlds which contains worlds where Sue doesn't sleep, the condition is not satisfied. For any assignment k that maps υ to #, the condition is undefined (it semantic value is #), because the reference of the argument cannot be determined and the consequent doesn't get a binary truth-value. For any assignment k that maps ϕ to some object W of type wt where $\langle w, \# \rangle \in W$ for some $w \in D_w$, the condition is undefined because proposition membership cannot be decided and the antecedent doesn't get a binary truth-value. **Table 3.3** provides an overview over the input states in the current discourse and their status wrt an update with the condition at hand.

An update with $[\mathsf{Sleep}_\phi\{\upsilon\}]$ filters the assignments in the discourse context: Out of all possible input values, the possible values for the output include those assignments which satisfy the condition. This is illustrated in **Figure 3.4**.

A potentially problematic aspect here is that any condition will be vacuously true of an assignment that maps a propositional argument to the empty set. This issue will

	ϕ	υ	$[\![\forall w[w \in \phi(k)](sleep_{e(wt)}(\upsilon(k))(w))]\!]^{M_3} = ?$
k_4	$\{w_1\}$	sue_e	Condition satisfied (= 1)
k_5	$\{w_1\}$	$mary_e$	
k_3	$\{w_2\}$	$mary_e$	
k_2	$\{w_1, w_2\}$	$mary_e$	
k_1	$\varnothing$	$mary_e$	Condition vacuously satisfied (= 1)
k_2	$\varnothing$	sue_e	
k_3	$\varnothing$	$\#_e$	
k_8	$\{w_2\}$	sue_e	Condition not satisfied (= 0)
k_9	$\{w_1, w_2\}$	sue_e	
k_{10}	$\{w_1\}$	$\#_e$	Condition undefined (= #)
k_{11}	$\{w_2\}$	$\#_e$	
k_{12}	$\{w_1, w_2\}$	$\#_e$	
k_{13}	$\{\langle w_1, \#\rangle, \langle w_2, \#\rangle\}$	$mary_e$	Condition undefined (= #)
k_{14}	$\{\langle w_1, \#\rangle, \langle w_2, \#\rangle\}$	sue_e	
k_{15}	$\{\langle w_1, \#\rangle, \langle w_2, \#\rangle\}$	$\#_e$	
k_{16}	$\{\langle w_1, 1\rangle, \langle w_2, \#\rangle\}$	$mary_e$	
k_{17}	$\{\langle w_1, 1\rangle, \langle w_2, \#\rangle\}$	sue_e	
k_{18}	$\{\langle w_1, 1\rangle, \langle w_2, \#\rangle\}$	$\#_e$	
k_{19}	$\{\langle w_1, 0\rangle, \langle w_2, \#\rangle\}$	$mary_e$	
k_{20}	$\{\langle w_1, 0\rangle, \langle w_2, \#\rangle\}$	sue_e	
k_{21}	$\{\langle w_1, 0\rangle, \langle w_2, \#\rangle\}$	$\#_e$	
k_{22}	$\{\langle w_1, \#\rangle, \langle w_2, 1\rangle\}]$	$mary_e$	
k_{23}	$\{\langle w_1, \#\rangle, \langle w_2, 1\rangle\}]$	sue_e	
k_{24}	$\{\langle w_1, \#\rangle, \langle w_2, 1\rangle\}]$	$\#_e$	
k_{25}	$\{\langle w_1, \#\rangle, \langle w_2, 0\rangle\}]$	$mary_e$	
k_{26}	$\{\langle w_1, \#\rangle, \langle w_2, 0\rangle\}]$	sue_e	
k_{27}	$\{\langle w_1, \#\rangle, \langle w_2, 0\rangle\}]$	$\#_e$	

Table 3.3: Overview over possible values of k and their status wrt $Sleep_{\phi}\{\upsilon\}$

be taken care of in the following section, as part of the pragmatics associated with the assertion operator I will be assuming. This will prevent a speaker's commitment set from being empty, and the propositional drefs introduced veridically will inherit this property.

3.2.3 Truth and Acceptability (in Context)

Updates specify relations between assignments. A formula $D_{s(st)}(i_s)(j_s)$ is TRUE in relation to a model M iff $[\![D(i)(j)]\!]^M = 1$. A full definition of the interpretation function

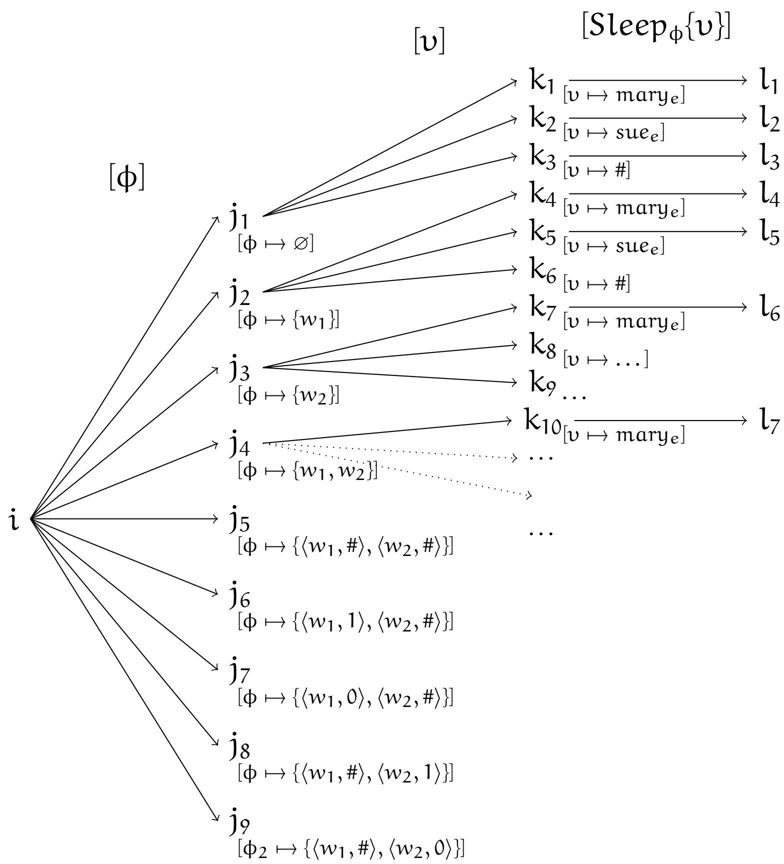


Figure 3.4: Graph illustrating update of i with $[\phi]$, $[\upsilon]$, then $[\mathit{Sleep}_{\phi}\{\upsilon\}]$

is given in **Appendix A.1.**

As a result of updates, including variable updates and updates with conditions, the output state only contains variable mappings (i.e. drefs) made salient in the discourse, which satisfy the conditions imposed by everything that is said about them. This makes Intensional CDRT suitable for analyzing natural language discourses in context.

To understand what it means for an utterance (i.e. an update) to be true in a context, let us be specific about the notion of context we are using here. In the introduction to this chapter, two notions of context were mentioned: A referential context (storing information about previously introduced drefs) and an epistemic context (storing propositional information about the interlocutors's discourse commitments). In order to interpret an utterance in context, we need both (in addition to a model).

The referential context has been formally defined in this section, as a variable assignment reflecting the current state of the discourse. Epistemic contexts will be formally introduced in the following section, as propositional drefs corresponding to the set of worlds compatible with an interlocutor's discourse commitments. Representing the epistemic contexts as propositional drefs allows us to represent them as part of the referential context. However, because they are conceived in relation to interlocutors in a conversation, we need to know who these conversational agents are.

We can determine the truth of an utterance in a context, if have a model of evaluation, a variable assignment, and access to the information about the interlocutors. We define a (CONVERSATIONAL) CONTEXT C is defined as a 3-tuple $C = \langle M, i_s, \textsc{Int} \rangle$, where M is a model, i_s is a variable assignment storing the current state of the discourse, and INT, a subset of D_e^M, is the set of interlocutors in C.

We say that an update $D_{s(st)}$ is DISCOURSE-TRUE in a context C iff there exists at least one output j_s s.t. $[\![D(i^C)(j)]\!]^{M^C} = 1$. D will be DISCOURSE-FALSE in C iff for all possible outputs j_s $[\![D(i^C)(j)]\!]^{M^C} = 0$. We can also capture the case where an utterance is pragmatically unacceptable due to unspecified reference. This happens if the update contains components which cannot be assigned an semantic value other than #: An update D is ACCEPTABLE in C iff there exists at least one output j_s, s.t. $[\![D(i^C)(j)]\!]^{M^C} \neq \#$

(and unacceptable otherwise).

3.2.4 Further Definitions

Dynamic Conjunction is used to conjoin updates of type $s(st)$. It is defined as relation composition. Conjoining two updates $D_{s(st)}, D'_{s(st)}$ therefore requires feeding the output context of the first forward as the input context of the latter, as defined in (11):

(11) Dynamic conjunction:

$$D_{s(st)}; D'_{s(st)} := \lambda i_s.\lambda j_s.\exists h_s.[D(i)(h)](D'(h)(j))$$

For example, with the same model M_3 and input state i as above, an update with the dynamic conjunction of the three updates ($[\phi]; [\upsilon]; [Sleep_\phi\{\upsilon\}]$) leads to the same output we get from applying the three updates separately, illustrated in **Figure 3.5**.

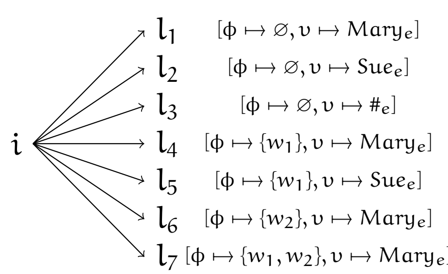


Figure 3.5: Update with $[\phi]; [\upsilon]; [Sleep_\phi\{\upsilon\}]$

With dynamic conjunction, we can also abbreviate the introduction of multiple drefs as follows:

(12) Introducing multiple drefs:

$$[\upsilon_1, \ldots, \upsilon_n] := [\upsilon_1]; \ldots; [\upsilon_n]$$

DRSs that are conjoined like $[\phi];[\upsilon];[\mathit{Sleep}_{\phi}\{\upsilon\}]$ can also be equivalently merged together, as shown in (13) for both inline CDRT and box-notation.

(13) a. $[\phi,\upsilon \mid \mathit{Sleep}_{\phi}\{\upsilon\}]$

b.

ϕ,υ
$\mathit{Sleep}_{\phi}\{\upsilon\}]$

A DRS that contains both variable updates and conditions abbreviates an update like in (14).

(14) $[\upsilon_1,\ldots,\upsilon_n \mid C_1,\ldots,C_n] :=$

$\lambda i_s.\lambda j_s.i[\upsilon_1,\ldots,\upsilon_n]j \wedge C_1(j) \wedge \cdots \wedge C_n(j)$

According to (14), a DRS, specifying a relation between an input context i_s and an output context j_s requires **1.** that j is the result of a variable update of i with the new discourse referents $\upsilon_1,\ldots,\upsilon_n$ and **2.** that the conditions $C_1,\ldots,C_n$ hold of the output context j.

3.3 The Interpretation of Natural Language Utterances

Section 3.2 introduced the framework needed to interpret dynamic updates of a discourse relative to a model and an input assignment. This section builds on that by laying out some assumptions required to apply this framework to the interpretation of natural language utterances. This will be illustrated by means of just one simple example involving declarative assertion and proper names.

Recall that the system in this chapter is not designed to handle quantification over individuals (or over anything other than possible worlds) but will be extended to do so in an upcoming chapter. For now, the focus is on the basic mechanisms needed to interpret anaphora, negation, and propositional anaphora in particular. The main point of this section is to present some assumptions about how the tools at hand can provide insight into the interpretation of simple utterances in discourse in terms of their truth-conditions and anaphoric potential.

A natural language UTTERANCE, is an utterance of a natural language sentence by a speaker. A natural language SENTENCE is a syntactically well-formed sentence of the object language. The meta-language, i.e., the intensional CDRT framework, is employed to provide TRANSLATIONS of utterances of object-language expressions, which can be semantically interpreted in a context. The (meta-language) translation for the utterance of a(n object language) sentence corresponds to an update of type $s(st)$. The associated update is represented as a DRS.

Let me illustrate some further assumptions about the semantics of simple declarative assertions using an example. Consider the utterance by a speaker S of a simple sentence of the object language (in this case, English), like (15).

(15) a. S: *Sue slept.*

b. $$\begin{array}{|c|}\hline \phi, \upsilon \\ \hline \phi_{DC_S} \Subset \phi \\ \upsilon \equiv \mathit{Sue}_{se} \\ \mathit{Sleep}_{\phi}\{\upsilon\} \\ \hline \end{array}$$

c. $[\phi \mid \phi_{DC_S} \Subset \phi]; [\upsilon \mid \upsilon \equiv \mathit{Sue}_{se}]; [\mathit{Sleep}_{\phi}\{\upsilon\}]$

Many parts of this representation should be familiar from the previous section. The upper row in this representation is interpreted s.t. two new drefs are introduced: A propositional dref ϕ and an individual dref υ. The main part of the DRS-box includes three conditions that are imposed on the output discourse state: The condition $\mathtt{Sleep}_{\phi}\{\upsilon\}$ ensures that *Sue slept* in all the ϕ-worlds, which in turn guarantees that the propositional dref ϕ corresponds to the proposition expressed by the utterance content. This condition is contributed by the interpretation of the verb *slept* in combination with the arguments for a sleeper υ and an intensional context ϕ.

Two further conditions are included here: The condition $\upsilon \equiv \mathtt{Sue}_{se}$ makes sure that υ refers to $\mathtt{sue}_{e}$. Like the update with the dref υ, this condition is contributed by the semantics of the proper name *Sue*. The condition $\phi_{DC_S} \Subset \phi$ involves a special dref ϕ_{DC_S}, which refers to a set of worlds that are compatible with S's discourse commitments. The condition states that ϕ is true in all the worlds where the S's discourse commitments are true. This condition restricts the possible values of ϕ_{DC_S} to only those values which are subsets of ϕ. This condition and the update with the dref ϕ are associated with declarative sentential mood. Let me discuss some of the formal definitions involved in these two conditions.

3.3.1 Proper Names

The proper name *Sue* introduces an individual dref υ and contributes the condition $\upsilon \equiv \mathtt{Sue}_{se}$. Like all conditions, this is a dynamic abbreviation of a property of a discourse state, as defined in (16):

(16) Dynamic equality:

$\alpha \equiv \beta :=$

$\lambda i_s.\alpha(i) = \beta(i)$

for $\alpha, \beta \in \textbf{Term}_{s\tau}$, and $\tau \in \{e, (wt)\}$

Accordingly, the condition $\upsilon \equiv Sue_{se}$ abbreviates the following:

(17) $\upsilon \equiv Sue_{se} =$

$\lambda i_s.\upsilon(i) = Sue_{se}(i)$

(where $Sue_{se} := \lambda i_s.sue_e$)

This condition holds of any discourse state i_s s.t. the value of $\upsilon(i)$ is equal to the value of $Sue_{se}(i)$. The dref Sue_{se} is a discourse constant associated with the proper name *Sue*. Like all constants associated with proper names, $Sue_{\mathbf{e}}$ is subject to the condition that it refers to the same static individual (in this case sue_e) across all variable assignments.

3.3.2 Assertion and Sentential Mood

Declarative Mood

I follow treatments of sentential mood after Stenius (1967), which suggest that the descriptive content of a sentence, a proposition, combines with sentence mood, which provides a modal component. This view is based on Wittgenstein's notion of a sentence radical. Wittgenstein (1952) suggests that sentences contain a sentence-radical which depicts a Fregean ‘assumption’, which can be asserted, negated, questioned, or used as an instruction. For our purposes, this will be interpreted as a proposition, which the speaker might assert to be true or false, or entertain as a hypothetical.

For now, we can limit our considerations to declarative mood, which is is associated with assertion, and formalized here as the update of the speaker’s commitments (Stalnaker (1978); Gunlogson (2004); Bittner (2009, 2011); Murray (2014)). In its cross-linguistic realization, some languages mark this morphologically (see e.g. Bittner (2009, 2011); Murray (2016)). In English, the contribution of assertion could rather be associated with falling intonation on declarative sentences (Steedman (2000); Gunlogson (2004); Rudin (2018)). I am going to gloss over this variation here and subsume the semantic contribution into one single assertive operator. Following Bittner (2009, 2011), I also assume that the semantics of declarative mood introduces a propositional dref for its prejacent and relates it to the global worlds of evaluation.

Accordingly, the assertion of a proposition fulfills two important functions: **(i)** It introduces a propositional dref corresponding to the asserted proposition (Bittner (2009, 2011), see also Snider (2017)). **(ii)** It imposes the condition that the propositional content of prejacent is entailed by the discourse commitments of the speaker (Gunlogson (2004)).

Commitment Update within Non-Deterministic Discourse Update

The update of the interlocutors' commitments states is part of the non-deterministic discourse update. Here, we are not assuming one context set as part of the discourse context which is updated by intersecting with asserted propositions in the sense of Stalnaker (1978). We are also not using an assertion mechanism like in Murray (2014), where a discourse referent for the epistemic context is retrieved, intersected with the main assertion, and then reintroduced as a new dref. This mechanism would reintroduce a dref for the epistemic context with each assertion.

However, we are going to make use of a kind of epistemic update in the spirit of Stalnaker, while incorporating the epistemic context as part of our referential context, like Murray. To do this, we will make use of designated propositional drefs for the interlocutors' commitment states (ϕ_{DC_x} for each $x \in \textsc{Int}$). These are discourse variables with special names. Their names are indexed to interlocutors, and they refer to the set of worlds associated with each interlocutor's commitments.

We assume here that these drefs are introduces implicitly at the beginning of a discourse via random variable update. Due to the non-deterministic nature of random variable update, there is inherent uncertainty in a discourse about the exact value for a given dref. The assertion mechanism we assume here will narrow down all possible values to just the ones that are compatible with the asserted proposition. (The updates involved will be illustrated in the upcoming **Subsection 3.3.3**, and further consequences of this approach will be discussed in **Section 3.5**.)

Narrowing down all possible values of ϕ_{DC_S} for a speaker S to just the ones that contain only worlds in which the assertion is true is achieved by an assertion operator which introduces a propositional dref ϕ for the asserted proposition and contributes the condition $\phi_{DC_S} \Subset \phi$. The dynamic $\Subset$-operator is defined as follows:

(18) Dynamic set inclusion:

$\phi_1 \Subset \phi_2 :=$

$\lambda i_s.\phi_1(i) \subseteq \phi_2(i)$

for $\phi_1, \phi_2 \in \mathbf{Term}_{s(wt)}$

Accordingly, the condition $\phi_{DC_S} \Subset \phi$ abbreviates the following static property:

(19) $\phi_{DC_S} \Subset \phi =$

$\lambda i_s.\phi_{DC_S}(i) \subseteq \phi(i)$

(where ϕ_{DC_S} is a discourse variable referring to the set of worlds compatible with the discourse commitments of speaker S)

This condition holds of any discourse state is s.t. the value of $\phi_{DC_S}(i)$ is a subset of the value of $\phi(i)$. ϕ_{DC_S} is a special discourse variable that corresponds to a set of worlds that is compatible with the discourse commitments of S and is constrained with every assertion by S. This condition specifies that the referent of the propositional dref ϕ associated with the asserted proposition is taken to be true in all the worlds where the speaker's discourse commitments are true. In an update, this condition ensures that the output assignment maps ϕ_{DC_S} to a subset of the possible values that ϕ_{DC_S} has in the input—the subset of values that are also in ϕ.

Because ϕ is the local intensional context of the assertion, other conditions contributed by the content of the utterance will be evaluated wrt ϕ (such as the condition $\mathit{Sleep}_{\phi}\{\upsilon_1\}$). Since conditions introduced by the utterance constrain ϕ, and the values of ϕ constrain the values of ϕ_{DC_S}, any condition referencing ϕ also indirectly constrains the value of ϕ_{DC_S}.

As a result, an update with the condition $\phi_{DC_S} \Subset \phi$ captures the pragmatic contribution of an assertion. At the same time, the introduction of a new dref corresponding to the matrix proposition accounts for the fact that any assertion of a declarative sentence will introduce a propositional dref.

The drefs associated with the interlocutors' commitment sets are formalized as discourse variables so that conditions can impose conditions on their values. I further assume that the drefs associated with a speaker's discourse commitments in any context satisfy the following implicit pragmatic assumptions:

1. **Accessibility:**

   For each member of the set of interlocutors $x \in \textsc{Int}$, a dref ϕ_{DC_S} is introduced implicitly at the beginning of a discourse. A possible referent for ϕ_{DC_S} is a set of worlds P s.t. the discourse-commitments of x are true throughout P. This allows for a representation of an interlocutor's discourse commitments, which can be referred to in discourse.

2. **Consistency:**

   The set of worlds associated with the discourse commitments of a speaker x is non-empty, for any assignment $i \in \textbf{Term}_s$. This prevents speakers from holding contradictory discourse commitments and requires that any speaker publicly entertains at least some possible worlds as candidates for the actual world.

   (20) $\forall x, i((x \in \textsc{Int} \wedge i \in \textbf{Term}_s) \rightarrow \phi_{DC_S}(i) \neq \varnothing)$

3. **Definedness:**

   The set of worlds associated with the discourse commitments of a speaker x is well-defined for any assignment $i \in \textbf{Term}_s$. That is, for each world $w \in D_w$, $\phi_{DC_x}(i)(w)$ maps to either 1 or 0, but not to #. This can be interpreted as stating that for each w, we can decide whether x's commitments are true or false in w.

   (21) $\forall x, w, i((x \in \textsc{Int} \wedge w \in D_w \wedge i \in \textbf{Term}_s) \rightarrow \langle w, \#_t \rangle \notin \phi_{DC_x}(i))$

Since assertion of a proposition ϕ requires ϕ_{DC_S} to be a subset of ϕ, the above pragmatic principles also force any asserted proposition to be non-empty and defined.

3.3.3 Update with a Declarative Sentence

Combining all components, the DRS associated with an utterance of *Sue slept* (15-b) (repeated here) abbreviates the update in (22).

(15-b) $\boxed{\begin{array}{c} \phi, \upsilon \\ \hline \phi_{DC_S} \Subset \phi \\ \upsilon \equiv Sue_{se} \\ Sleep_{\phi}\{\upsilon\} \end{array}}$

(22) Update associated with (15-b)

$$\lambda i_s.\lambda j_s.i[\phi, \upsilon]j \;\wedge\; \phi_{DC_S}(j) \subseteq \phi(j) \;\wedge\; \upsilon(j) = sue_e \wedge$$
$$\forall w[w \in \phi(j)](sleep_{e(wt)}(\upsilon(j))(w))$$

To illustrate the effect of this update more concretely, we consider a conversational context $C_1 = \langle M, \textsc{Int}, i\rangle$ for interpreting an update, which consists of a model, a set of interlocutors, and a discourse state.

The Models we will consider to illustrate the effect of an update will be toy models. The domain of individuals will be the smallest set including individuals that can be the referents of any individual drefs evoked in the update, the interlocutors in the conversation, and the dummy individual #. For M^{C_1} (the model in context C_1), this is $D_e = \{sue_e, S_e, \#\}$, if we take $\textsc{Int} = \{S\}$ (i.e. one person talking to themself).

The domain of possible worlds will be the smallest set that includes enough worlds to include all possible cominations of truth/falsity for each static proposition referenced in the conditions contributed by lexical predicates. For M^{C_1}, $D_w = \{w_1, w_2\}$ includes two worlds. One where Sue slept, and another where she didn't. Accordingly, we assume that the semantic value of $sleep_{e(wt)}$ in M^{C_1} is such that

$[\![sleep(sue)(w_1)]\!]^{M^{C_1}} = 1$ and $[\![sleep(sue)(w_2)]\!]^{M^{C_1}} = 0$.

Other properties of the model will not be relevant and not be specified. Of course, the same update could be interpreted in relation to a richer model, e.g., one that includes additional individuals or worlds.

Discourse-initial states. Considering the dynamic effects of an update like (15-b), which uses the dref names ϕ, ϕ_{DC_S} and υ, we only care about the values that possible input and output assignments assign to these drefs. The values that an assignment might give to other drefs are irrelevant in the interpretation of (15-b). However, for purposes of specificity, we assume an idealized discourse-initial state, which does not hold any information about any dref names and is one where none of the interlocutors have any public discourse commitments.

We can construct such an object in two steps: First, let's assume an assignment i_0, which doesn't specify any information wrt any individual or propositional discourse variable. As such, i_0 satisfies the statements in (23).

(23) Properties of i_0:

a. No individual drefs:

$\forall \upsilon_{se}[\mathbf{dVar}(\upsilon)](\upsilon(i_0) = \#)$

b. No propositional drefs:

$\forall \phi_{s(wt)}[\mathbf{dVar}(\phi)](\phi(i_0) = \{\langle w, \#_t \rangle \mid w \in D_w\})$

(23-a) states that every individual discourse variable refers to # in i_0. (23-b) states that every propositional discourse variable contains no information in relation to i_0. Here, this means that it refers to $\{\langle w, \#_t\rangle \mid w \in D_w\}$ a relation of type $\{w, t\}$, which maps every world to the indeterminate truth value $\#_t$. This object can be interpreted as a set of worlds, s.t. for any possible world, it cannot be determined whether it belongs to the set or not.

These axioms ruling out discourse-initial reference apply to discourse variables, but not to constants, like Mary_{se} or Sue_{se}. Discourse constants are not affected by the statements in (23) and will rigidly refer throughout variable assignments.

Based on this input, we assume that for each interlocutor, a dref corresponding to their commitment set is introduced implicitly at the beginning of a discourse:

(24) For all members $x, x', x'', \ldots$ of Int:

$$i_0[\phi_{DC_x}, \phi_{DC_{x'}}, \phi_{DC_{x''}}, \ldots]i$$

Due to the random nature of variable update, there will be different possible output values i, mapping each of the commitment set drefs to different elements of D_{wt}. At the same time, commitment sets need to satisfy the pragmatic constraints of consistency and definedness. No dref for a commitment set can be mapped to the empty set or a $w \mapsto t$ function that maps any world to $\#_t$. Therefore, the possible values for i may map a commitment set to D_w or any of its non-empty, well-defined subsets. This is a formal way of ensuring that none of the interlocutors have any commitments in i.[2].

[2]This choice is made for concreteness only. If a speaker S enters the discourse with previous commitments, an implicit condition will be added that ϕ_{DC_S} satisfies S's discourse commitments. The possible values will thereby be constrained to any (non-empty, well-defined) subset of D_w in which the previous commitments are true.

The result of these considerations is that a discourse-initial state i contains no information wrt any individual discourse variables and contains no information wrt any propositional discourse variables that do not refer to a speaker's commitment set. In addition, i doesn't specify any interlocutors' discourse commitments: Possible values for i include those, which map commitment-set drefs to any non-empty, well-defined subset of D_w.

The set of interlocutors assumed will be the smallest set including as many interlocutors as there are different speakers in a discourse. For concreteness, a discourse involving only utterances by a single speaker will be interpreted in a context where $\textsc{Int} = \{S\}$, assuming that a single person is talking to themself. Of course, the same update could be interpreted in a context with a larger set of interlocutors (and given that we are only considering simple assertions without local pronouns, this should be possible without a change in meaning).

Our context. The properties of C_1 that are relevant for interpreting the update in (15-b) are summarized in **Table 3.4**.

Based on our definition of a a discourse-initial state, i^{C_1} specifies mappings only for drefs corresponding to the commitment state of S. i^{C_1} is compatible with mapping ϕ_{DC_S} to any non-empty, well-defined subset of D_w. As a result, i^{C_1} is compatible with the possible mappings given in **Table 3.5**.

Context for interpreting the update:	$C_1 = \langle M, \textsc{Int}, i \rangle$
Relevant properties of M_3:	$D_e = \{sue_e, S_e, \#\}$ $D_w = \{w_1, w_2\}$ $[\![sleep(sue)(w_1)]\!]^{M_3} = 1$ $[\![sleep(sue)(w_2)]\!]^{M_3} = 0$
Set of Interlocutors	$\textsc{Int} = \{S\}$
i	Is a discourse-initial state, as defined above

Table 3.4: An overview of the properties of C_1 relevant for interpreting the update (15-b)

	ϕ_{DC_S}
i_1	$\{w_1, w_2\}$
i_2	$\{w_1\}$
i_3	$\{w_2\}$

Table 3.5: Possible values for the initial assignment i^{C_1}

The effect of the update

We say that an possible output assignment j after updating $D_{s(st)}$ in a context C is an assignment j_s, s.t. $[\![D(i^C)(j)]\!]^{M^C} = 1$. We illustrate the effect of an update by investigating the properties of the output.

Let us now consider the effect of the update associated with (15-b), i.e. (22) which is repeated here, in our context C_1.

(22) Update associated with (15-b)

$$\lambda i_s.\lambda j_s.i[\phi, \upsilon]j \;\wedge\; \phi_{DC_S}(j) \subseteq \phi(j) \;\wedge\; \upsilon(j) = sue_e \wedge$$

$$\forall w[w \in \phi(j)](sleep_{e(wt)}(\upsilon(j))(w))$$

As discussed briefly above, the interpretation of S's utterance of *Sue sleeps* specifies a relation between an input assignment i_s and an output assignment j_s, that updates i to j by introducing two new drefs: a propositional dref ϕ and an individual dref υ. It will further impose conditions on the interpretation of these discourse referents in relation to the output assignment j. The possible values of $\phi(j)$ constrain the set of worlds $\phi_{DC_S}(j)$, representing the set of worlds compatible with S's discourse commitments. Further, $\upsilon(j)$ is equated with the value of the discourse-constant $Sue_{\mathbf{e}}(j)$, which rigidly refers to the entity sue_e across all variable assignments. The relation $sleep_{et}(\upsilon(j))(w)$ is required to hold at all $w \in \phi(j)$, thereby constraining possible values of $\phi(j)$ to sets of possible worlds where the proposition expressed by the sentence is true.

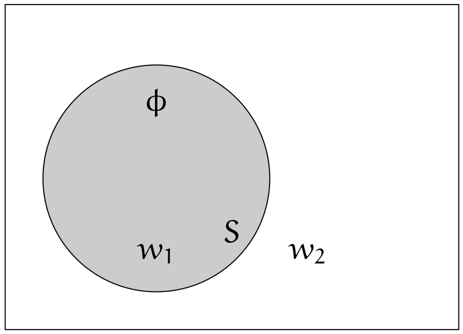
Since the input i has several possible values for ϕ_{DC_S}, we might expect the output j to have several possible values as well. However, there is only one possible discourse state that is an update of i, s.t. υ and ϕ are introduced, $\upsilon(j) = sue_e$, $\phi(j) = \{w \in D_w \mid sleep(sue)(w)\}$, and $\phi_{DC_S}(j) \subseteq \phi(j)$, given in **Table 3.6**.

	ϕ_{DC_S}	ϕ	υ
j	$\{w_1\}$	$\{w_1\}$	sue_e

Table 3.6: The mappings in the output

To illustrate the relations between the discourse referents graphically, the relationship between the propositional dref and the epistemic state of the speaker is depicted in **Figure 3.6**.

The location of the label S represents the speaker's epistemic location, which is within $\phi(j_1)$ in this case. That illustrates that the discourse commitments of the speaker $\phi_{DC_S}(j_1)$ are a subset of $\phi(j_1)$. Additionally, the gray shading indicates the set of worlds



$$\begin{aligned} \phi_{DC_S}(j) &= \{w_1\} \\ \phi(j) &= \{w_1\} \\ \upsilon(j) &= sue_e \end{aligned}$$

Figure 3.6: Diagram of the propositional drefs after updating with *'Sue slept.'*

associated with the proposition that the speaker is committed to being true.

The representation in (15), and the introduction of the two drefs, can accurately account for the anaphoric potential of the utterance in (15). In particular, it will allow us to account for the fact that both the referent of the DP *Sue*, as well as the proposition expressed by the full sentence, may be picked up anaphorically in the subsequent discourse, illustrated in (25).

(25) [Sue$^{\upsilon}$ slept]$^{\phi}$.

a. She$_{\upsilon}$ was very tired earlier. (*she* $\approx$ Sue)

b. I know that$_{\phi}$. (*that* $\approx$ that Sue slept)

3.4 Anaphora

In classic DRT (Kamp (1981); Kamp and Reyle (1993)), DRSs represent sentence meanings, but they also provide a representation of the discourse. Sentence representations are merged into the discourse representation, and semantic operators introduce structural relations over DRSs. The accessibility of drefs for subsequent anaphora is constrained by structural relations in a hierarchical DRS representation of the discourse. I am adopting a different notion of accessibility, which is based on flat update (Stone (1999); Stone and Hardt (1999); Brasoveanu (2006)). Following Muskens (1996), the representation is of the discourse is stored in current discourse state, i.e. a variable assignment, like in DPL (Groenendijk and Stokhof (1991)). Variable assignments provide a representation of the discourse state that stores information about semantic relations between the drefs and does not involve hierarchical structure. As a result, the interpretation of subsequent utterances, and anaphora included in them, is not directly influenced by previous utterances' syntactic structure but only by semantic relations that they introduce.

This section illustrates how the interpretation of anaphora works in this system, illustrated by means of a few simple examples. Anaphora are formalized as discourse variables that are interpreted in relation to the discourse state.

3.4.1 Anaphoric Co-reference

Here, I am assuming that an anaphor is co-referential with its antecedent if it introduce a variable with the same name as its antecedent, in other words: by means of a pre-semantic co-indexation as part of the object language. The anaphoric use of an index is indicated as a subscript in the NL sentence and as a red highlight in the CDRT translation:

	$\lambda w.\mathtt{sleep}(\mathtt{sue})(w)$	$\lambda w.\mathtt{tired}(\mathtt{sue})(w)$	$\lambda w.\mathtt{tired}(\#)(w)$
w_1	1	1	#
w_2	1	0	#
w_3	0	1	#
w_4	0	0	#

Table 3.7: Some relevant mappings specified by the semantic values of $\mathtt{sleep}_{e(wt)}$ and $\mathtt{tired}_{e(wt)}$ in M^{C_2}.

(26) a. *She$_\upsilon$ was tired.*

b.
$$\begin{array}{|c|}\hline \phi_2 \\ \hline \phi_{DC_S} \Subset \phi_2 \\ \mathrm{Tired}_{\phi_2}\{\upsilon\} \\ \hline \end{array}$$

c. $[\phi_2 \mid \phi_{DC_S} \Subset \phi_2]; [\mathrm{Tired}_{\phi_2}\{\upsilon\}]$

To illustrate how the interpretation of anaphora works, let us consider the following discourse, where (26) follows the utterance of *She slept*. The superscript on the DP *Sue* indicates that it introduces υ as a dref.

(27) a. *Sue$^\upsilon$ slept. She$_\upsilon$ was tired.*

The pronoun *she* anaphorically refers to the previously mentioned *Sue*. To illustrate the effect of the updates involved, consider a context $C_2 = \langle M, \{S\}, i\rangle$, where M^{C_2} is s.t. $D_e = \{\mathtt{sue}_e, S_e, \#\}$ and $D_w = \{w_1, w_2, w_3, w_4\}$. Also assume a semantic value for $\mathtt{sleep}_{e(wt)}$ and $\mathtt{tired}_{e(wt)}$, for which I am specifying some relevant mappings in **Table 3.7**.

Assume that i is a discourse-initial state. That's a way of expressing that in C_2, no prevous assertions have been made, and no drefs have been introduced. **Table 3.8.** gives all 15 possible values[3] for the input i.

	ϕ_{DC_S}
i_1	$\{w_1, w_2, w_3, w_4\}$
i_2	$\{w_1, w_2, w_3\}$
i_3	$\{w_1, w_2, w_4\}$
i_4	$\{w_1, w_3, w_4\}$
i_5	$\{w_2, w_3, w_4\}$
i_6	$\{w_1, w_2\}$
i_7	$\{w_1, w_3\}$
i_8	$\{w_1, w_4\}$
i_9	$\{w_2, w_3\}$
i_{10}	$\{w_2, w_4\}$
i_{11}	$\{w_3, w_4\}$
i_{12}	$\{w_1\}$
i_{13}	$\{w_2\}$
i_{14}	$\{w_3\}$
i_{15}	$\{w_4\}$

Table 3.8: Possible values for the initial state i in C_2

Now, for our discourse in (27-a), the sentence containing the anaphor *She was tired* is interpreted following an update with *Sue slept* ($[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\upsilon \mid \upsilon \equiv \mathsf{Sue}]; [\mathsf{Sleep}_{\phi_1}\{\upsilon\}]$). This first update proceeds as discussed in the previous section. For the update in C_2 with input i, the possible output values j are given in **Table 3.9**.

Table 3.9: Possible values for the discourse state after updating C_2 with $[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\upsilon \mid \upsilon \equiv \mathsf{Sue}]; [\mathsf{Sleep}_{\phi_1}\{\upsilon\}]$

	ϕ_{DC_S}	ϕ_1	υ
j_1	$\{w_1, w_2\}$	$\{w_1, w_2\}$	sue_e
j_2	$\{w_1\}$	$\{w_1, w_2\}$	sue_e
j_3	$\{w_2\}$	$\{w_1, w_2\}$	sue_e
j_4	$\{w_1\}$	$\{w_1\}$	sue_e
j_5	$\{w_2\}$	$\{w_2\}$	sue_e

In this updated context (let's call it $C_2' = \{M_4, \{S\}, j\}$), the drefs υ and ϕ_1 can be referred

[3] $\phi_{DC_S}(i)$ can refer to any non-empty, well-defined subset of D_w (/functions from $D_w \mapsto D_t$). The number of functions $|D_w \mapsto D_t| = |D_t|^{|D_w|}$. The number of well-defined functions from $D_w \mapsto D_t$ is $|\{0,1\}|^{|D_w|}$. The number of non-empty, well defined subsets (a.k.a. functions that map any member in the domain to $\{0,1\}$, but not all to 0). is $\sum_{k=1}^{|D_w|} \binom{|D_w|}{k}$. In this case: $\sum_{k=1}^{4} \binom{4}{k} = 15$.

to anaphorically, because the discourse state assigns them a referent. An update of C_2' with (26) in this context has the following effects: **(i)** It introduces a dref ϕ_2 referring to a set of worlds where υ *was tired,* **(ii)** The condition $\phi_{DC_S} \Subset \phi_2$ further restricts S's commitments to be compatible with ϕ_2. **(iii)** It relies on the discourse state to provide a referent for υ and requires that υ sleeps in all worlds in ϕ_1. The update abbreviated by the DRS (26) is given in (28).

(28) Update associated with (26):

$$\lambda j_s.\lambda k_s.j[\phi_2]k \ \wedge \ \phi_{DC_S}(k) \subseteq \phi_2(k) \wedge$$
$$\forall w[w \in \phi_2(k)](\mathtt{tired}_{e(wt)}(\upsilon(k))(w))$$

Because the discourse state in $C_2\prime$, namely j, is an assignment, which maps υ to $\mathtt{sue}_e$, the variable name υ can be used subsequently with the same reference. Anaphora are modelled by reusing a previously introduced discourse variable. Because k is an successor of j, and $\upsilon(j) = \mathtt{sue}$, the same holds for $\upsilon(k)$. The update (28) in C_2' therefore holds of any output k that maps the drefs ϕ_{DC_S}, ϕ_2 and υ in a way, s.t. $\phi_{DC_S}(k)$ is a subset of $\phi_2(k)$, and $\upsilon(k)$ sleeps in each world in $\phi(k)$. The possible values for k are given in **Table 3.10**.

The update (28) in C_2' is discourse-acceptable because there exists at least one output k s.t. $[\![[\phi_2 \mid \phi_{DC_S} \Subset \phi_2]; [\mathtt{Tired}_{\phi_2}\{\upsilon\}]]\!]^{M_4}(j)(k) \neq \#$. In fact, this is the case for all possible values of k listed in **Table 3.10**, and any other possible successor l of j (i.e. any l s.t. $l = j$, or for any dref name δ: $j[\delta]i$).

	ϕ_{DC_s}	ϕ_1	υ	ϕ_2
k_1	$\{w_1\}$	$\{w_1, w_2\}$	sue_e	$\{w_1, w_3\}$
k_2	$\{w_1\}$	$\{w_1, w_2\}$	sue_e	$\{w_1\}$
k_3	$\{w_1\}$	$\{w_1\}$	sue_e	$\{w_1, w_3\}$
k_4	$\{w_1\}$	$\{w_1\}$	sue_e	$\{w_1\}$

Table 3.10: Possible values for the discourse state after updating C_2' with $[\phi_2 \mid \phi_{DC_S} \Subset \phi_2, \mathtt{Tired}_{\phi_2}\{\upsilon\}]$

The update (28) in C_2' is discourse-true, because there exists at least one output k s.t. k s.t. $[\![[\phi_2 \mid \phi_{DC_S} \Subset \phi_2]; [\mathtt{Tired}_{\phi_2}\{\upsilon\}]]\!]^{M_4}(j)(k) = 1$. This is the case for the possible output values listed above.

3.4.2 Licensing

On the flipside, we can account for why the use of anaphora fails when no suitable referents have been established. To illustate this, consider an update with (26) in the discourse-initial context C_2 from above.

Specifically, the update with (26) in C_2 fails for the interpretation of the condition $\mathtt{Tired}_{\phi_1}\{\upsilon\}$. To localize the update failure, let's consider the equivalent representation as a series of conjunctive atomic updates:

(29) $[\phi_1]; [\phi_{DC_S} \Subset \phi_1]; [\mathtt{Tired}_{\phi_1}\{\upsilon\}]$

Updating with $[\phi_1]$ in C_2 leads to a random explosion of possible output values. Possible values for an output j include all possible combinations of some value for ϕ_{DC_S} already present in i, and a possible mapping for ϕ_1 (at this point, that can be any object of

	ϕ_{DC_S}	ϕ_1
k_1	$\{w_1, w_2, w_3, w_4\}$	$\{w_1, w_2, w_3, w_4\}$
k_1	$\{w_1, w_2, w_3\}$	$\{w_1, w_2, w_3, w_4\}$
k_1	$\{w_1, w_2, w_3\}$	$\{w_1, w_2, w_3\}$
...		

Table 3.11: Some possible values for k

type wt). As a result, there are 1215 possible values for j,[4] which aren't listed here. A subsequent update with $[\phi_{DC_S} \Subset \phi_1]$ restricts possible values for an output state k to those combinations of mappings s.t. $\phi_{DC_S}(k)$ is a subset of $\phi_1(k)$. This substantially cuts down our the number of possible outputs k, but at 65[5], they are still too many to list here. **Table 3.11** gives a few possible values. Possible values for k are those that map ϕ_{DC_S} and ϕ_1 to subsets of D_w, s.t. $\phi_{DC_S}(k) \subseteq \phi_1$.

Crucially, k does not specify a referent for υ, and $\upsilon(k) = \#$, for any value of k. That is why the attempted update with $[\mathtt{Tired}_{\phi_1}\{\upsilon\}]$ fails. This DRS abbreviates the following update:

(30) $\lambda i_s.\lambda j_s.i = j \wedge \forall w((w \in \phi_1(j) \to (\mathtt{sleep}_{e(wt)}(\upsilon(j))(w)))$

Let's see what properties an output l needs to have to satisfy this update, given the input k.

(31) a. Update of k with (30):

$[\lambda i_s.\lambda j_s.i = j \wedge \forall w((w \in \phi_1(j) \to (\mathtt{sleep}_{e(wt)}(\upsilon(j))(w)))](k)(l) =$

$k = l \wedge \forall w[w \in \phi_1(l)](\mathtt{sleep}(\upsilon(l))(w))$

[4]Possible combinations of 15 possible values for $\phi_{DC_S}(i)$ and $|D_w \mapsto D_t| = |D_t|^{|D_w|} = 3^4 = 81$ possible values for $\phi_1(j)$, i.e. $15 \times 81 = 1215$ possible values.

[5]Number of possible outputs k: $\sum_{n=1}^{|D_w|} \binom{|D_w|}{n} (2^n - 1) = \sum_{n=1}^{4} \binom{4}{n} (2^n - 1) = 65$

b. Truth and definedness-conditions:

$$[\![k = l \wedge \forall w[w \in \phi_1(l)](sleep_{e(wt)}(\upsilon(l))(w))]\!]^{M^{C_2}} =$$

•1 iff

–$[\![k = l]\!]^{M^{C_2}} = 1$ and for all $w \in D_w$:

$[\![\phi_1(l)]\!]^{M^{C_2}}(w) = 0$ or

$[\![sleep(\upsilon(l))]\!]^{M^{C_2}}(w) = 1$

•# iff

–$[\![k = l]\!]^{M^{C_2}} = \#$ or for some $w \in D_w$:

$[\![\phi_1(l)]\!]^{M^{C_2}}(w) = \#$ or

$[\![sleep(\upsilon(l))]\!]^{M^{C_2}}(w) = \#$

(31-b) states that all possible values for l are those that were also possible values for k, and in addition satisfy the condition that for each world w in the domain of worlds D_w, w is not a member of the proposition $\phi(l)$, or a world, in which it is true that $\upsilon(l)$ sleeps. For concreteness, consider the above k_1. Here, the interpretation function will return #, because $\upsilon(l) = \#$, and for any world w in D_w, $[\![sleep]\!]^{M^{C_2}}(\#)(w) = \#$.

(32) For $k = k_1$:

$[\![k_1 = l \wedge \forall w[w \in \phi_1(l)](sleep_{e(wt)}(\upsilon(l))(w))]\!]^{M_4} = \#$, because

•For all $d \in D_w$: $[\![sleep]\!]^{M_4}(\#)(d) = \#$

We would get the same result for any other possible value of k, because none of them map υ to anything other than #. As a result, the update (26) is unacceptable in C_2

because for any possible value of l: $[\![[\phi_1];[\phi_{DC_S} \in \phi_1];[\mathrm{Tired}_{\phi_1}\{\upsilon\}](i)(l)]\!]^{M^{C_2}} = \#$.

The framework guarantees that an update D involving a pronoun with an index υ will fail when no discourse referent υ was introduced previously and is stored in a current discourse state i. Any such an update in i will be 'undefined', and any non-υ-successor j of i (i.e. any state j s.t. $j = i$, or for any dref name $\delta : \delta \neq \upsilon \wedge i[\delta]j$) will be s.t. $D(i)(j) = \#$. The use of an indeterminate third value here captures the intuition that the use of pronouns without a suitable antecedent leads to unacceptability.

3.5 Assertion and Maximization

The models used above to illustrate the effects of updates are toy models with a quite restricted domain of worlds. They were designed to yield only one compatible output for the illustrated updates, in which the dref associated with the asserted proposition refers to a singleton set. When we consider somewhat richer domains of worlds, things get more complicated.

In cases where there are several possible output values, and the dref associated with an asserted proposition may refer to a proposition containing multiple worlds, some of the possible outputs may have undesirable properties. Specifically, some of them will be truth-conditionally too restrictive. This subsection illustrates and addresses this issue.

The possible outputs that are truth-conditionally too restrictive come with unwarranted inferences based on the information available in the discourse. This section illustrates this issue and suggests maximization over the discourse commitment sets as a possible solution. However, this runs into undesired consequences as well because this move would disallow any future assertion. The solution will involve a defeasible mechanism of maximization.

An overview of the issues and suggested solutions is given in **Table 3.12**, while the rest of this section illustrates them in some detail.

Cause	**Problem**	**Fix**	**Problem Solved?**
Non-Deterministic Implicit Epistemic Update	Unwarranted inferences possible	Maximization	We have a new problem now
Maximization over commitment sets	No subsequent assertions possible	Pragmatic defeasible maximization	Yes

Table 3.12: Overview of the issues and suggested solutions in this section

3.5.1 Unwarranted Inferences and Maximization

Assume a context $C_3 = \langle M_3, \{S\}, i \rangle$, where i is a discourse-initial state and M_3 is s.t. $D_e = \{\mathsf{sue}, S, \#\}$, $D_w = \{w_1, w_2, w_3, w_4\}$, and $[\![\mathsf{sleep}]\!]^{M_3}$ is s.t. it characterizes the mappings in **Table 3.13**.

w	$\lambda w.\mathsf{sleep}(\mathsf{sue})(w)$	$\lambda w.\mathsf{sleep}(\#)(w)$
w_1	1	#
w_2	1	#
w_3	0	#
w_4	0	#

Table 3.13: Truth-values for $\lambda w.\mathsf{sleep}(\mathsf{sue})(w)$, $\lambda w.\mathsf{sleep}(\mathsf{mary})(w)$ for the worlds in D_w

Now, consider an update with an assertion by S of *Sue slept*. As stated above, the associated update, repeated here, will have the effect that the discourse commitments of the speaker in the output are restricted to a non-empty well-defined set of worlds, s.t. Sue slept in that world.

(33)
$$\boxed{\begin{array}{c} \phi, \upsilon \\ \hline \phi_{DC_S} \Subset \phi \\ \upsilon \equiv Sue_{se} \\ Sleep_{\phi}\{\upsilon\} \end{array}}$$

The problem that we are facing here is that this update is compatible with an output j that maps ϕ_{DC_S} *any* non-empty well-defined set of worlds, s.t. Sue slept in that world. That includes the maximal set of all worlds where that is the case, but also any non-empty well-defined subset. The possible mappings for the drefs in j are given in **Table 3.14,** for all possible values of j.

	ϕ_{DC_S}	ϕ	υ
j_1	$\{w_1, w_2\}$	$\{w_1, w_2\}$	sue
j_2	$\{w_1\}$	$\{w_1, w_2\}$	sue
j_3	$\{w_2\}$	$\{w_1, w_2\}$	sue
j_4	$\{w_1\}$	$\{w_1\}$	sue
j_5	$\{w_2\}$	$\{w_2\}$	sue

Table 3.14: Possible output values for j after updating C_3 with (15-b)

The table shows that there is a possible output j_1, s.t. $\phi_{DC_S}(j) = \{w_1, w_2\}$, the set including *all* worlds where Sue slept. However, there are also possible outputs j_2 and j_3, where $\phi_{DC_S}(j)$ specifies that Sue slept, but in addition Mary slept, or Mary didn't sleep, respectively. While it is a welcome result that S is committed that Sue slept, for any output j, it is undesirable that we also get possible outputs that accidentally commit S to a proposition that wasn't expressed in the utterance. The result that we do want is that

S's commitment set is the maximal set of worlds that is compatible with the update (and any previous commitments). We need some additional assumptions to achieve this.

We might ensure that ϕ_{DC_S} is the maximal set of worlds compatible with the update (and any previous commitments) semantically or pragmatically: The semantic route might involve including a maximization operator into the representation of an assertive update. This would mean that the only possible mapping for ϕ_{DC_S} is the maximal set of worlds compatible with the previous commitments and the update. The pragmatic route might involve introducing an assumption that assertion is epistemically conservative, and that among all possible interpretations, the intended one will be one that involves as little new information as possible.

Let's consider the first option, the introduction of a maximization operator, which might be formalized as in (34), modeled after the selective maximization operator in Brasoveanu (2006).

(34) $\mathbf{max}_{\phi}(D) :=$

$\lambda i.\lambda j.D(i)(j) \wedge \forall k[D(i)(k)](\neg(\phi(j) \subset \phi(k)))$

(34) takes a propositional dref ϕ and an update D and makes sure that (1.) the input i is updated with D to yield the output j, and (2.) that for any other possible update of i with D yielding some other possible output k, no $\phi(k)$ is strictly larger than $\phi(j)$. This means the max-operator takes an update (and a dref to maximize over) and returns a conditional update. Placing conditions on the update is different from DRS-conditions,

which place conditions of the output of an update rather than an update itself. The operator is externally dynamic since the variable updates and conditions specified in D still apply globally. The condition placed on the update states that the output is such that no other possible output maps ϕ to a strictly larger set.

Now, we might add this operator to our representation of (15-b) to state that we only want the output that applies this update and maps $\phi_{\mathrm{DC_S}}$ to the largest possible set.

(35) $$\mathbf{max}_{\phi_{\mathrm{DC_S}}}\left(\begin{array}{|c|}\hline \phi, \upsilon \\ \hline \phi_{\mathrm{DC_S}} \Subset \phi \\ \upsilon \equiv \mathsf{Sue}_{se} \\ \mathsf{Sleep}_{\phi}\{\upsilon\} \\ \hline \end{array}\right)$$

The update with (35) only has one possible output, i.e. the output value j_1, where $\phi_{\mathrm{DC_S}}(j_1) = \{w_1, w_2\}$. Any output mapping $\phi_{\mathrm{DC_S}}$ to a subset of this set will not survive.

3.5.2 No Subsequent Assertions

Although using maximization in this way allows us to eliminate the implicitly possible discourse values for $\phi_{\mathrm{DC_S}}$ that led to the problem of having undesired unreasonably strong inferences. However, the complete absence of these possibilities also causes a problem. As a result of maximization over the commitment set $\phi_{\mathrm{DC_S}}$, its value will be fixed to this one value for the entire discourse. This is an undesirable result since we will not be able to further narrow down the associated set with any subsequent utterance. Assume that following the above utterance of *Sue slept*, S also wants to assert that *Mary* *slept*.

Considering the update without maximization (15-b) with the possible output values in **Table 3.14**, we can see that this is possible: The set where it is true that *Sue slept* and that *Mary slept* $\{w_1\}$ is still a possible value for the set of S's discourse commitments. Therefore, a subsequent update can further constrain the dref ϕ_{DC_S} to refer only to a set of this kind. If we rule out all non-maximal values entirely, a future update with *Mary slept* (or any future update further narrowing down S's commitments) would fail because the new conditions would conflict with the previously enforced maximality-requirement. The semantic route, therefore, seems to be the wrong approach here. We want the semantic representation to allow for possible future interpretations where the set of the speaker's commitments is smaller than the maximal set satisfying a current one. In fact, the presence of multiple possible values for ϕ_{DC_S}, which are exactly the largest possible set compatible with S's commitments, and all its subsets, is what allows us to model the kind of epistemic update assumed to go along with assertion. That is because future utterances will allow for ϕ_{DC_S} to be constrained to one of these subsets.

Therefore, we do want to stick to a semantic representation that allows us to keep all the possible output values associated with an update like (15-b), and add a mechanism by which we might conversationally privilege one possible output over the others in a manner that is defeasible.

The pragmatic route seems intuitively reasonable from a Gricean perspective, and at first glance, it seems like a Gricean account could be pretty straightforward. The Maxim of Quantity (Grice (1975)) states that a speaker will make their utterance as informative as possible, given the information that is available to them. Therefore, if the information about whether or not Mary slept is not expressed explicitly, the interlocutors will assume that S has made no commitment about whether Mary slept. Accordingly, it looks like we might derive this inference as a quantity-implicature.

A way to treat implicatures in this system might be as defeasible conditions on the output state, which privilege those of the possible output states that satisfy them.

(36) Epistemically conservative update:

For an utterance by a Speaker S in context C, the update D:

a. Yields possible output values j (Assertion):

$j \in \lambda j_s.[\![C(i^C)(j)]\!]^{M^C}$

b. And, by implicature, privileges one of the possible output states:

$\forall k_s.[[\![C(i^C)(k)]\!]^{M^C}](\neg(\phi_{DC_S}(j) \subset \phi_{DC_S}(k)))$

(37) Accordingly, an assertion of *Sue slept* with the semantic representation (15-b) by a Speaker S in context C:

a. Yields possible output values j (Assertion):

$$j \in \boxed{\begin{array}{c} \phi, \upsilon \\ \hline \phi_{DC_S} \Subset \phi \\ \upsilon \equiv Sue_{se} \\ Sleep_{\phi}\{\upsilon\} \end{array}}(i^C)$$

b. And, by implicature, privileges one of the possible output states:

j_i is conversationally privileged iff:

$$\forall k. \left[\!\!\left[\begin{array}{|c|}\hline \phi, \upsilon \\ \hline \phi_{DC_S} \Subset \phi \\ \upsilon \equiv Sue_{se} \\ Sleep_{\phi}\{\upsilon\} \\ \hline \end{array} \right]\!\!\right](i^C)(k) \; (\neg \phi_{DC_S}(j_1) \subset \phi_{DC_S}(k))$$

(36-b) states that j_i is a privileged output state, iff for any other output k obtained by updating with D in C, $\phi_{DC_S}(k)$ is not a subset of $\phi_{DC_S}(j_i)$. In other words, no other possible output maps ϕ_{DC_S} to a larger set.

The max-implicature in (36-b) has a similar effect to the semantic max-operator defined above: Following an assertive update, it compares the output with other possible outputs of the same update and checks whether any of them assign a strictly larger set to ϕ_{DC_S}. The main difference here is that it does not prohibit outputs where that is the case. Since this implicature is defeasible, it does not semantically constrain the possible output values for the update. Instead, it privileges one output state as the preferred one for the purposes of the conversation. This is illustrated in **Table 3.15**, where the privileged state is indicated with a star symbol, and non-privileged states are in gray text.

A mechanism like this has some desirable features: It allows the implicature to be defeated. The non-privileged outputs are still available; they are just conversationally dispreferred. Further, relying on Gricean reasoning allows for an explanation based on a theory rooted in basic intuitions about conversation.

	ϕ_{DC_S}	ϕ	υ
$\star j_1$	$\{w_1, w_2\}$	$\{w_1, w_2\}$	sue
j_2	$\{w_1\}$	$\{w_1, w_2\}$	sue
j_3	$\{w_2\}$	$\{w_1, w_2\}$	sue
j_4	$\{w_1\}$	$\{w_1\}$	sue
j_5	$\{w_2\}$	$\{w_2\}$	sue

Table 3.15: Possible output values for j after updating C_4 with (15-b)

There are some challenges to this Gricean kind of approach, like the question of how the kind of analysis roughly sketched here might interact with other implicatures.[6] The goal of presenting some formalization here is to show that a concrete formalization is possible, rather than giving a complete account of implicatures or quantity-based reasoning or providing an argument for one or another analysis. We might imagine

[6]One thing that I might point out is this: If there are other implicatures (formalized as defeasible conditions on the output state), the maximality-inference would need to take scope over them. Let's say an utterance of *It's warm today* in context C implicates that it's not hot today. Say $D_w^C = \{w_1, w_2, w_3, w_4\}$ (It is cold in w_1, neither cold nor warm in w_2, warm in w_3, w_4, and hot in w_4.) Without the max-implicature taking scope over the first one, we get:

(i) a. Assertion:

$$j \in \begin{array}{|c|}\hline \phi \\ \hline \phi_{DC_S} \Subset \phi \\ \mathrm{warm}_\phi \\ \hline \end{array}(i^C)$$

b. Implicature (not hot today) privileges some j_i iff:

$$j_i \text{in} \sim \begin{array}{|c|}\hline \phi' \\ \hline \mathrm{hot}_{\phi'} \\ \hline \end{array}$$

c. Implicature (Max) privileges some j_i iff:

$$\forall k. \left[\begin{array}{|c|}\hline \phi \\ \hline \phi_{DC_S} \Subset \phi \\ \mathrm{warm}_\phi \\ \hline \end{array}(i^C)(k) \right] (\neg \phi_{DC_S}(j_1) \subset \phi_{DC_S}(k))$$

This results in conflicting privileged output states:

	ϕ_{DC_S}	
$\star j_1$	$\{w_3, w_4\}$	Privileged due to (i)c
$\star j_2$	$\{w_3\}$	Privileged due to (i)b
j_3	$\{w_4\}$	

Only if the statement in (c) takes scope over the implicature in (b), this can yield good results:

(ii) c. Implicature (Max) privileges some j_i iff:

$$\forall k. \left[\begin{array}{|c|}\hline \phi \\ \hline \phi_{DC_S} \Subset \phi \\ \mathrm{warm}_\phi \\ \hline \end{array} ; \begin{array}{|c|}\hline \sim \begin{array}{|c|}\hline \phi' \\ \hline \mathrm{hot}_{\phi'} \\ \hline \end{array} \\ \hline \end{array}(i^C)(k) \right] (\neg \phi_{DC_S}(j_1) \subset \phi_{DC_S}(k))$$

	ϕ_{DC_S}	
j_1	$\{w_3, w_4\}$	
$\star j_2$	$\{w_3\}$	Privileged due to (i)b+(ii)c
j_3	$\{w_4\}$	

This footnote is meant to show how the max-implicature might interact with other implicatures. If it cannot take scope over other implicatures, it will always privilege the largest set of worlds that satisfies the explicit content of the update. This cannot explain why the implicature *It is not hot today* can be derived.

an alternative analysis that does not rely on formalizations of implicature but instead conversationally privileges those possible values for commitment states that contain the least amount of information compared to other possible values. The amount of information contained in a set of worlds might be quantified in terms of Shannon's (1948) information-theoretic notion of information-content (or information entropy) calculated in relation to the domain of all possible worlds.

Exploring the consequences of theories of the pragmatic reasoning involved in (defeasibly) maximizing the set of worlds representing an interlocutor's commitments or a comparison of these theories falls out of the scope of the work presented here. The points from this discussion that will be relevant are **(i)** that this kind of maximization-inference is necessary to get the desired results regarding the representations of assertive updates, and **(ii)** that the inference is defeasible and should be dealt with on the level of pragmatic reasoning, rather than semantic representation.

Having discussed this, I will leave the maximization-inference associated with assertion implicit in the following and assume that some pragmatic mechanism privileges the possible output that maps any commitment set to a maximal set that otherwise satisfies the conditions placed on it in the discourse.

3.6 Compositionality

The compositional semantics, i.e., the mapping between the form of the utterance and its meaning, characterizes the possible CCPs associated with the utterance of a sentence with a given syntactic representation (based on Muskens (1996); Brasoveanu (2006, 2010a)).

Following Brasoveanu's (2006, 2010a) conception of Muskens's (1996) CDRT, subclausal semantic composition in a Montegovian/Fregean sense follows from the underlying type logic. However, there are different possible assumptions about how exactly the compositional semantics could be explicated. For instance, different assumptions could be made for the mode of composition and scope-taking mechanisms of quantifiers, or modes of composition more generally. The semantic system presented in the previous sections could be associated with a compositional system somewhat independently of the particular choices.

In this section, I will make explicit one way of providing a compositional system, in order to provide concrete semantic translations for natural language expressions, and localize of the semantic contribution of particular NL expressions.

A Montegovian compositional system makes use of semantic types for individuals and sentences, which we might abbreviate using the 'meta-types' **e** and **t** (following Brasoveanu (2006)). An intensional system in addition makes reference to intensional content, i.e. specifications of possible worlds, which we might abbreviate as **w**. In an intensional static logic, **e** corresponds to type e, **w** corresponds to type w, and **t** to type t. The English verb *sleeps* can be analyzed as an intensional property of individuals (type **e(wt)**). Accordingly the NL expression *sleeps* might receive a translation into a static intensional meta-language as in (38):

(38) Static intensional translation of unary predicates:

$$\text{sleeps} \rightsquigarrow \lambda x_e.\lambda w_w.sleep_{e(wt)}(x)(w).$$

The dynamic system presented here differs from this static intensional compositional system in two ways: First, in (38) the compositional type **w**, the intensional context of evaluation, is treated as a singleton world of evaluation, which may differ from the global world of evaluation. Instead, we will use sets of worlds to provide an intensional context of evaluation. This treatment is in line with semantic theories that treat local intensional contexts as sets of worlds that can differ from the global context of evaluation (following Heim (1992)). This is based on a conception of the global context of evaluation as the context set, i.e., an epistemic representation associated with the speech participants. The original proposal of this kind of context set comes from Stalnaker (1978), who assumed that it was a representation of the interlocutors' common ground, i.e., their publically shared beliefs. Here, it is taken to represent each speaker's public communicative commitments in the sense of Gunlogson (2004).

The result of this adjustment is that the compositional meta-type **w** stands in for sets of worlds of type wt. Whereas a relation in an intensional system with singleton worlds of evaluation (38) will be true iff the relation holds at the world of evaluation, now a relation needs to hold at each world in the intensional context set. The updated translation of a unary predicate is given in (39).

(39) Static version with intensional context sets:

$$\text{sleeps} \rightsquigarrow \lambda x_e.\lambda W_{wt}.\forall w[w \in W](sleep_{e(wt)}(x)(w)).$$

The intensions in our dynamic system are based on sets of worlds, analogously to (39), but the types are lifted to dynamic types. The compositional meta-type **t** will correspond to dynamic utterance meanings, i.e. type $s(st)$, **e** will correspond to individual drefs of type se, and **w** to propositional drefs of type $s(wt)$. NL expressions corresponding to unary predicates may now receive translations as in (40).

(40) Dynamic version with intensional context sets:

$$\text{sleeps} \rightsquigarrow \lambda \upsilon_{se}.\lambda \phi_{s(wt)}.\lambda i_s.\lambda j_s.i = j \wedge \forall w_w[w \in \phi(j)](sleep_{e(wt)}(\upsilon(j))(w)).$$

Using the dynamic abbreviations defined in this chapter, we can abbreviate (40) as (41), for increased readability:

(41) Dynamic intensional translation of unary predicates:

$$\text{sleeps} \rightsquigarrow \lambda \upsilon_{\mathbf{e}}.\lambda \phi_{\mathbf{w}}.[Sleep_{\phi}\{\upsilon\}]$$

I will use the notion of sentence radical to refer to the semantic representation of the smallest syntactic constituent in a clause that includes all lexical material, i.e., the verb and its arguments. Formally, that is the smallest constituent in a clause which is translated to a dynamic proposition of type **wt**.

The verb *sleep* is formalized as a dynamic intensional property of entities (type **e**(**wt**)). It takes an individual dref υ, and a propositional dref ϕ as arguments and predicates that the static relation $sleep_{\text{et}}$ holds of the individual referred to by υ in all the worlds in the set referred to by ϕ.

The translation of a DP (in this version of a system the only DPs are proper names) compositionally functions as a dynamic quantifier over individuals (**e**(**wt**),(**wt**)).

(42) $\text{Sue}_{\mathbf{e(wt),(wt)}} \rightsquigarrow$

$\lambda P_{\mathbf{e(wt)}}.\lambda\phi_{\mathbf{w}}.[\upsilon_{\mathbf{e}} \mid \upsilon \equiv Sue_{\mathbf{e}}]; P(\upsilon)(\phi)$

(where $Sue_{\mathbf{e}} := \lambda i_s.sue_e$)

A DP is translated as a function taking a dynamic intensional property $P_{\mathbf{e(wt)}}$ and a dref for a set of worlds $\phi_{\mathbf{w}}$ as arguments. A proper name in particular introduces an individual dref, which is equal to the discourse constant associated with the name. In addition, it introduces the condition that that the property P holds of υ in the ϕ-worlds.

The lexical translations for intransitive verbs and DPs, and the assumption of a binary branching syntactic structure allow for deriving the meaning associated with the sentence radical in *Sue slept*, as in (43).

(43) Compositional derivation of the dynamic proposition expressed in (15)

$\lambda\phi_{\mathbf{w}}.[\upsilon_{\mathbf{e}} \mid \upsilon \equiv Sue_{\mathbf{e}}]; [Sleep_{\phi}\{\upsilon\}]$

Sue — $\lambda P_{\mathbf{e(wt)}}.\lambda\phi_{\mathbf{w}}.[\upsilon_{\mathbf{e}} \mid \upsilon \equiv Sue_{\mathbf{e}}]; P(\upsilon)(\phi)$

sleep — $\lambda\upsilon_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[Sleep_{\phi}\{\upsilon\}]$

The translation of the sentence radical, i.e., the propositional content of the clause expressed by the smallest constituent containing all the lexical content, is a dynamic proposition of type **wt**. This dynamic proposition can combine with sentential mood to derive a dynamic formula of type $s(st)$.

Compositionally, declarative sentential mood is of type **(wt),t**, and therefore maps a dynamic proposition to a dynamic formula. This is needed to derive the type of an update from a dynamic proposition by saturating the intensional argument.

(44) Declarative mood:

$$\textsc{dec}_S := \lambda \mathcal{P}_{\mathbf{wt}}.[\phi_{1\mathbf{w}} \mid \phi_{DC_S} \Subset \phi_1]; \mathcal{P}(\phi_1)$$

The declarative mood operator is indexical to the speaker S and takes a dynamic proposition $\mathcal{P}_{\mathbf{w,t}}$ as its argument to return an update. It introduces a propositional dref $\phi_{\mathbf{w}}$, and the condition that ϕ is entailed by S's commitments. Further, it states that this dref ϕ refers to a proposition satisfying the conditions specified in the update $\mathcal{P}(\phi)$.

Combining the dynamic proposition from (43) with declarative mood, we yield our familiar representation, the dynamic formula (/DRS) in (45).

(45) a. $\textsc{dec}_S\ (\lambda\phi_{\mathbf{w}}.[\upsilon \mid \upsilon \equiv Sue_{\mathbf{e}}]; [Sleep_{\phi}\{\upsilon\}]) =$

$[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\upsilon \mid \upsilon \equiv Sue_{\mathbf{e}}]; [Sleep_{\phi_1}\{\upsilon\}]$

b.

ϕ_1, υ
$\phi_{DC_S} \Subset \phi_1$ $\upsilon \equiv Sue_{\mathbf{e}}$ $Sleep_{\phi_1}\{\upsilon\}$

While this should give the reader an impression of what my assumptions about the compositional system are, a complete list of the compositional translations for NL expressions assumed in this chapter is given in **Appendix A**.

3.7 Conclusion

Let's take stock of what we have accomplished so far. This chapter has introduced a dynamic semantics for the interpretation of natural language utterances suitable for analyzing the truth conditions and anaphoric potential of simple natural language utterances. **Section** 3.2 introduced the dynamic semantic framework of intensional CDRT. **Section** 3.3 introduced the linking assumptions needed to apply this framework as an analysis of the truth-conditions and anaphoric potential of natural language utterances. **Section** 3.4 illustrated the treatment of anaphora in a system like this. **Section** 3.5 introduces maximization over propositional drefs, which operates defeasibly on a pragmatic level for the drefs that are associated with interlocutors' commitment states. **Section** 3.6 briefly sketched how the assumed semantic representations can be derived compositionally. All formal definitions, dynamic abbreviations, and assumed lexical representations can be found in **Appendix** A.

We are now in a position to address propositional anaphora. The system presented in this chapter already includes a propositional dref for the main assertion and a way of relating it to the speaker's commitments: A declarative assertion comes with the condition that the asserted proposition is a subset of the speaker's commitments. The following chapter will give a semantics for propositional operators and discourse referents introduced in embedded contexts along similar lines. Propositional operators introduce a dref for their prejacent and specify its relation to the superordinate context. That enables an analysis of propositional drefs that captures their relation to the speaker's commitment set and thereby provides a basis for stating the constraints on anaphoric accessibility in terms of veridicality.

Chapter 4

Propositional Operators and Propositional Drefs

4.1 Introduction

All analyses in this work are based on the framework introduced in the previous chapter. To account for propositional anaphora, we need to add a semantics for propositional operators, a mechanism of introducing propositional anaphora, and a way of capturing their veridicality in relation to the speaker. We will add these components in this chapter, proceeding in the following way:

Section 4.2 outlines two proposals in the literature about the natural language expressions introducing propositional discourse referents. I discuss some reasons for adopting Snider's (2017) account, which assumes that propositional operators introduce discourse referents referring to the propositions in their scope.

Section 4.3 discusses the semantics of propositional operators assumed here. They introduce a local intensional context and lexically specify a relation between this local context and the embedding context, as in Heim (1992). This relation will be formalized as a relation between propositional drefs. This provides the main ingredients for the proposed treatment of propositional anaphora and an explanation of the constraints on their availability in terms of veridicality. This section also discusses a maximization operator assumed for propositional operators and briefly discusses some possible assumptions about the compositional derivation of semantic representations of utterances with propositional operators.

Section 4.4 discusses some additional issues in the treatment of propositional drefs: It discusses why the system in its current form cannot handle quantification over individuals in the scope of propositional operators. The limitation is due to a problem that arises from the interaction of propositional and individual drefs. The problem will be illustrated for unspecific indefinites in the scope of negation. This problem will be addressed in **Chapter 7**, which extends the analysis to individual anaphora.

Section 4.5 concludes this chapter by summarizing its main points.

Since the version presented in this chapter cannot handle quantification over individuals, all examples here involve predication over individual expressions whose referent is constant, like proper names. That allows us to focus on what the current system can do: It includes a representation of propositional drefs and relations between them, based on which we may understand the (im-)possibilities for propositional anaphora in terms of veridicality.

4.2 Introducing Propositional Drefs

While it is well established in dynamic semantics that individual drefs are introduced by DPs, there has been some discussion about how propositional drefs are introduced. I am reviewing two proposals here from Krifka (2013) and Snider (2017) and some arguments for the latter account, which is the one I will be adopting here. It holds that propositional drefs are introduced by propositional operators for the content of their prejacent.

4.2.1 Krifka (2013): Anaphora to Propositions in the Scope of Negation

Discussing propositional drefs introduced under negation, Krifka (2013) assumes that (syntactically) negation selects for a TP and projects a NegP, where both TP and NegP introduce a propositional dref into the discourse. This is illustrated in (1)[1]:

[1]This is a simplified representation of the structure assumed by Krifka. He also assumes that drefs are introduced by the *v*P, associated with an event representation, and by a speech-act-level phrase, associated with a representation of the speech act. For our present purposes, only the propositional drefs associated with the negated proposition and the proposition in the scope of negation are relevant.

(1) Anaphoric potential of negative sentences, based on Krifka (2013)

$[_{\text{NegP}}$*Ede didn't* $[_{\text{TP}}$ t_{Ede} t_{did} *steal the cookie*]

$\hookrightarrow \phi_1$ $\hookrightarrow \phi_2$

Here, the dref ϕ_1 is associated with the proposition *Ede didn't steal the cookie* and the dref ϕ_2 with the proposition *Ede stole the cookie.* A subsequent propositional anaphor could be interpreted as picking up either of these drefs.

Krifka draws attention to the fact that negated propositions introduce propositional drefs into the discourse and makes use of it to analyze the semantics of polarity particles as propositional anaphora. As mentioned above, I am adopting the assumption that a negative sentence introduces a propositional dref for the negated content. Krifka suggests that the dref is introduced by the TP expressing the proposition. Following the discussion in Snider (2017), I will dub this assumption *TP-hypothesis.*

(2) **TP-hypothesis**:

Propositional drefs are introduced by constituents that are associated with propositional interpretations, such as TPs or larger constituents.

The presented account builds on the kind of semantic representations assumed for negative utterances by Krifka (2013), along the lines of (1), as well as his analysis of polarity particles as propositional anaphora. However, I am departing from the assumption in (2) about how the representations relate to the form of these utterances. Instead, I follow Snider (2017) in assuming that propositional drefs are introduced by propositional operators (as opposed to clauses). A summary of Snider's arguments is provided in the following subsection.

4.2.2 Snider (2017): Anaphora to Propositions in the Scope of Propositional Operators

In his work on propositional anaphora, Snider (2017) investigates the introduction of drefs corresponding to embedded propositions more generally. He concludes that the question when subsentential constituents introduce propositional drefs cannot be addressed by a generalization about the syntactic category of these constituents but rather by a semantic generalization about the operators embedding them.

Snider discusses Krifka's TP-hypothesis, stating that any TP or larger constituent introduces a propositional dref and argues that it could not capture when an embedded proposition is available for subsequent anaphoric reference. Instead, he suggests that a propositional dref is introduced by an utterance precisely when the utterance involves a propositional operator (along with its argument). The result is a generalization in terms of semantic type rather than syntactic category. He suggests that propositional operators introduce propositional drefs for their prejacents.

The arguments against the syntactic TP-hypothesis are based on **(i)** cases where constituents smaller than a TP introduce propositional drefs and **(ii)** cases where constituents of size TP or larger do not. I will summarize some of these arguments here. Some cases that Snider brings up of propositional drefs introduced by constituents that are smaller than TP involve epistemic small clauses (c.f. Snider (2017), see also Wilder (1992)), as well as epistemic DP or VP adverbials:

(3) Epistemic small clause construction:

Context: Francine and Rosa were wed in an airport casino, officiated by an Elvis impersonator. [Snider (2017): 152]

The clerk considered Francine and Rosa married, but I don't think that's true.

a. #that the clerk considered Francine and Rosa married

b. ✓that Francine and Rosa are married

(3) shows that the propositional anaphor *that* can be interpreted as *that Francine and Rosa are married,* the proposition expressed by the small clause *Francine and Rosa married*. Snider chose *but I don't think that's true* as a follow-up to the first clause to rule out reference of the propositional pronoun to the complete first conjunct. This effect relies on Moore's paradox (Moore (1942)), which states that an assertion of the form 'p, but I don't think that p' is unacceptable. As a result, *that* in (3) cannot refer to *that the clerk considered Francine and Rosa married.* This highlights the interpretation under consideration, where *that* refers to the proposition *that Francine and Rosa are married*, associated with the sub-sentential small clause embedded by *consider.*

Other sub-sentential constituents that can be antecedents for propositional anaphora when in the scope of a propositional operator include DPs and VPs:

(4) Epistemic DP adverbials: [after Snider (2017): (162)]

Dustin moved a seemingly heavy box, but I didn't believe that.

a. #that Dustin moved a box

b. #that the box was seemingly heavy

c. ✓that the box was heavy

(5) Epistemic VP adverbials: [Snider (2017): (168)]

Steve allegedly fled, but I don't believe that.

a. #that Steve allegedly fled

b. ✓That Steve fled

(4) shows that the epistemic adverbial *seemingly* makes available a propositional dref for the phrase it modifies, and (5) illustrates the same for the adverb *allegedly.*

Altogether, the examples in (3)–(5) illustrate that antecedent sentences including the embedders *consider, seemingly,* and *allegedly* make available propositional drefs associated with the constituents in their scope. This shows that the TP-hypothesis is too restrictive since it predicts that propositional anaphora to sub-sentential antecedents should not be possible here.

Snider argues that instead, the propositional anaphora are licensed here due to the embedders taking a propositional argument. That is supported by the contrast with the small clauses, DP and VP adverbials in (6)–(8). These examples are structurally parallel to (3)–(5), but involve embedders that do not take a propositional argument:

(6) Resultative small clause construction: [Snider (2017): 153]

Context: Francine and Rosa were wed in a Baptist church.)

The pastor pronounced Francine and Rosa married, # but I don't think that's true.

(7) *Joyce's recently single nephew went on a date, and she told Hopper that.*

[Snider (2017): 156]

a. ✓that Joyce's nephew went on a date

b. #that Joyce's nephew is recently single

c. #Joyce's nephew is single

(8) *Steve immediately fled, but Nancy doesn't believe that.* [Snider (2017): 166]

a. ✓that Steve immediately fled

b. #That Steve fled

In (6)–(8), the propositional anaphor can only refer to the proposition introduced by the full antecedent clause, and not to a proposition corresponding to any smaller constituent. The contrast between (6)–(8) and (3)–(5) can be associated with the semantics of the embedding predicate: while *consider, seemingly,* and *allegedly* select for a propositional argument, *pronounce, recently,* and *immediately* do not.

The other part of Snider's argument against the TP-hypothesis involves TP-size (or larger) constituents, which do not introduce propositional drefs. Among these are embedded non-finite TPs in subject-raising constructions.

(9) Subject control with *try*: [Snider (2017): 246]

Nancy tried to be at the party, but that wasn't true.

a. #that Nancy tried to be at the party

b. #that Nancy was at the party.

(9) involves an example of a TP that does not license subsequent propositional anaphora. The TP-hypothesis is too broad to predict the unavailability of propositional anaphora here. Snider's generalization about propositional operators, on the other hand, can predict this: Since *try* does not select for a propositional argument, no propositional dref is introduced here. This is supported by the contrast with (10), a structurally parallel example of a subject control clause embedded under the verb *claim*, which does take a propositional argument:

(10) Subject control with *claim*: [Snider (2017): 245]

Nancy claims to be at the party, but that isn't true.

a. #that Nancy claims to be at the party

b. ✓that Nancy is at the party.

In conclusion, Snider provides compelling arguments for the assumption that the introduction of propositional drefs is part of the meaning contribution of propositional operators: Semantic operators that select for propositional arguments introduce a dref for the proposition in their scope. For a detailed discussion of further examples and arguments, see Chapter 3 of Snider (2017).

4.2.3 Stone and Hardt (1999): Externally Dynamic Negation

Based on the kinds of considerations discussed in the previous subsection, Snider adopts a semantics for negation along the lines of Stone and Hardt (1999), and we will do the same. According to Stone and Hardt, negation introduces a dref for a counterfactual scenario and a relation between this dref and the scenario corresponding to the embedding contexts (see also Stone (1999)). This formalization of negation is externally dynamic because it makes available a propositional dref for the content in its scope.

Drefs for scenarios are of type $(sw), (wt)$. They map to a set of worlds, given a variable assignment, and in relation to a set of possible worlds. This provides a way to identify a set of worlds based on a possible world. As such, we can think of scenario drefs as propositional drefs that are relativized to possible worlds or as drefs storing modal accessibility relations (see Brasoveanu's (2006) discussion of accounts along these lines). This allows for scenario drefs to store dependencies between possible worlds.

For the purposes discussed here, we will go with a slightly simpler representation and assume that negation (and other propositional operators) introduce propositional drefs of type $s(wt)$ (like it is also assumed in Brasoveanu (2006); Snider (2017)). We already have a way of storing information about the relations over drefs in the discourse: Through imposing conditions on the discourse state.

That will be sufficient for the purpose of capturing the relationship between propositional drefs and the speaker's epistemic state and characterizing the possibilities for propositional anaphora. The rest of this chapter illustrates this mechanism, while the following chapter illustrates its application to analyze the patterns of propositional anaphora. A more detailed discussion of drefs relativized to possible worlds will be provided in **Chapter 7**, in which we will be using Stone-style drefs for individuals, which are relativized to possible worlds.

4.3 Semantics of Propositional Operators

Following Stone (1999); Stone and Hardt (1999); Snider (2017), I assume that propositional operators introduce a dref for their propositional argument(s), while the propositional dref associated with the full clause is introduced by sentential mood. Propositional operators introduce intensional arguments of evaluation that may be distinct from the global intensional argument of evaluation. In dynamic semantics this has often been treated as a set of worlds that is distinct from the matrix context set (following Heim (1992)).

These contributions are intricately connected: The propositional dref introduced, storing a set of worlds, provides the context of evaluation for the material of the prejacent. This allows for a local intensional context that is distinct from the global context, and also guarantees that the dref refers to the set of worlds specified by the prejacent. The truth-functional relation associated with the operator is formalized as a relation between propositional drefs: The embedded one and a superordinate one.

In summary, propositional operators are assumed to have the following contributions: They **(i)** introduce propositional drefs for their arguments, **(ii)** introduce a relation between the introduced drefs and the dref corresponding to the superordinate context, and **(iii)** provide the newly introduced drefs as a local context of evaluation for their prejacent(s).

Propositional operators differ in how many propositional arguments they take, and what relation they specify between embedded propositions and the matrix context of evaluation.

In the following, the semantics of negation, disjunction, and conjunction is discussed. I will discuss the anaphoric potential of sentences including these operators, to illustrate that we in fact want them to introduce propositional drefs for their arguments, the assumed semantic representations, and the effects of the updates associated with them, and the veridicality of the material in their scope, in particular in relation to our extended notion of veridicality for discourse referents. We do this here for these operators to illustrate the overall mechanism, and how updates with veridical, non-veridical and anti-veridical operators works. The effect of updating with utterances containing each of these operators will be important for the analyses in **Chapters 5&7**.

4.3.1 Negation

The introduced version of intensional CDRT with propositional drefs will allow us to give a formal semantic analysis of the interaction between negation and propositional anaphora. A significant consideration throughout the investigations of these interactions is the anaphoric potential of utterances of sentences including negation, and the effect that negation has on subsequent anaphoric possibilities.

Anaphoric Potential and Truth-Conditions

Like with a positive sentence *'Sue slept'*, an utterance of the negative counterpart *Sue didn't sleep* allows for subsequent anaphoric pronominal reference to the proposition expressed by the whole sentence (11-a), and to the prejacent of negation (11-b)

(11) [Sue didn't[~~Sue~~ sleep]$^{\phi_2}$]$^{\phi_1}$.

a. I know that$_{\phi_1}$. (*that* $\approx$ that Sue slept)

b. Even though she told me that$_{\phi_2}$. (*that* $\approx$ that Sue slept)

The anaphoric possibilities for negative utterances are captured by the assumption that sentences with negation introduce two propositional drefs: One corresponding to the whole negated sentence, and another corresponding to the proposition that is in the scope of negation (following Stone and Hardt (1999); Krifka (2013) and Snider (2017)).

What is central to the interpretation of negation and anaphora to drefs introduced in the scope of negation is the semantic relation between the two propositional drefs. Since one is the negation of the other, the propositions that these drefs refer to are complements of each other. The truth-functional meaning of negation will be captured as a relation between these two discourse referents. In particular, negation introduces the condition that they are complements.

Semantic Representation

An utterance of the sentence *Sue didn't sleep* by a speaker S (12-a), receives the discourse representation in (12-b):

(12) a. S: *Sue didn't sleep.*

b.
$$\begin{array}{|c|}\hline \phi_1, \phi_2, \upsilon \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \phi_1 \equiv \overline{\overline{\phi_2}} \\ \upsilon \equiv \mathsf{Sue}_{\mathbf{e}} \\ \mathsf{Sleep}_{\phi_2}\{\upsilon\} \\ \hline \end{array}$$

c. $[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid \phi_1 \equiv \overline{\overline{\phi_2}}]; [\upsilon \mid \upsilon \equiv \mathsf{Sue}_{\mathbf{e}}]; [\mathsf{Sleep}_{\phi_2}\{\upsilon\}]$

Three things are different here in comparison to the DRS associated with an affirmative sentence in the previous subsection: An additional propositional dref ϕ_2 is introduced, the condition $\mathsf{Sleep}_{\phi_2}\{\upsilon\}$ is interpreted wrt this new dref ϕ_2, and the condition $\phi_1 \equiv \overline{\overline{\phi_2}}$ is added.

This representation will be interpreted s.t. three new drefs are introduced: two propositional drefs ϕ_1, and ϕ_2, and an individual dref υ. Also, four conditions are imposed on the output discourse state: The condition $\upsilon \equiv \mathsf{Sue}_{\mathbf{e}}$ makes sure that υ refers to Sue. The condition $\mathsf{Sleep}_{\phi_2}\{\upsilon\}$ ensures that Sue slept in all the ϕ_2-worlds so that the propositional dref ϕ_2 corresponds to the proposition in the scope of negation. The condition $\phi_1 \equiv \overline{\overline{\phi_2}}$ states that the sets of worlds stored in the drefs ϕ_1 and ϕ_2 are complements of each other, with the result that ϕ_1 refers to the set of worlds in which it is false that *Sue slept*.[2] That introduces the semantics of negation as a relation between propositional drefs and ensures that the dref ϕ_1 corresponds to the proposition expressed by the utterance content. The condition $\phi_{DC_S} \Subset \phi_1$ ensures that ϕ_1 is true in all the worlds where the S's discourse commitments are true.

The condition $\phi_1 \equiv \overline{\overline{\phi_2}}$ makes use of dynamic complementation, a dynamic operator over drefs introducing set complementation of their referent.

(13) a. Dynamic complementation:

$$\overline{\overline{\phi}} := \lambda i_s.\overline{\phi(i)}$$

Accordingly, the condition guarantees that the set of worlds in which the prejacent holds and the set of worlds corresponding to the embedding context are disjoint and exhaust the domain of possible worlds, introducing the semantics of contrary negation.

In line with the assumptions laid out in this chapter, the presence of negation contributes a new propositional dref (ϕ_2), and a condition that specifies its relation to the proposition of the embedding context ($\phi_1 \equiv \overline{\overline{\phi_2}}$). The newly introduced dref ϕ_2 provides a local intensional context for the interpretation of the prejacent. That is, is provides the intensional argument in the condition $\mathtt{Sleep}_{\phi_2}\{v\}$. This also has the effect of constraining ϕ_2 to worlds where the condition holds. This guarantees that it stores the propositional content of the prejacent.

[2]To be precise, maximization over the set of worlds that ϕ_2 refers to is needed to get the right truth conditions. At this point, there is no mechanism to guarantee this yet, but this will be discussed and added later in this chapter.

The Effect of the Update

To illustrate how the update in (12-b) is interpreted in context, we assume a context $C_1 = \langle M_1, \{S\}, i \rangle$, where M_1 is s.t. $D_e = \{sue, S, \#\}$, $D_w = \{w_1, w_2\}$, $[\![sleep(sue)(w_1)]\!]^{M_1} = 1$ and $[\![sleep(sue)(w_2)]\!]^{M_1} = 0$. i^{C_1} is a discourse-initial state.

The semantic representation in (12-b) abbreviates the dynamic update in (14).

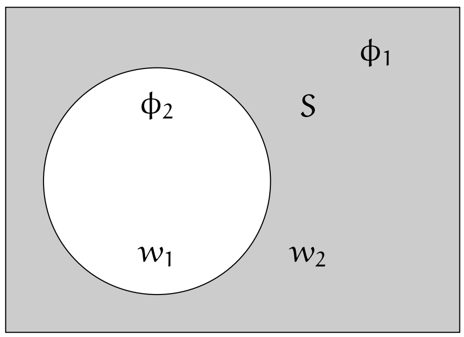
(14) Update associated with (12-b):

$$\lambda i_s.\lambda j_s.i[\phi_1, \phi_2, \upsilon]j \ \wedge \ \phi_{DC_S}(j) \subseteq \phi_1(j) \ \wedge \ \phi_1(j) = \overline{\phi_2(j)} \ \wedge \ \upsilon(j) = sue_e$$
$$\wedge \, \forall w[w \in \phi_2(j)](sleep_{e(wt)}(\upsilon(j))(w))$$

The interpretation of S's utterance of *Sue didn't sleep* specifies a relation between an input assignment i_s and an output assignment j_s, that updates i to j by introducing three new drefs: two propositional drefs ϕ_1 and ϕ_2 and an individual dref υ. It will further impose conditions on the interpretation of these discourse referents in relation to the output assignment j. The possible values of $\phi_1(j)$ constrain the set of worlds $\phi_{DC_S}(j)$, representing the set of worlds compatible with S's discourse commitments, to contain only worlds that are also in $\phi_1(j)$. The contribution of negation constrains $\phi_1(j)$ to contain the worlds that are not in $\phi_2(j)$, i.e., the worlds where the proposition expressed in the prejacent is false. The referent of υ at j is equated with the value of the discourse-constant Sue_{se}, which rigidly refers to the entity sue_e across all variable assignments.

Finally, the value of $\phi_2(j)$ is constrained to sets of possible worlds s.t. the relation $\mathtt{sleep}(\upsilon(j))(w)$ holds at all $w \in \phi_2(j)$.

Updating (14) in C_1 has a possible output j, which assigns static referents to drefs as in in **Figure 4.1**.



$$
\begin{aligned}
\phi_{DC_S}(j) &= \{w_2\} \\
\phi_1(j) &= \{w_2\} \\
\phi_2(j) &= \{w_1\} \\
\upsilon_1(j) &= \mathtt{sue}_e
\end{aligned}
$$

Figure 4.1: Diagram of the propositional drefs after updating with *‘Sue didn’t sleep.’*

In this diagram of the drefs and relations between them stored in j, the location of the worlds w_1 and w_2 within the area corresponding to $\phi_1(j)$ or $\phi_2(j)$ illustrates which propositions contain which sets in this toy model. The diagram indicates that $\phi_1(j)$ is the complement of the proposition expressed by the prejacent of negation $\phi_2(j)$. The gray shading indicates that $\phi_1(j)$ is the proposition expressed by the assertion, and coincides with the location of the label S represents the epistemic location of the speaker, which here is within $\phi_1(j)$, which illustrates that the discourse commitments of the speaker $\phi_{DC_S}(j)$ are a subset of $\phi_1(j)$.

The semantics of negation introduces a propositional dref which provides the intensional argument for the conditions in its scope. It also introduces a relation between the introduced propositional dref and the dref provided by the local embedding context. In the case of negation, this relation is one of non-overlap.

Veridicality

Empirically, we established that NL negation is anti-veridical. At this point, we can formally define veridicality for propositional drefs. Veridicality for propositional drefs is defined in relation to the epistemic state of an interlocutor in context. Intuitively, a propositional dref ϕ is *veridical* wrt an interlocutor A in a context C, iff the proposition referred to by ϕ is true wrt the discourse commitments of A in C.

(15) Veridicality for propositional drefs:

a. A propositional dref ϕ is veridical in $C = \langle i, M, \textsc{Int} \rangle$ wrt $A \in \textsc{Int}$ iff $\phi_{DC_A}(i) \subseteq \phi(i)$, and non-veridical otherwise.

b. A propositional dref ϕ is anti-veridical in $C = \langle i, M, \textsc{Int} \rangle$ wrt $A \in \textsc{Int}$ iff $\phi_{DC_A}(i) \subseteq \overline{\phi(i)}$.

In the example above, the updated the context is $C'_1 = \langle M_1, \{S\}, j \rangle$, where $[\![[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid \phi_1 \equiv \overline{\overline{\phi_2}}]; [\upsilon \mid \upsilon \equiv Sue_{\mathbf{e}}]; [Sleep_{\phi_2}\{\upsilon\}](i^{C_1})(j^{C'_1})]\!]^{M_1}$. ϕ_1 is veridical wrt S in C'_1, because $\phi_{DC_S}(j) \subseteq \phi_1(j)$. ϕ_2 is anti-veridical wrt S in C'_1, because $\phi_{DC_S}(j) \subseteq \overline{\phi_2(j)}$. The fact that negation is anti-veridical is reflected in the analysis because negation introduces an anti-veridical propositional dref.

4.3.2 Conjunction

Anaphoric Potential and Truth-Conditions

We want to give a Semantics for natural language conjunction, which captures its effect on the anaphoric potential and its truth-functional meaning like we did for negation. Let's consider the anaphoric possibilities for propositional anaphora under English conjunction with *and*. Conjunctive sentences license anaphora to the whole conjunction, as well as to each of the two conjuncts (Snider (2017)). This is shown in (16).

(16) Conjunction:

a. Anaphoric reference to the conjunction: [Snider (2017): 330]

[A water main burst and they closed off the highway.]$^{\phi_1}$

The announcer said that$_{\phi_1}$ *on the radio.*

b. Anaphoric reference to the second conjunct: [Snider (2017): 331]

John fell, and [it was Christina who pushed him]$^{\phi_2}$.

He didn't know that$_{\phi_2}$*, though*

c. Anaphoric reference to the first conjunct: [Snider (2017): 335]

[Steve got an A on the exam]$^{\phi_3}$*, and he told his mom that*$_{\phi_3}$.

Accordingly, we assume that a sentence with conjunction introduces three propositional drefs: One for the entire clause and one for each conjunct. Like with negation, the truth-functional component of the meaning of conjunction is captured as a relation over these drefs. In the case of conjunction, the proposition referred to by the matrix dref is the intersection of the propositions associated with the two conjuncts. This guarantees that the worlds in which the matrix conjunction holds include all worlds in which both conjuncts are true.

Semantic Representation

The utterance of a conjunction (17-a) receives the interpretation in (17-b)

(17) a. S: *Sue slept and Mary danced.*

b.
$$\begin{array}{|c|}\hline \phi_1, \phi_2, \phi_3, \upsilon_1, \upsilon_2 \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \phi_1 \equiv \phi_2 \Cap \phi_3 \\ \upsilon_1 \equiv \mathsf{Sue} \\ \mathsf{Sleep}_{\phi_2}\{\upsilon_1\} \\ \upsilon_2 \equiv \mathsf{Mary} \\ \mathsf{Dance}_{\phi_3}\{\upsilon_2\} \\ \hline \end{array}$$

c. $[\phi_1 \mid \phi_{DC_S} \Subset \phi_]; [\phi_2, \phi_3 \mid \phi_1 \equiv \phi_2 \Cap \phi_3]; [\upsilon_1 \mid \upsilon_1 \equiv \mathsf{Sue}]; [\mathsf{Sleep}_{\phi_2}\{\upsilon_1\}];$
$[\upsilon_2 \mid \upsilon_2 \equiv \mathsf{Mary}]; [\mathsf{Dance}_{\phi_3}\{\upsilon_2\}]$

In this representation, the contribution of conjunction has the effect that the two embedded drefs ϕ_2 and ϕ_3 are introduced. Further, the condition $\phi_1 \equiv \phi_2 \Cap \phi_3$ ensures that the matrix proposition ϕ_1 refers to the intersection of the propositions associated with the embedded propositions. The conditions $\mathsf{Sleep}_{\phi_2}\{\upsilon_1\}$ and $\mathsf{Dance}_{\phi_3}\{\upsilon_2\}$, representing the truth-conditional content of the first and second disjunct, are interpreted in relation to ϕ_2 and ϕ_3, respectively. The condition $\phi_1 \equiv \phi_2 \Cap \phi_3$ makes use of dynamic intersection, which is defined as in (18):

(18) Dynamic intersection:

$$\phi_1 \Cap \phi_2 := \lambda i_s.\phi_1(i) \cap \phi_2(i)$$

for $\phi_1, \phi_2 \in \textbf{Term}_{s(wt)}$

The Effect of the Update

To illustrate the effect of this update more concretely, assume a context $C_2 = \langle M_2, \{S\}, i\rangle$, where i^{C_2} is a discourse-initial state and M_2 is s.t. $D_e = \{\mathsf{mary}, \mathsf{sue}, S, \#\}, D_w = \{w_1, w_2, w_3, w_4\}$, and $[\![\mathsf{sleep}]\!]^{M_2}$ and $[\![\mathsf{dance}]\!]^{M_2}$ characterize the mappings in **Table 4.1**.

	$\lambda w.\mathsf{sleep}(\mathsf{sue})(w)$	$\lambda w.\mathsf{dance}(\mathsf{mary})(w)$
w_1	1	1
w_2	1	0
w_3	0	1
w_4	0	0

Table 4.1: Some relevant mappings for $[\![\mathsf{sleep}]\!]^{M_2}$ and $[\![\mathsf{dance}]\!]^{M_2}$.

The semantic representation in (17-b) abbreviates the dynamic update in (19).

(19) Update associated with (17-b)

$$\lambda i_s.\lambda j_s.i[\phi_1, \phi_2, \phi_3, \upsilon_1, \upsilon_2]j \wedge$$

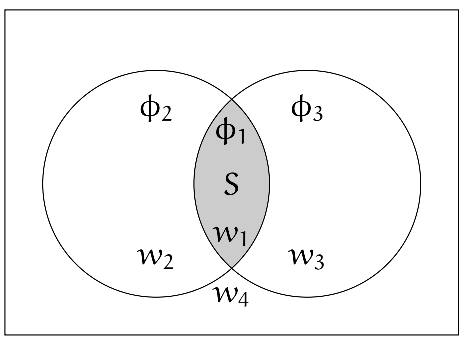
$$\phi_{DC_S}(j) \subseteq \phi_1(j) \ \wedge\ \phi_1(j) = \phi_2(j) \cap \phi_3(j) \wedge$$

$$\upsilon_1(j) = \mathsf{sue}_e \ \wedge\ \forall w[w \in \phi_2(j)](\mathsf{sleep}_{e(wt)}(\upsilon_1(j))(w)) \wedge$$

$$\upsilon_2(j) = \mathsf{mary}_e \ \wedge\ \forall w[w \in \phi_3(j)](\mathsf{dance}_{e(wt)}(\upsilon_2(j))(w))$$

S's utterance of *Sue slept and Mary danced.* is interpreted as an update of the discourse state i_s with the output j_s. (19) introduces three new propositional drefs for the conjunction and its conjuncts, and two individual ones for *Mary* and *Sue*. It also requires that the output j is such that the discourse commitments of the speaker entail $\phi_1(j)$, the set of worlds in which both conjuncts are true, that υ_1 is $mary_e$, and υ_2 is sue_e. Additional conditions on the output j are that in all ϕ_2-worlds, $sleep(mary_e)$ is true (the first conjunct), and in all ϕ_3-worlds, $dance(sue_e)$ is true.

The update with (19) in C_2 results in an output assignment j that assigns static referents to drefs as in **Figure 4.2**.



$$
\begin{aligned}
\phi_{DC_S}(j) &= \{w_1\}\\
\phi_1(j) &= \{w_1\}\\
\phi_2(j) &= \{w_1, w_2\}\\
\phi_3(j) &= \{w_1, w_3\}\\
\upsilon_1(j) &= sue_e\\
\upsilon_2(j) &= mary_e
\end{aligned}
$$

Figure 4.2: Diagram of the propositional drefs after updating with *'Sue slept and Mary danced.'*

The proposition expressed by the utterance's main clause at j ($\phi_1(j)$) is shaded in gray. It is the intersection of the propositions expressed by the two conjuncts $\phi_2(j)$ and $\phi_3(j)$. The location of the label S represents the speaker's epistemic location, which here is within $\phi_1(j)$, indicating that the discourse commitments of S are a subset of $\phi_1(j)$.

The truth-functional contribution of conjunction is captured because the proposition corresponding to the clause is an intersection of its disjuncts, the set of worlds where both are true. The anaphoric potential of conjunction is captured because the utterance of a conjunction introduces three propositional drefs.

Veridicality

NL conjunction is a veridical operator. In a semantics based on a static logic, this is reflected in the fact that static conjunction entails each conjunct:

(20) a. $(\phi_t \wedge \psi_t) \rightarrow \phi$

b. $(\phi_t \wedge \psi_t) \rightarrow \phi$

In our dynamic system, the veridicality of conjunction is reflected in the fact that an utterance of a sentence with conjunction introduces two veridical drefs for its conjuncts.

In the example above, the updated context is $C_2' = \langle M_2, \{S\}, j\rangle$, where $[\![(19)(i^{C_2})(j^{C_2'})]\!]^{M_2} = 1$. ϕ_1 is veridical wrt S in C_2', because $\phi_{DC_S}(j) \subseteq \phi_1(j)$. ϕ_2 and ϕ_3 are also veridical wrt S in C_2', because $\phi_{DC_S}(j) \subseteq \phi_2(j)$ and $\phi_{DC_S}(j) \subseteq \phi_3(j)$. The fact that conjunction is veridical is reflected in the analysis because it introduces two veridical propositional drefs for its conjuncts.

4.3.3 Disjunction

Anaphoric Potential and Truth-Conditions

Disjunctive sentences license propositional anaphora in a similar way to conjunctions: A propositional anaphor could pick up the proposition corresponding to the whole disjunction, as well as either disjunct, as illustrated in (21).

(21) Disjunction:

a. Anaphoric reference to the disjunction: [Snider (2017): 336]

[Nancy's birthday is approaching]

[Jonathan will buy Nancy flowers or he'll buy chocolates.]$^{\phi_1}$

She doesn't know that$_{\phi_1}$*, though.*

b. Anaphoric reference to the second disjunct: [Snider (2017): 337]

Steve cheated on the test, or [he got really lucky]$^{\phi_2}$.

He told the whole class that$_{\phi_2}$*, but I don't quite believe him.*

c. Anaphoric reference to the first disjunct: [Snider (2017): 339]

Either [Joyce won the lottery]$^{\phi_3}$*, or she wants everyone to believe that*$_{\phi_3}$.

We assume that disjunction introduces three propositional drefs: One for the full disjunction and one for each disjunct. The truth-functional meaning component of disjunction is contributed as a relation over these drefs. In the case of disjunction, the proposition referred to by the matrix dref is the union of the propositions associated with the disjuncts. This guarantees that the worlds in which the whole disjunction is true include all worlds in which either disjunct is true.

Semantic Representation

The utterance of a disjunction (22-a) receives the interpretation in (22-b)

(22) a. S: *Mary slept or Sue danced.*

b.
$$\begin{array}{|c|}\hline \phi_1, \phi_2, \phi_3, \upsilon_1, \upsilon_2 \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \phi_1 \equiv \phi_2 \Cup \phi_3 \\ \upsilon_1 \equiv \mathsf{Mary}_{\mathbf{e}} \\ \mathsf{Sleep}_{\phi_2}\{\upsilon_1\} \\ \upsilon_2 \equiv \mathsf{Sue}_{\mathbf{e}} \\ \mathsf{Dance}_{\phi_3}\{\upsilon_2\} \\ \hline \end{array}$$

c. $[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2, \phi_3 \mid \phi_1 = \phi_2 \Cup \phi_3]; [\upsilon_1 \mid \upsilon_1 \equiv \mathsf{Mary}_{\mathbf{e}}]; [\mathsf{Sleep}_{\phi_2}\{\upsilon_1\}];$
$[\upsilon_2 \mid \upsilon_2 \equiv \mathsf{Sue}_{\mathbf{e}}]; [\mathsf{Dance}_{\phi_3}\{\upsilon_2\}]$

In this representation, the contribution of disjunction has the effect that the two embedded drefs ϕ_2 and ϕ_3 are introduced. Further, the condition $\phi_1 = \phi_2 \Cup \phi_3$ ensures that the matrix proposition ϕ_1 refers to the union of the propositions associated with the embedded propositions. The conditions $\mathsf{Sleep}_{\phi_2}\{\upsilon_1\}$ and $\mathsf{Dance}_{\phi_3}\{\upsilon_2\}$, representing the truth-conditional content of the first and second disjunct, are interpreted in relation to ϕ_2 and ϕ_3, respectively.

The condition contributed by disjunction relies on dynamic union ($\Cup$) and abbreviates a property of a discourse state in the following way:

(23) Dynamic union:

$$\phi_1 \Cup \phi_2 := \lambda i_s.\phi_1(i) \cup \phi_2(i)$$

for $\phi_1, \phi_2 \in \mathbf{Term}_{s(wt)}$

The Effect of the Update

To illustrate the effect of the update in (22-b), we will consider an update with (22-b) in the previously defined context C_2. The semantic representation in (22-b) abbreviates the dynamic update in (24).

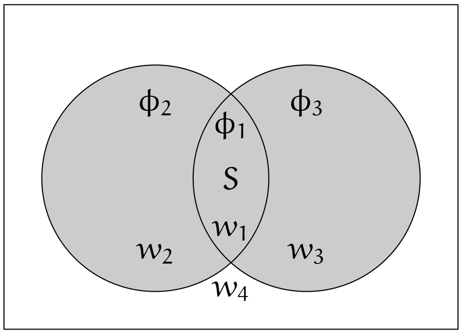
(24) Update associated with (22-b)

$$\begin{aligned}&\lambda i_s.\lambda j_s.i[\phi_1, \phi_2, \phi_3, \upsilon_1, \upsilon_2]j \wedge\\ &\quad \phi_{DC_S}(j) \subseteq \phi_1(j) \ \wedge \ \phi_1(j) = \phi_2(j) \cup \phi_3(j) \wedge\\ &\quad \upsilon_1(j) = mary_e \ \wedge \ \forall w[w \in \phi_2(j)](sleep_{e(wt)}(\upsilon_1(j))(w)) \wedge\\ &\quad \upsilon_2(j) = sue_e \ \wedge \ \forall w[w \in \phi_3(j)](dance_{e(wt)}(\upsilon_2(j))(w))\end{aligned}$$

Updating with (24) in C_2 results in an output assignment j that assigns static referents to drefs as in **Figure 4.3**.

In **Figure 4.3**, the proposition expressed by the main clause of the utterance at j $\phi_1(j)$ is shaded gray. It is the union of the propositions expressed by the two disjuncts $\phi_2(j)$ and $\phi_3(j)$. The discourse commitments of S are a subset of $\phi_1(j)$.

The truth-functional contribution of disjunction is captured because the proposition corresponding to the clause is a union of its disjuncts, the set of worlds where at least



$$
\begin{aligned}
\phi_{DC_S}(j) &= \{w_1, w_2, w_3\} \\
\phi_1(j) &= \{w_1, w_2, w_3\} \\
\phi_2(j) &= \{w_1, w_2\} \\
\phi_3(j) &= \{w_1, w_3\} \\
\upsilon_1(j) &= sue_e \\
\upsilon_2(j) &= mary_e
\end{aligned}
$$

Figure 4.3: Diagram of the propositional drefs after updating with *'Mary slept or Sue danced.'*

one of the disjuncts is true. The anaphoric potential of disjunction is captured because the utterance of a disjunction introduces three propositional drefs.

Veridicality

Natural language disjunction is non-veridical. In a semantics based on a static logic, this is reflected in the fact that static disjunction does not entail its disjuncts:

(25) a. $\neg((\phi_t \vee \psi_t) \rightarrow \phi)$

b. $\neg((\phi_t \vee \psi_t) \rightarrow \psi)$

Further, disjunction is not anti-veridical. In a static semantics, this is reflected in the fact that disjunction does not entail the negation of the disjuncts:

(26) a. $\neg((\phi_t \vee \psi_t) \rightarrow \neg\phi)$

b. $\neg((\phi_t \vee \psi_t) \rightarrow \neg\psi)$

In our dynamic system, the fact that disjunction is non-veridical, but not anti-veridical is reflected by the fact that an utterance of a sentence with disjunction introduces two drefs with these properties. In the example above, the updated context is $C_2'' = \langle M_2, \{S\}, j \rangle$, where $[\![(19)(i^{C_2})(j^{C_2''})]\!]^{M_2} = 1$. ϕ_1 is veridical wrt S in C_2'', because $\phi_{DC_S}(j) \subseteq \phi_1(j)$. ϕ_2 and ϕ_3 are non-veridical wrt S in C_2'', because $\phi_{DC_S}(j) \not\subseteq \phi_2(j)$ and $\phi_{DC_S}(j) \not\subseteq \phi_3(j)$. ϕ_2 and ϕ_3 are not anti-veridical wrt S in C_2'', because $\phi_{DC_S}(j) \not\subseteq \overline{\phi_2(j)}$ and $\phi_{DC_S}(j) \not\subseteq \overline{\phi_3(j)}$.

4.3.4 Maximization

Embedded Propositions

Despite the fact that the maximization over epistemic commitment states associated with assertion needs to be accounted for pragmatically (this was discussed in the previous chapter), we will have some use for a semantic maximization-operator. It is needed to derive the correct truth-conditions for utterances of sentences where propositional operators introduce propositional drefs for embedded propositions.

I will illustrate this here for the interpretation of negation. An utterance of the negated sentence *Sue didn't sleep* receives the representation in (12-b), repeated here:

(12-b)
$$\begin{array}{|c|}\hline \phi_1, \phi_2, \upsilon \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \phi_1 \equiv \overline{\overline{\phi_2}} \\ \upsilon \equiv Sue_e \\ Sleep_{\phi_2}\{\upsilon\} \\ \hline \end{array}$$

An update with (12-b) in the above defined context C_3 yields possible outputs, that have some value for the drefs ϕ_1, ϕ_2 and υ, where $\upsilon(j) = sue$, $\phi_2(j)$ is a set of worlds where

Sue slept and $\phi_1(j)$ is a set of worlds that is the complement of ϕ_2. Among the possible values for j are the one given in **Table 4.2**.

	ϕ_{DC_S}	ϕ_1	ϕ_2	υ
j_1	$\{w_3, w_4\}$	$\{w_3, w_4\}$	$\{w_1, w_2\}$	sue
j_2	$\{w_3\}$	$\{w_3, w_4\}$	$\{w_1, w_2\}$	sue
j_3	$\{w_4\}$	$\{w_3, w_4\}$	$\{w_1, w_2\}$	sue
j_4	$\{w_2, w_3, w_4\}$	$\{w_2, w_3, w_4\}$	$\{w_1\}$	sue
j_5	$\{w_1, w_3, w_4\}$	$\{w_1, w_3, w_4\}$	$\{w_2\}$	sue
j_6	$\{w_1, w_2, w_3, w_4\}$	$\{w_1, w_2, w_3, w_4\}$	$\varnothing$	sue
…				

Table 4.2: Some relevant possible values for the output of updating with (12-b) in C_3.

j_1 is the output that we want. It maps ϕ_2 to $\{w_1, w_2\}$, i.e. the largest set of worlds where *Sue sleeps*. It maps ϕ_1 to $\{w_3, w_4\}$, i.e. the largest set of worlds where *Sue doesn't sleep*. And it maps ϕ_{DC_S} to $\{w_3, w_4\}$, i.e. the largest set of worlds that commits S to ϕ_1, the proposition expressed in the assertion.

j_2 and j_3 are possible output values that map ϕ_2 to $\{w_1, w_2\}$, i.e. the largest set of worlds where *Sue sleeps*. In addition, they map ϕ_{DC_S} to a subset of $\phi_{DC_S}(j_1)$, and can therefore be disregarded based on the considerations laid out in the previous subsection.

The possible output values j_4–j_6, where ϕ_2 is mapped to a subset of $\{w_1, w_2\}$ present a problem. The problem is that this allows mapping ϕ_{DC_S} to a superset of $\phi_{DC_S}(j_1)$. j_4–j_6 should therefore be conversationally preferred over j_1. But at the same time, they relations between the drefs contradict what we think the truth-conditions of *Sue didn't sleep* should be: ϕ_{DC_S} contains w_3, w_4 here, i.e. the worlds where *Sue didn't sleep.* In addition, for all of these possible output states, it also contains at least some world where

Sue didn't sleep is false.

In particular j_6, mapping ϕ_{DC_S} to $\{w_1, w_2, w_3, w_4\}$, will be the conversationally preferred output value. No new commitments for S are added in this output state, so the truth-conditions of *Sue didn't sleep* are mischaracterized here. Further, ϕ_1 and ϕ_2 do not make available the the sets of worlds where *Sue didn't sleep* and *Sue slept*, respectively, so the anaphoric potential is mischaracterized as well. As a result, allowing ϕ_2 to a subset of $\{w_1, w_2\}$ causes problems.

In principle, we do want to maintain the possibility for the proposition in the scope of negation to be empty. Specifically, we want this to be possible in utterances of sentences where a contradiction is negated (e.g. *It's not the case that two plus two equals five.*) Crucially, though, in these cases there is no other possible (larger) set of worlds that satisfies the prejacent of negation. Only in this case do we want an embedded propositional dref to be empty. In general, we want the embedded propositional dref to be the largest possible one satisfying the prejacent. For embedded propositions in particular, anything else will derive the incorrect truth-conditions and anaphoric potential.

This is another instance, where we want to make use of maximization. However, in this case, maximization is required to get the right truth-conditions in the first place, and it is not a defeasible inference. Therefore, we are taking the semantic route in this case, and assume that the explicit introduction of a propositional dref by means of some linguistic expression, a propositional operator, is associated with semantic maximization, which we introduce in form of the max-operator (as defined in the previous chapter, repeated here).

(27) $\textbf{max}_{\phi}(D) :=$

$\lambda i.\lambda j.D(i)(j) \wedge \forall k[D(i)(k)](\neg(\phi(j) \subset \phi(k)))$

Based on (27), we assume that negation, which introduces a propositional dref for its prejacent, in addition places the condition that the dref introduces is a maximal dref satisfying the update corresponding to the prejacent.

Since **max** operates over updates, including it requires us to split the DRSs apart. The updated semantic representation of *Sue didn't sleep* is given in (28).

(28) a. $\begin{array}{|c|}\hline \phi_1, \phi_2 \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \phi_1 \equiv \overline{\phi_2} \\ \hline \end{array}\ ;\textbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \upsilon \\ \hline \upsilon \equiv \mathrm{Sue} \\ \mathrm{Sleep}_{\phi_2}\{\upsilon\} \\ \hline \end{array}\right)$

b. $[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid \phi_1 \equiv \overline{\phi_2}]; \textbf{max}_{\phi_2}([\upsilon \mid \upsilon \equiv \mathrm{Sue}]; [\mathrm{Sleep}_{\phi_2}\{\upsilon\}])$

Any output j derived with this update will map ϕ_2 to a maximal set s.t. *Sue slept* in $\phi_2(j)$. As a result $\phi_1(j)$ will contain only worlds where *Sue didn't sleep*, and the same holds of $\phi_{DC_S}(j)$. Updating the input state of the context C_4 with this yields the following possible outputs given in **Table 4.3**.

	ϕ_{DC_S}	ϕ_1	ϕ_2	υ
$\bigstar j_1$	$\{w_3, w_4\}$	$\{w_3, w_4\}$	$\{w_1, w_2\}$	sue
j_2	$\{w_3\}$	$\{w_3, w_4\}$	$\{w_1, w_2\}$	sue
j_3	$\{w_4\}$	$\{w_3, w_4\}$	$\{w_1, w_2\}$	sue

Table 4.3: The possible values for the output of updating with (28) in C_4.

Now, based on our pragmatic assumptions about conservative epistemic update, the preferred value will be j_1, and we do get the correct truth-conditions and anaphoric potential wrt propositional drefs for sentences with negation.

Asserted Propositions

We need to add maximization over propositional drefs in one more place: The propositional dref associated with the asserted content needs to be maximal as well. This is related to the issue regarding negated propositional drefs, and arises due to the fact that a propositional dref introduced by the asserted content can be picked up anaphorically and negated.

(29) A: [*Sue slept.*]$^{\phi}$

B: *That*$_{\phi}$*'s not true!*

The previous example showed that propositional drefs that are negated require maximization. Because the commitment set of the speaker is a subset of the asserted proposition ϕ, many pragmatic pressures that constrain ϕ_{DC-S} also apply to ϕ, in particular the constraints that it cannot be empty, it contain any undefined mappings, and that contains as little information as possible given the current state of the discourse will apply to the dref corresponding to the asserted proposition as well.

This last requirement also has the effect that the possible output value that maps ϕ to the maximal set of worlds satisfying the conditions will be privileged, however in a defeasible manner. In the case of negation of an anaphorically retreived proposition as in (29), the same issues arise as in the monosentential negation case–on a truth-conditional level. As a result, we want ϕ_1 to be maximal, non-defeasibly, and introduce semantic maximization over asserted propositions. As a result, we change the semantic representation for assertions as follows:

(30) $$\begin{array}{|c|}\hline \phi \\ \hline \phi_{\mathrm{DC}_s} \Subset \phi \\ \hline\end{array} \; ; \mathbf{max}_{\phi}\left(\begin{array}{|c|}\hline \upsilon \\ \hline \upsilon \equiv \mathit{Sue} \\ \mathit{Sleep}_{\phi}\{\upsilon\} \\ \hline\end{array}\right)$$

Since any asserted propositional dref can be anaphorically picked up and negated, we need maximization for asserted propositional drefs as well. Since this seems to be the case for any propositional dref that is introduced linguistically, we will add maximization for all linguistically introduced propositional drefs.

4.4 Limitations of ICDRT with Propositional Anaphora: Non-Veridical Indefinites

As noted above, this dynamic system has an issue with quantification over individuals. It mischaracterizes the anaphoric potential and, in some cases, the truth-conditions for sentences with non-specific indefinites. In particular, we made negation externally dynamic without changing our representation of individual drefs. In other words: The constraints on individual anaphora that other systems impose by making negated content dynamically inert are not enforced here. As a result, our system, as of now, can only capture (scopally) specific indefinites. This section illustrates some of these issues and discusses how they might be addressed by placing constraints on the interaction of individual drefs and the propositional drefs that constitute their local intensional context.

The solution proposed here is implemented until **Chapter 7**, which offers an analysis of the interaction of individual anaphora and negation.

For illustration of the problem, consider utterances of the following sentences with indefinites:

(31) a. *Sue owns a car.*

b. *Sue doesn't own a car.*

Traditionally, indefinites are analyzed as introducing existential quantification. In dynamic semantics, this is captured by means of random variable update (Heim (1982); Groenendijk and Stokhof (1991); Kamp and Reyle (1993)). (31-a) can be treated equivalently to the paraphrase 'There exists a car, s.t. Sue owns that car.' The dynamic semantic interpretation captures this by asking if there is at least one way in which we can update the discourse state with a variable dref υ, s.t. υ refers to an element $d \in D_e$ and d is a car that Sue owns.

Some issues arise as soon as we are in a system where both individual and propositional drefs can be introduced by linguistic expressions. Assuming that indefinites introduce existential quantification by means of random variable assignment, an utterance of 'There exists a car, s.t. Sue owns the car.' is true iff there is at least one way in which we can update the current discourse state with a propositional dref ϕ and an individual dref υ, s.t. υ refers to an element $d \in D_e$ and d is a car that Sue owns in each world in the proposition referred to by ϕ.

The issue is that any possible output picks out a single value for ϕ and υ and maintains that this value of υ is a car that Sue owns in all worlds of ϕ. As a result, the value of υ is the same for any world in ϕ. So the dref ϕ will store propositions that correspond to an interpretation of (31-a) with a specific indefinite. It does not allow us to yield an output where the proposition ϕ refers to the set of all worlds where Sue owns any car υ, independent of the specific value of υ. This leads to an incorrect characterization of the anaphoric potential.

4.4.1 Distributive and Collective Treatments of Possible worlds

Asher and McCready (2007) discuss a version of this problem for the related phenomenon of modal subordination. They point out the issue of how possible worlds (points of evaluation in their terms) are treated in the semantics. On the one hand, an account of the truth-functional component of modal operators requires a collective interpretation of sets of worlds. On the other hand, Asher and McCready argue that an account of the anaphoric potential of individual drefs introduced in the scope of modals (or propositional operators more broadly) requires a distributive component: It needs to define the effects of a sentence on each element of the context, the relevant component of the context being possible worlds (or world-assignment pairs in A&M).

Their prime example for the collective kind of interpretation over possible worlds is Veltman's (1987, 1996) test semantics for *might*, which works as a test on a(n epistemic) discourse context: It checks whether an epistemic information state would be compatible with a potential propositional update with the prejacent. As such, it dynamically specifies a relation between the input state and the prejacent of *might*. The crucial component here is the collective interpretation for sets of worlds, though, which is a property that is also satisfied by the dynamic system presented here by making reference to static relations between propositional discourse, which ultimately relate static sets of worlds to each other, although, crucially, collectively.

The prime examples A&M give for distributive dynamic update over possible worlds are the kind of distributive semantics for quantification found in DPL and DRT. The example they give for a distributive update over sets of worlds, in particular, makes reference to Roberts's (1989) analysis of modal operators in her account of modal subordination. This account compares all elements in the context of the antecedent with the elements of the local context of the prejacent of *would* to check if they satisfy certain properties. In the case of modal subordination, a crucial component of these properties involves the question of whether or not there is an available dref.

In summary, the truth-functional semantics of propositional operators is captured as a propositional relationship between the prejacent of the propositional operator and its superordinate context and requires a collective treatment of possible worlds. In contrast, the interpretation of anaphora and its antecedents in intensional contexts requires a component that is distributive over possible worlds.

The semantic system proposed here does combine collective and distributive operations on possible worlds. It involves relations between sets of worlds that are interpreted collectively: The truth-functional component of propositional operators is captured in terms of relations between sets of possible worlds. This is central for capturing the epistemic relationship between the prejacent and embedding context. Using Stone's (1999) mechanism of encapsulating universal quantification over possible worlds, the evaluation of DRT conditions is interpreted distributively: It looks at all members of the local context set of worlds and checks whether the condition is satisfied (for the given discourse state) in all of these worlds (see also the interpretation of relations in the final chapter of Dekker, 1993).

As A&M point out for modal subordination, one crucial component that needs to be checked distributively is the availability of a dref. Empirically, an antecedent in a non-veridical context is not available in a veridical context. So the correct interpretation requires an individual variable update that is distributed over the worlds in the local propositional context of the non-veridical operator. This is where the issues lie. The introduction of drefs and their subsequent anaphoric reuse should be evaluated distributively over the elements of the local propositional context (see also the selectively distributive variable update in accounts of discourse subordination in Brasoveanu, 2006, 2010a).

The system introduced here in its current state does interpret DRS-conditions distributively over possible worlds. However, individual drefs are introduced and evaluated wrt the discourse state, i.e., a variable assignment. This mechanism does not relate individual drefs to the possible worlds in their intensional context—distributively or otherwise. This gives rise to problems for the interaction of propositional and individual drefs, which are illustrated in some detail in this section.

4.4.2 No Scopally Unspecific Indefinites

I illustrate the problem here for utterances of the sentences in (31). While it manifests for simple cases of individual quantification with (epistemically[3]) unspecific indefinites in affirmative sentences (31-a), it is especially pervasive in the case of scopally unspecific indefinites in the scope of non-veridical operators, such as negation.

In such cases, the propositional dref associated with negated prejacent proposition cannot store a proposition with a non-specific indefinite interpretation. This results in a mischaracterization of the anaphoric potential and the truth conditions. The problem, illustrated in detail for this account, is a problem for any account involving quantification over propositions and individuals. For illustration, let us consider the effect of the updates associated with utterances of the sentences in (31) in context.

A context

Consider a context $C_4 = \{M_4, \{S\}, i\}$, where i is a discourse-initial state, and in M_4: $D_e = \{sue, car_1, car_2, S, \#\}$, $D_w = \{w_1, w_2, w_3, w_4\}$. **Table 4.4** gives the truth-values for $\lambda x.\lambda w.car(x)(w)$, for $x \in D_e, w \in D_w$ in M_4. Here, car_1 and car_2 are cars in every world and sue is not a car in any world.

Table 4.5 gives the truth values of $\lambda w.own(sue, car_1)(w)$ and $\lambda w.own(sue, car_2)(w)$

[3]For the distinction between epistemically vs. scopally unspecific indefinites see Farkas (1994) on (un)specific indefinites

w	$\lambda w.\mathsf{car}(\mathsf{car}_1)(w)$	$\lambda w.\mathsf{car}(\mathsf{car}_2)(w)$	$\lambda w.\mathsf{car}(\mathsf{sue})(w)$	$\lambda w.\mathsf{car}(\#)(w)$
w_1	1	1	0	#
w_2	1	1	0	#
w_3	1	1	0	#
w_4	1	1	0	#

Table 4.4: Truth-values for $\lambda x.\lambda w.\mathsf{car}(x)(w)$, for $x \in D_e^{M_4}$, $w \in D_w^{M_4}$ in M_4

for each $w \in D_W^{M_4}$ in M_4. Here, Sue owns both car_1 and car_2 in w_1. w_2 and w_3 are worlds where Sue owns just car_1 or car_2, respectively. In w_4, Sue owns neither car.

w	$\lambda w.\mathsf{own}(\mathsf{sue}, \mathsf{car}_1)(w)$	$\lambda w.\mathsf{own}(\mathsf{sue}, \mathsf{car}_2)(w)$
w_1	1	1
w_2	1	0
w_3	0	1
w_4	0	0

Table 4.5: Truth-values for $\lambda w.\mathsf{own}(\mathsf{sue}, \mathsf{car}_1)(w)$ and $\lambda w.\mathsf{own}(\mathsf{sue}, \mathsf{car}_2)(w)$ for $w \in D_W^{M_4}$ in M_4

For the sentences in (31), we want to capture the fact that that *Sue owns a car* is true in w_1 through w_3, and *Sue doesn't own a car* is true in w_4. This fact should be reflected on two levels: The discourse-relative truth conditions of the update, and the static propositions made available by the update.

Updating with (31-a)

Representing indefinites in terms of existential quantification modeled as random variable update, we get the following representation for (31-a):

(32) a. S: *Sue owns a car.*

b. $\boxed{\begin{array}{c}\phi \\ \hline \phi_{DC_S} \Subset \phi\end{array}}$; $\mathbf{max}_{\phi}\left(\boxed{\begin{array}{c}\upsilon_1, \upsilon_2 \\ \hline \upsilon_1 \equiv Sue \\ car_{\phi}\{\upsilon_2\} \\ own_{\phi}\{\upsilon_1, \upsilon_2\}\end{array}}\right)$

An update with (32-b) in C_4 yields the possible outputs j given in **Table 4.6**.

	ϕ_{DC_S}	ϕ	υ_1	υ_2
j_1	$\{w_1, w_2\}$	$\{w_1, w_2\}$	sue	car_1
j_2	$\{w_1, w_3\}$	$\{w_1, w_3\}$	sue	car_2

Table 4.6: (Pragmatically privileged) outputs for updating C_5 with (32-b).

An update with (32-b) characterizes the desired truth-conditions correctly: (32-b) is discourse true, given an input i, and a model just in case there is some assignment j, s.t. j is an update of i, s.t. ϕ, υ_1, and υ_2 are introduced, $\phi_{DC_S}(j)$ is a subset of $\phi(j)$, $\upsilon_1(j)$ is Sue, and $\phi(j)$ is the maximal set of worlds s.t. $\upsilon_2(j)$ is a car, and Sue owns the car in all ϕ-worlds. The truth-conditions are independent of the particular value of $\upsilon_2(j)$. Given context C, (32-b) is true, as long as there is at least one output j satisfying the update in relation to the input i and the model in the context. In C_4, this is the case.

Some Problematic Results

The static propositional content of the asserted propositional dref is mischaracterized: It becomes apparent that the possible output values in **Table 4.6** differ wrt the value assigned to the asserted proposition ϕ. In j_1, that's $\{w_1, w_2\}$, i.e. the proposition that Sue owns car_1. In j_2, that's $\{w_1, w_3\}$, i.e. the proposition that Sue owns car_2. Neither of these possible outputs yield the existential interpretation we might aim for when representing unspecific indefinites, namely the one where the asserted proposition is $\{w_1, w_2, w_3\}$, i.e. the maximal set of worlds where Sue owns some car in each world (indecently of which car that might be).

The maximization operator introduced in the previous subsection cannot help us here. The possible output values for ϕ are local maxima, because there is no other possible output value that is a superset of either of them. The interaction of propositional and individual drefs leads to a mischaracterization of the content of the dref for the asserted proposition, even for simple cases of unspecific indefinites. A correct characterization would require reference to a set of worlds, where the value of υ_2 to may vary by world. In a sense, we would like the universal quantification over the worlds that are elements of the proposition to scope over the individual quantification. In this system, that's not possible, because we are existentially quantifying over sets of worlds, and the universal quantification over their members is implicit in the interpretation of conditions. Since conditions need to be interpreted wrt assignment functions that already store a value for any of their arguments, any type of quantification introduced in terms of variable update will always scope over quantification that is implicit in the interpretation of conditions.

Negation: Even Worse

This problem persists, in an even worse form in the case of unspecific indefinites under negation, for which we might provide the following semantic representation:

(33) a. S: *Sue doesn't own a car.*

b. $\boxed{\begin{array}{c}\phi_1\\ \hline \phi_{DC_S} \Subset \phi_1\end{array}}\,;\, \mathbf{max}_{\phi_1}\left(\boxed{\begin{array}{c}\phi_2\\ \hline \phi_1 \equiv \overline{\overline{\phi_2}}\end{array}}\,;\, \mathbf{max}_{\phi_2}\left(\boxed{\begin{array}{c}\upsilon\\ \hline \upsilon \equiv \mathsf{Sue}\\ \mathsf{Sleep}_{\phi_2}\{\upsilon\}\end{array}}\right)\right)$

An update with (33-b) in C_4 yields the possible outputs j given in **Table 4.7**.

	ϕ_{DC_S}	ϕ_1	ϕ_2	υ_1	υ_2
j_1	$\{w_2, w_4\}$	$\{w_3, w_4\}$	$\{w_1, w_2\}$	sue	car_1
j_2	$\{w_2, w_4\}$	$\{w_2, w_4\}$	$\{w_1, w_3\}$	sue	car_2

Table 4.7: (Pragmatically privileged) outputs for updating C_5 with (33-b).

Again, the propositional dref ϕ_2, which we want to store the worlds where Sue has a car only characterizes a specific interpretation for the indefinite. This is the same issue as before, now concerning the propositional dref introduced by negation. The anti-veridical proposition ϕ_2 refers either to the set of worlds where Sue owns car_1 or to the set of worlds where Sue owns car_2, it cannot refer to a set of worlds where Sue owns either car in each of the worlds.

In addition, for either of two possible outputs, the asserted proposition ϕ_1 will contain worlds where Sue owns the other car. Although I'm not giving an additional context to illustrate this explicitly, I assume that the reader can see that the truth conditions are too liberal here: Since both possible outputs characterize situations where there is a specific car that Sue doesn't own, the update can be true in a context with a model where (none of the speaker's commitments contradict the information that) there is some car that Sue doesn't own, even if there is a different car that Sue does own.

Again, this problem arises from the interaction of propositional and individual drefs:

since the referent of an individual dref will be the same across all worlds of a proposition, only specific interpretations are possible. Like before, this is due to the fact that we quantify over sets of worlds, and quantification over worlds only happens in the interpretation of conditions.

4.4.3 Distributing over Possible Worlds—Relativizing Drefs

We might think that this issue is an artifact of the particular instantiation of this system. In the following, I am discussing some possible alternative systems to illustrate that the problem applies more broadly. The goal is to illustrate the point made by Asher and McCready (2007): That any intensional dynamic system will run into problems, unless individual discourse referents are interpreted intensionally as well, and more specifically distubutively over possible worlds, and we can also refer to propositions, i.e. collective sets of possible worlds. I will not explicate all the formal details for these alternatives here, but discuss them in the abstract, somewhat informally.

Dynamic Quantification over Possible Worlds

Consider what happens if, instead of quantifying over sets of worlds, we dynamically quantify over singleton worlds, and thereby allow the value of an individual dref to vary by world. Such a dynamic semantics might yield the possible output values in **Table 4.8** for an update with the affirmative case (31-a).

What does that gain us? Well, not too much. We are adding the possibility that the

	ϕ_{DC_S}	ϕ	υ_1	υ_2
j_1	w_1	w_1	sue	car_1
j_2	w_1	w_1	sue	car_2
j_3	w_2	w_2	sue	car_1
j_4	w_3	w_3	sue	car_2

Table 4.8: Possible outputs for updating C_4 with (31-a) if we had a system that uses quantification over singleton worlds

value of υ_2 can vary with the worlds associated with ϕ. We can characterize the truth conditions correctly, like before. It doesn't immediately help us with the content stored in propositional drefs: Since now the dref ϕ refers to single worlds, it cannot be picked up subsequently to refer to a set of worlds.

Still, all the information about which worlds are members of the proposition in question is stored in the output values. While it is not stored in one single possible output state, the combination of all possible outputs does give us the complete set.

Plural Dynamic Systems

At this point, it seems like a natural solution to switch to a plural dynamic system as in van den Berg (1996), particularly one where plural drefs over worlds are treated as propositional drefs, like the one in Brasoveanu (2006, 2010a).

In Van den Berg's Dynamic Plural Logic and Brasoveanu's Intensional Plural CDRT, discourse states are sets of assignments. That allows storing multiple assignments in one discourse state and simultaneously updating and imposing conditions on the whole set. Plural variable update and plural conditions work in the following way:

Like in the present system, variable update introduces new drefs. But in a plural system, it does so across all elements of a plural discourse state, i.e., a set of assignments.

(34) Plural variable update: [Brasoveanu (2010a): (18)]

$$[\upsilon] := \lambda I_{st}.\lambda J_{st}.\forall i[i \in I](\exists j(j \in J(i[\upsilon]j))) \wedge \forall j[j \in J](\exists i(i \in I(i[u]j)))$$

(34) states that plural variable update is a generalization of single variable update where any assignment $i \in I$ corresponds to an assignment $j \in J$ s.t $i[\upsilon]j$, and the other way round (Brasoveanu (2006, 2010a)).

Plural DRT conditions are properties of discourse states, like in the present system, with the difference that discourse states are sets of assignments.

(35) Abbreviation for plural conditions: [Based on Brasoveanu (2010a): (25)]

$$Own_{\phi}\{\upsilon_1, \upsilon_2\} := \lambda I_{st}.I \neq \varnothing \wedge \forall i_s[i \in I](own(\upsilon_1(i), \upsilon_2(i))(\phi(i)))$$

This states that for any assignment in the discourse state i, the value of the arguments $\upsilon_1(i), \upsilon_2(i)$ satisfies the own-relation in the world at $\phi(i)$. Therefore, we evaluate the condition wrt triplets of a world and two individuals, stored as values in relation to a value i. As a result, the update characterizes a discourse state containing only assignments where ϕ, υ_1 and υ_2 are in the proper relation.

In a system like this, the update with (31-a) in a discourse-initial context with M_7 might have the output in **Table 4.9**. It looks like our previous collection of output states, but here they are members of the set constituting a plural discourse state.

Now it is possible to access the proposition as a plural anaphor. When we encounter a propositional anaphor, we might assume that it ranges over the assignments within a

J	ϕ_{DC_S}	ϕ	υ_1	υ_2
j_1	w_1	w_1	sue	car_1
j_2	w_1	w_1	sue	car_2
j_3	w_2	w_2	sue	car_1
j_4	w_3	w_3	sue	car_2

Table 4.9: A possible output for updating C_5 with (31-a) if we had a plural system that uses quantification over singleton worlds

J	ϕ_{DC_S}	ϕ_1	ϕ_2	υ_1	υ_2
j_1	w_4	w_4	w_1	sue	car_1
j_2	w_4	w_4	w_1	sue	car_2
j_3	w_4	w_4	w_2	sue	car_1
j_4	w_4	w_4	w_3	sue	car_2

Table 4.10: A possible output for updating C_5 with (31-b) if we had a plural system that uses quantification over singleton worlds

plural discourse state, and we get a correct characterization of the static content associated with the asserted proposition ϕ. However, this still does not solve all of our issues with capturing the anaphoric potential of negative sentences.

Assuming that negation imposes a condition on the relation between propositional drefs, we might formulate this condition in the following way in a plural system:

(36) Condition imposed by negation:

$$\text{NOT}\{\phi_1, \phi_2\} := \lambda I.\forall i, i'[i, i' \in I](\phi_1(i) \neq \phi_2(i'))$$

This condition requires that there is no overlap between the worlds stored in $\phi_1(i)$ and $\phi_2(i')$, for any two assignments i, i' in the discourse state I.[4] Based on this, an update with (31-b) in a discourse-initial context with M_7 yields the output.

Now, this move does help: We get the correct characterization for the dref of the

[4]In combination with maximization, this essentially has the same effect as the complementation assumed for negation earlier.

asserted proposition ϕ_1, which doesn't include any worlds where *Sue owns a car* is true. As a result, it also allows us to capture the correct truth-conditions for negated sentence with unspecific indefinites. However, some issues remain. While the drefs ϕ_1 and ϕ_2 store the content we want them to store, this cannot capture the constraints on accessibility for the drefs under negation ϕ_2 and υ_2. This hypothetical version of a dynamic system allows for subsequent anaphoric reference to *'a car'*, even in a veridical context.

Back to the Original Problem

As the drefs are added to the discourse state, we predict that they can be picked up by subsequent anaphora. This leaves us with our initial problem: We need to characterize the conditions under which propositional and individual drefs introduced in the scope of negation cannot license anaphora (37), and the cases in which they can (38)+(39).

(37) *Sue doesn't [~~Sue~~ have [a car]$^{\upsilon_2}$]$^{\phi_2}$.*

a. *#It$_{\upsilon_2}$ is parked outside.*

b. *#That$_{\phi_2}$ is surprising.*

Intended: 'That Sue has a car is surprising.'

(38) A: *Sue doesn't have [a car]$^{\upsilon_2}$.*

a. *It$_{\upsilon_2}$ would be parked outside.*

b. B: *What are you talking about? I saw it$_{\upsilon_2}$ parked outside.*

c. *Or maybe she just doesn't bring it$_{\upsilon_2}$ to work.*

(39) A: *Sue doesn't [~~Sue~~ have a car]*$^{\phi_2}$.

a. *Even though she claimed that*$_{\phi_2}$.

‘Even though Sue claimed that she has a car.’

b. B: *Or maybe you just don't know this*$_{\phi_2}$ *about her.*

'Or maybe you just don't know about Sue that she has a car.'

We want to rule out anaphoric reference in (37) while at the same time allowing our negated drefs to be accessible in certain contexts; notably, in cases of modal subordination, disagreement, and in disjunctive contexts (illustrated in (38)+(39)), as well as cases where the antecedent is under double negation.

Epistemic Status and Intensional Context

The empirical generalization that I argued for in the first chapter is this: A dref introduced in an anti-veridical context can be the antecedent for an anaphor, only if the anaphor is in a non-veridical (or anti-veridical) context. This is reminiscent of generalizations about modal subordination (40).

(40) *A wolf might walk in* [after Roberts (1989)]

a. *#It is grey.*

b. *It would eat you first.*

c. *#It would have eaten you first.*

Here, the generalization is that a dref in a non-veridical context can be the antecedent for an anaphor, only if the anaphor is in a non-veridical, but not anti-veridical, context. The generalizations about drefs in negated and modal contexts are part of a broader constraint on accessibility: The proposition most local to the antecedent is true in the local context of the anaphor. What does this mean for individual anaphora?

It means that the accessibility of an individual dref should be stated in relation to the intensional context of the antecedent and the anaphor. When an anaphor is interpreted, we need to know the intensional context of the antecedent. While the information about the content of previous utterances is stored in the discourse state, we need the discourse state to store the information about the antecedent's intensional content.

Ways of Intensionalizing Mappings from Discourse Variables to Referents

The literature on modal subordination in dynamic semantics offers a few different ways to formalize this: We might choose a dynamic system in which discourse states are world-assignment pairs, and the mappings stored in variable assignments are accessed in relation to a world.

We might choose a plural system with distributive update and structured inclusion between sets of worlds. Brasoveanu (2006, 2010a) suggested this as an analysis of modal anaphora and modal subordination. While Brasoveanu's system would require additional assumptions to capture anti-veridical operators like negation and counterfactuals, it offers a quite general account of anaphoric reference to quantificational dependencies, independent of the types being quantified over, and independent of how many referents are dependent on one another.

Another possibility involves quantificational encapsulation over possible worlds, and storing dependencies relations between possible worlds and the individual referents in a world in the dref itself, as suggested for modal subordination in Stone (1999); Stone and Hardt (1999). This involves formalizing individual drefs as drefs for individual concepts, i.e. functions from worlds to individuals. Accessing an individual dref amounts to accessing this function, and allows for accessing potentially different individuals in relation to a world.

While I believe that any of these three general approaches could in principle be modified to capture our generalizations for anaphoric reference, and the problems outlined in this subsection, I am going to adopt an encapsulated quantification account for individual anaphora with individual drefs relativized to possible worlds, a simplified version of the system proposed in Stone (1999); Stone and Hardt (1999).

4.4.4 Conclusions about the Interaction of Propositional and Individual Anaphora

We discussed two problems here: Our overall issue of constraining anaphoric relations along the lines of our veridicality-generalization and the problems with the interaction of non-specific indefinites and propositional drefs discussed in this subsection. Treating individual drefs as drefs to individual concepts in a quantificational encapsulation account will address both of them.

The individual concepts will allow us to store individual referents in combination with the information about their local context and access this information to impose constraints on the relation between the local context of the anaphor and a possible antecedent.

Further, if individual drefs store values relative to possible worlds, the referenced individual can vary by world. In a sense, the dref stores a scope-relationship between quantification over possible worlds and quantification over individuals. That is, whenever we introduce or access an individual dref, quantification over worlds is implicit, and the information about this quantification, i.e., which individuals satisfy the conditions in which worlds, is stored in the individual dref itself. The quantificational dependency between worlds and individuals is stored in the representation of the individual dref.

Now that we understand the limits of the system introduced in this chapter, I will postpone a formal implementation of its solution to **Chapter 7**. There, we will update the representation of individual drefs in the way just discussed and illustrate how it addresses the question of accessibility for individual drefs and the problem of the proper representation of the content for propositional drefs for utterances with unspecific indefinites under negation. In the meantime, we will work with the system introduced in this Chapter, setting individual quantification aside. For now, our examples will involve drefs introduced by proper names, which are discourse constants and do not vary by assignment or world. We do this to focus on the things that this version of the system can do: Giving an analysis for accessibility of propositional anaphora.

4.5 Conclusion

This chapter focused on propositional drefs, particularly ones that are introduced in embedded contexts. In order to proceed with an analysis of propositional anaphora, we needed an understanding of how they are introduced. **Section 4.2** discussed two approaches to this and some of the reasons for adopting Snider's (2017) approach, which assumes that propositional operators introduce drefs referring to the proposition in their scope.

Section 4.3 discussed the assumed semantics for propositional operators. This component is central for the treatment of propositions in this system. Importantly for the treatment of anaphora to propositional content, propositional operators introduce a propositional dref storing the content of its prejacent, and specify a relation between that embedded dref and their own local embedding context. This section shows that we need a mechanism of maximization the propositional drefs that are introduced by linguistic expressions. Unlike the previously discussed mechanism of defeasible maximization associated with assertion, maximization for linguistically introduced propositional dref happens indefeasibly on a semantic level.

Section 4.4 addressed some limitations of the current system regarding the relationship between propositional and individual drefs: While the introduction of propositional drefs has the effect of existential quantification over sets of worlds, we cannot directly quantify over worlds. As a result, we are fatally mischaracterizing several aspects of the meaning: The content of propositional drefs in sentences with unspecific indefinites, as well as the truth-conditions and anaphoric potential for sentences with non-specific indefinites under negation in particular.

The presented system includes propositional operators specifying relations between propositional drefs. That leads to all propositional drefs being related to the speaker's discourse commitments, allowing us to assess their veridicality. Now, adding the assumption that propositional anaphora presuppose the truth of their referent in their local context, the generalizations from **Chapters 2** fall out from the analysis based on the general requirement that a discourse be consistent.

Chapter 5

Propositional Anaphora and Veridicality

5.1 Introduction

To capture the systematic interaction between propositional anaphora and veridicality, we need to keep track of the epistemic status of propositional drefs, i.e., their relationship to the speaker's commitments. **Chapter 4** showed how we can implement this in ICDRT. This chapter builds on that by illustrating how this allows us to give an analysis of the (im)possibilities for propositional anaphora. The analysis incorporates the intuition behind the generalization (1), developed in **Chapter 2**.

(1) **Generalization about accessibility for propositional anaphora**

A (proposition-denoting) expression is a possible antecedent for a propositional anaphor just in case the proposition expressed by the antecedent is true in the local context of the anaphor.

A consequence of this generalization is that the veridicality of the embedding context constrains anaphoric accessibility for propositions; and that whether the speaker is committed to the local context directly affects the availability of an anaphoric relation.

This chapter more systematically explores the consequences of this using some simple examples from English. We see that the possibilities for propositional anaphora systematically pattern along the lines of veridicality.

By themselves, the patterns and generalizations for propositional anaphora are quite straightforward and not unexpected. However, I suggest that the generalizations are useful and interesting, because they can be explained in terms of basic principles of discourse interpretation, and they relate to other ways in which anaphora interact with veridicality, in turn offering an explanation of these phenomena in term of the same basic principles. Specifically:

1. This chapter develops an analysis of the generalizations for propositional anaphora. The analysis is based on the framework of in intensional CDRT presented in the previous **Chapters 3 & 4**, and—in addition to the representations assumed there—only relies on two very basic assumptions: That the interpretation of a discourse requires that **(i)** propositions are true in their own local context, and that **(ii)** a discourse needs to be consistent.

2. The following **Chapter 6** shows how an account of these generalizations can provide an understanding of anaphoric polarity-sensitivity and discourse-negativity. account for anaphoric negativity-tags (i.e. expressions which are anaphoric to propositional content and require a negative antecedent) and the relevant notion of negativity in terms of anti-veridicality, and the introduction of an anti-veridical propositional dref.

3. Further, **Chapter 7** shows how individual anaphora pattern in parallel ways. the chapter then argues that, assuming the appropriate representations for individual drefs in this framework allows us to capture accessibility for individual anaphora similarly in terms of basic principles of discourse consistency

Keeping these future extensions in mind, we first develop an extension of the framework presented in the previous chapter, which can account for the interaction of veridicality distinctions and propositional anaphora. The chapter proceeds in the following way:

Section 5.2 provides an overview of interaction of propositional anaphora and veridicality. It discusses how the generalizations developed in **Chapter 2** lead us to expect the patterns we see and how they can be understood in terms of a more general requirement that a discourse be consistent. Because of this, the same kinds of constraints do no apply in these kinds of discourse relations which do not require the discourse segments to be consistent. However, this can still be understood in terms of our veridicality-based generalizations, because these are exactly the kinds of discourse relations which Asher

(1993); Asher and Lascarides (2003) characterizes as non-veridical discourse relations.

Section 5.3 illustrates how the formal analysis works to capture the veridicality-based generalizations for propositional anaphora illustrated in **Section 5.2**. Based on the assumptions of ICDRT and the constraint (1), we derive the patterns from **Section 5.2** in cases where the discourse segments are interpreted in conjunction, and need to be consistent with each other.

Section 5.4 shows how the account can explain how the same constraints do not apply when the two discourse segments are in a non-veridical discourse relation like disjunction or disagreement.

Section 5.5 concludes the chapter by summarizing its main points.

5.2 Patterns of (Non-)Accessibility

This section aims to give a more systematic overview of the possible combinations for propositional anaphora and antecedents in (non-/anti-/)veridical contexts. It will show that, in order to yield a consistent discourse, propositional anaphora impose a veridicality-matching requirement: Propositional anaphora require that their antecedent have the same epistemic status. An overview of this basic pattern is given in **Table 5.1**. The pattern applies only to discourses where the discourse segments containing antecedent and anaphor are interpreted in a way that requires them to be compatible with

Epistemic context of anaphor	Epistemic context of antecedent	Propositions
Veridical	Veridical	✓
	Non-veridical	✗
	Anti-veridical	✗
Non-veridical	Veridical	?
	Non-veridical	✓
	Anti-veridical	✗
Anti-veridical	Veridical	✗
	Non-veridical	✗
	Anti-veridical	✓

Table 5.1: Overview of (Im-)possible Combinations of Veridicality of Antecedent and Anaphor for Propositional Anaphora

each other. For example, they might be in a conjunction or conditional. The patterns are different in cases where the discourse relations are not required to be consistent, for example, in disjunction or inter-speaker disagreement. However, this could also be related to the notion of veridicality: disjunction and disagreement belong to the discourse relations that are characterized as non-veridical discourse relations in SDRT Asher (1993); Asher and Lascarides (2003).

In the following, I illustrate these patterns for propositional anaphora with some simple examples from English. I will first do so for veridical discourse relations and anaphora in veridical, non-veridical, and anti-veridical contexts. Then I will briefly discuss some exceptions to these generalizations that arise under non-veridical discourse relations and summarize the main results.

This discussion aims to illustrate how the generalization about anaphoric accessibility can be reduced to the fundamental constraint that discourses should be consistent. That applies to the propositional anaphora discussed here, as well as the individual anaphora (discussed in **Chapter 7**). While both come with their respective presuppositions that the referent exist/be true, the veridicality of the embedding context determines how these presuppositions are interpreted relative to the belief state of the speaker.

5.2.1 Anaphora in Veridical Contexts

The Basic Pattern

If an anaphor is in a veridical context, it can only have an antecedent that is compatible with a veridical interpretation.

In (2)–(4), the propositional anaphor *that* is embedded in a veridical embedding context *'that's the truth'*. It needs a propositional antecedent which itself is compatible with a veridical interpretation.

(2) [$^{\phi_1}$ *I know* [$^{\phi_2}$ *that Mary is sick*]] *and that*$_{\phi_1/\phi_2}$*'s the truth.*

a. ✓that$_{\phi_1}$: that I know that Mary is sick

b. ✓that$_{\phi_2}$: that Mary is sick

(3) [$^{\phi_1}$ *John believes* [$^{\phi_2}$ *that Mary is sick*]] *and that*$_{\phi_1/\phi_2}$*'s the truth.*

a. ✓that$_{\phi_1}$: that John believes that Mary is sick

b. ✓that$_{\phi_2}$: that Mary is sick

(4) [$^{\phi_1}$ *It's not the case* [$^{\phi_2}$ *that Mary is sick*]] *and that*$_{\phi_1/\phi_2}$*'s the truth.*

a. ✓that$_{\phi_1}$: that it's not the case that Mary is sick

b. #that$_{\phi_2}$: that Mary is sick

This pattern is expected under the assumption that a propositional anaphor requires that the antecedent proposition is true in the local context of the anaphor. In general, the assertion of any proposition introduces a veridical dref corresponding to the asserted proposition. Therefore, the initial clauses in (2)–(4) all provide an accessible antecedent corresponding to the first conjunct.

In addition (2) also involves a proposition embedded under the veridical embedding *I know*. As a result, this embedded proposition is an accessible antecedent for the propositional anaphor in the second conjunct. This option is not available for an anti-veridically embedded proposition in (4).

Here, the embedded proposition in an anti-veridical context and its falsity is asserted. That is in direct contradiction to the interpretation of the subsequent veridical anaphor, which requires that the referent is true (according to the speaker). An interpretation of the second utterance where *that* refers to the previous anti-veridical proposition is inconsistent with the first.

Finally, an antecedent in the non-veridical context (3) is not immediately accessible as an antecedent for the veridical anaphor *that.* This is because the interpretation of an anaphor in a veridical context use requires commitment to the truth of the antecedent proposition. However, like in the accommodation cases discussed in the previous chapter, non-veridical embeddings could be pragmatically strengthened to a veridical or anti-veridical interpretation. For example, suppose it is known that John knows Mary better than the speaker, and the interlocutors take John to be an epistemic authority in the context. In that case, we might interpret *John believes* as a veridical embedding in (3). In that case, we do expect an anaphoric relation to be possible.

Licensing vs. Interpretation On the flip side, the interpretation of (3), where the proposition *Mary is sick* is anaphorically referred to as a veridical proposition comes with the global inference that the proposition *Mary is sick* is true. The result is that interpretation of the antecedent is strengthened to a veridical interpretation. This claim is supported in contrast with non-veridical embeddings under *might*, which seem to be less amenable to pragmatic strengthening.

(5) [$^{\phi_1}$ *It might be the case* [$^{\phi_2}$ *that Mary is sick*]] *and that*$_{\phi_1/\phi_2}$*'s the truth.*

a. ✓that$_{\phi_1}$: that it might be the case that Mary is sick

b. ✗that$_{\phi_2}$: that Mary is sick

The contrast between (3) and (5) suggests that antecedents in non-veridical contexts *can* be antecedents for anaphora in veridical contexts, but only if the interpretation of the antecedent segment can be strengthened to a veridical interpretation. This suggests that it will be useful to distinguish between licensing and interpretation: While antecedents in non-veridical contexts may license anaphora in veridical contexts, the interpretation of the anaphor leads to a veridical interpretation for the antecedent.

An antecedent may also be accessible if the previous discourse is compatible with the accommodation of the requirement of the truth of its referent, like with certain non-veridical contexts. If the truth of the referent has been negated, co-reference will result in contradictory requirements on the interpretation and an inconsistent discourse. In contrast, a propositional antecedent in a non-veridical context does not commit the speaker to the truth or falsity of the proposition. In that case, a subsequent anaphor is available only if the interpretation of the local context of the anaphor is compatible with a veridical interpretation. Although some non-veridical antecedents allow for veridical anaphora, that is the case only if they can subsequently be interpreted as veridical.

An overview of the patterns for veridical anaphora, taking into account the difference between licensing and interpretation, is given in **Table 5.2**.

Epistemic context of anaphor	Epistemic context of antecedent	Propositions	
		Licensing	Interpretation
Veridical	Veridical	✓	✓
	Non-veridical	✓	✗
	Anti-veridical	✗	✗

Table 5.2: Overview of (Im-)possible Combinations for Veridical Propositional Anaphora

5.2.2 Anaphora in Anti-Veridical Contexts

Anti-veridical propositional anaphora is the most restricted among propositional anaphora. If the anaphor is in an anti-veridical context, the antecedent has to be non-veridical.

(6) [$^{\phi_1}$ *I know* [$^{\phi_2}$ *that Mary is sick*]] *and/but that*$_{\phi_1/\phi_2}$ *is not the case.*

a. ✗that$_{\phi_1}$: that I know that Mary is sick

b. ✗that$_{\phi_2}$: that Mary is sick

(7) [$^{\phi_1}$ *John believes* [$^{\phi_2}$ *that Mary is sick*]] *and/but that*$_{\phi_1/\phi_2}$ *is not the case.*

a. ✗that$_{\phi_1}$: that John believes that Mary is sick

b. ✓that$_{\phi_2}$: that Mary is sick

(8) [$^{\phi_1}$ *John pretended* [$^{\phi_2}$ *that Mary is sick*]] *and/but that*$_{\phi_1/\phi_2}$ *is not the case.*

a. ✗that$_{\phi_1}$: that John pretended that Mary is sick

b. ✓that$_{\phi_2}$: that Mary is sick

(8) is the most straightforward case here because the antecedent and anaphor are both anti-veridical. The first sentence comes with the inference that it is not true that *Mary is sick*, and the second sentence consistently asserts the same. (6) leads to a contradiction: Since the first sentence presupposes that *Mary is sick*, an assertion to the contrary is inconsistent here. The pattern is illustrated in **Table 5.3**.

Epistemic context of anaphor	Epistemic context of antecedent	Propositions: Licensing	Propositions: Interpretation
Anti-veridical	Veridical	✗	✗
	Non-veridical	✓	✗
	Anti-veridical	✓	✓

Table 5.3: Overview of (Im-)possible Combinations for Anti-veridical Propositional Anaphora

5.2.3 Anaphora in Non-Veridical Contexts

A pronoun in a non-veridical context does not place any restrictions on the veridicality of its antecedent in terms of licensing. The requirement of the anaphor is evaluated in relation to a non-veridical intensional context. The non-veridical embedding does not commit the speaker to its content's truth (or falsity). So even if the interpretation of the anaphor requires that its antecedent exists in its local context, there will be no inconsistency on a global level. As a result, any possible antecedent is accessible here.

(9) [$^{\phi_1}$ *I know* [$^{\phi_2}$ *that Mary is sick*]] *and she has been suspecting that*$_{\phi_1/\phi_2}$ *for a while.*

   a. ✓that$_{\phi_1}$: that I know that Mary is sick

   b. ✓that$_{\phi_2}$: that Mary is sick

(10) [$^{\phi_1}$ *John believes* [$^{\phi_2}$ *that Mary is sick*]] *and she has been suspecting that*$_{\phi_1/\phi_2}$ *for a while.*

   a. ✓that$_{\phi_1}$: that John believes that Mary is sick

   b. ✓that$_{\phi_2}$: that Mary is sick

(11) [$^{\phi_1}$ *It's not the case* [$^{\phi_2}$ *that Mary is sick*]] *but she has been suspecting that*$_{\phi_1/\phi_2}$ *for a while.*

   a. ✓that$_{\phi_1}$: that it's not the case that Mary is sick

   b. ✓that$_{\phi_2}$: that Mary is sick

A propositional anaphor in a non-veridical embedding under *suspect* does not necessarily impose any requirements on a global level. Mary suspecting that she is sick is consistent with it being true that she is sick, but also with it being false or uncertain. The antecedent context can constrain the veridicality of the discourse referent, but the context of the anaphor does not impose any additional restrictions here.

Again, though, this seems to be due to the fact that these non-veridical contexts are compatible with a strengthened (anti-)veridical interpretation. If we compare (9)-(11) to the following cases with propositional anaphora embedded under *might*, we again see a pattern where the veridicality of the anaphor and antecedent need to match. The case with an anti-veridical propositional antecedent (14) is contradictory (see Veltman, 1996), while the case with a veridical antecedent is (at least) uninformative (Asher and McCready, 2007, see e.g.).

(12) ? [$^{\phi_1}$ *I know* [$^{\phi_2}$ *that Mary is sick*]] *and that*$_{\phi_1/\phi_2}$ *might be the truth.*

  a. ***?*** that$_{\phi_1}$: that I know that Mary is sick

  b. ***?*** that$_{\phi_2}$: that Mary is sick

(13) # [$^{\phi_1}$ *John believes* [$^{\phi_2}$ *that Mary is sick*]] *and that*$_{\phi_1/\phi_2}$ *might be the truth.*

  a. ***?*** that$_{\phi_1}$: that John believes that Mary is sick

  b. ✓that$_{\phi_2}$: that Mary is sick

(14) [$^{\phi_1}$ *It's not the case* [$^{\phi_2}$ *that Mary is sick*]] *but that*$_{\phi_1/\phi_2}$ *might be the truth.*

  a. ***?*** that$_{\phi_1}$: that it's not the case that Mary is sick

  b. ✗that$_{\phi_2}$: that Mary is sick

In summary, a non-veridical anaphor requires that its antecedent be consistent with a non-veridical interpretation. In this section, we are considering a particular case of non-veridical contexts, namely ones that are not anti-veridical. These kinds of contexts do not commit the speaker to the truth or falsity of the embedded proposition either way. The English possibility modal *might* further requires that its prejacent is an epistemic possibility (it cannot be an anti-veridical proposition) and would be uninformative in case its prejacent is already ruled out as a possibility (it normally would not be a veridical proposition).

The additional inferences associated with the semantics of *might* show us that a non-veridical propositional anaphor can have veridical and anti-veridical antecedents only if the context of the anaphor is compatible with a strengthened interpretation. Similar to the above cases, an anaphor in a semantically non-veridical embedding context does not require its antecedent to match its veridicality in terms of licensing. However, in terms of interpretation, only a non-veridical proposition is a suitable antecedent for a pronoun whose context requires a non-veridical interpretation.

The pattern for non-veridical propositional anaphora is indicated in **Table 5.4**.

Epistemic context	Epistemic context	Propositions	
of anaphor	of antecedent	Licensing	Interpretation
Non-veridical	Veridical	✓	?
	Non-veridical	✓	✓
	Anti-veridical	✓	✗

Table 5.4: Overview of (Im-)possible Combinations for Non-veridical Propositional Anaphora

5.2.4 Interim Summary

This section presented an argument that the possibilities for propositional anaphora systematically pattern along the lines of veridicality. It further suggested that this is due to requirements that the interpretation of an anaphor places on its local context and the more general one that a discourse cannot be inconsistent. These claims are supported by the empirical patterns that the accessibility constraints give rise to:

Veridical Anaphora require antecedents that are compatible with a veridical interpretation. That includes veridical antecedent, or non-veridical antecedents that can be pragmatically strengthed to a veridical interpretation, but not anti-veridical ones.

Non-verdical Anaphora do not commit the speaker about the truth of their referent. They are compatible with antecedents of any veridicality.

Anti-veridical Anaphora show a reversal of the pattern for veridical anaphora: An anti-veridical propositional anaphor requires an antecedent that is compatible with an anti-veridical interpretation.

An overview of the (im-)possibilities for propositional anaphora along the lines of veridicality distinctions is given in **Table 5.5**. We may describe this pattern by stating that there is a weak veridicality-matching requirement[1]. While the interpretation requires

[1]Thanks to an anonymous reviewer for this suitable description.

that a (anti-)veridical propositional anaphor has an antecedent with the same epistemic status,

Epistemic context of anaphor	Epistemic context of antecedent	Propositions: Licensing	Propositions: Interpretation
Veridical	Veridical	✓	✓
	Non-veridical	(✓)	✗
	Anti-veridical	✗	✗
Non-veridical	Veridical	✓	?
	Non-veridical	✓	✓
	Anti-veridical	✓	✗
Anti-veridical	Veridical	✗	✗
	Non-veridical	(✓)	✗
	Anti-veridical	✓	✓

Table 5.5: Overview of (Im-)possible Combinations of Veridicality of Antecedent and Anaphor in Licensing and Interpretation for Propositional Anaphora

5.2.5 Discourse Consistency

The generalizations and patterns discussed in this section so far can be explained in terms of the requirement that a discourse be consistent. Cases where an anaphora are impossible can be ruled out on the basis that their interpretation would introduce contradiction. However, all of the above cases involve semantic conjunction, a veridical discourse relation in Asher and Lascarides's (2003) terms. The same restrictions do not apply under non-veridical discourse relations such as disjunction or disagreement, which do not require that the respective discourse segments be consistent.

The above discussion suggests that veridical anaphora with non-veridical antecedents, or anti-veridical propositional anaphora with propositional antecedents which are not possible. However, in discourse relations that do not require consistency, these combinations are possible. This contrast is illustrated for conjunction vs. disjunction in (15) and (16).

(15) Veridical propositional anaphor with anti-veridical antecedent: .

a. *Either Mary didn't [~~Mary~~ get infected]$^{\phi}$, or she's already aware of that$_{\phi}$.*

b. *# Mary didn't [~~Mary~~ get infected]$^{\phi}$, and she's already aware of that$_{\phi}$.*

(16) Anti-veridical propositional anaphor with veridical antecedent:

a. *Either it's true that Mary is sick, or that's not the case.*

b. *# It's true that Mary is sick and that's not the case.*

The same contrasts can be observed for agreement vs. disagreement:

(17) Veridical propositional anaphor with anti-veridical antecedent:

A: *Mary didn't [~~Mary~~ get infected]$^{\phi}$,*

a. B: *(That's not true.) I am sure of it$_{\phi}$.*

b. B: *# (Indeed.) I am sure of it$_{\phi}$.*

(18) Non-veridical propositional anaphor with veridical antecedent:

A: *It's true that Mary is sick*

a. B: *(No.) That's not the case.*

b. B: *# (Indeed.) That's not the case.*

In these cases, the discourse relations connecting the discourse segments, including the antecedent and anaphor, are non-veridical. Therefore, these cases are not counterexamples but rather an extension of the generalization. The truth requirement of the anaphor is satisfied in its local context. The difference here is that there are fewer constraints on the relationship between the antecedent's local context and the pronoun's local context.

5.3 Analysis: Weak Veridicality-Matching in a Consistent Discourse

5.3.1 Local Truth-Requirement

We can explain these patterns by assuming that the interpretation of an anaphoric pronoun requires that its antecedent is true in the local context of the anaphor. Depending on whether it is evaluated in a veridical, hypothetical, or counterfactual local context, this requirement will impose different felicity conditions on the relationship between the referent and the speaker's commitments, i.e., on its veridicality. Based on this assumption, the relevant constraints on accessibility for propositional anaphora fall out from the veridicality of the local context of antecedent and anaphor and the requirement that a discourse be consistent.

The local truth-requirement is related to Kroll's (2019) notion of local givenness, which was developed to account for the interpretation and licensing of polarity-reversal in sluicing. This involves cases of sluicing in which the ellipsis site is interpreted in a way that is paraphrased as a sentence with a polarity opposite from its antecedent, illustrated in (19).

(19) Polarity-reversing sluices:

a. Under neg-raising: [Kroll (2019): 2]

I don't think that [California will comply]$_A$, but I don't know why [$_{TP}$ ~~California won't comply~~]$_E$.

b. Under disjunction: [Kroll (2019): 3]

Either [John$_j$ didn't do an extra credit problem]$_A$, or he$_j$ didn't mark which one$_i$ [~~he$_j$ did do t$_i$~~]$_E$.

Local givenness is defined as entailment of a proposition by the local context and invoked as a condition on the licensing of sluicing. Kroll suggests that TP-deletion in sluicing is licit just in case the proposition expressed by the antecedent (modulo F-closure, see also Merchant (2001)) is entailed by the local context of the ellipsis site. (This semantic condition on sluicing is understood in conjunction with a syntactic identity requirement that applies to heads in the *v*P-domain, see Kroll and Rudin (2017).) Dissociating a Merchant-style e-givenness licensing requirement from the global context and allowing for it to be interpreted in relation to a local context allows Kroll to account for the interpretation of polarity-reversing sluices because this allows the process of anaphoric resolution to incorporate reasoning about (non-/anti-)veridicality, and the epistemic status of the local proposition.

For propositional anaphora, we impose a similar but slightly weaker requirement. The content of the requirement itself is the same: The proposition expressed by the antecedent is entailed by the local context of the anaphor (i.e., it is true throughout the worlds in the local context). However, we impose this as a condition on interpretation rather than licensing. The interpretation of a propositional anaphor requires entailment if its referent by the local context. Therefore, licensing requires that the context is compatible with this interpretation, imposing a slightly weaker condition on the context.

The local truth-requirement is also related to Stone's (1999) idea of discourse referents being interpreted in relation (potentially hypothetical) scenarios, proposed to account for modal subordination. Formally, scenarios are functions picking out sets of worlds given a world, allowing us to store information about relationships between sets of possible worlds. We can use a less complex representation for our purposes, treating local contexts as sets of worlds (propositions), while information about their relationships is contributed by lexical relations.

5.3.2 Veridical Antecedents

Let us consider how the system can account for the following contrasts for propositional anaphora with veridical antecedents.

(20) *A test proved* [*that Mary is sick*]$^{\phi}$

  a. ✓*so that*$_{\phi}$*'s true.*

  b. ✓*and she has been suspecting that*$_{\phi}$ *for a while.*

  c. # *but/so/and that*$_{\phi}$*'s not the case.*

The propositional anaphor *that* in the second clauses in (20) can refer to the previous veridically embedded proposition (ϕ: *that Mary is sick*) if the anaphor is embedded in a veridical or non-veridical context (20-a) + (20-b), but not if it's introduced in an anti-veridical one (20-c).

In the following, I illustrate how the semantic assumptions made here allow us to explain why. We will consider the semantics of the utterance containing the antecedent, and how updates involving the different follow-ups might work.

The representation for the Antecedent Sentence can be given as in (21):

(21) S: *A test proved that Mary is sick.* $\rightsquigarrow$

$$\boxed{\begin{array}{c} \phi_1 \\ \hline \phi_{DC_S} \Subset \phi_1 \end{array}}\,; \mathbf{max}_{\phi_1}\left(\boxed{\begin{array}{c} \upsilon_1, \phi_2 \\ \hline \mathsf{Test}_{\phi_1}\{\upsilon_1\} \\ \mathsf{Prove}_{\phi_1}\{\upsilon_1, \phi_2\} \end{array}}\,; \mathbf{max}_{\phi_2}\left(\boxed{\begin{array}{c} \upsilon_2 \\ \hline \upsilon_2 \equiv \mathsf{Mary}_{\mathbf{e}} \\ \mathsf{Sick}_{\phi_2}\{\upsilon_2\} \end{array}}\right)\right)$$

(21) represents an update introducing two propositional drefs ϕ_1, ϕ_2 and two individual drefs υ_1, υ_2. ϕ_1 is the (maximal) proposition corresponding to the assertion, υ_1 is a *test* in all ϕ_1-worlds, and the verb *proved* is treated as a propositional operator, introducing the dref ϕ_2 and a condition specifiying a relation between ϕ_1 and ϕ_2.

In particular, the condition introduced by the verb ($\mathsf{Prove}_{\phi_1}\{\upsilon_1, \phi_2\}$) states that υ_1 proved ϕ_2 in all ϕ_1 worlds. ϕ_2 is the maximal proposition consistent with the results of the test, and υ_2 refers to the individual mary. The condition $\mathsf{Sick}_{\phi_2}\{\upsilon_2\}$ states that υ_2 is sick in all ϕ_2 worlds.

The relationship between ϕ_1 and ϕ_2 is provided by the interpretation of the condition

$\mathsf{Prove}_{\phi_1}\{\upsilon_1, \phi_2\}$. Here, I am treating *prove* as a veridical assertive attitude verb in a Hintikkan semantics for propositional attitudes, based on the discussion in Anand and Hacquard (2014).[2] As such, it makes reference to an assertive information state, i.e. the set of worlds compatible with an assertion, and states that the proposition expressed by its propositional argument holds true throughout. Following Heim (1992), the assertive information state (i.e. a set of worlds) is derived as a function from a world and an attitude holder, in this case a communicative agent (22-a). Based on this, the static predicate $\mathsf{prove}_{\langle e,\langle\langle wt\rangle,wt\rangle}$ is defined as specifying a relation between a communicative agent x, a proposition p and a world w: p holds throughout the assertive information state that is compatible with the commitments of x in w (22-b).

(22) a. $[\![\textsc{Assertive}_x(w)]\!]^{M,g} =$

$\{w' \in D_w : w'$ conforms to the (communicative) commitments of $[\![x]\!]^{M,g}$ in $[\![w]\!]^{M,g}\}$

b. Interpretation of $\mathsf{prove}_{\langle e,\langle\langle wt\rangle,wt\rangle}$ (first version):

$[\![\mathsf{prove}(x_e)(p_{wt})(w_w)]\!]^{M,g} = [\![\textsc{Assertive}_x(w) \subseteq p]\!]^{M,g}$

I should point out that this is a simplified representation of the semantics of assertive

[2] While they do not claim that this is the appropriate analysis for verbs like *prove* (it might be more complicated), this is consistent with their overall claims, and sufficient for our present purposes.

attitudes[3], but one that is sufficient for our present purposes. Based on this, a condition containing the corresponding dynamic predicate is defined in the usual way:

(23) $\mathsf{Prove}_{\phi_1}\{\upsilon, \phi_2\} :=$

$\lambda i_s.\forall w \in \phi_1.\mathsf{prove}(\upsilon(i))(\phi_2(i))(w)$

By evoking the static predicate, this dynamic condition specifies a relation between ϕ_1 and ϕ_2: ϕ_2 holds throughout any assertive information state that is compatible with the commitments of the subject in any world in ϕ_1.

(22-b) characterizes *prove* as an assertive attitude verb, but it doesn't yet take into account its veridicality. We characterize the property of introducing the inference that the propositional content be true in its superordinate context by adding a conjunct to the definition that states exactly that:

(24) $[\![\mathsf{prove}(x_e)(p_{wt})(w_w)]\!]^{M,g} = [\![\textsc{Assertive}_x(w) \subseteq p \wedge p(w)]\!]^{M,g}$

Now, this has the desired effect for the interpretation of (21): With the update, (the speaker is committed that) the proposition that *Mary is sick* is true in all worlds that are consistent with the test results (relative to the ϕ_1-worlds), and (the speaker is committed

[3] I suppose that it doesn't really make sense to talk about the commitments of a test in a world. The semantics assumed here is an oversimplification. We would want to understand an assertive information state in relation to a communicative event. For example, Anand and Hacquard (2014) suggest the relevant information state is a projected common ground (in the sense of Farkas and Bruce (2010)) associated with the reported communicative event. Further, an inanimate test couldn't really qualify as a communicative agent with commitments. However, a test (result) could be characterized as a *repository of information* (ROI), which Anand and Hacquard (2008) point out as a possible subject for assertive attitudes. Presumably, these cases are interpreted in a way s.t. the ROI is used by an unmentioned intentional agent in order to communicate a point. Using a simplified pragmatics for assertion (making reference to commitments and omitting a projected common ground), we can think of the commitments of the test as the information that is consistent with the test results. In this case, they will contain all the worlds in which Mary is sick.

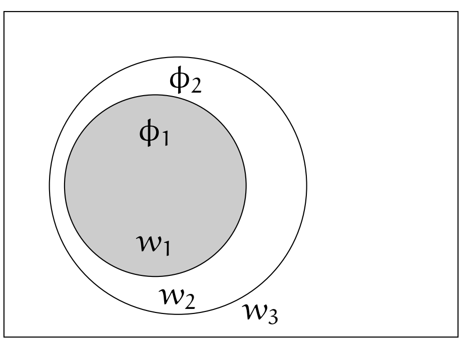
that) *Mary is sick*.

Update with the Antecedent Sentence. We illustrate the effect of this update concretely by assuming a toy context $C_1 = \langle M_1, \{S\}, i\rangle$, where i is a discourse-initial context, and M_1 is s.t. $D_e = \{mary, S, t, \#\}$, $D_w = \{w_1, w_2, w_3\}$, and $[\![test]\!]^{M_1}$, $[\![sick]\!]^{M_1}$, and $[\![prove]\!]^{M_1}$ characterize the propositions in **Table 5.6**.

	$\lambda w.test(t)(w)$	$\lambda w.sick(mary)(w)$	$\lambda w.prove(t)(\lambda w'.sick(mary)(w'))(w)$
w_1	1	1	1
w_2	1	1	0
w_3	1	0	0

Table 5.6: Truth-values for $\lambda w.test(t)(w)$, $\lambda w.sick(mary)(w)$, $\lambda w.prove(t)(\lambda w'.sick(mary)(w'))(w)$, for the worlds in D_w.

In M_1, t is a test in all worlds, *Mary is sick* in w_1, w_2, and *A test proved that Mary is sick* is true in w_1. Since the latter proposition entails the former, there is no possible world where the latter is true while the former isn't. Updating i_{C_1} with (21) yields an output j, which maps the drefs as depicted in **Figure 5.1**.



$$\begin{aligned} \phi_{DC_S}(j) &= \{w_1\} \\ \phi_1(j) &= \{w_1\} \\ \phi_2(j) &= \{w_1, w_2\} \\ \upsilon_1(j) &= t_e \\ \upsilon_2(j) &= mary_e \end{aligned}$$

Figure 5.1: Drefs after updating with *'A test proved that Mary is sick'* in C_1

This illustrates the veridicality of the dref ϕ_2. The area shaded in gray represents the asserted proposition ϕ_1: The commitments of the speaker will be a subset of it.

Due to the veridical embedding, the embedded proposition ϕ_2 is a superset of ϕ_1. The proposition that *Mary is sick* will be true throughout ϕ_1, and in turn throughout the speakers commitments.

A Veridical Anaphor

Let's see why the update with the assertion of a speaker S of *A test proved [that Mary is sick]*$^{\phi_2}$. is compatible with a subsequent update with the assertion of *that*$_{\phi_2}$*'s true.*

The semantic representation is given in (25).

(25) S: *That*$_{\phi_2}$*'s true.* $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_3 \\ \hline \phi_{DC_S} \Subset \phi_3 \\ \hline \end{array}; \mathbf{max}_{\phi_3}\left(\boxed{\phi_3 \Subset \phi_2}\right)$$

Setting aside the pragmatic effects of *it's true that* ϕ[4], I am treating this embedding as semantically vacuous with the assumption that asserting *it's true that* ϕ amounts to asserting ϕ. Accordingly, we might characterize this operator as introducing the truth-functional relation over two drefs ϕ, ϕ', where ϕ' is true in ϕ as the condition in (26).

(26) $\phi \Subset \phi'$

Accordingly, (25) represents an update that introduces the propositional dref ϕ_3, corre-

[4] I looks like one of these effects is that *it's true that* ϕ requires that ϕ is given in some way, as the below contrast suggests.

(i) a. A: *Sue said that Mary might be sick.*
B: *It's true that Mary is sick.*
b. A: *Sue was asking how Mary is doing.*
B: *It's true that Mary is sick.*

sponding to the (maximal) proposition corresponding to the assertion. The propositional embedder *be true* is taken to be semantically vacuous, with the assumption that asserting *it's true that* ϕ amounts to asserting ϕ. Accordingly, this operator is characterized as introducing the condition $\phi_3 \Subset \phi_2$. The anaphoric discourse variable ϕ_2 is introduced by the pronoun *that*$_{\phi_2}$. Importantly, because $\phi_3 \Subset \phi_2$ and $\phi_{DC_S} \Subset \phi_3$, this will also require that $\phi_{DC_S} \Subset \phi_2$. That means that ϕ_2 is a veridical discourse-variable.

The update. We want to see what happens when we update with (25) after updating with (21). We consider the context $C'_1 = \langle M_1, \{S\}, j\rangle$, where $[\![(21)(i)(j)]\!]^{M_1} = 1$.

Recall that $\phi_2(j)$ is a veridical proposition, i.e. one that is true in all the worlds associated with S's commitments ($\phi_{DC_S}(j) \subseteq \phi_2(j)$). A subsequent update with (25) imposes its own conditions on the relationship between ϕ_{DC_S} and ϕ_2 in its output k.

The condition $\phi_3 \Subset \phi_2$ requires that $\phi_3(k)$ is a subset of ϕ_2, and the assertion of ϕ_3 introduces the condition $\phi_{DC_S}(k) \subseteq \phi_3(k)$. In combination of the two, we get that $\phi_{DC_S}(k) \subseteq \phi_2(k)$. Since this relation already holds for the input j, we can consistently update with (25), i.e. there is some k_s s.t. $[\![(25)(j)(k)]\!]^{M_1} = 1$.

Discussion. The interpretation of the context of the anaphor requires that it refers to a proposition that is veridical, i.e. one that is a superset of (or equal to) the speaker's commitment set. Since ϕ_2 is a veridical dref after updating with the antecedent sentence, it is a viable antecedent for the propositional anaphor here.

The proposition asserted by the antecedent sentence ϕ_1 is also a veridical dref, and also a possible antecedent for *it*, by the same reasoning. As a result, assertions of sentences with veridical embeddings systematically make available two possible veridical antecedents. While both of them are (correctly) predicted to be accessible antecedent due to their veridicality, this gives rise to the further question how the referential ambiguity of the anaphoric pronoun might be resolved in these cases. Two factors seem to play an role here: The choice of pronoun, and the discourse relation between the discourse segment containing antecedent and anaphor. This is illustrated in (27).

(27) [$^{\phi_1}$*A test proved* [$^{\phi_2}$ *that Mary is sick*]]

a. *so it's/that's*$_{\{\phi_2/\#\phi_1\}}$ *true.*

b. *and it's*$_{?}$ *true.*

c. *and that's true*$_{\{\phi_1/??\phi_2\}}$.

An Anti-Veridical Anaphor

If an initial utterance by a speaker S of (28-a) is followed up by an utterance of (28-b) by the same speaker, the result is an inconsistent discourse.

(28) a. *A test proved [*$^{\phi_2}$ *that Mary is sick],*

b. *but that*$_{\phi_2}$*'s not the case.*

The embedding in (28-b) is non-veridical. As such it entails the falsity of its prejacent which in turn is provided by the propositional anaphor *that*. As a result, *that* could only anaphorically refer to a proposition that is compatible with an anti-veridical interpretation, i.e. one where the proposition is false. Since (28-a) already commits the speaker to the truth of ϕ (that Mary is sick), this proposition is not a suitable antecedent for the propositional anaphor in a non-veridical context in (28-b).

In line with common assumptions, I assume that the embedding *it's not the case that* is semantically analogous to sentential negation. Here, I will assume the semantics introduced earlier for propositional negation. Accordingly, the assumed representation for (28-b) will be (29).

(29) $$\begin{array}{|c|}\hline \phi_3 \\ \hline \phi_{DC_S} \Subset \phi_3 \\ \hline \end{array}\,;\ \mathbf{max}_{\phi_3}\left(\boxed{\phi_3 \equiv \overline{\overline{\phi_2}}}\right)$$

Since the non-veridical embedding entails the falsity of the embedded content, this representation requires that the discourse commitments of the speaker are a subset of the complement set of the proposition associated with ϕ_2.Therefore $\phi_{DC_S} \Cap \phi_2 = \varnothing$, they need to be disjoint.

In a context that is the result of updating C_1 with the antecedent (21) (i.e. $C_1' = \langle M_1, \{S\}, j\rangle$, where $[\![(21)(i)(j)]\!]^{M_1} = 1$), ϕ_{DC_S} is a subset of ϕ_2.

As a result, (28-a) cannot be followed up by an utterance of (28-b) (in which the pronoun *that*$_{\phi_2}$ refers to the proposition *that Mary is sick*) by the same speaker without contradiction, since (28-a) doesn't make available any appropriate antecedents which would be compatible with a anti-veridical interpretation.

5.3.3 Anti-Veridical Antecedents

Now we illustrate how we can account for the following contrasts for propositional anaphora with anti-veridical antecedents.

(30) *John was mistaken* [*that Mary is sick*]$^{\phi}$

a. # *but/and that*$_{\phi}$*'s true.*

b. ✓*but she has been suspecting that*$_{\phi}$ *for a while.*

c. ✓*so that*$_{\phi}$*'s not the case.*

The semantic representation for the antecedent sentence is given in (31):

(31) S: *John was mistaken that Mary is sick.* $\rightsquigarrow$

$$\boxed{\begin{array}{c}\phi_1\\ \hline \phi_{DC_S} \Subset \phi_1\end{array}}\,;\mathbf{max}_{\phi_1}\left(\boxed{\begin{array}{c}\upsilon_1,\phi_2\\ \hline \upsilon_1 \equiv \mathsf{John_e}\\ \mathsf{Mistaken}_{\phi_1}\{\upsilon_1,\phi_2\}\end{array}}\,;\mathbf{max}_{\phi_2}\left(\boxed{\begin{array}{c}\upsilon_2\\ \hline \upsilon_2 \equiv \mathsf{Mary_e}\\ \mathsf{Sick}_{\phi_2}\{\upsilon_2\}\end{array}}\right)\right)$$

Here, I treat *be mistaken* as an anti-veridical assertive attitude ascription, analogously to the semantics that Anand and Hacquard (2014) suggest for *be wrong*. The interpretation of the static predicate invoked in the condition $\mathsf{Mistaken}_{\phi_1}\{\upsilon_1,\phi_2\}$ is given in (32).

(32) $[\![\mathsf{mistaken}(x_e)(p_{wt})(w_w)]\!]^{M,g} = [\![\textsc{Assertive}_x \subseteq p \wedge \neg p(w)]\!]^{M,g}$

(32) characterizes *be mistaken* as an anti-veridical operator: The propositional argument is false in the worlds provided as intensional argument. Due to encapsulated quantification over possible worlds, the interpretation of the condition $\mathsf{Mistaken}_{\phi_1}\{\upsilon_1,\phi_2\}$ will

check whether the proposition referred to by ϕ_2 is false in each world in ϕ_1. As a result, any output j will be s.t. $\phi_1(j) \cap \phi_2(j) = \varnothing$.

This discussion also assumes that S's utterance of *John was mistaken that Mary is sick* presupposes that John previously asserted that Mary is sick. So the '$\textsc{Assertive}_x \subseteq p$'-part of (32) should be taken to be presupposed rather than entailed. However, our system is not designed to provide an analysis of presupposition. Instead of giving an account of the presupposition here, we will consider models where the domain of worlds contains only such worlds where John communicated that Mary is sick. As a result, any output j will be s.t. $\phi_1(j) = \overline{\phi_2(j)}$, and because $\phi_{DC_S}(j) \subseteq \phi_1(j)$, ϕ_2 is an anti-veridical dref.

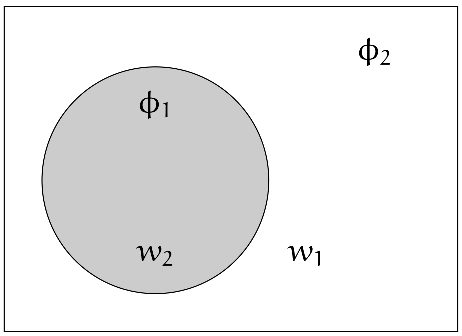
Update. Assume a context $C_2 = \langle M_2, \{S\}, i\rangle$, where i is an initial state and M_2 is s.t. $D_e = \{john, mary, S, \#\}$, $D_w = \{w_1, w_2, w_3\}$, and $[\![sick]\!]^{M_2}$, $[\![mistaken]\!]^{M_2}$ characterize the propositions in **Table 5.7**.

	$\lambda w.sick(mary)(w)$	$\lambda w.mistaken(john)(\lambda w'.sick(mary)(w'))(w)$
w_1	1	0
w_2	0	1

Table 5.7: Truth-values for $\lambda w.sick(mary)(w)$ and $\lambda w.mistaken(john)(\lambda w'.sick(mary)(w'))(w)$ for the worlds in D_w.

Mary is sick in w_1 and John was mistaken that Mary is sick in w_2. D_w does not include a world in which Mary is sick and John was mistaken that Mary is sick, because the latter entails the negation of the former. D_w does not include a world where Mary is not sick and John was not mistaken that Mary is sick, because such a world where John didn't communicate that Mary is sick, i.e. one where the respective presupposition of the utterance is not satisfied.

An update with (31) in C_2 yields an output j, which maps drefs as depicted in **Figure 5.2**.



$$\begin{aligned} \phi_{DC_S}(j) &= \{w_2\} \\ \phi_1(j) &= \{w_2\} \\ \phi_2(j) &= \{w_1\} \\ \upsilon_1(j) &= john_e \\ \upsilon_2(j) &= mary_e \end{aligned}$$

Figure 5.2: Drefs after updating with *'John was mistaken that Mary is sick'* in C_2

This illustrates that ϕ_2 is an anti-veridical dref. The area shaded in gray represents the asserted proposition ϕ_1. The commitments of the speaker are a subset of ϕ_1. As a result, $\phi_{DC_S}(j) \subseteq \overline{\phi_2}$, i.e. ϕ_2 is an anti-veridical dref whose referent is false throughout S's commitment state.

An Anti-Veridical Anaphor

The fact that an anti-veridical anaphor (33-b) can be used with an anti-veridical antecedent (33-a) follows from the assumed representations, because it does not cause a contradictory discourse.

(33) a. *John was mistaken* [*that Mary is sick*]$^{\phi_2}$

b. *so that*$_{\phi_2}$*'s not the case.*

The semantic representation of (33-b) is repeated here:

(29) $\boxed{\begin{array}{c}\phi_3\\ \hline \phi_{DC_S} \Subset \phi_3\end{array}}\,; \mathbf{max}_{\phi_3}\left(\boxed{\phi_3 \equiv \overline{\overline{\phi_2}}}\right)$

Since the non-veridical embedding entails the falsity of the embedded content, this representation requires that the discourse commitments of the speaker are a subset of the complement set of the proposition associated with ϕ_2. Therefore ϕ_2 is an anti-veridical dref.

In a context that is the result of updating C_2 with the antecedent (31) (i.e. $C_2' = \langle M_2, \{S\}, j\rangle$, where $[\![(31)(i)(j)]\!]^{M_2} = 1$), ϕ_2 is an anti-veridical dref and $\phi_2(j) \cap \phi_{DC_S}(j) = \varnothing$. Since this relation already holds for the input j, j(33-b)k is a consistent update, and (33-b) is true in C_2'.

A Veridical Anaphor

The fact that a veridical anaphor (34-b) cannot have an anti-veridical antecedent (34-a) is explained because with the assumed updates, this will have the effect that S is committed to a contradiction.

(34) a. *John was mistaken* [*that Mary is sick*]$^{\phi_2}$

b. # *but/and that*$_{\phi_2}$*'s true.*

The semantic representation of (28-b) is repeated in (25).

(25) S: *That*$_{\phi_2}$*'s true.* $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_3 \\ \hline \phi_{DC_S} \Subset \phi_3 \\ \hline \end{array}; \mathbf{max}_{\phi_3}\left(\boxed{\phi_3 \Subset \phi_2}\right)$$

The condition $\mathtt{True}_{\phi_3}\{\phi_2\}$ requires that $\phi_3(k) \subseteq \phi_2$, and the assertion of ϕ_3 introduces the condition $\phi_{DC_S}(k) \subseteq \phi_3(k)$. In combination, we get that $\phi_{DC_S}(k) \subseteq \phi_2(k)$.

In a context that is the result of updating C_2 with the antecedent (31) (i.e. $C_2' = \langle M_2, \{S\}, j\rangle$, where $[\![(31)(i)(j)]\!]^{M_2} = 1$), ϕ_2 is an anti-veridical dref and $\phi_2(j) \cap \phi_{DC_S}(j) = \varnothing$.

As a result, (33-a) cannot be followed up by an utterance of (28-b) (in which the pronoun *that*$_{\phi_2}$ refers to the proposition *that Mary is sick*) by the same speaker without contradiction, since ϕ_2 is introduced as an anti-veridical dref and then anaphorically referred to as a veridical one.

5.3.4 Non-Veridical Antecedents

This part is concerned with showing how the analysis captures the fact that a non-veridical antecedent can be picked up by propositional anaphora regardless of their veridicality.

(35) *Sue believes* [*that Mary is sick*]$^{\phi_2}$

a. ✓*and that*$_{\phi_2}$*'s true.*

b. ✓*and she has been suspecting that*$_{\phi_2}$ *for a while.*

c. ✓*but that*$_{\phi_2}$*'s not the case.*

The semantic representation of the antecedent is given in (36).

(36) S: *Sue believes that Mary is sick.* $\rightsquigarrow$

$$\boxed{\begin{array}{c}\phi_1\\ \hline \phi_{DC_S} \Subset \phi_1\end{array}}\,;\mathbf{max}_{\phi_1}\left(\boxed{\begin{array}{c}\upsilon_1,\phi_2\\ \hline \upsilon_2 \equiv \mathsf{Sue}_{\mathbf{b}}\\ \mathsf{Believe}_{\phi_1}\{\upsilon_2,\phi_2\}\end{array}}\,;\mathbf{max}_{\phi_2}\left(\boxed{\begin{array}{c}\upsilon_2\\ \hline \upsilon_2 \equiv \mathsf{Mary}_{\mathbf{e}}\\ \mathsf{Sick}_{\phi_2}\{\upsilon_2\}\end{array}}\right)\right)$$

The relationship between the assertion ϕ_1 and the embedded context ϕ_2 is provided by the interpretation of the condition $\mathsf{Believe}_{\phi_1}\{\upsilon_2, \phi_2\}$. I am treating *believe* as a doxastic attitude verb in a Hintikkan semantics for propositional attitudes, following Heim (1992). It makes reference to a doxastic information state, i.e. the set of worlds compatible with the beliefs of an epistemic agent in a given world.

(37) a. $[\![\mathrm{Dox}_x(w)]\!]^{M,g} =$

$\{w' \in D_w : w'$ conforms to what $[\![x]\!]^{M,g}$ believes in $[\![w]\!]^{M,g}\}$

b. $[\![\mathsf{believe}(x_e)(p_{wt})(w_w)]\!]^{M,g} =$

$[\![\mathrm{Dox}_x(w) \subseteq p]\!]^{M,g}$

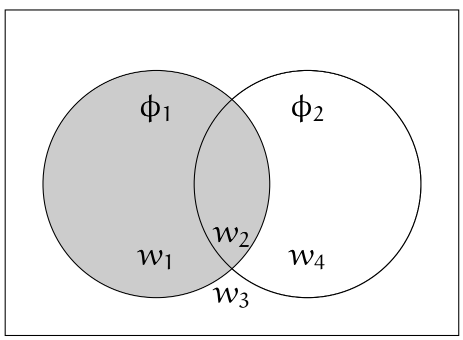
This is a non-veridical embedding. Since the the relationship between ϕ_1 and ϕ_2 is indirect, mediated via Sue's beliefs, the two propositions are independent of each other. Whether or not (S is committed to) *Sue believes that Mary is sick* has no bearing on whether (S is committed to) *Mary is sick*.

The Update. Assume a context $C_3 = \langle M_3, \{S\}, i\rangle$, where i is a discourse-initial context, and M_3 is s.t. $D_e = \{\mathsf{sue}, \mathsf{mary}, S, \#\}, D_w = \{w_1, w_2, w_3, w_4\}$, and $[\![\mathsf{sick}]\!]^{M_3}$, $[\![\mathsf{believe}]\!]^{M_3}$ characterize the propositions in **Table 5.8**.

	$\lambda w.\mathsf{sick}(\mathsf{mary})(w)$	$\lambda w.\mathsf{believe}(\mathsf{sue})(\lambda w'.\mathsf{sick}(\mathsf{mary})(w'))(w)$
w_1	1	1
w_2	1	0
w_3	0	1
w_4	0	0

Table 5.8: Truth-values for the propositions $\lambda w.\mathsf{sick}(\mathsf{mary})(w)$ and $\lambda w.\mathsf{believe}(\mathsf{sue})(\lambda w'.\mathsf{sick}(\mathsf{mary})(w'))(w)$ for the worlds in D_w.

An update with (36) in C_3 yields an output j, which maps drefs as shown in **Figure 5.3**. The asserted proposition ϕ_1 is shaded in gray, and $\phi_{DC_S}(j) \subseteq \phi_1(j)$. However, since ϕ_2 is introduced non-veridically, it can contain worlds where $\phi_1(j)$ is true, and ones where it isn't. As a result, $\phi_{DC_S}(j) \subseteq \phi_2(j)$, i.e. ϕ_2 is a non-veridical dref. The update is compatible with ϕ_{DC_S} containing worlds that are in ϕ_2, and ones that aren't.



$$\begin{aligned}
\phi_{DC_S}(j) &= \{w_1, w_2\}\\
\phi_1(j) &= \{w_1, w_2\}\\
\phi_2(j) &= \{w_2, w_4\}\\
\upsilon_1(j) &= \mathsf{sue}_e\\
\upsilon_2(j) &= \mathsf{mary}_e
\end{aligned}$$

Figure 5.3: Drefs after updating with *'Sue believes that Mary is sick'* in C_3

A Veridical Anaphor

Let's turn to (35-a) to see why the update with the assertion of a speaker S of *Sue believes [that Mary is sick]*$^{\phi_2}$. is compatible with a subsequent update with the assertion of *It*$_{\phi_2}$*'s true*. The semantic representation of *It*$_{\phi_2}$*'s true* is shown again in (25).

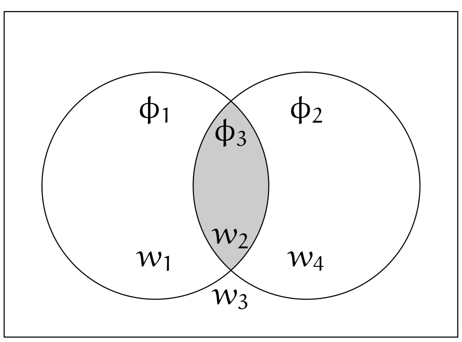
(25) S: *That*$_{\phi_2}$*'s true.* $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_3 \\ \hline \phi_{DC_S} \Subset \phi_3 \\ \hline \end{array}\,; \mathbf{max}_{\phi_3}\left(\begin{array}{|c|}\hline \phi_3 \Subset \phi_2 \\ \hline \end{array}\right)$$

As discussed above, (25) requires that an output k is s.t. $\phi_3(k) \subseteq \phi_2$, while due to the main assertion $\phi_{DC_S}(k) \subseteq \phi_3(k)$. In turn, the combination of the two requires that $\phi_{DC_S}(k) \subseteq \phi_2(k)$.

Consider the context $C_3' = \langle M_3, \{S\}, j\rangle$, where j is the result of updating with the antecedent (36) in C_3, i.e $[\![(36)(i_{C_3})(j_{C_3'})]\!]^{M_3} = 1$. The relation $\phi_{DC_S} \Subset \phi_2$ doesn't already hold for j. However (for the conversationally privileged value of j) $\phi_{DC_S}(j)$ contains worlds that are in $\phi_2(j)$ and ones that are not, and can be narrowed down consistently.

An update with (25) in C_3' yields an output k, which maps drefs as shown in **Figure 5.4**.



$$\begin{aligned} \phi_{DC_S}(k) &= \{w_2\} \\ \phi_1(k) &= \{w_1, w_2\} \\ \phi_2(k) &= \{w_2, w_4\} \\ \upsilon_1(k) &= \mathit{sue}_e \\ \upsilon_2(k) &= \mathit{mary}_e \\ \phi_3(k) &= \{w_2\} \end{aligned}$$

Figure 5.4: Drefs after updating with *'Sue believes that [Mary is sick]*$^{\phi_2}$*'*; *That*$_{\phi_2}$*'s true* in C_3

The two subsequent assertions of ϕ_1 and ϕ_3 are interpreted in conjunction and the speaker is committed to both $\phi_{DC_S}(k) \subseteq \phi_1(k) \cap \phi_2(k)$. As a result, S is committed to both propositions. Since $\phi_3(k) \subseteq \phi_2(k)$, the speaker is now committed to ϕ_2 as well. The update of (25) in C_3' does not commit the speaker to a contradiction, and the dref ϕ_2, which has been introduced non-veridically receives a strengthened interpretation.

An Anti-Veridical Anaphor

We see a similar effect of strengthening for anti-veridical anaphora with non-veridical antecedents.

(35) *Sue believes* [*that Mary is sick*]$^{\phi_2}$

(35-c)*but that*$_{\phi_2}$*'s not the case.*

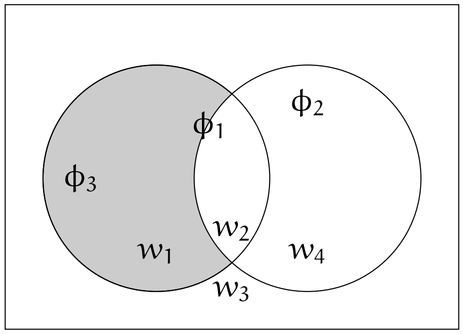
The semantic representation of (35-c) is given again in (29).

$$(29) \quad \boxed{\begin{array}{c} \phi_3 \\ \hline \phi_{DC_S} \Subset \phi_3 \end{array}} ; \mathbf{max}_{\phi_3} \left(\boxed{\phi_3 \equiv \overline{\overline{\phi_2}}} \right)$$

Importantly, (29) requires that an output k is s.t. $\phi_3(k) = \overline{\phi_2(k)}$, while due to the main assertion $\phi_{DC_S}(k) \subseteq \phi_3(k)$. In turn, the combination of the two requires that $\phi_{DC_S}(k) \subseteq \overline{\phi_2(k)}$, i.e. ϕ_2 is an anti-veridical anaphor.

This relation doesn't already hold for $j_{C_3'}$ (in $C_3' = \langle M_3, \{S\}, j\rangle$, where j is s.t. $[\![(36)(i_{C_3})(j_{C_3'})]\!]^{M_3} = 1$). However (for the conversationally privileged value of j) $\phi_{DC_S}(j)$ contains worlds that are in $\phi_2(j)$ and ones that are not, and can be narrowed down consistently.

An update with (29) in C_3' yields an output k, which maps drefs as shown in **Figure 5.5**.



$$
\begin{aligned}
\phi_{DC_S}(k) &= \{w_1\} \\
\phi_1(k) &= \{w_1, w_2\} \\
\phi_2(k) &= \{w_2, w_4\} \\
\upsilon_1(k) &= sue_e \\
\upsilon_2(k) &= mary_e \\
\phi_3(k) &= \{w_1\}
\end{aligned}
$$

Figure 5.5: Drefs after updating with *'Sue believes* [*that Mary is sick*]$^{\phi_2}$*'*; *'but that*$_{\phi_2}$*'s not the case'* in C_3

An update with (36) in C_3 introduces the dref ϕ_2 as a non-veridical dref that is not anti-veridical. The commitments of the speaker contain worlds that are in ϕ_2, and ones that are not. A subsequent update with (29) narrows down the speaker's commitments in a way s.t. it only contains worlds where ϕ_2 is false. The dref that was introduced non-veridically is strengthened and now interpreted anti-veridically.

5.4 Non-Veridical Discourse Relations

All of the above cases are ones where the discourse segment containing the antecedent and the discourse segment containing the anaphor are required to be consistent with each other. In the above cases, that is because they are both assertions by the same speaker. If the two discourse segments are not required to be consistent with each other, the same restrictions do not apply. Consider for example, a case of disjunction (38), where a propositional anaphor in the indicative second disjunct can pick the proposition introduced under negation in the first. Similarly, (39) is a case of inter-speaker disagreement, where first A introduces the proposition *Mary was sick* in an anti-veridical context, and then B makes reference to it veridically.

(38) Disjunction:

[Context:] You are planning a party with your friends. Your friend is concerned that Mary might have been exposed to a dangerous virus and infect others without knowing. You know that Mary tests herself regularly. You tell your friend: *Either Mary is not sick, or she knows it already.*

(39) Disagreement:

A: *John was mistaken [that Mary is sick]*$^{\phi}$.

B: *Actually, it*$_{\phi}$*'s true.*

Let us see how we can account for the availability of these anaphora in these cases.

5.4.1 Disjunction

Disjuncts do not need to be consistent with each other. On a global level, the discourse variable is non-veridical in both the antecedent and anaphoric contexts.

(40) a. *Mary is not [~~Mary~~*$^{\upsilon}$ *sick]*$^{\phi_4}$*, or she*$_{\upsilon}$ *knows it*$_{\phi_4}$.

b.

$$\begin{array}{|c|}\hline \phi_1 \\ \hline \phi_{\mathrm{DC}_S} \Subset \phi_1 \\ \hline \end{array}; \mathbf{max}_{\phi_1}\left(\begin{array}{|c|}\hline \phi_2, \phi_3 \\ \hline \phi_1 \equiv \phi_2 \Cup \phi_3 \\ \hline \end{array};\right.$$

$$\mathbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \phi_4 \\ \hline \phi_2 \equiv \overline{\overline{\phi_4}} \\ \hline \end{array}; \mathbf{max}_{\phi_4}\left(\begin{array}{|c|}\hline \upsilon \\ \hline \upsilon \equiv \mathsf{Mary}_{\mathbf{e}} \\ \mathsf{Sick}_{\phi_4}\{\upsilon\} \\ \hline \end{array}\right)\right);$$

$$\left.\mathbf{max}_{\phi_3}\left(\begin{array}{|c|}\hline \mathsf{Know}_{\phi_3}\{\upsilon, \phi_4\} \\ \hline \end{array}\right)\right)$$

c. $[\phi_1 \mid \phi_{\mathrm{DC}_S} \Subset \phi_1]; \mathbf{max}_{\phi_1}([\phi_2, \phi_3 \mid \phi_1 \equiv \phi_2 \Cup \phi_3]; \mathbf{max}_{\phi_2}([\phi_4 \mid \phi_2 \equiv \overline{\overline{\phi_4}}]; \mathbf{max}_{\phi_4}([\upsilon \mid \upsilon \equiv \mathsf{Mary}_{\mathbf{e}}]; [\mathsf{Sick}_{\phi_4}\{\upsilon\}])); \mathbf{max}_{\phi_3}([\mathsf{Know}_{\phi_3}\{\upsilon, \phi_4\}]))$

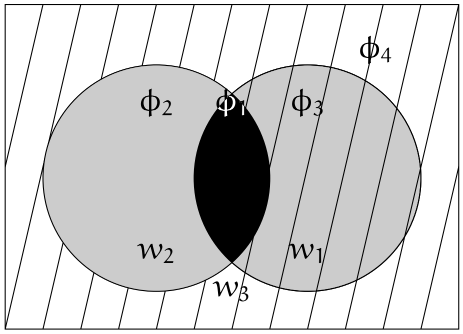
The two sequential assertions in the previous examples were interpreted in conjunction, leading to a requirement that the intersection between the propositions expressed by the two clauses couldn't empty. The same doesn't apply under disjunction. Since disjunction is a non-veridical operator, both the antecedent and the anaphor are in a non-veridical context. We do not predict constraints for this combination, and the anaphoric relation is predicted to be possible.

The update. We illustrate the effect of this update concretely by assuming a context $C_5 = \langle M_5, \{S\}, i\rangle$, where i is a discourse-initial state, and M_5 is s.t. $D_e = \{\mathsf{mary}, S, \#\}$, $D_w = \{w_1, w_2, w_3\}$, and $[\![\mathsf{sick}]\!]^{M_5}$ and $[\![\mathsf{know}]\!]^{M_5}$ characterize the propositions in **Table 5.9**.

	$\lambda w.\mathsf{sick}(\mathsf{mary})(w)$	$\lambda w.\mathsf{know}(\lambda w'.\mathsf{sick}(\mathsf{mary})(w'))(\mathsf{mary})(w)$
w_1	1	1
w_2	1	0
w_3	0	0

Table 5.9: Truth-values for $\lambda w.\mathsf{sick}(\mathsf{mary})(w)$ and $\lambda w.\mathsf{know}(\lambda w'.\mathsf{sick}(\mathsf{mary})(w'))(\mathsf{mary})(w)$ for the worlds in D_w.

Updating i_{C_5} with (40) yields an output j, which maps the variables as depicted in **Figure 5.6**.



$$\begin{aligned}
\phi_{DC_S}(j) &= \{w_1, w_2\}\\
\phi_1(j) &= \{w_1, w_2\}\\
\phi_2(j) &= \{w_2\}\\
\phi_3(j) &= \{w_1\}\\
\phi_4(j) &= \{w_1, w_3\}\\
\upsilon(j) &= mary_e
\end{aligned}$$

Figure 5.6: Drefs after updating with *'(Either) Mary is not sick, or she (already) knows it'* in C_5

In **Figure 5.6**, the anaphorically retrieved proposition ϕ_4 is depicted as the hatched area. All the relevant relationships between propositions are satisfied here: ϕ_1 is the union of ϕ_2 and ϕ_3, the asserted proposition. ϕ_4, a proposition which is introduced as the complement of ϕ_2 is also true in the context of ϕ_3. This is given here, because $\phi_3 \subset \phi_4$. Because the intersection of ϕ_2 and ϕ_3 is allowed to be empty, this output is consistent with all conditions placed on the relationships between the drefs, and an anaphoric relation is possible here.

5.4.2 Disagreement

The disagreement-case works similarly to disjunction. Because both discourse segments are assertions by different speakers (who are disagreeing with each other), their utterances can be interpreted as contradictory and the intersection between the asserted propositions can be empty. Let's see how this works for (41).

(41) Disagreement:

$\mathcal{A}$: *John was mistaken [that Mary is sick]*$^{\phi_2}$.

B: *It*$_{\phi_2}$*'s true.*

The antecedent utterance

Consider the following representation for utterance containing the anti-veridical antecedent utterance, repeated from (31), above.

(31) $\mathcal{A}$: *John was mistaken that Mary is sick.* $\rightsquigarrow$

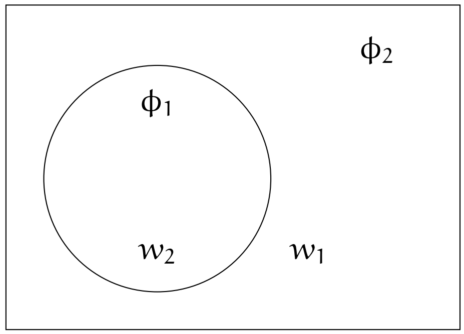
$$\begin{array}{|c|}\hline \phi_1 \\ \hline \phi_{DC_{\mathcal{A}}} \Subset \phi_1 \\ \hline \end{array} ; \mathbf{max}_{\phi_1}\left(\begin{array}{|c|}\hline \upsilon_1, \phi_2 \\ \hline \upsilon_1 \equiv \mathsf{John}_{\mathbf{e}} \\ \mathsf{Mistaken}_{\phi_1}\{\upsilon_1, \phi_2\} \\ \hline \end{array} ; \mathbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \upsilon_2 \\ \hline \upsilon_2 \equiv \mathsf{Mary}_{\mathbf{e}} \\ \mathsf{Sick}_{\phi_2}\{\upsilon_2\} \\ \hline \end{array}\right)\right)$$

The effect of this update is illustrated here for the context $C_6 = \langle M_6, \{\mathcal{A}, B\}, i\rangle$, where i is a discourse-initial state, and M_6 is s.t. $D_e = \{\mathsf{john}, \mathsf{mary}, \mathcal{A}, B, \#\}$, $D_w = \{w_1, w_2\}$, and $[\![\mathsf{sick}]\!]^{M_6}$ and $[\![\mathsf{mistaken}]\!]^{M_6}$ characterize the propositions in **Table 5.10**.

	$\lambda w.\mathsf{sick}(\mathsf{mary})(w)$	$\lambda w.\mathsf{mistaken}(\mathsf{john})(\lambda w'.\mathsf{sick}(\mathsf{mary})(w'))(w)$
w_1	1	0
w_2	0	1

Table 5.10: Truth-values for $\lambda w.\mathsf{sick}(\mathsf{mary})(w)$ and $\lambda w.\mathsf{mistaken}(\mathsf{john})(\lambda w'.\mathsf{sick}(\mathsf{mary})(w'))(w)$ for the worlds in D_w.

This update with (31) in C_6 proceeds analogously to the update with the anti-veridical antecedent in the previous section (**5.3.3**) and its output j maps variables as depicted in **Figure 5.7**.



$$\begin{aligned} \phi_{DC_A}(j) &= \{w_2\} \\ \phi_1(j) &= \{w_2\} \\ \phi_2(j) &= \{w_1\} \\ \upsilon_1(j) &= \mathrm{john}_e \\ \upsilon_2(j) &= \mathrm{mary}_e \end{aligned}$$

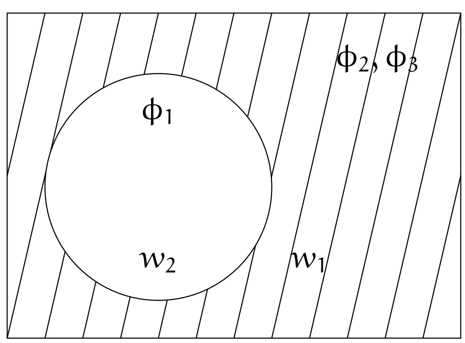
Figure 5.7: Drefs after updating with A's utterance of *John was mistaken [that Mary is sick]*$^{\phi_2}$ in C_6

The utterance containing the anaphor

Now the discourse state j_{C_6} is compatible with B's assertion (42):

(42) B : *It*$_{\phi_2}$*'s true.* $\rightsquigarrow \boxed{\begin{array}{c} \phi_3 \\ \hline \phi_{DC_B} \in \phi_3 \end{array}}\,; \mathbf{max}_{\phi_3}\left(\boxed{\phi_3 \in \phi_2}\right)$

Updating j_{C_6} with (42) yields an output k with the mappings shown in **Figure 5.8**.



$$\begin{aligned} \phi_{DC_A}(k) &= \{w_2\} \\ \phi_{DC_B}(k) &= \{w_1\} \\ \phi_1(k) &= \{w_2\} \\ \phi_2(k) &= \{w_1\} \\ \phi_3(k) &= \{w_1\} \\ \upsilon_1(k) &= \mathrm{john}_e \\ \upsilon_2(k) &= \mathrm{mary}_e \end{aligned}$$

Figure 5.8: Drefs after updating with A's utterance of *John was mistaken [that Mary is sick]*$^{\phi_2}$; then B's utterance of *It*$_{\phi_2}$*'s true* in C_6

Similarly to the disjunction case, the two propositions asserted by disagreeing speakers A and B are interpreted as inconsistent. As a result, the intersection of the respective asserted propositions, ϕ_1 and ϕ_2 may be empty. Since ϕ_3 was introduced by A as an anti-veridical proposition, the opposite of ϕ_1, this empty intersection is what allows B to pick up ϕ_3 veridically in their assertion of ϕ_2.

5.5 Conclusion

This chapter has made two main points about propositional anaphora: First, the empirical patterns show that anaphoric accessibility for propositional anaphora patterns systematically along the lines of veridicality distinctions. Second, as long as our dynamic semantic theory allows for: 1. externally dynamic non-veridical operators, and 2. keeping track of the relationship between propositional drefs and the interlocutors' commitments. An analysis of the patterns can be derived based on 1. a requirement that propositional anaphora are interpreted as true in their local context, and 2. the general requirement that a discourse be consistent.

Chapter 6

Discourse Negativity and Anaphoric Polarity-Sensitivity

6.1 Introduction

The previous chapters' claims and results about the interaction between propositional anaphora and veridicality have consequences for the relationship between anaphora and negation more broadly. The view represented here suggests that polarity-sensitivity for anaphora can be explained based on veridicality. This chapter gives a pragmatic account of anaphoric polarity-sensitivity in relation to speakers' commitments along the lines of the previous chapters: We characterize negative polarity for the purposes of the discourse (discourse-negativity) on the level of discourse-representation, as a context as one that introduces an anti-veridical propositional dref.

The chapter shows that this assumption allows us to give an analysis for negativity-tags. I use this term for expressions which have been characterized as anaphoric to propositions, and sensitive to the polarity of the antecedent. These are the negativity-tags introduced by Klima (1964), polarity particles (Pope, 1972; Ginzburg and Sag, 2000; Kramer and Rawlins, 2009; Farkas and Bruce, 2010; Cooper and Ginzburg, 2012; Holmberg, 2013; Roelofsen and Farkas, 2015), certain cases of polarity ellipsis (Kramer and Rawlins, 2009; Hofmann, 2018), as well as anti-veridical propositional anaphora.

The observation that some anaphoric expressions are sensitive to the polarity of their antecedent clause goes back to Klima (1964), who first identified a class of negativity-tags as diagnostics for a heterogeneous class of negative sentences.

(1) Klima's negativity-tags: [Klima (1964)]

a. *Writers won't accept suggestions,*

(i) *...and neither will publishers*

(ii) *...will they?*

(iii) *...not even reasonable ones.*

(iv) *...and publishers won't accept them, either.*

b. *Writers will accept suggestions,*

(i) #*...and neither will publishers*

(ii) #*...will they?*

(iii) #*...not even reasonable ones.*

(iv) #*...and publishers will accept them, either.*

Klima used these diagnostics to show that the class of contexts that we might count as ‘negative’ is heterogeneous. While Klima’s negativity-tags pick out negative antecedents from positive ones, they are licensed in a variety of contexts, besides just ones involving the negative marker *not*. In English, negativity can be introduced by anti-additive quantifiers (*never, nobody, nothing, no NP*), downward entailing quantifiers (*rarely, few NP*), or negative proximatives (*hardly, barely, hardly any NP*) in adverbial positions, or argument positions, in addition to the overt negative marker *not*.

We are faced with two interrelated questions: **(i)** what does it mean for a context to be negative in a sense that is relevant for the interaction with anaphora? That is, how can the negativity of the antecedent for anaphoric purposes (DISCOURSE-NEGATIVITY) be characterized?, and **(ii)** how can the polarity-sensitivity of the anaphora be accounted for?

Discourse polarity is determined in complex ways. Some of them are morphosyntactic and overt, while some of them aren’t. The discussion of these mechanisms in this chapter will focus on what renders a sentence negative in the sense that it is a suitable antecedent for a negativity-tag.

In addition to the semantic and pragmatic variation of discourse-negative contexts discussed in the previous literature, this chapter discusses some additional variation on a pragmatic dimension, providing evidence that suggests that negativity may be introduced implicitly in discourse. Examples of contexts in which this is the case are the contexts which Kroll (2019, 2020) identifies as contexts licensing polarity-reversals in sluicing, i.e., neg-raising and disjunction, as well as contexts involving sarcasm or accommodation.

I argue that the implicit contextual introduction of discourse-negativity suggests that it is not a property of sentences or the syntactic or semantic representation of clauses. Instead, the new data points towards a characterization on a dynamic semantic level, which allows for interaction with contextual factors.

In particular, the chapter suggests that we can suitably characterize a discourse-negative utterance in terms of the introduction of an anti-veridical propositional dref in addition to the dref corresponding to the entire assertion. This is based on the assumptions from Stone (1999); Stone and Hardt (1999); Krifka (2013); Snider (2017)) about the anaphoric potential of sentences including negation, and in line with the generalizations in the previous chapters, about the interaction between anaphora and negation, as well as other non-veridical operators. The current chapter shows how this characterization can be applied to the diverse landscape of discourse-negativity and aims to provide an argument that it is more suitable for the analysis of discourse-negativity than other accounts of the polarity-sensitivity of negativity-tags and expressions that pattern like them.

The chapter proceeds as follows:

Section 6.2 provides a brief discussion of the negativity-tags introduced by Klima, and some additional anaphoric expressions, which exhibit the same kind of polarity-sensitivity. This section aims to establish a few negativity-tags that could serve as diagnostics of discourse negativity for the upcoming discussion.

Section 6.3 provides an overview of the syntactically and semantically diverse landscape of environments that are negative for the purposes of the discourse, i.e., ones that license negativity tags.

Section 6.4 provides an overview of different possible characterizations of discourse-negativity: On the level of a syntactic or semantic representation of a clause, its pragmatic effect, or its anaphoric potential in a dynamic semantic framework. The section also discusses how an account on any of these levels of analysis could explain the diversity of negative licensing expressions discussed in **Section 6.3.**

Section 6.5 introduces new data that can help distinguish between the approaches. We will look at data where discourse-negativity is introduced contextually, which provides evidence for a characterization that is amenable to contextual manipulation. I suggest that our generalization, that a discourse-negative utterance is one that updates the discourse to include an anti-veridical propositional dref, is a suitable generalization for the data we have seen. The section goes on to discuss some data that seems to favor a characterization that does make reference to syntactic structure and how the challenges posed by these data points can be addressed.

Section 6.6 reports an experiment that supports the view that discourse-negativity does not rely on sentential negation. It compares the availability of negativity-tags (in particular *'Why not'*-questions) with antecedents that involve different kinds of anti-veridical embeddings. The results show that cases of neg-raising that are uncontroversially derived pragmatically, as well as anti-veridical attitudes, all show proportions of *'why not'* follow-ups that would not be expected if they are interpreted like a non-negative antecedent.

Section 6.7 illustrates how the negative antecedent requirement and other facts about the interaction with negation of the negativity-tag *'why not'* may be analyzed in our system with only minimal extensions.

Section 6.8 concludes the chapter by summarizing its main points.

6.2 Negativity-Tags

The anaphoric expressions we use as diagnostics of discourse-negativity in English are Klima's *neither*-tags, agreeing uses of negative PolPs, canonical-question uses of *'Why not?'*, and anti-veridical propositional anaphora.

Neither-tags

Among Klima's negativity-tags are *neither*-tags, which involve the negative polar additive *neither* in a sentence with verb-phrase ellipsis (VPE). Requiring a negative antecedent, *neither*-tags can be used with a negative antecedent utterance, but not with a positive one.

(2) *Neither*-tags:

a. *Camels don't store water in their humps. Neither do andromedarys.*

('Andromedarys don't store water in their humps either.')

b. #*Camels store water in their humps. Neither do andromedarys.*

Klima's diagnostics fall into a larger class of cross-clausal anaphoric expressions that are sensitive to the polarity of their antecedent. Some other anaphoric expressions which require a negative antecedent are agreeing uses of negative polarity particles *'no'*, canonical question uses of reduced reason-interrogatives *'Why not?'*, and anti-veridical uses of propositional anaphora. These contrasts are illustrated below.

Polarity Particles

The interpretation of polarity particles (like English *yes/no*) is also sensitive to the polarity of their antecedent clause (Pope (1972); Ginzburg and Sag (2000); Kramer and Rawlins (2009); Farkas and Bruce (2010); Cooper and Ginzburg (2012); Holmberg (2013); Brasoveanu et al. (2013); Roelofsen and Farkas (2015)). In particular, agreeing uses of *no* are possible with negative antecedents, but not positive ones.

(3) Agreeing uses of *no*:

a. A: *Camels don't store water in their humps.*

B: *No, Camels don't store water in their humps.*

b. A: *Camels store water in their humps.*

B: *#No, Camels store water in their humps.*

The contrast reported in the literature (3) is supported by experimental evidence from Brasoveanu et al. (2013), showing that, in agreeing responses to positive assertions, *yes* is overwhelmingly preferred over *no* and in agreeing responses to negative assertions, either *yes* or *no* can be used. Based on this contrast, Brasoveanu et al. (2013) suggest that agreeing uses of *no* can be used as a diagnostic of sentential negativity in the sense of Klima.

***'Why not'* Questions**

A further diagnostic we will be considering here comes from canonical question uses of reduced *why not* questions (Hofmann (2018); Anand et al. (2021)). We will call these uses factive uses of *'Why not'*.[1] These involve clausal ellipsis and a remnant including

[1]Canonical question / factive uses of *why not* include information-seeking question uses or factive embedded uses. These uses require a negative antecedent, in contrast to modal interpretations of *why not*, which arise under rhetorical question uses or in non-veridical embeddings.

(i) a. A: *Let's go to the movies!*
B: *Sure, why not!*
(We should go to the movies)

negation and the question-word *why*, and are available with negative antecedents, but not positive ones.

(4) Factive *why not*:

a. *Camels don't store water in their humps, but I don't know why not.*

('I don't know why camels don't store water in their humps.')

b. #*Camels store water in their humps, but I don't know why not.*

As a result, we might similarly consider these as diagnostics of negativity.

Anti-veridical Propositional Anaphora

Based on the discussion in Part I of this dissertation, we also consider propositional anaphora in anti-veridical embeddings. **Chapter 5** suggested that these require an antecedent that is also anti-veridical (in case both discourse segments are uttered by the same speaker). As such, they may be consistently used with a negative antecedent sentence but not with a positive one (5).

(5) a. *Camels don't store water in their humps. That is a misconception.*

(that ≈ that camels store water in their humps)

b. #*Camels store water in their humps. That is a misconception.*

Based on these contrasts, we will use the four diagnostics discussed here—*neither*-tags,

b. *Sue wants to have a barbeque and I don't really see why not.*
(I don't see why we shouldn't/wouldn't have a barbeque)

agreeing uses of negative polarity particles, factive *'Why not'*, and discourse-consistent uses of anti-veridical propositional anaphora—to distinguish between negative and affirmative antecedent utterances in discourse.

6.3 Negative Environments

Klima's negativity tags and the additional anaphora requiring a negative antecedent listed here are not available with positive antecedents. Since Klima's seminal paper, it is known that the class of possible antecedent sentences, which pattern as discursively negative, is heterogeneous.

The variation can be characterized along a syntactic and a semantic dimension (Klima (1964), see also discussion in Brasoveanu et al. (2014)). The syntactic variation arises because discourse negativity can be introduced by the negative marker *not*, negative quantifiers in argument position, and negative sentential adverbials. The semantic variation arises because negative operators introducing sentential negativity can be characterized as having varying semantic strength in the sense of Zwarts (1995): They include the anti-morphic negative marker *not* (as shown above)[2], anti-additive negative quantifiers

[2]A function $f : A \to B$, where A, B are boolean algebras is anti-morphic (cf. Zwarts (1995)) iff it satisfies both sentences of De Morgan's laws, i.e. iff

- $f(a \vee b) = f(a) \wedge f(b)$ and
- $f(a \wedge b) = f(a) \vee f(b)$

- Jasmine doesn't drink and smoke. $\Leftrightarrow$ Jasmine doesn't drink or Jasmine doesn't smoke.
- Nobody drinks and smokes. $\not\Leftrightarrow$ Nobody drinks or nobody smokes.
- I hardly drink and smoke. $\not\Leftrightarrow$ I hardly drink or I hardly smoke.

(*never, nobody, nothing, no NP*), certain (although not all) downward-entailing quantifiers (*rarely, few NP, not all NP*), and negative proximatives (*hardly, barely, hardly any NP*, which can be characterized as downward-asserting (i.e. downward-entailing in their at-issue content, cf. Horn (2002, 2011, 2016)). All of these operators also license NPIs, which suggests that they are all 'negative' in some way.

6.3.1 Anti-additive Quantifiers

Negativity-tags are licensed following utterances including anti-additive quantifiers[3].

(6) Negativity-tags following negative quantifiers:

a. Sentence adverbials:

Jasmine never exercises

(i) *Neither does Sue.*

(ii) *No, she never does.*

(iii) *She never told me why not, though.*

(iv) *That was a lie.*

[3]A function $f : A \to B$, where A, B are boolean algebras is anti-additive (cf. Zwarts (1995)) iff it satisfies at least the first sentence of De Morgan's laws, i.e. iff

- $f(a \vee b) = f(a) \wedge f(b)$
- Jasmine doesn't drink or smoke. $\Leftrightarrow$ Jasmine doesn't drink and Jasmine didn't smoke.
- Nobody drinks or smokes. $\Leftrightarrow$ Nobody drinks and nobody smokes.
- I hardly drink or smoke. $\not\Leftrightarrow$ I hardly drink and I hardly smoke.

b. Subject position:

No student read the paper.

(i) *Neither did the instructor.*

(ii) *No, no student did.*

(iii) *I need to find out why not.*

(iv) *Even though they claimed that.*

c. Object position:

The PR-agents gave no comments at this time

(i) *Neither did the chair of the board.*

(ii) *No, they didn't.*

(iii) *but they didn't say why not.*

(iv) *That was just a rumor.*

The observation that negative quantifiers in adverbial and argument positions license subsequent negativity-tags is documented in Klima (1964), for *neither*-tags (as well as the other Klima-tags). It is further supported experimentally by the studies in Brasoveanu et al. (2013) and Brasoveanu et al. (2014). Both studies present experimental data from forced-choice tasks showing that English utterances including negative quantifiers, lead to a preference for a negativity tag over a follow-up that is compatible with a positive antecedent. The studies used agreeing uses of negative PolPs (Brasoveanu et al. (2013)),VPE in agreeing utterances[4], and question-tags (Brasoveanu et al. (2014)) as diagnostics of negativity and compared the proportion of follow-ups that indicate a negative antecedent and ones that indicate a positive one.

Experiment 2 (with PolPs) and experiment 3 (with agreeing sentences including VPE) presented in Brasoveanu et al. (2013), as well as Brasoveanu et al. (2014) suggest that anti-additive neg-words in subject position robustly pattern with the controls involving negative markers (*not*). This experiment confirms Klima's observations: Sentences with negative quantified subjects pattern like negative sentences rather than positive ones.

For PolPs, sentences with neg-words allow for both *yes* and *no* in agreeing responses, but exhibit a preference for *no*. Overall, the results suggest that anti-additive subjects and adverbs license negativity tags to the same extent as utterances with negative markers.

Negative quantifiers in object position also received a high proportion of follow-ups licensed by negativity in Brasoveanu et al. (2014). However, this proportion was lower than in the case of negativity introduced by adverbs and subjects.

While the results show that neg-words pattern as negative for the purposes of licensing negativity-tags in discourse, object neg-words do so less robustly than those realized in adverb or subject position.

---

[4]For example:

(i) Brasoveanu et al. (2013): (10+11)

a. A: *Mary visited some of the children.*
   B: *I agree, she did / * didn't.*

b. A: *Mary didn't visit any of the children.*
   B: *I agree, she * did / didn't.*

The polarity-sensitivity of configurations like this are further discussed in chapter 10.

6.3.2 Downward-Entailing Quantifiers and Negative Proximatives

Further, Klima points out that negativity can be introduced by different kinds of (often weakly) negative expressions. Besides the anti-additive negative markers and negative quantifiers above, sentential negativity may be introduced by negative proximatives like *hardly* or *barely* (Klima (1964)), as well as downward-entailing[5] quantifiers (Brasoveanu et al. (2014)). For negative proximatives, this is illustrated in (7).

(7) Negative proximatives

a. Sentence adverbials:

Jim hardly/barely exercises

(i) *Neither does Sue.*

(ii) *No, he hardly/barely does.*

(iii) *but he never told me why not.*

(iv) *that was just a rumor.*

b. Subject position:

Hardly/barely any student read the paper

(i) *Neither did the instructor.*

[5] A function $f : A \rightarrow B$, where $\langle A, \leq \rangle, \langle B, \leq \rangle$ are boolean algebras is downward-entailing (cf. Zwarts (1995)) iff

- $(f(a) \wedge b \leq a) \rightarrow f(b)$
- Jasmine didn't eat. $\rightarrow$ Jasmine didn't eat breakfast.
- Nobody ate. $\leftrightarrow$ Nobody ate breakfast.
- I barely ate. $\not\leftrightarrow$ I barely ate breakfast.

(ii) *No, hardly any students did.*

(iii) *and I need to figure out why not.*

(iv) *that was just a rumor.*

c. Object position:

My PR-agent gave hardly/barely any comments about this

(i) *Neither does my PR-agent.*

(ii) *No, she gave hardly/barely any comments.*

(iii) *and I won't tell you why not.*

(iv) *that was just a rumor.*

Klima (1964) reports that there is some variation regarding these contexts. While many English speakers accept the negativity-tags in (7), some do not. Horn (2002, 2011) characterizes negative proximatives as downward-asserting, i.e., downward-entailing in their asserted content, while the component that prevents negative proximatives from being full downward-entailing is part of a backgrounded meaning dimension, which is 'assertorically inert'. Horn made this distinction to account for the fact that negative proximatives behave like downward-entailing operators for the purposes of NPI licensing.

Negative proximatives and downward-entailing quantifiers show similar patterns, as well. Besides the negative proximatives discussed by Klima, Brasoveanu et al. (2014) show that certain downward-entailing quantifiers (in particular English *few* and *rarely*) also introduce negativity into the discourse. This is illustrated in (8).

(8) Downward-entailing quantifiers:

a. Sentence adverbials:

Jim rarely exercises.

(i) *Neither does Sue.*

(ii) *No, he rarely does.*

(iii) *but he never told me why not.*

(iv) *that was just a rumor.*

b. Subject position:

Few students read the paper.

(i) *Neither did the instructor.*

(ii) *No, hardly any students did.*

(iii) *and I need to figure out why not.*

(iv) *that was just a rumor.*

c. Object position:

I have few comments at this time.

(i) *Neither does my PR-agent.*

(ii) *No, she gave hardly/barely any comments.*

(iii) *and I won't tell you why not.*

(iv) *that was just a rumor.*

Although these weaker negative expressions introduce sentential negativity to some extent and the follow-ups in (7) and (8) are acceptable for some speakers, they seem to do so less strongly than anti-additive negative operators. Klima's remarks about the degree to which negative proximatives introduce sentential negativity in English suggest that it is subject to some dialectal variation.

6.3.3 Brasoveanu et al. (2014): Scales of Negativity

Further testament to a varying degree of sentential negativity introduced by these weakly negative expressions is provided by the experimental results presented in Brasoveanu et al. (2014). The study used question-tags as a diagnostic of discourse-negativity. In a forced-choice task, participants were presented with an utterance and asked to choose between two different follow-ups, indicating a negative or positive antecedent, respectively. For negative operators in the antecedent utterance, they manipulated the negative strength (i.e., anti-additives and downward-entailing ones) and syntactic position, i.e., quantifiers in subject, object, and sentence-adverb position.

Compared the proportion of negativity-tags following utterances including negative expressions of these different kinds, the study finds that the ratio of choices in favor of a negativity-tag is modulated by the two factors introduced as the independent variables. There was a cline between anti-morphic, anti-additive, and downward-entailing negative expressions, but only the difference between anti-additive negative expressions and downward-entailing ones was found to be statistically significant. The other significant difference they found was whether or not the negative expression is in an overt position in the I-domain or introduced in an object DP, so whether it is scoping over sentence radical in its surface position or whether this scope-relation needs to be derived. Their experimental findings give rise to the following hierarchy of sentential negativity, where linguistic devices higher on the scale introduce sentential negativity more robustly than those lower on the scale:

(9) Hierarchy of sentential negativity: [Brasoveanu et al. (2014)]

{NEG, N-ADV, N-SUBJ} >> N-OBJ >> {DE-ADV, DE-SUBJ} >> DE-OBJ

In particular, they find that anti-additive negative expressions are followed by a higher proportion of negativity-tags than downward-entailing operators. Also, sentential negation, sentential adverbs, and subjects are followed by a higher proportion of negativity-tags than objects. A possible conclusion is that these expressions introduce discourse-negativity more robustly.

The experimental data reported in Brasoveanu et al. (2013, 2014) experimentally confirms the observations reported in the literature, i.e., that question-tags, polarity particles, and agreeing utterances with VPE are licensed with negative antecedents, but not positive ones. The studies also support the generalization that anti-additive quantifiers and downward-entailing quantifiers introduce discourse-negativity. While these are clearly negative compared to a positive baseline, the gradience found in the forced-choice data shows that different negative expressions are associated with different proportions, suggesting that they might be characterized as more or less negative in some sense.

The gradience found here suggests that some of these expressions are in some sense ‘more/less negative’ than others. This raises the question of whether the gradience should be explained as part of a theory of linguistic competence and representations or linguistic performance and sentence processing. While this question lies outside of the scope of this work, I will briefly return to it in **Chapter 8**, the conclusion, where I briefly outline some avenues for future research.

6.4 Approaches to Discourse-Polarity

The sensitivity of cross-clausal anaphora to the polarity of their antecedent poses a theoretical challenge: It calls for a representation of the antecedent polarity and a mechanism by which anaphoric processes may be sensitive to it. This section (and the one after this one) addresses the question of what makes an utterance discourse-negative, thereby allowing for subsequent negativity-tags. Here, we compare different possible characterizations of discourse-negativity:

Ones that make reference to a syntactic or semantic notion of sentential negation, and ones that appeal to the pragmatic effect of negative sentences. It also discusses an commitment-based discourse-level approach based on the theoretical claims made in **Chapters 2& 5**.

6.4.1 (Morpho)-Syntactic Characterizations

In his seminal paper, Klima conceives of sentential negativity as a syntactic property of sentences. While the syntactic heterogeneity quickly rules out the possibility of a surface generalization, there is a rich literature on covert syntactic mechanisms which analogize anti-additive neg-words to a syntactic notion of sentential negation in the sense of Zanuttini (1997), i.e., projecting a PolP. An analysis of discourse-negativity in the spirit of Klima would allow all of our licensing expressions to introduce sentential negation, which is characterized syntactically.

The assumption would be that the various expressions introducing negativity express negation, and because of that, the expression projecting a polarity-phase may be non-overt, while a syntactic relationship between the negative exponent and this PolP is assumed. Work in this area has mainly focused on how sentences including anti-additive neg words could be characterized as negative.

The NEG-criterion approach presented by Haegeman and Zanuttini (1991) suggested that neg-words bear negative features and need to (possibly covertly) move to Spec PolP to get these features checked. Ladusaw (1992) suggested that anti-additive quantifiers are themselves non-negative indefinites and trigger the presence of an abstract negative operator, which needs to bind the variables introduced by the neg-words.

Penka (2007, 2010) provides arguments for a decompositional approach under which neg-words have separate indefinite and negative semantically active components by noting the availability of split-scope readings.

(10) After Penka (2007): (3)

Du musst keine Jacke anziehen.
you must no jacket put.on

'You don't have to wear a jacket'

a. It is not obligatory that you wear a jacket $\neg >$ musst $> \exists$

b. There is no jacket that you must wear $\neg > \exists >$ musst

c. It is obligatory that you don't wear a jacket musst $> \neg > \exists$

Penka's arguments give rise to decompositional analyses of split scope readings, which take neg-words to be semantically indefinites and syntactically licensed by an abstract negation operator. Different decompositional analyses make different assumptions about the syntactic status of the abstract negative operator, the morphology of neg-words, and the syntactic dependency between them (see, e.g., Zeijlstra (2011); Iatridou and Sichel (2011), Tubau (2008), or Zeijlstra (2016) for an overview). What is crucial for our purposes is that they syntactically analogize clauses including negative quantifiers to clauses including sentential negation.

This property is what any syntactic characterization of discourse-negativity would rely on: The assumption of a morphosyntactic realization of sentential negation as part of the syntactic representation of the clause.

The previous section noted that certain downward-entailing quantifiers also introduce discourse negativity. De Swart (2000) shows that downward-entailing quantifiers may also exhibit split-scope readings:

(11) De Swart (2000): (13) [Dutch]

Ze hoeven weinig verpleegkundigen te ontslaan.
they need few nurses to fire

a. For a group Y consisting of few nurses y, it is the case that it is necessary for them to fire each individual member of Y (de re)

b. #It is necessary for them to fire few nurses (de dicto)

(Note that this is ruled out because *hoeven* ('need') is an NPI)

c. It is not necessary for them to fire more than a small number of nurses (split)

So long as split readings can be taken as evidence for a decompositional account[6], this suggests that DE-quantifiers like *few* should be analyzed analogously to negative quantifiers and receive a decompositional analysis which involves an upward monotone proportional quantifier which tends towards the upper end of the scale (roughly paraphrasable as *many*), and sentential negation.

In summary, syntactic characterizations of sentential negativity require the presence of a morphosyntactic expression of negation within the clause and a decompositional analysis of anti-additive neg-words, as well as downward-entailing quantifiers.

6.4.2 Characterizations based on the Semantic Representation

We can also conceive of a semantic characterization of discourse-negativity.

Jackendoff: Semantic Sentential Negation

For a semantic characterization of negativity, we look to semantic characterizations of sentential negation.

[6]DeSwart's paper actually presents an argument against decompositional approaches, and in favor of a generalized-quantifier view of negative quantifiers, which will be further discussed below.

Jackendoff (1969) gave a semantic revision of Klima's syntactic characterization of sentential negation. He suggested that a sentence with a surface form [X-neg-Y] includes sentential negation iff there exists a paraphrase *It is not the case that [X-Y]*. This characterization is discussed in Horn (1989), who notes that it is essentially equivalent to assuming that (semantic) sentential negation involves a representation where contradictory negation is the highest scoping operator.

This semantic characterization is hard to distinguish from some of the syntactic accounts since the generalizations for some of the syntactic accounts could also be stated on the level of LF, i.e., a syntactic representation of semantic content. However, the semantic generalization applies to a semantic representation, independently of whether it is assumed to be a part of syntax or not (and therefore applies to LF-based accounts, but also Montegovian indirect interpretation approaches). Horn (1989) (p. 187) notes that the two kinds of characterizations come apart when considering the scope of negation: While Jackendoff's semantic characterization requires negation to take scope over the whole sentence, syntactic characterizations of sentential negation may allow for other scope configurations. This is discussed in more detail in the comparison of the approaches below.

Crucially, the characterization in terms of high-scoping contradictory negation also relies on a decompositional analysis of anti-additive neg-words (the Aristotelian contradictory of *No S is P* is *Some S is P*) and negative downward-entailing quantifiers (*Few S are P* could be paraphrased as *It's not the case that many S are P*).

A Possible Alternative: Negative Quantifiers

A different possible semantic generalization noted by Brasoveanu et al. (2013) might characterize discourse-negative utterance as involving downward-entailing operators. This generalization would be compatible with a non-decompositional analysis of negative quantifiers as generalized quantifiers (e.g. Geurts (1996); De Swart (2000)). Zwarts (1995) characterizes negative markers (like *not*), anti-additive quantifiers (like *never, no*), and downward-entailing quantifiers (like *rarely, few*) all under the umbrella of downward-entailingness, a weaker notion of negativity for operators. No anti-morphic or contradictory negation is required to satisfy this characterization, and a decompositional account would not be necessary in this case. Zwarts suggests that it is this property due to which all of these operators introduce negative contexts and license (weak) NPIs. Although downward-entailingness characterizes a class of expressions licensing weak NPIs, the hypothesis that all downward-entailing expressions license negativity-tags is quickly dismissed when considering counterexamples to this generalization, including *at most n NP*:

(12) Brasoveanu et al. (2013):

a. *At most six volunteers signed up for free housing.*

b. *Yes / # No, at most six of them did.*

Brasoveanu et al. (2013) point out the possibility of another semantic generalization that might not require a decompositional account. At the same time, they suggest that in the absence of such a theory, the fact that anti-additive neg-words pattern with anti-morphic contradictory negative markers in their licensing of negativity-tags should be taken as evidence for a decompositional account of neg-words.

Upshot

A semantic generalization requires the presence of a sufficiently 'negative' operator in the semantic representation. In particular, Jackendoff's (1969) characterization, in combination with a decompositional account of negative quantifiers seems promising. This kind of account differs from the syntactic account because the negative operator, while assumed as part of the semantic representation, does not necessarily have a sentential negation counterpart in the syntax. Further, Jackendoff's (1969) characterization states that negation needs to be the highest-scoping operator in the interpretation of the sentence under consideration, which is not necessarily the case for a syntactic characterization.

6.4.3 Pragmatic / Psychological Characterization

Moxey et al. (2001) point out a connection between sentential negativity and complement anaphora: Certain (but not all) downward-entailing quantifiers license complement anaphora (i.e., anaphora to the set that is the result of subtracting the nuclear scope from the restrictor), like *they* in (13).

(13) Complement anaphora Moxey et al. (2001): (1+2)

a. *Few football fans went to the match.*

b. *They watched it at home on TV instead.*

(they: the football fans who didn't go to the match)

Moxey et al. (2001) point out that those DE-quantifier that license negativity-tags also license complement anaphora. They suggest that certain downward-entailing quantifiers (like *few*) involve denials while others (like *at most three*) involve affirmations and that the licensing of complement anaphora is related to the denial associated with negation. More specifically, the experimental results in Sanford, Dawydiak, and Moxey (2007) show a near-perfect correlation between the proportion of complement anaphora following a given quantifier and the proportion of negativity tags following the same quantifier (measured in terms of a denial index, which is computed based on the proportions of anaphoric neither-tags, polar-question tags, and same-sentence either-tags).

While Moxey et al. (2001) take the Klima tags to be diagnostics of denial by assumption, for us, of course, it is a question of whether the assumption of expectation denial can explain the licensing of negativity-tags.

Proponents of the presupposition denial and inference theory of anaphoric complement set reference (Moxey et al. (2001); Moxey (2006); Sanford et al. (2007)) suggest that the licensing contexts are characterized by the denial of a previous expectation. The term 'presupposition' is used here referring to a previous supposition or expectation, rather than presupposed implications commonly discussed in linguistic pragmatics.

Due to the potential for confusion, I will speak of expectation denial in the following.

Moxey et al. (2001) motivate the theory based on the pragmatic generalization that negative statements (like, e.g., *John didn't go to the cinema.*) are often quite uninformative (in fact, there are many things which John might not have done). However, negated statements become informative if they deny a previous expectation. As a result, negation often expresses the denial of some supposition, assumption, or expectation. Further, there is experimental evidence that contexts in which this is the case facilitate the comprehension of negative sentences (Clark (1976); Horn (1989); Wason (1965))

Moxey et al. (2001) take the denial of an expectation as a pragmatic and psychological effect or function associated with contradictory negation and further suggest that expectation denial is core to the observed effects for CA. This is combined with the inference theory of complement anaphora (Sanford et al. 1996; Paterson, Sanford, Moxey, & Dawydiak 1998), which suggests that when an expectation is invoked and denied, the processor searches for reasons why the expectation was not met. This, in turn, places a focus on the members of the set that witness the expectation denial, i.e., the complement set of the negative DE-quantifier. Moxey and Sanford (1993) present experimental evidence suggesting that the quantifiers licensing CA (not many, few, very few) lead listeners to expect that more might have been the case (as opposed to the positive a few).

While the denial of a supposition or expectation is argued to be functionally associated with contradictory negation, there are some attempts of separating the expectation-denial theory from one that only considers negative quantifiers. Manipulating expectations introduced by the previous discourse, Moxey (2006) provide evidence that whether or not quantified sentences contradict previous expectations influences the proportion of subsequent complement anaphora in production (investigated for sentence completion tasks). A slight increase of the (otherwise negligible) proportion of complement anaphora was even found for 'positive' quantifiers (*a few / a small number*), which generally do not license complement anaphora or negativity-tags.

Moxey, Filik, and Paterson (2009) present similar findings for an online comprehension task using eye-tracking while reading. The argument made in these studies is that negative quantifiers license complement anaphora because they imply an expectation, which is denied by the assertion.

The strong correlation between the proportion of complement anaphora and negativity tags for the various quantifiers tested in Sanford et al. (2007) suggests that quantifiers license complement anaphora to the extent that they license negativity-tags. In light of this, the studies in Moxey (2006); Moxey et al. (2009) give rise to the question of whether we might want to think about discourse negativity in terms of expectation denial as well. We would want to address this question by asking whether expectation-denials would also influence the availability of negativity tags.

The work needed to address this question falls outside of the scope of the present investigation. However, this characterization is compatible with a decompositional analysis of negative quantifiers. Since it is based on the empirical observation that among downward-entailing quantifiers, the same set that licenses complement anaphora also licenses negativity-tags, the account analyzes assertions with negative downward-entailing quantifiers as negations of the expectation of a larger quantity.

6.4.4 Anti-veridical Propositional Drefs

Here, I argue for a characterization on the level of discourse representation that does not rely on the presence of a negation operator in a semantic or syntactic representation of the clause under consideration. Instead, I suggest that sentential negativity could be characterized as involving the assertion of the falsity of a contextually salient proposition, i.e., an anti-veridical propositional dref.

This dynamic semantic characterization of discourse-negativity aligns with the analyses and generalizations from the previous chapters. It is also consistent with the expectation-denial account. It seems like the basic assumptions are not about the operator of contradictory negation itself but rather about asserting the falsity of a proposition.

Importantly, this characterization is consistent with the various previously discussed negative contexts under a decompositional account of negative quantifiers.

That is because following Krifka (2013), we assume that the basic distinction responsible for a contrast between simple positive and negative utterances is related to a difference in their anaphoric potential. Utterances involving negation introduce a propositional dref corresponding to the prejacent of negation. In addition, this work stressed that crucially, this dref refers to an anti-veridical proposition.

Further, we follow Snider (2017), who showed that propositional operators introduce propositional anaphora. The availability of anti-veridical propositional anaphora following utterances with negative quantifiers, illustrated in (14), therefore provides an additional argument for a decompositional analysis under which negative quantifiers involve propositional negation.

(14) a. *No students handed in the assignment early, although I would have been happy about that.*

(I would have been pleased if some students had handed in the assignment early)

b. *The dossier presented no new information of significance. That was just a rumor.*

(It was a rumor that the dossier presented some new information of significance)

c. *Nobody fantasizes about a corporate desk job compiling spreadsheets. but Mary seems to believe that.*

(Mary seems to believe that some people fantasize about a corporate desk job compiling spreadsheets)

d. *Signing online petitions achieves nothing. That's a pipe dream.*

(That signing online petitions achieves some things is a pipe dream.)

6.5 Comparing the Approaches

This section presents arguments for a dynamic semantic account of discourse-negativity based on some data for which various approaches would make different predictions. First, we will consider contexts that license discourse-negativity implicitly, without an overt expression of negation in the clause that introduces the antecedent. This kind of data requires an account that does not tie negativity to the linguistic representation of a clause itself, such as our dynamic semantic analysis under which a discourse-negative context is one in which an anti-veridical propositional dref is made available. Then, we consider some licensing contexts for negativity tags, which may be regarded as arguments for an account of discourse-negativity, which does make reference to syntactic structure. We will discuss why this data is challenging for a semantic (or non-syntactic) account and why our dynamic semantic account can nonetheless account for these data points. The last part of this section considers a few additional contexts illustrating constraints on the implicit introduction of discourse-negativity.

6.5.1 Pragmatic Variation

The heterogeneous landscape of discourse-negative contexts discussed in **Section 6.3** cannot distinguish between a syntactic account, a semantic account that relies on the presence of negative operators in the representation of a clause, an account in terms of expectation-denial, or the dynamic account proposed here, in terms of anti-veridical propositional drefs. That is because, under a decompositional analysis of negative quantifiers, any of these contexts can be analogized to a sentence that includes a negative marker. A sentence with an instance of a (both semantically and syntactically) negative operator could satisfy any of these characterizations, as the discussion in the previous section shows.

However, cases in which negativity is not overtly realized as part of the sentence introducing a proposition but instead contextually supplied provide evidence against a view under which negativity is characterized on the level of a clausal representation (be it a syntactic or semantic representation). These are contexts which Kroll (2019) identifies as contexts licensing polarity-reversals in sluicing, i.e., contexts involving neg-raising or disjunction.

Kroll (2019) shows that neg-raising contexts and disjunctive contexts allow for polarity-reversing sluices.

(15) Polarity reversal contexts [Kroll (2019)]

a. Neg-raising contexts:

I don't think that California will comply but I don't know why ~~California wont comply~~

b. Pos-neg Reversal under Disjunction:

Either the Board grants the license by December 15 or it explains why (~~the Board did not grant the license by December 15~~).

c. Neg-pos Reversal under Disjunction:

Either John didn't do an extra credit problem, or he didn't mark which one

(~~he did which one~~).

In (15-a), the clause including the ellipsis can be paraphrased as *'I don't know why California won't comply'*, while the antecedent is the embedded clause in the first sentence *'that California will comply'*. Although the antecedent clause does not include negation, the ellipsis is interpreted as if it did.

A similar case is shown in (15-b). Here, the clause including the ellipsis can be paraphrased as *'why the Board did not grant the license by December 15'*, where the antecedent is the clause *'Either the Board grants the license by December 15'*, which does not include negation. Again, the ellipsis is interpreted as if the antecedent included negation.

In (15-c), we see a polarity-reversal in the opposite direction: While the antecedent is expressed in the clause *'Either John didn't do an extra credit problem'* which includes negation, the clause including the ellipsis can be paraphrased as *'he didn't mark which extra credit problem he did'*., i.e., in a way that is compatible with a positive antecedent.

Kroll uses the examples in (15) to show that the pragmatic licensing requirement on clausal ellipsis is sensitive to the local context of the anaphoric expression. The polarity-mismatches between the ellipses in (15) and their antecedents can be explained when we take into account the effect of neg-raising and disjunction on the local context of the ellipsis site and the antecedent.

Kroll suggests that the proposition that provides the 'negative' interpretation for the sluice in the above cases is given in the local context because the context supports an interpretation s.t. that the proposition expressed by the antecedent clause is false. Specifically, Kroll argues that sluicing is licensed only if the proposition expressed by the elided constituent is locally given. Kroll's requirement of local givenness requires that the proposition expressed by the ellipsis (i) is expressed by an antecedent, and (ii) is entailed by the local context (modulo F-closure). This analyzes clausal ellipsis as anaphoric to a proposition and is similar to what I suggested for pronominal anaphora in **Chapter 5**. For pronominal propositional anaphora, we found that they need to be compatible with the local contexts, resulting in a slightly weaker requirement.

Under this view, the sluices in (15-a) and (15-b) are interpreted in this way because the embedded 'positive' clause expresses a proposition, and the 'negative' interpretation, i.e., its falsity is supported in the local context.

Accordingly, we are dealing with local contexts in which a previously introduced propositional dref is locally anti-veridical, i.e., false in the local context. Importantly for our purposes, the same contexts allow for negativity-tags. Here, though the antecedent is introduced by a positive clause, and the negativity is introduced dynamically in the local context.

Neg-Raising

The embedded clause in an utterance with first-person neg-raising is interpreted as non-veridical in relation to the speaker. Based on the generalizations from part I, we expect it to be an antecedent for an anti-veridical propositional anaphor.

(16) *I don't think that Trump will comply with the debate requirements, even though he claimed that beforehand.*

In addition, neg-raising contexts license *neither*-tags and canonical *why not*.

(17) *I don't think that Trump will comply with the debate requirements,*

  a. *but neither will Biden.*

  b. *but I don't know why not.*

  (✓I don't know why Trump won't comply with the debate requirements)

Agreeing uses of *no* are not a good test for contextually introduced negativity, since these might just as well be analyzed as contradicting / reversing uses of *no*, which are compatible with non-negative antecedents. Instead, we should look at uses of *doch*, which does require a negative antecedent, and seems quite natural with a neg-raising antecedent.

(18) A: *Ich glaube nicht, dass die Baustelle vor nächstem Jahr fertig wird.*
I believe not that the construction site before next year finished gets
'I don't believe that the construction will be finished before next year.'

B: *Doch, das wird sie bestimmt.*
DOCH that will it certainly
'Yes, it certainly will.'

Of course, this argument for contextual implicit introduction of negativity could be undermined if the neg-raising contexts that license negativity-tags are derived by syntactic mechanisms as argued for some cases of neg-raising in Collins and Postal (2018). Under these kinds of approaches, the embedded clause in a neg-raising antecedent utterance could be characterized negative on a syntactic level.

However, we can look at neg-raising with non-canonical neg-raising predicates (Kroll uses *remember* to make this point for polarity-reversing sluices), or negativity-tags with neg-raising antecedents where the embedded clause is in an island, i.e. cases which even Collins and Postal (2018) would characterize as pragmatically derived.

(19) Embeddings under *remember*:

I don't remember being scared, …

a. …but I don't know why not. (why I wasn't scared)

b. …and neither was my dog. (my dog also wasn't scared)

make sure with nouns; NP islands: neg-raising with nouns taking nominal complement clauses

(20) Collins and Postal (2018) (18, 19)

a. It is not my opinion that Mars can be colonized.

b. I am not of the opinion that Mars can be colonized.

(21) *I don't get the impression that Mars can be colonized*

a. *and scientists can explain why.*

(✓why Mars cannot be colonized)

b. *and scientists can explain why not.*

(✓why Mars cannot be colonized)

c. *and neither can the moon.*

d. *That is a misconception.*

(that Mary can be colonized)

wh-islands:

(22) Context: I talked to everyone in the department trying to settle the question whether donuts should be considered breakfast food.

I haven't found anyone who thinks that donuts are breakfast food

a. *and according to most, neither are crullers.*

b. *but there are different options about why (not).*

c. *that just doesn't seem to be what most people believe.*

(that ≈ that donuts are breakfast food)

While many speakers support the judgments reported here, and rate the semantic neg-raising cases highly, there are other speakers who did not rate some of the semantic neg-raising very highly. (For example (21-b) or (19-b) received a couple of low ratings). I interpret the relative availability of negativity-tags with uncontroversially pragmatic neg-raising as an indication that these contexts license negativity-tags, and introduce discourse-negativity. This generalization is also tested experimentally in **Section 6.6**.

In case of neg-raising, the embedded clause does not include overt negation but nonetheless expresses an anti-veridical proposition. In particular in cases where we can be sure that this interpretation is derived pragmatically, based on pragmatic assumptions about speaker knowledge, an analysis of discourse-negativity based on the epistemic status of propositions in discourse is the only one which can capture this.

For syntactic accounts, this data is problematic: If discourse-negativity were a property of clauses, tied to an instance of morphosyntactic negation, neg-raising contexts would not be expected to pattern like a negative context. For semantic accounts which require a negative operator as part of the semantic representation of a proposition, these data are problematic for similar reasons.

Disjunction

Kroll shows that polarity-reversing sluices are licensed under disjunction for positive-negative reversals (23-a), and for negative-positive reversals (23-b).

(23) a. After Kroll (2019): (28)

Either John didn't do an extra credit problem, or he didn't mark which one.

(or he didn't mark which one he did)

b. After Kroll (2019): (30)

Either the Board grants the license by December 15th, or it explains why.

(why the Board did not grant the license by December 15th)

As discussed in **Chapter 4**, disjunction is a non-veridical discourse relation, which doesn't require that the propositions expressed by the disjuncts are compatible with each other. As a result, pronominal propositional anaphora under anti-veridical embeddings in the second disjunct may have antecedents that are veridical in the context of the first disjunct (24-a), or the other way round (24-b).

(24) a. *Either the instructor granted an extension to the deadline, or that is just a rumor.*
(that $\approx$ that they granted an extension to the deadline)

b. *Either Mary didn't get infected, or she is aware of it already.*
(it $\approx$ that Mary is sick)

Both *why not*, and *neither* seem less than ideal in these contexts:

(25) a. Context: Students of a class are discussing a recent class announcement. knowing that their instructor tries to be maximally transparent and beneficient to the students.
? *Either the instructor granted an extension to the deadline, or they will explain why not.*
(why they did not grant an extension to the deadline)

b. Context: It is known that Mary and Sue will always come to parties together. The hosts are discussing who RSVP'd.
?? *Either Mary is coming, or neither is Sue.*

However, the simple *why*-sluice also doesn't seem great in this kind of context:

(26) ? *Either the instructor granted an extension to the deadline, or they will explain why.*

(why they did not grant an extension to the deadline)

But Kroll showed that, at least, polarity-reversing sluices under disjunction are attested. However, it seems that the reported cases involve future-oriented modals or imperatives.

It seems like negativity-tags are supported by contexts that support polar alternatives, such as polar questions or future-oriented modals and imperatives, which introduce the possible options of doing or not doing an action.

(27) (COCA Davies (2008)):

Do you think this was a good idea? Why or why not?

(28) a. *Do it by Monday or explain why not.*

b. *Love yourself, or neither will anybody else.*

The fact that this kind of polar QUD supports this kind of situation raises questions about the impact of the QUD on ellipsis licensing, but the disjunction cases provide additional support for the idea that discourse-negativity can be introduced contextually.

Discussion

Contexts involving neg-raising or disjunctionare contexts in which discourse-negativity is introduced implicitly, either by means of pragmatic reasoning or due to the dynamic effects of disjunction on the local contexts of its disjuncts. What these cases have in common is that negativity is not realized in the linguistic representation of the clause that provides the antecedent, neither on a syntactic level, nor a semantic one.

This provides evidence for a dynamic semantic view, which allows us to capture the relevant generalizations without making reference to the syntactic or semantic representation of a clause. The key representational component determining whether a negativity-tag is licensed has to be on the level of discourse rather than a linguistic representation that is syntactic or limited to the asserted content.

The particular dynamic semantic analysis of negative contexts proposed here can account for these cases, as propositional drefs can be interpreted as anti-veridical by virtue of the context, or based on pragmatic reasoning that goes beyond what is asserted.

However, there are certain challenges of this kind of explanation that we haven't discussed yet. Importantly, this account predicts that any anti-veridical embedding, not just ones introduced by negation, should license negativity tags. If this is the case, it seems like it is certainly less strongly so than in utterances with negation.

(29) a. *It's a misconception that the published communications discussed the pressing problems this week, but I can't imagine why not.*

b. *It's a fallacy that the monks said their prayers at noon today, and neither did the apprentices.*

c. *It's a lie that the children in my Arabic class studied the difficult vocabulary items this week, and I'm not sure why not.*

We know that negativity is gradient and influenced by various factors. We might ask the question, if negativity-tags are in principle possible in ((29)), but there is a further constraint decreasing the acceptability here. For the above examples of disjunction, sarcasm, and accommodation, we saw that the licensing of negativity-tags in the implicit cases seems to be aided in case the content of the anaphor addresses a polar QUD. It is conceivable that addressing a polar QUD could be another pragmatic factor in determining discourse-negativity. If there is a polar QUD, in case of simple negation, the embedded proposition automatically addresses the same QUD as the full assertion. In case of anti-veridical propositional attitudes, this is not necessarily the case.

Further research exploring the predictions of the propositional anaphor account of discourse-negativity should ask two kinds of questions: **(i)** Do anti-veridical propositional attitudes license negativity-tags, potentially to a smaller extent than negation? **(ii)** If they do introduce negativity, but less robustly compared to negation, is there an additional constraint explaining the difference? If we could find evidence in favor of affirming both of these questions, this would point towards anti-veridical propositional drefs as the key representational component in determining whether a negativity-tag is acceptable.

Question **(i)**, is addressed by the experiment presented in **Section 6.6**. The work needed to address question **(ii)** and test the polar QUD hypothesis will be left for future research.

Before reporting an experiment testing these predictions, the following subsection will address some data which might be interpreted as a challenge to any non-syntactic account of negativity. This involves some generalizations about the scope of negation in a discourse-negative utterance, and also the fact that discourse-negativity can only be introduced by contextually supported pragmatic inferences in the kinds of contexts discussed in this subsection. It cannot be introduced by entailment. While these kinds of contexts may present challenges for a non-syntactic account, I also show how a propositional anaphor account can capture the same generalizations.

6.5.2 Why a Syntactic Account Seems Good, Actually

A few pieces of evidence and generalizations seem like they might provide arguments for characterizing discourse-negativity on a syntactic level of representation. This subsection addresses some arguments of this kind and suggests reasons why the generalizations and facts considered here may still be compatible with characterizing discourse-negativity in terms of the introduction of an anti-veridical dref.

Sentential vs. Constituent Negation

In cases where negation is realized overtly in the antecedent, negativity-tags have been pretty reliably used as a diagnostic between constituent and sentential negation. This could be an indication that they can make reference to a sentential negation on a syntactic level. However, we will see that the same generalization can be explained in terms of the introduction of an anti-veridical dref.

The cases in which discourse-negativity is introduced by the negative marker *not* in the inflectional domain, contrast with constituent negation (30), which doesn't create a context for *'Why not'* or *neither*-tags.

(30) a. VP constituent negation:

A: *Jack insisted on [not going to school today]*$_{\mathrm{VP}}$

B: *#Neither did I.*

#Why not?

b. AdvP constituent negation: [Klima (1964)]

A: *He found something interesting there not long ago.*

B: *#Neither did I.*

#Why not?

The contrast could be described in syntactic terms: it seems that the English negative marker *not* introduces sentential negativity (in simplex sentences) if it is introduced in the inflectional domain (using the terminology from Grimshaw (2000)), but not if it is introduced in the lexical domain (30-a) or in sub-sentential adjuncts (30-b).

Klima notes that sentences involving constituent negation (30) can be characterized as involving truth-conditional negation in their semantic representation: (30-a) entails that Jack didn't go to school today, (30-b) entails that he didn't find the interesting thing there long ago. In spite of this negative semantics, *neither*-tags aren't licensed in these contexts. Klima takes this to be an argument against an account of sentential negativity in meaning-related terms and in favor of a purely formal account.

Abandoning a purely syntactic view of sentential negativity, I recast this generalization in semantic terms: If negation is overt, it needs to scope over the sentence radical to introduce a propositional dref.

This generalization is supported by the fact that the only expressions in the lexical domain which can introduce negativity are negative quantifiers. Plausibly, object quantifiers, unlike other phrases with constituent negation, can scope higher than their surface position and take scope over the proposition expressed by the sentence radical. We know that this can be the case for objects, but negative quantifiers in adjuncts also may take high scope in a similar way. (31) in particular shows that this is optional.

(31) *John would be happy with no job.*

  a. 'John would be happy with not having any job.'

  b. 'It's not the case that there exists any job that John would be happy with.'

The experimental results from Brasoveanu et al. (2014) show that negative quantifiers in object position introduce negativity less robustly than subjects and sentence adverbs also support this generalization. We might hypothesize that the lower proportion of negativity-tag follow-ups is because the representation that licenses them is derived by an additional derivational step of quantifier scope-taking.

The acceptability contrast of negativity-tags in the context of sentential vs. constituent negation is compatible with an analysis that the relevant restriction is semantic and based on scope, although it clearly has syntactic consequences.

Further, the contrast falls out from the assumptions made here, particularly from Snider's (2017) generalization that propositional drefs are introduced by propositional operators for their prejacent.

Since propositional operators may combine with sentence radicals, a sentence radical in the scope of negation will introduce a counterfactual propositional dref. Constituent negation, on the other hand, is not a propositional operator and could not introduce a propositional dref. For Snider, this fact, illustrated in (32) is an argument in favor of the generalization that propositional drefs are introduced precisely in the cases where there is a propositional operator.

(32) Snider (2017): 173

a. *It's not the case that Nancy has been able to go skiing, though people believed that.*

b. *Nancy has not been able to go skiing, though people believed that.*

c. #*Nancy has been not able to go skiing, though people believed that.*

d. #*Nancy has been able not to go skiing, though people believed that.*

e. #*Nancy has been able to not go skiing, though people believed that.*

f. #*Nancy has been able to go not skiing, though people believed that.*

Snider's generalization that only utterances of sentences with sentential negation introduce a propositional drefs for the prejacent of negation allows us to make sense of the sentential/constituent negation contrast within a dynamic semantic account.

Negation Outscoped

Jackendoff's (1969) suggests that a sentence is semantically negative just in case contradictory negation is the highest-scoping operator in the clause. In their analysis of PolPs, Roelofsen and Farkas (2015) assume that this characterization of sentential negation determines whether a proposition is negative for discourse purposes.[7]

This generalization could include all the above cases where sentential negation was introduced by negative markers, adverbials, and negative quantifiers and scopes over the entire proposition. It also includes the cases where negation is contextually introduced because the negative component also scopes over the proposition in question there. There are counterexamples to this generalization in which sentential negation is outscoped while still introducing negative contexts that license a negativity-tag (33):

(33) a. High-scoping modals:

A: *You shouldn't go alone.* $(\Box > \neg), (\neg \not> \Box)$

B: ✓*Why not?* (Why should you not go alone?)

✓*Neither should you.* (you shouldn't go alone either $(\Box > \neg)$)

✓*No, you really shouldn't.*

✓*Even though you may believe that.* (that $\approx$ that you should alone?)

b. Quantified pronouns:

A: *We all didn't sleep.* $(\forall > \neg), (\neg \not> \forall)$

B: ✓*Why not?* (Why did you all not sleep?)

[7] Many insights of their account do not necessarily hinge on this particular choice of how discourse-negativity is determined.

✓*Neither did I.*

✓*No, you truly didn't*

✓*Even though you may believe that.* (that ≈ that we all slept)

While we established that in overt negation cases, negation needs to scope over the sentence radical at least, it doesn't seem like it needs to scope over the full sentence.

Horn (1989) notes that cases where another scope-taking element takes scope over negation are a challenge for purely semantic accounts of sentential negation that is meant to have syntactic consequences. For our anaphoric cases, in particular, these cases are challenging.

We characterize a discourse-negative utterance as one introducing a propositional dref corresponding to the content in the scope of negation. Consequently, the material that is picked up anaphorically should amount to that negated content. However, the interpretation of the negativity-tags in (33) involves the same scope relationships as the antecedent. If the modal *should* or the universal quantification in (33-b) are not part of the negated content in the antecedent sentence, we might not expect their meaning contribution to be part of the dref that is being picked up anaphorically either. Accordingly, the cases where negation is outscoped by another operator are a challenge for the account presented here. However, there may be room for a semantic explanation that is compatible with it.

Something to note is that different quantifiers seem to vary with respect to how negative a context still is when they scope over negation. Contexts where negation scopes below *should*, like (33-a), seem to give rise to sentential negativity pretty robustly. On the other hand, there seems to be more inter-speaker variation with regards to the acceptability of the follow-ups in (33-a). Brasoveanu et al. (2013) show quantificational subjects which command negation on the surface provide strong statistically significant preferences against a negativity-tag, in contrast to referential subjects. This effect is even larger for *at most* n and *exactly* n, compared to the indefinite *some*.

We see variation between different quantifiers between the impressionistic individual judgments collected for *should* and *all*, and in the experimental results from Brasoveanu et al. (2013). The question of why some quantifiers prevent the material in their scope from introducing discourse-negativity more than others should be addressed on a semantic rather than syntactic dimension.

A possible way of addressing the data in (33) from a semantic perspective might be an explanation stating that in the interpretation of the anaphoric expressions *should* and *all* do not take high scope by usual scope-taking mechanisms, but that the interpretation represents a pragmatically derived pseudoscope.

One possibility of a semantic generalization over quantifiers might have to do with what Löbner (2000) calls a presupposition of indivisibility for distributive definite plurals. Löbner argues that distributive definites are often interpreted as universals but that a presupposition of indivisibility gives rise to all-or-nothing effects for distributive plurals:

(34) Löbner (2000):

a. *The books are written in Dutch.*

(all the books are written in Dutch)

b. *The books are not written in Dutch.*

(all of the books are s.t. they are not written in Dutch)

The cases in (34) pattern as if there was a universal operator taking scope over negation. However, Löbner argues that this effect can be derived from a presupposition that the restrictor set of the quantifier is indivisible wrt the nuclear scope: either the restrictor set is completely contained within the nuclear scope set, or it is completely outside of it. If we can find that the puzzling interpretation in these cases is derived by presupposition and not by the quantifier taking scope over negation, that might alleviate Horn's (1989) worry that they might fundamentally challenge purely semantic accounts of sentential negation.

For a related problem, Zaitsu (to appear) appeals to Homer's (2015) analysis of the modal *should* and its interaction with negation. Homer suggests that *should* is both a PPI and a neg-raiser, which means that it can achieve wide scope over negation by syntactic movement or through a Gajewski (2007)-style homogeneity inference.

Since *should*, like *must*, is a PPI, neither of the two can be interpreted in the scope of negation. Both need to be interpreted as scoping over negation. According to Homer, as mobile PPIs, they can be moved to a position above negation and take scope from there.

(35) Homer (2015):

a. *John mustn't jog.* (MUST $>\neg$, $\neg \not> $ MUST)

b. *John shouldn't jog.* (SHOULD $>\neg$, $\neg \not> $ SHOULD)

In a sentence where *should* appears with clause-mate negation (35), it takes high scope via canonical scope-taking mechanisms (in this case movement), due to its PPI nature, a property that it shares with *must.*

Things are different when *should* is part of a clause that expresses an anti-veridical proposition, neg-raising under *think* in (36). In this kind of context *must* does not take scope over negation, but *should* does.

(36) Homer (2015):

a. *I don't think that John should marry Susan.* (THINK $>$ SHOULD $>\neg$)

Paraphrasable as: I think that John shouldn't marry Susan.

b. *I don't think that John must marry Susan.* (✗THINK $>$ MUST $>\neg$)

Not paraphrasable as: I think that John mustdeonn't marry Susan.

Since *should* and *must* behave differently here, the scope-taking behavior of *should* could not be due to their shared PPI nature. Homer presents arguments that this is due to a homogeneity-inference associated with *should*, which leads to a neg-raising interpretation.

Zaitsu points out a similar contrast between *should* and *must* in a case where the proposition expressed by the sentence containing the modal is negated on a contextual level, i.e., with disagreeing uses of the negative PolP *no*.

(37) a. *Should Amy leave?*

b. *No.* (Amy shouldn't leave) (should > not)

(38) a. *Must Amy leave?*

b. *No.* (Amy doesn't have to leave) (not > must)

This contrast, as well as the interpretation of negativity-tags following sentences including *should* can be understood semantically, based on Homer's characterization of *should* as a neg-raiser.

The cases of negativity-tags being licensed following an utterance of a sentence where negation is not the highest scoping operator (or does not seem to be) present a challenge. However, this section presented some preliminary arguments suggesting that there is variation between different scope-taking operators wrt how strongly they allow for a negative operator in their scope to introduce negativity and that this variation should be thought of on a semantic dimension. The subsection presented some independently motivated sources of pragmatic inference that could give rise to the reported interpretation pragmatically.

6.5.3 Interim Summary

We discussed what kind of properties a characterization of discourse-negativity would need to have, if it were to make reference to a syntactic or semantic level of the linguistic representation of a clause vs. to semantic properties of a representation of the discourse, which integrates the information from the linguistic content with information from the discourse context and the epistemic commitments of the speaker.

I suggest that the arguments presented here point in the direction of the latter kind of characterization: What has been referred to as 'sentential negativity' since Klima's introduction of the topic is a property of the discourse-effect of an utterance, rather than its representational features.

This is suggested by the fact that negativity-tags can be licensed via contextual introduction of negativity. Further, I show that the account proposed here, which assumes that the relevant discourse effect associated with negativity amounts to providing an anti-veridical propositional dref can address some data which have previously been interpreted as arguments for a syntactic account.

6.6 Experiment: Anti-Veridical Contexts

This section reports a third forced-choice study assessing how robustly different contexts license negativity tags with the goal of experimentally testing the predictions of different kinds of approaches to discourse-negativity. It also outlines some possible follow-ups addressing the same question in other ways.

Based on the above discussion of possible characterizations of negativity on various levels of representation, here is an overview of the central properties for the different kinds of accounts:

1. **Morphosyntactic licensing:**

   Discourse negativity is characterized by *a morphosyntactic realization* of negation in the sentence introducing the antecedent.

2. **Semantic Polarity:**

   Discourse negativity is characterized by contradictory negation as the highest-scoping operator in the semantic representation of the sentence introducing the antecedent.

3. **Anti-veridicality:**

   Discourse negativity is characterized by the introduction of an anti-veridical propositional dref.

The first two kinds of analyses of discourse-negativity would need to make referent to the representation of the clause includes negation. The dynamic semantic account based on anti-veridicality also allows for negativity to be introduced in contexts where negativity is introduced implicitly. In particular, these are contexts where a non-negative embedded clause is interpreted as anti-veridical due to the semantically or pragmatically contributed inferences associated with its embedding context, like anti-veridical attitudes or cases of neg-raising that are uncontroversially considered to be semantically/pragmatically derived, even by theorists who suggest that canonical examples of neg-raising (under *not think/not believe*) are derived syntactically.

If one of the first two accounts is the more suitable characterization of negativity, we expect contexts with semantic neg-raising and anti-veridical propositional attitudes to pattern more like positive contexts wrt the licensing of negativity-tags. If the third is more accurate, we expect contexts with semantic neg-raising and anti-veridical propositional attitudes to pattern more like negative contexts.

This section reports a forced-choice study addressing this question. It compares the availability of negativity-tags in different kinds of contexts. The antecedents are all bi-clausal structures where the negativity tags will target the embedded clause. The study manipulates how the veridicality of the proposition expressed by the embedded clause is determined. This involves five conditions: In the POS-condition, the embedded clause is affirmative and expresses a veridical proposition. For NEG, there is overt negation expressed by *not*. The NR- and PNR-conditions both involve neg-raising, where the second involves noun complements to ensure a pragmatic derivation of the neg-raising interpretation. For AV, the embedded clause is embedded under an anti-veridical propositional attitude. The manipulation is illustrated in (39).

(39) a. POS:

I think that the party planning committee really made an effort this year

b. NEG:

I think that the party planning committee didn't make an effort this year

c. NR:

I don't think that the party planning committee made an effort this year

d. SNR:

I don't get the impression that the party planning committee made an effort *this year*

e. AV:

I heavily doubt that the party planning committee made an effort this year

It is methodologically modeled after Experiments 1 and 2 from Brasoveanu et al. (2013), which take advantage of the polarity-sensitivity of polarity particles and the generalization that agreeing responses with *no* are only possible with negative antecedents. As a result, agreeing uses of *no* are used as a diagnostic of sentential negativity in their, Experiment 2, which provides evidence that negative quantifiers (*never, nobody*) are negative in the relevant sense.

However, agreeing uses of *no* don't quite work as a diagnostic here, and as a result, we need to adjust the methodology and diagnostics slightly. The experiment reported here assesses the negativity of various anti-veridical embeddings in English using *why not* questions as a diagnostic of negativity instead. Before reporting the design and results of this experiment, the following explains the methodological considerations of choosing this diagnostic over agreeing *no*-uses as a minimal necessary change to the methodology from Brasoveanu et al. (2013).

Methodological Considerations

The experiments reported in Brasoveanu et al. (2013) used a forced-choice task asking participants to choose between *yes* and *no* in order to agree with the previous utterance. This is a suitable diagnostic for monoclausal antecedents. However, in (39), some of the embedded clauses are morphologically positive.

Consider (39-e) and suppose a participant were given a choice between following up with *Yes / no, they didn't.* If a speaker were to choose a *no*-response, different theories would lead to different interpretations. A theory based on clausal negation might characterize the antecedent as positive, and the follow-up as rejecting a positive. A theory based on veridicality would characterize the antecedent as anti-veridical, and the response as an agreeing response. Therefore, English *no*-responses are not a useful diagnostic here.

Therefore, the planned English studies use *why not* questions as diagnostics of negativity. Canonical question uses of *why not* can be used as diagnostic of discourse-negativity. Canonical question uses of *why not* are matrix uses as information-seeking question, as well as factive embedded uses. The non-negative counterpart of reduced *why not* clauses are positive *why*-sluices. For our case with anti-veridical embeddings, the choice would look as shown in (40).

(40) a. *I heavily doubt that the party planning committee made an effort this year*

b. *but they didn't explain why / why not.*

While there is a possibility that the positive *why*-sluice could receive an interpretation that under which the full antecedent sentence is part of the ellipsis ('they didn't explain why I heavily doubt that the party planning committee made an effort'), the other interpretation is much more plausible, under which the ellipsis antecedent is the embedded clause and 'they' (i.e. the party-planning-committee) might explain their own actions rather than the speaker's thoughts.

As a result, we are really testing, for the embedded clause, will speakers choose a negativity-tag to refer back to it, or not.

6.6.1 Design

The experimental task presents the participants with an antecedent sentence, followed by an incomplete sentence, and a choice between two possible completions. The binary forced-choice task is followed by a five-point confidence rating.

The experimental manipulation involves the five conditions described above ({NEG(ative), N(eg) R(aising), S(emantic) N(eg) R(aising), A(nti-)V(eridical), POS(itive)}) for different kinds of anti-veridical attitude embeddings in the antecedent sentences.

The bi-clausal antecedent was followed by the beginning of another bi-clausal sentence, specifically a context which may embed a *why*-question. The experiment tested the participants' preference for completing the sentence with *why* or *why not* in a binary forced-choice task, followed by a five-point scale confidence rating.

We assess the proportion of *why not*-choices in relation to *why*-choices, based on the kind of anti-veridical propositional embedding used in the antecedent sentence. Since the use of *why not* is a diagnostic of negativity, we expect the proportion of *why not*-uses we see with a given embedding context to be an indicator of how robustly the embedding context introduces negativity.

6.6.2 Predictions

Because polarity-ellipsis, like English PolPs, exhibits a pattern of negative neutralization, this study is very similar to the PolP studies in Brasoveanu et al. (2013): While a positive antecedent only allows for *why*, negative antecedents allow for either *why* or *why not*. As a result, we expect both options to be chosen to some extent if the antecedent is interpreted as negative.

Based on the generalizations from the literature, characterizing sentences with *not* as strongly discourse-negative, and based on the results from experiments 1 and 2 in Brasoveanu et al. (2013), we expect that the NEG-condition robustly introduces negativity. Brasoveanu et al. (2013) found a preference for *no*-responses in the corresponding condition, and a clear difference to the positive baseline.

While this preference for agreeing with *no* for negative antecedents is found for English, Meijer, Claus, Repp, and Krifka (2015); Claus, Meijer, Repp, and Krifka (2017) suggest that the preference patterns for agreeing PolPs with negative antecedents are different in German. While they report a higher overall ratings for *ja* ('no') over *nein* in agreeing responses to negative antecedents, they also report inter-speaker variation with respect to the relative ratings (which they interpret as preferences) between *ja/nein* in this context. For English *why not*, it is therefore unclear what kind of proportion we should expect. What we do expect is that the NEG-condition marks the higher end in a scale of proportion of *why not* completions. Our positive baseline POS on the other hand is expected to have a negibly low proportion of *why not* choices.

If the anti-veridical propsitional anaphor approach is right, we expect the conditions in which negativity is introduced contextually to pattern like negatives. This means that we would expect the conditions NR, SNR, and AV to show a significant portion of *why not* choices.

If our initial consideration is on the right track, that the introduction of an anti-veridical propositional anaphor is the crucial kind of representation, and in addition some notion of QUD-givenness, may additionally constrain the availability of negativity-tags, we expect a cline between NEG > NR > SNR > AV wrt the proportion of negativity-tag responses. This is because the anti-veridical proposition in a sentence with negation is automatically relevant to the same polar QUD as the matrix proposition, i.e. its opposite. Cases of canonical neg-raising are often discourse-transparent in the sense that they are predicates that a speaker may use to self-ascribe commitment (i.e. the same class of negated propositional attitudes that also allow for negative slifting in English). As such, they can also often be interpreted as parentheticals, with the embedded content addressing the QUD.

On the other hand, if discourse negativity is associated with some representation of the main asserted content of a clause, we expect the conditions SNR and AV to pattern like positives, and might go either way on NR. The latter would depend on whether the mechanism which derives the neg-raising interpretation renders the embedded clause negative in the relevant sense.

6.6.3 Method

Participants

The experiment was run as an online-questionnaire, in which 49 self-identified native English speakers participated. The participants were recruited from Prolific Academic and compensated at a rate of 6.20 GBP ($\approx$ 8.30 USD at the point of writing) per hour. They took about 22 minutes on average to complete the questionnaire.

Materials

The experiment included 45 items and 50 fillers. Each item was passed through the five conditions and the participants were rotated through 5 lists generated following a latin square design.

The fillers included different kinds of bi-clausal antecedent sentences (10 with attitude embeddings and 40 with speech report embeddings; 10 with a first-person matrix subject and 40 with third-person definite matrix subjects; 10 with negation in the matrix clause, 10 with negation in the embedded clause, 30 without negation), and participants were asked to make a choice between *why* and *why not*, between *why* and *why + OBJ* stripping, *why* and *why + SUBJ* stripping, *why not* and *why not + OBJ* stripping, as well as *why not* and *why not + SUBJ* stripping for 10 fillers each.

Procedure

Each participant responded to the 110 stimuli in randomized order. For each stimulus, a participant read a target and choose one of two continuation. On the next screen, they were asked to rate their confidence in the preceding binary choice on a scale between 1 (‘not at all confident’) and 5 (‘very confident’). Before starting with the trial period containg the 110 experimental stimuli, all participant was introduced to the task by step-by-step instructions and a 5-stimulus practice period. A trial was concluded by a debriefing survey.

Three participants were excluded, because they satisfied both of the following criteria **(i)** they have two or more reaction times for the forced-choice task below $1000\mathrm{ms}$; **(ii)** they have three or more *why not*-responses out of ten observations in the POS baseline control condition, i.e. the one where we only expect *why*-answers to be possible. Based on unusually short reaction times and low accuracy in the control condition, we can safely assume that these three participants completed the survey without paying attention to the experiment, in order to receive the associated payment. With 47 participants remaining, the final number of observations is $47 \times 45 = 2115$, i.e. 423 observations per condition.

6.6.4 Forced-Choice Data

Data Summary

The counts of *why*/*why not* responses by condition are given in **Table 6.1**, and visualized in **Figure 6.1**.

	why	why not
pos	406	17
av	315	108
snr	250	173
nr	207	216
neg	206	217

Table 6.1: Cross-tabulation of responses and conditions

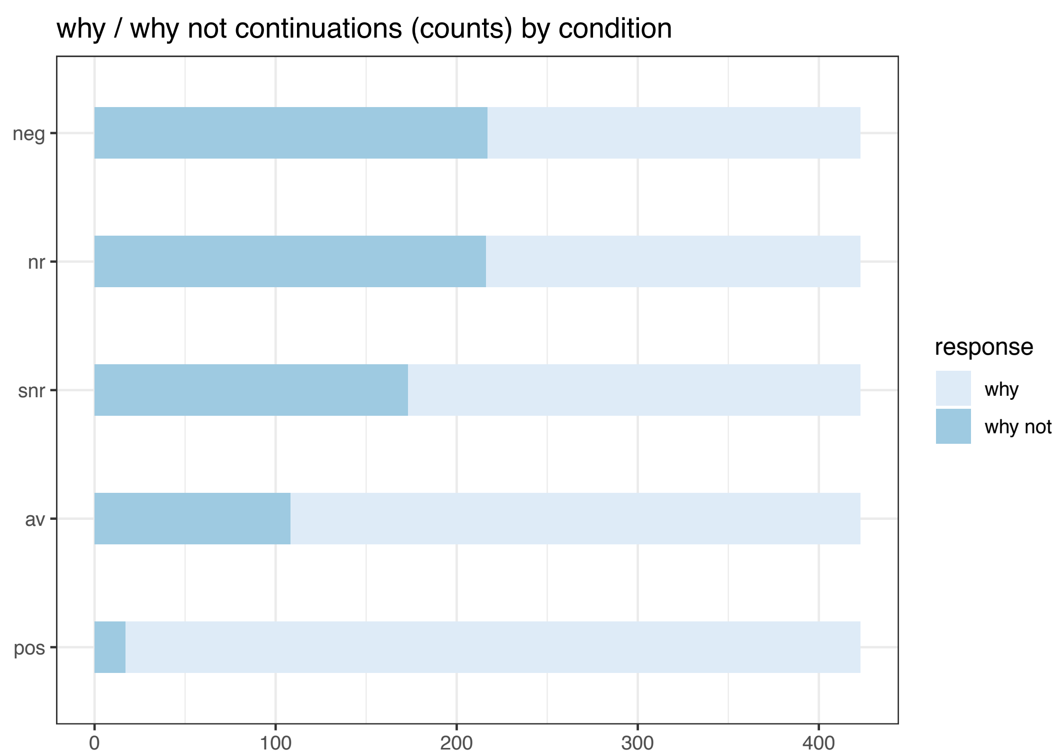


Figure 6.1: Barplots for responses in the five conditions

Based on the raw counts, the following generalizations emerge:

- In order to ask for the reason of the truth of a previous embedded clause that contains overt negation (NEG-condition), the choice between *why* and *why not* is roughly at chance.

- Canonical cases of neg-raising with *think* and *believe* (NR-condition) behave just like overtly negated embedded clauses in the discourse.

- Veridical embeddings without negation (POS-condition) lead to a strong preference for *why* over *why not* (only 4% *why not* responses).

- Both antecedent sentences in the SNR- and AV-conditions exhibit a preference for *why* over *why not*. However, as we expected on the basis of the veridicality-based theory of discourse negativity, both of these contexts show higher proportions of *why not* compared to the positive baseline.

- Further, there is a difference between anti-veridical embeddings without negation and semantic cases of neg-raising. The proportion of *why not* choices is lower in the latter.

To test how confident we can be in these generalizations, we use a statistical analysis of our data.

Analysis and Results

In the following, I present a statistical analysis, which suggests that the above generalizations, are based on statistically significant differences between the conditions, although some of them are quite small. In order to measure the uncertainty for the above suggested generalizations for our binary response variable, we use a mixed-effects logistic regression. The predictor (/fixed effect) is the main effect of experimental condition ({NEG, NR, SNR, AV, POS}), where POS is the reference level.

For an initial pass at the data, we use a frequentist logistic regression.[8] Following Barr, Levy, Scheepers, and Tily (2013), we use the maximal random effects structure that converges. In this case, the model includes fully crossed random intercepts for items and subjects. Attempts to estimate models where these intercepts are correlated with a random slope (for conditions) did not converge.

The model obtained in this way supports the generalization that anti-veridical embeddings license *why not* negativity-tags more robustly than the positive baseline, whether the anti-veridicality is introduced by negation, neg-raising or anti-veridical attitudes. For all non-baseline conditions, the estimated coefficients (presented as the log odds for a 'why not'-response) are negative, indicating that 'why not'–responses are more likely than in the positive baseline.

Maximum likelihood estimates (MLEs) for the coefficients are the following, in or-

[8] The model was estimated in `R` (Team (2014)) using the `lme4`-package (Bates, Maechler, Bolker, Walker, Christensen, Singmann, Dai, Scheipl, Grothendieck, Green, Fox, Bauer, and simulate.formula) (2021)).

der of decreasing log-odds for a 'why not'–response: POS (/intercept): -3.4194, AV: 2.2324, SNR: 3.0197, NEG: 3.4544, NR: 3.4806. In all cases, the (absolute) z-values are larger than 8.0, and all p-values (for the Z-statistic) are $< 1.15 \times 10^{-15}$, indicating statistically highly significant differences between the non-baseline conditions and the positive baseline.

While this analysis supports the generalization that all non-baseline conditions pattern as negative (or at least don't pattern as positive) in discourse, we also want to ascertain the statistical significance of the differences between the various non-baseline conditions. For this part of the analysis, we follow Brasoveanu et al. (2014), getting at these differences by quantifying the uncertainty about the estimated coefficients. This can be done by examining posterior distributions for these values using Bayesian modeling with low-information priors.

For this purpose, a Bayesian logistic regression is used[9] with fully crossed random intercepts for subjects and items, correlated with random slopes by condition.

The (mean) estimates for the model coefficients in this model are (in log odds of a *yes*-response): POS (/intercept): -3.36, AV: 2.18, SNR: 2.94, NEG: 3.41, NR: 3.47.

While the posterior distributions cointain the information about the uncertainty associated with possible parameter values, we consider their median values for the model coefficients, along with their 95% credible intervals (CRIs), scaled as probabilities of a

[9]The model is estmated in `R` using the parckages `brms` (Bürkner (2017)), `RStan` (Team (2020)), and `rstanarm` (Goodrich, Gabry, Ali, and Brilleman (2020)). These `R`-packages provide an interface to the statistical modeling platform `Stan` (Carpenter, Gelman, Hoffman, Lee, Goodrich, Betancourt, Brubaker, Guo, Li, and Riddell (2017)). The model was estimated using the `brms` default priors.

'why not'-response. These are plotted in **Figure 6.2**.

This graphic representation shows a clear separation between the POS and AV conditions, as well as between SNR and SNR, NR, and POS. In contrast, the CRIs for SNR, NR, and NEG overlap, while all other CRIs do not.

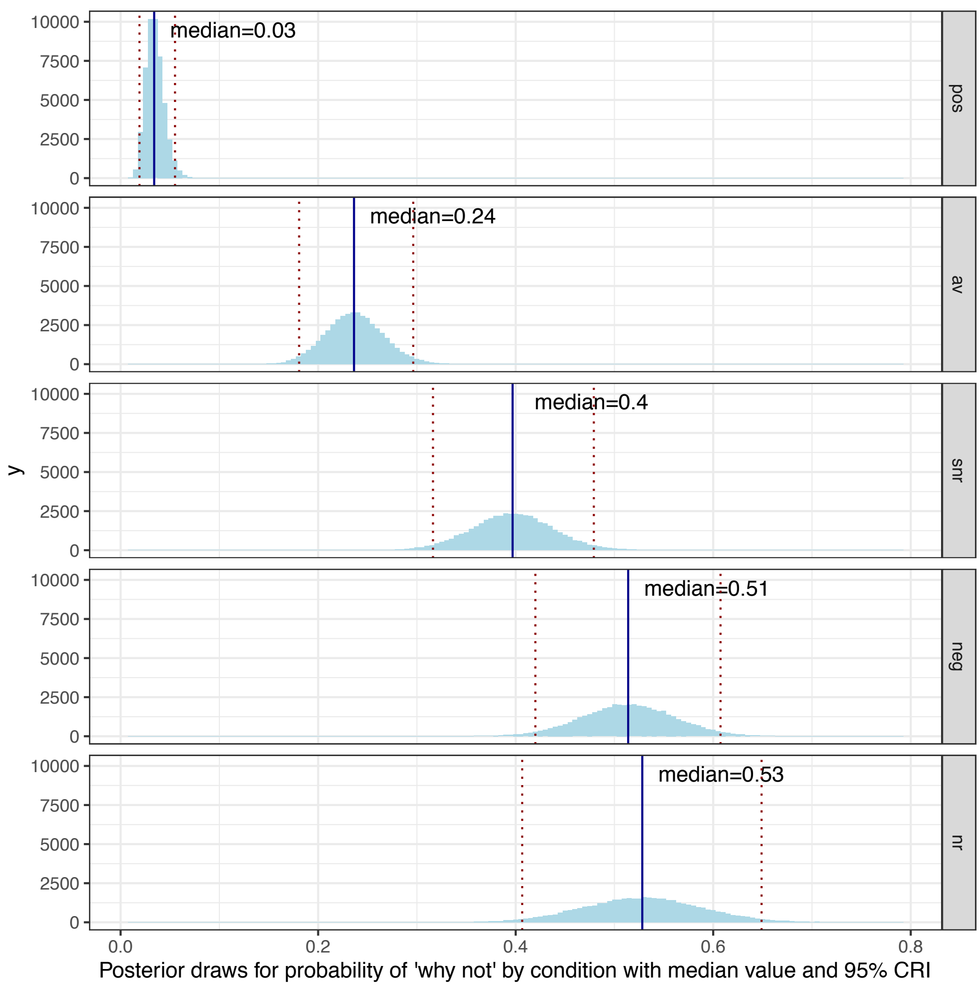


Figure 6.2: Median probabilites for a *'why not'*-response and 95% CRIs by condition

Considering (pairwise) differences between the (scaled) posteriors for two coefficients provides a measure of the (un)certainty about their difference. They are plotted in **Figure 6.3**.

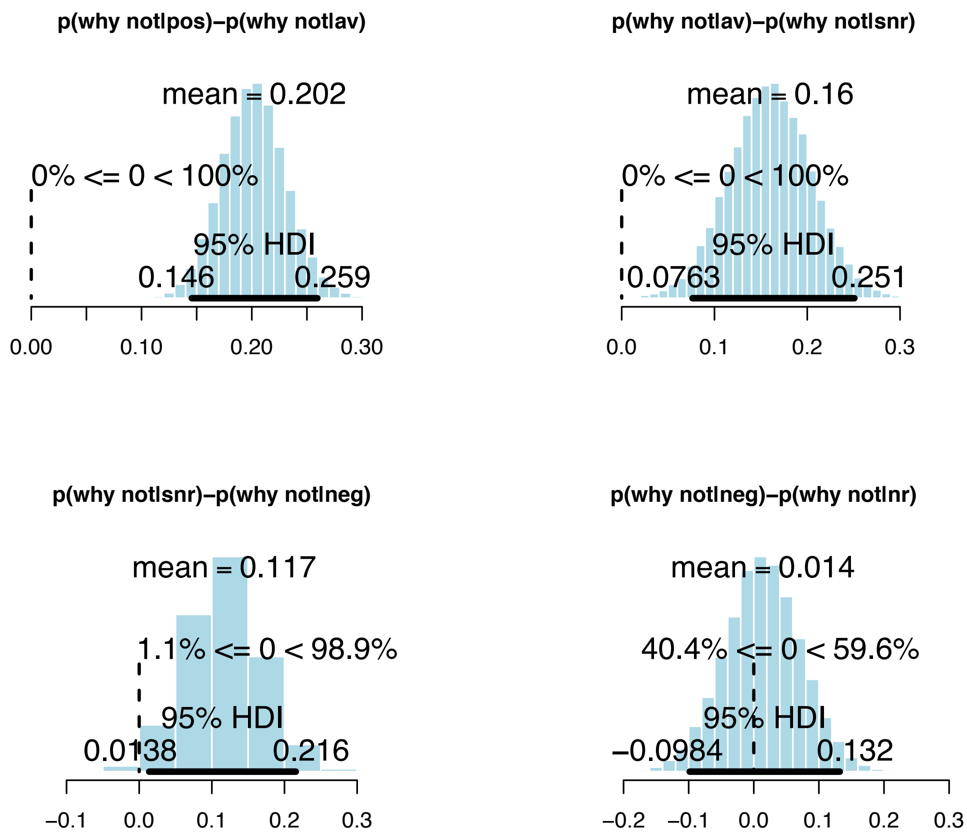


Figure 6.3: Differences between the posterior distributions for the model coefficients, scaled as probabilities for a *'why not'*-continuation, with 95% HDI and indication of how much probabiliy mass is associated with a difference $x <= 0$.

As is expected based on the graphical representation in **Figure 6.2**, the pairwise comparison between the conditions NEG, NR show a 95% HDI which includes a 0 difference. For the difference between these conditions we cannot say with certainty that they are statistically significant, for a significance level of $\alpha = 0.05$.

The difference between SNR, NEG does show that a 0 difference numerically lies outside of a 95% HDI. Accordingly, for a significance level of $\alpha = 0.05$, we characterized these conditions as significantly different. However, while the mean difference in probability between the two conditions is around 0.117, a 0-difference only lies slightly outside of the 96% HDI. As a consequence, we cannot say with certainty that the difference between these conditions is as large as the mean suggests.

The differences are much clearer for the pairwise comparisons between POS, AV and AV, SNR. Here, the difference between probability samples for the conditions does not include 0 in the respective 95% HDIs. In both cases, we can be at least 95% confident that the conditions are associated with population means that are truly different. The mean difference is quite substantive in the case of POS, AV (with a mean difference of 0.202 and a 95% HDI between 0.146 and 0.259), but also non-negligible for AV, SNR (mean difference: 0.16 and 95% HDI: 0.076, 0.251).

6.6.5 Discussion

We find that anti-veridical attitudes and semantic neg-raising can be characterized as pretty discourse-negative, but less negative than antecedents with overt negation or prototypical neg-raising. This pattern was roughly what we expected, on the basis of our veridicality-based characterization of the discourse-representation associated with discourse-negativity.

Accordingly, I interpret the increased proportion of *'why not'* continuations in these cases compared to the positive baseline as evidence that negativity-tags are licensed by anti-veridical embeddings (even ones that do not involve a negative operator). On the other hand, the reduced proportion of *'why not'* continuations compared to the NEG-condition suggests that additional factors constrain the availability of *'why not'* in the case of anti-veridical attitude embeddings. The question of what this additional constraint may be is an avenue for future research. Interestingly we also see that canonical neg-raising (with *think/believe*) is just as negative as NEG. Whatever kinds of constraints affect anti-veridical attitudes do not apply here.

6.7 Why not

This section presents an analysis of elliptical *'Why not'*-interrogatives. Let us start by more explicitly distinguishing two readings of *'Why not?'*. The observation emerged out of the annotation work in the Santa Cruz sluicing group (Anand et al., 2021): Negative reduced *why*-questions can have two different kinds of readings, which we will call the *factive* reading and the *modal* reading, illustrated in (41) and (42).

(41) Factive reading:

A: *Sue didn't sleep.*

B: *(Really?) Why not?*

Paraphrase: Why didn't Sue sleep?

(42) Modal reading:

A: *Should we sleep?*

B: *(Sure,) Why not?*

Paraphrase: Why shouldn't we sleep?

RQ: There is no reason we shouldn't sleep.

The factive reading in (41) is our negativity-tag. This interpretation arises in information-seeking question uses or veridical/factive embeddings, where the inference is that the interrogative has a true answer. Negative why-questions that are assumed to have a true answer presuppose their prejacent (in (41): *Sue didn't sleep*), and if they include ellipsis in addition, require a discourse-negative antecedent. The modal reading, on the other hand (42), is interpreted as a modalized interrogative with the suggestion that there is no true answer, and often used to take up on a suggestion or offer. This reading arises in rhetorical question uses, and certain non-veridical embeddings. In English, it characteristically receives a paraphrase involving *would* or *should.*

Setting aside the modal reading for now, this section contributes three points about the interpretation of factive uses of elliptical *'why not'*-questions:

1. It explains the negativity-tag-like behavior of factive *'Why not'* within the analysis presented here by showing that the negative-antecedent requirement comes about as a conspiracy of a configuration of multiple anaphoric elements and negation: First, the clausal ellipsis in the scope of negation requires an overtly realized propositional antecedent which can negated in the context of the question (following Kroll (2019)). Second, the factivity of the question requires presuppositional commitment to the prejacent *not* ϕ. As a result of this configuration, the requirement on the context is that it supply a veridical proposition *not* ϕ and an anti-

veridical proposition ϕ, exactly the kind of configuration we see in a discourse-negative context.

All three components are needed to derive the negative antecedent requirement: both ellipsis, the factive presupposition, and negation, as it is schematically shown in (43). Particularly, negation scopes over the ellipsis (β) and is in the scope of the presupposition trigger (α).

(43) α … [NEG [… β …]]

2. It shows how the same assumptions can also account for another feature of the interpretation of negation in factive *'Why not'*: negative neutralization.

6.7.1 Factive *'Why not'*: The Interaction With Negation

Besides the fact that factive *'Why not'* requires a negative antecedent (see **Section 6.2**, and discussion throughout this chapter), this negativity-tag exhibits a pattern that is familiar from the literature on polarity particles: negative neutralization (Pope (1972); Ginzburg and Sag (2000); Kramer and Rawlins (2009); Farkas and Bruce (2010); Roelofsen and Farkas (2015)).

This refers to the pattern that emerges when comparing the negative PolP *no* to its positive counterpart. Given a positive antecedent, only *yes*, but not *no* can be used to agree with the antecedent (44). In the negative context, this contrast is neutralized, and either polarity particle can be used to agree.[10]

(44) Agreement with English polarity particles (*yes/no*)

[after Farkas and Bruce (2010); Roelofsen and Farkas (2015)]

a. Mary slept.
   (i) Yes, she slept.
   (ii) #No, she slept.

b. Mary didn't sleep.
   (i) Yes, she didn't sleep.
   (ii) No, she didn't sleep.

We can think of positive *why*-sluices as the positive counterpart of *why not*-tags. Comparing the two, we also show see a pattern of negative neutralization: With a positive antecedent, only *why*-sluices are available, whereas in the negative context, both are acceptable and receive the same interpretation.

(45) Clausal ellipsis in *why*-questions

a. *Jasmin slept.*
   (i) *Why?*
   (why did Jasmin sleep?)
   (ii) *# Why not?*

b. *Jasmin didn't sleep.*
   (i) *Why?*
   (why didn't Jasmin sleep?)
   (ii) *Why not?*
   (why didn't Jasmin sleep?)

[10] Disagreement with English PolPs exhibits a similar pattern: To disagree with a positive, only *no* is acceptable, and *yes* isn't. In the negative context this contrast is neutralized, and either particle may be used to disagree.

(i) Disagreement with English polarity particles (*yes/no*)

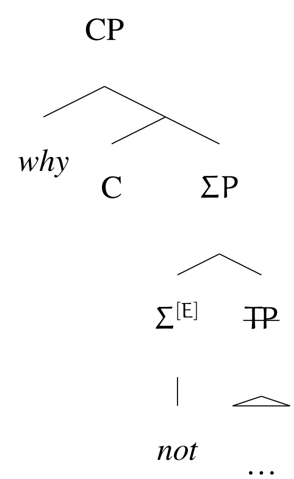
[after Farkas and Bruce (2010); Roelofsen and Farkas (2015)]

a. Mary slept.
   1) #Yes, she didn't sleep.
   2) No, she didn't sleep.

b. Mary didn't sleep.
   1) YES, she slept.
   2) No, she slept.

Based on earlier work (Hofmann, 2018, as well as suggestions in Kramer and Rawlins (2009)), I suggest that *'Why not'* involves ellipsis of the complement of a CP-level polarity head Σ (polarity ellipsis).[11]

(46)



As such, *'Why not'* is similar to other cases of polarity ellipsis under *if* or epistemic sentence adverbials. Kramer and Rawlins (2009) suggest that polarity ellipsis (PE) may be a suitable analysis for these cases as well, and note that they also show negative neutralization patterns in contrast with their positive versions (47) + (48).

(47) a. *Maybe Jasmine is coming to the party*

(i) *If so, you can meet her there.*

(ii) *If not, you'll meet her some other time.*

[11] Some Precedence in the literature (among others) for a high polarity phrase (ΣP) above TP can be found in e.g. Laka (1990); Zanuttini (1997); McCloskey (2017b) and for English more specifically in e.g. Ladusaw (1980, 1992); Merchant (2004); for polarity ellipsis of the complement of high Σ: see e.g. Kramer and Rawlins (2009); Holmberg (2013); Gribanova (2015); McCloskey (2017a).

b. *Maybe Jasmine isn't coming to the party*

(i) *If so, you'll meet her some other time.*

(ii) *If not, you'll meet her some other time.*

(48) a. *Maybe Jasmine is coming to the party.*

(i) *Probably/maybe/of course.* (Probably/maybe/of course she is.)

(ii) *Probably/maybe/of course not.* (Probably/maybe/of course she isn't.)

b. *Maybe Jasmine isn't coming to the party.*

(i) (?) *Probably/maybe/of course.* (Probably/maybe/of course she isn't.)

(ii) *Probably/maybe/of course not.* (Probably/maybe/of course she isn't.)

You may recall that agreeing uses of negative PolPs, which also exhibit negative neutralization also have a requirement for a negative antecedent, which is why we were able to use them as a diagnostic of discourse-negativity in **Section6.2**. For the cases of PE with *if* or epistemic sentence adverbials, this is different: Although they exhibit a pattern of negative neutralization (the contrast between the positive and negative forms is leveled with negative antecedents), the negative form does not require a negative antecedent in these cases (47-a-ii) + (48-a-ii).

(49) Maybe Mary slept,

a. maybe not (maybe she didn't sleep)

b. but probably not (probably she didn't sleep)

So while factive *'Why not'*, agreeing uses of *no*, and other cases of PE all exhibit negative neutralization, only the former two have a requirement for a negative antecedent. An overview of the ways in which the interpretation of these anaphoric expressions interacts with negation is given in **Table 6.2**.

	Factive Why not	**Agreeing PolP uses**	**PE w/ *if*/adv**
Negative Neutralization	✓	✓	✓
Negative Antecedent Requirement	✓	✓	✗

Table 6.2: Overview of how the interpretation of relevant expressions interacts with negation

6.7.2 The pragmatics of *why*-questions and factivity

Possible Answers and the Factivity of *why*

Bromberger suggests that the answer to a *why*-question is an *explanation*: A proposition r, which can count as a reason for the prejacent p. Accordingly, we can roughly characterize the meaning of a *why*-interrogative as characterizing a set of propositions that may count as the reason for p (Hamblin, 1976; Karttunen, 1977; Groenendijk and Stokhof, 1984).

(50) Semi-formal characterization of *why*-questions

a. Why did Mary sleep?

$\{r$: Mary slept because $r\}$

b. Why didn't Mary sleep?

$\{r$: Mary didn't sleep because $r\}$

It is often assumed in the literature that *why*-questions are always factive and cannot receive a rhetorical question interpretation under which the set of answers is assumed to be empty ((Lawler, 1971; Bromberger, 1992; Gutiérrez-Rexach, 1997; Giannakidou, 1998)). Fitzpatrick (2005) shows that while that is true for indicative non-modalized *why*-questions, modalized *why-would* questions can be used rhetorically with the inference that all possible answers are false, in contrast to *how-come-would* questions.

(51) Fitzpatrick (2005):

a. *Why did John leave?*

 # He has been here the whole time, and won't move an inch.

b. *Why would John leave?*

✓*He has been here the whole time, and won't move an inch.*

c. *How come John would leave?*

 # He has been here the whole time, and won't move an inch.

Based on this contrast, Fitzpatrick argues that *why*, unlike *how come*, is not a true (semantic) factive. Rather, the presupposition in information-seeking why questions is a pragmatic presupposition in the sense of Stalnaker, Munitz, and Unger (1977) that arises, because all possible answers to the question license an inference to the truth of the prejacent (Keenan and Hull (1973), see also Abusch (2010)). This allows for the presupposition to be absent in a RQ interpretation, under which none of the possible answers are true.

The relationship between the existence of a true answer and the truth of the prejacent follows from the semantics Bromberger (1992) assumes for *why*-questions. Bromberger suggests that the answer to a question *why* p: r is a premise which in conjunction with a special abnormic law warrants the conclusion that p is true, and if r were false, would lead to the conclusion that p is false. Bromberger defines several conditions on abnormic laws, the context, and use of *why*-questions, which I will gloss over here. Bromberger's theory of *why*-questions, which is intended as a theory of what constitutes an explanation. I am giving here are strongly simplified version of his analysis, which nonetheless captures the basic intuition, and the relationship between the existence of a true answer and the truth of the prejacent:

(52) Special abnormic law (simplified):

A true law-like statement of the form:

$\forall w(\mathrm{p}(w) \leftrightarrow (\mathrm{r}_1 \vee \cdots \vee \mathrm{r}_n))$

The intuition behind this is that a statement like this describes 'abnormal circumstances'. (52) could be paraphrased as follows: 'p is false, unless at least one of the $\mathrm{r}_1, \ldots, \mathrm{r}_n$ propositions is also true' (or: 'p is true just in case one of the $\mathrm{r}_1, \ldots, \mathrm{r}_n$ propositions is also true'.) For our case, we can assume the following:

(53) $\forall w(\mathrm{p}(w) \leftrightarrow (\mathrm{r}_1(w) \vee \cdots \vee \mathrm{r}_n(w)))$

For each world w*: Sue slept in* w*, unless…*

a. r_1 : Sue had too much coffee in w

b. r_2 : Sue was kept up by the neighbors having a party in w

c. $\mathfrak{r}_3$: Sue was too nervous too sleep in w

d. …

Now, any one proposition on the right-hand side of a true statement of the form (52) could count as an explanation for the proposition on the left side. That is, a reason $\mathfrak{r}$ for $\mathfrak{p}$ is a sufficient condition for $\mathfrak{p}$, and at the same time it is necessary that at least one possible reason $\mathfrak{r}$ is true (again, Bromberger's definition is quite a bit more complicated than this, but this is sufficient for our current purposes).

From this simplified definition, it follows that if any reason $\mathfrak{r}$ for $\mathfrak{p}$ is true, $\mathfrak{p}$ is true as well (as does from Bromberger's more complicated version).

Bromberger's semantics for *why*-questions can provide an explanation for why the factivity of *why* is a pragmatic presupposition, as Fitzpatrick (2005) notes, and how the existence of a true answer leads to an inference that the prejacent of the *why*-question is true. On the other hand, if the implication is that there is no true reason $\mathfrak{r}$ for $\mathfrak{p}$, then there is an inference that $\mathfrak{p}$ is false.

Factivity and the Negative Antecedent Requirement

The factive reading of *'Why not'* arises with canonical question uses. In matrix questions, this means being interpreted as an information seeking question (ISQ) inquiring about the truth of a salient negative proposition (54-a). In embedded contexts, it receives an interrogative interpretation in relation to the embedding context (54-b).

(54) Factive reading:

a. A: *Sue didn't sleep.*

B: *(Really?) Why not?*

Paraphrase: Why didn't Sue sleep?

b. *Jasmine didn't sleep and I know why not.*

Paraphrase: I know why Sue didn't sleep.

Based on the discussion above, we can now understand why canonical question uses, where the interrogative is assumed to have a true answer come with a factive presupposition.

Now, based on the generalization that factive *'Why not'*, which comes with a negative antecedent requirement, also comes with a presupposition of the (negative) prejacent, suggests that the two facts are related.

Above, we saw that PE alone does not lead to a requirement that the antecedent be negative. That this requirement is in fact related to the presupposition is suggested by the following observation: Not only does the antecedent need to be be negative, but also do we need previous commitment to the proposition that would be expressed by the prejacent (55).

(55) a. A: *I wonder if Jasmine didn't sleep last night.*

B: #*Really, why not?*

b. A: *It's a possibility that Mary didn't go to school today.*

B: #*Really, why not?*

c. A: *I'm sure that Jim doesn't surf.*

B: *Really, why not?*

Despite the fact that the requirement for previous commitment points towards the factivity being part of the explanation for the negative-antecedent requirement, the factivity in negative information-seeking *why*-questions by itself does not lead to a negative antecedent requirement. Only in combination with clausal ellipsis do we see this effect:

(56) A: *Mary was awake all night.*

a. B: *Why didn't Mary sleep?*

b. B: #*Why didn't Mary sleep?*

An overview of the kinds constructions discussed here, and the constraints they place in the previous discourse is shown in **Table 6.3**.

	Factive *'Why not'*	**Negative *why*-Q's w/o ellipsis**	**Modal *'Why not'***
Presuppose Prejacent	✓	✓	✗
Negative Antecedent Requirement	✓	✗	✗

Table 6.3: Overview of how the interpretation of different uses of *why*-questions interacts the previous discourse

6.7.3 Towards an Analysis: Factive *'Why not'*

Dissecting the Negative Antecedent Requirement

Factive *'Why not'* involves two anaphoric expression: The clausal ellipsis in the scope of negation and the factive presupposition trigger *why*, in addition to the negation. All three components are needed: both anaphora and negation. Particularly, negation scopes over one anaphor and is in the scope of another as shown in the schematic representation (43), repeated here.

(43) α … [NEG [… β …]]

In the following, we will see that this requires that a veridical antecedent and its anti-veridical opposite are both available in the discourse. This is further evidence for a pragmatic view of discourse-negativity: The negative antecedent requirement comes about as a a semantic conspiracy, not by an anaphoric mechanism that makes reference to sentential negation.

Clausal Ellipsis

Following Kroll and Rudin (2017); Kroll (2019), assume that clausal ellipsis **(i)** requires an overt linguistic antecedent and (a weak form of) syntactic parallelism over the *v*P-domain, and **(ii)** local givenness of the proposition expressed by the elided TP.

For concreteness, I am adopting the following assumptions:[12]

(57) The E-feature for polarity ellipsis (modified from Merchant (2004), updated with Kroll's (2019) local givenness):

a. Syntax: $[\mathrm{E}_{\Sigma}]$

b. Phonology: $\varphi_{\mathrm{TP}} \mapsto \varnothing \mid \mathrm{E}\_\_\_\_$

[12]The only part of this that is crucial for my analysis is that the ellipsis is licensed in relation to an antecedent proposition.

c. Semantics: ϕ is locally given

(58) **Local givenness** [after Kroll (2019)]

A TP α can be deleted *iff* ExClo($[\![\alpha]\!]^g$) expresses a proposition ϕ, such that ϕ is entailed by the local context and ϕ is maximally salient.

With this licensing requirement, we understand clausal ellipsis as anaphoric to propositions, and closely related to pronominal propositional anaphora. **Chapter 5** introduced a closely related licensing on the interpretation (as opposed to licensing) of pronominal anaphora. Based on this, we can turn to Krifka's (2013) analysis negative neutralization patterns, which was developed for PolPs.

Krifka (2013) suggests an analysis of negative neutralization patterns, which exploits the fact that negative sentences introduce two propositional drefs. This account assumes a minimal semantics of *yes* and *no* as realizing affirmative and negative polarity, respectively. They are propositional anaphora, s.t. *yes* asserts its antecedent (i.e. a veridical propositional anaphor) and *no* asserts the negation of its antecedent (i.e. an anti-veridical one). The patterns for English and German PolPs are analyzed by assuming that PolPs may be anaphoric to either the proposition introduced by the matrix clause or the one embedded under negation.

(59) A: $^{\phi_1}$ Ede didn't$^{\phi_2}$ steal the cookie.

a. B: Yes$_{\phi_1}$, he didn't. (agreement)
b. B: Yes$_{\phi_2}$, he did. (disagreement)
c. B: No$_{\phi_1}$, he didn't. (agreement)
d. B: No$_{\phi_2}$, he did. (disagreement)

While this analysis would work straightforwardly with the analysis of anaphora and pragmatic account of discourse-negativity proposed here, our current goal is not to make claims about how to analyze PolPs. However, we can draw on this to develop our analysis for negative neutralization patterns with negative polarity ellipsis. Negative neutralization could be analyzed similarly, by appealing to the two propositions made salient by a negative antecedent. While the anaphoric mechanism of ellipsis under negation may retrieve the anti-veridical proposition ϕ_2, in the case without negation, the veridical proposition ϕ_1 is picked up.

(60) A: $^{\phi_1}$ Ede didn't$^{\phi_2}$ steal the cookie.

a. B: Maybe ~~TP~~$_{\phi_1}$. (maybe he didn't)

b. B: Maybe not ~~TP~~$_{\phi_2}$. (maybe he didn't)

c. B: Why ~~TP~~$_{\phi_1}$? (why didn't he?)

d. B: Why ~~TP~~$_{\phi_2}$? (why didn't he?)

Here is a simple way of implementing this in our system, for *maybe (not)*, for example: Assume that ellipsis is licensed by associating an [E]-feature with the syntactically lowest overt remnant Merchant (2001). While the [E]-feature signals towards the phonology that the complement shouldn't be pronounced, it signals towards semantics that the complement should be locally given. Let's say we ensure givenness by assuming that the [E]-feature on an expression makes sure that that expression doesn't introduce a dref, but instead retrieves one anaphorically.

(61) a. Negation (non-anaphoric):

$\lambda \mathcal{P}_{\mathbf{wt}}.\lambda \phi_{1\mathbf{w}}.[\phi_2 \mid \phi_1 = \overline{\phi_2}]; \mathcal{P}(\phi_2)$

b. *maybe* (simplified, non-anaphoric):

$\lambda \mathcal{P}_{\mathbf{wt}}.\lambda \phi_{1\mathbf{w}}.[\phi_2 \mid \phi_1 \not\subseteq \phi_2]; \mathcal{P}(\phi_2)$

(62) a. Negation (anaphoric):

$\lambda \mathcal{P}_{\mathbf{wt}}.\lambda \phi_{1\mathbf{w}}.[\phi_1 = \overline{\phi_2}]; \mathcal{P}(\phi_2)$

b. *maybe* (anaphoric):

$\lambda \mathcal{P}_{\mathbf{wt}}.\lambda \phi_{1\mathbf{w}}.[\phi_1 \not\subseteq \phi_2]; \mathcal{P}(\phi_2)$

Negative Antecedent Requirement

We get the right interpretation if the antecedent is the proposition introduced under negation:

(63) a. A: [Sue didn't [~~Sue~~ sleep]$^{\phi_2}$]$^{\phi_1}$

b. B: why not ~~[Sue sleep]~~$^{\phi_2}$

Intuitively, this is a licit discourse, because it is interpreted in the following way: In (63-a), A introduces ϕ_2 as an anti-veridical proposition. In (63-b), B asks $\{p : (\text{not } \phi_2) \text{ because } p\}$ (i.e. *why not* ϕ_2). By asking this question, B presupposes $(\textit{not}\ \phi_2)$ (i.e. ϕ_1). Because this is the proposition that A has just asserted, the presupposition is satisfied.

So how can reference to the matrix proposition be ruled out in a case like (64)?

(64) a. A: [Sue slept]$^{\phi_1}$

b. B: why not ~~[Sue slept]~~$_{\phi_1}$

This is interpreted in the following way: In (64), A asserts ϕ_1 (i.e. *Sue slept*). In (64), B asks $\{p : (\text{not } \phi_1) \text{ because } p\}$ (i.e. *why not* ϕ_1). By asking this question, B presupposes $(\textit{not } \phi_1)$. Neither A nor B are committed to $(\textit{not } \phi_1)$. In fact, A is committed to ϕ_1 and the presupposition fails.

Main Points

The assumptions made here and the pragmatic characterization of discourse-negativity allow us to account for the negative antecedent requirement and negative neutralization property of factive *'Why not'* using only minimal extensions to the current framework.

The properties of the polarity-ellipsis and the factive presupposition are both independently motivated, and only in combination conspire with negation to require that the antecedent be an anti-veridically introduced proposition. This section showed that the requirement for a negative antecedent in information-seeking *'Why not'*-questions comes about as a conspiracy of negation scoping over a propositional anaphor and in turn be in the scope of another propositional anaphor. Because the negativity-tag requires both a veridical and an anti-veridical salient proposition, this is an additional argument for a pragmatic account of sentential negativity of the kind proposed here rather than one making reference to sentential negation.

6.8 Conclusion

Through careful empirical investigation, I established that discourse-polarity is influenced by complex factors, some of them syntactic, semantic and pragmatic in nature. This chapter presented arguments for a characterization of discourse-negativity on a level of representation that is amenable to contextual manipulation and pragmatic inference, and against one that relies on the expression of negation within a clause. I further suggested that the generalization developed in Part I of this work is a suitable characterization: A discourse-negative utterance introduces an anti-veridical propositional dref. With this generalization in mind, I further addressed some possible arguments for an account of discourse-negativity which makes reference to syntactic structure. I showed how these data could be addressed within the account assumed here. Further, I outlined a way of experimentally testing the predictions of the proposed account and comparing them against other possible characterizations of discourse-negativity.

Chapter 7

Individual Anaphora

7.1 Introduction

This chapter extends the framework introduced in **Chapters 3 & 4** to account for the interaction of individual anaphora and veridicality. It does this by extending the analysis of modal subordination presented in Stone (1999); Stone and Hardt (1999) to other cases of non-veridically and anti-veridically introduced drefs with the goal of understanding the circumstances under which a dref is available for subsequent anaphoric reference.

As suggested in **Chapter 2**, the analysis requires that the discourse representation keeps track of the epistemic status of individual drefs, i.e. their relationship to the speaker's commitments. We do this by incorporating the analytical intuition behind the generalization (1), developed in **Chapter 2**.

(1) **Generalization about accessibility for individual anaphora:**

An (individual-denoting) expression is a possible antecedent for an individual anaphor just in case the referent of the antecedent exists in the local context of the anaphor.

The central claim behind this generalization is that a dref introduced under negation can be the antecedent for a pronoun, only if the interpretation does not require that the assertion of its non-existence and the existence-requirement of the pronoun be true wrt the same set of worlds. Assuming an appropriate intensional and epistemic representation for individual drefs, this will result from the general requirements that the referent of a DP exist in its local context, and that a discourse be consistent.

The presented analysis of individual anaphora is therefore in many ways analogous to the analysis of propositional anaphora developed in **Chapter 5**, and builds on it. In order to track the relationship of drefs to the speaker's commitments (which is represented intensionally—as a set of possible worlds), we assume an intensional representation for discourse referents. Propositional drefs are already an intensional representation, as they map to sets of possible worlds. Individual drefs (previously formalized as se-mappings) require some additional mechanism in order to be interpreted intensionally, i.e., relative to possible worlds.

The analysis of individual anaphora is based on relativizing individual drefs to the worlds in which their referent exists (Stone (1999); Stone and Hardt (1999); Brasoveanu (2006, 2010a)), and sentential operators introducing propositional drefs that provide a local context for the interpretation of their prejacent (Karttunen (1973); Heim (1992)). This allows us to understand individual drefs in relation to the speaker's commitments. It extends accounts of modal subordination as involving dependencies between propositional drefs and individual drefs (Stone (1999); Stone and Hardt (1999); Brasoveanu (2006, 2010a)) to new empirical cases involving negation. The analysis is framed in intensional CDRT (developed in **Chapters 3 & 4**, for references see Muskens, 1996; Brasoveanu, 2006, 2010a).

This chapter includes a more systematic overview of the possible combinations for individual anaphora and antecedents in (non-/anti-)veridical contexts. An overview of the combinations is given in **Table 7.1**.

Epistemic context of anaphor	Epistemic context of antecedent	Individuals
Veridical	Veridical	✓
	Non-veridical	✗
	Anti-veridical	✗
Non-veridical	Veridical	✓
	Non-veridical	✓
	Anti-veridical	✗
Anti-veridical	Veridical	✓
	Non-veridical	✓
	Anti-veridical	✓

Table 7.1: Overview of (Im-)possible Combinations for Individual Anaphora

As discussed in **Chapter 2**, classic dynamic semantic theories (Kamp, 1981; Heim, 1982, 1983; Groenendijk and Stokhof, 1991; Kamp and Reyle, 1993) require a potential antecedent to be in a veridical context in order to be picked up by subsequent anaphora.

This is motivated by the contrast in (2).

(2) a. *Mary has [a car]*$^{\upsilon_1}$. *It*$_{\upsilon_1}$ *is parked outside.*

b. *Mary doesn't have [a car]*$^{\upsilon_2}$. *#It*$_{\upsilon_2}$ *is parked outside.*

In fact, **Table 7.1** shows that veridical antecedents are the most flexible with regards to the veridicality of a subsequent anaphor: Potential antecedents in veridical contexts can be picked up by anaphora of any veridicality.

However, the table shows that anaphora to non-veridical and anti-veridical antecedents are also possible. Veridical anaphora can pick up potential antecedents in non-veridical contexts if the antecedent context is compatible with a strengthened veridical interpretation. But even if that is not possible, non-veridical individuals can be antecedents for anaphora in non- and anti-veridical contexts.

Here, I focus on anaphora to non-specific indefinites in the scope of negation. Many of these are anti-veridical and can only be picked up by non-(/anti-)veridical anaphora. They include modal subordination with negative antecedents, such as the example discussed in **Chapter 2**, repeated here as (3-a). They also included cases of what Kibble (1994) calls 'negative subordination', where both the antecedent and the anaphor are in the scope of negation (3-b).

(3) a. Modal subordination:

There isn't [a bathroom]$^{\upsilon_1}$ *in this house. It*$_{\upsilon_1}$ *would be easier to find.*

b. ‘Negative subordination:’

There isn’t [a bathroom]$^{\upsilon_2}$ in this house, so you won’t be able to find it.$_{\upsilon_2}$

Anaphora to indefinites in the scope of negation are also possible when they are in a veridical context, as is the case in double-negation sentences. The example discussed in **Chapter 2** is repeated here as (4).

(4) Double negation:

It’s not true that there isn’t [a bathroom]$^{\upsilon}$ in this house. It$_{\upsilon}$ is just in a weird place.

Anaphora in anti-veridical contexts can be picked up by veridical anaphora, in case they are in discourse segments that are in a non-veridical discourse relation, such as disjunction (5-a) or disagreement (5-b) (examples also repeated from **Chapter 2**).

(5) a. Disjunction:

Either there isn’t [a bathroom]$^{\upsilon_1}$ in this house, or it$_{\upsilon_1}$ is in a weird place.

b. Disagreement:

A: *There isn’t [a bathroom]$^{\upsilon_2}$ in this house.*

B: *(What are you talking about?) It$_{\upsilon_2}$ is just in a weird place.*

The availability of anaphora to non-specific indefinites in the scope of negation, especially the fact that they are systematically available in certain contexts, suggests that we cannot treat negation as externally static. **Chapter 2** argued that these cases provide evidence for a flat-update dynamic semantics under which any dref is relevant for the discourse as a whole, while storing the information about dependencies and relations to other drefs (Stone, 1999; Stone and Hardt, 1999; Brasoveanu, 2006, 2010a,b).

This chapter provides a more detailed overview of the systematic patterns that arise from the interaction of individual anaphora, veridicality distinctions, and speaker commitments. It also develops a version of a flat-update dynamic semantics that can give a unified account of the veridicality-based constraints on anaphora, as well as the classically problematic exceptional cases discussed here. The chapter also gives a more detailed comparison of the presented analysis to other accounts, which adds to the more broad-level comparison of flat-update approaches to accommodation-type approaches and conditional variable update approaches in **Chapter 2**. The chapter proceeds in the following way:

Section 7.2 gives an overview of interaction of individual anaphora and veridicality. It discusses how the generalizations developed in **Chapter 2** lead us to expect the patterns we see and how they can be understood in terms of the generalization in (1), and a general requirement that a discourse be consistent.

Section 7.3 introduces a mechanism of introducing individual drefs intensionally, in relation to the possible wolds in which their referent exists. The section then shows how it is used to interpret individual-denoting expressions relative to their intensional context.

Section 7.4 shows how intensionalizing individual drefs in ICDRT allows us to give an analysis of anaphora to negated content, which can explain the (im)possibilities in terms of a local existence requirement and discourse consistency.

Section 7.5 compares the analysis presented here to other accounts of anaphora to negated and non-veridical content.

Section 7.6 concludes the chapter by summarizing its main points.

7.2 Patterns of (Non-)Accessibility

This section gives a more systematic overview of the possible combinations for individual anaphora and antecedents along the lines of veridicality distinctions. To yield a consistent discourse, individual anaphora in veridical contexts require antecedents in veridical contexts. In contrast, anaphora in non-veridical or anti-veridical contexts place no restrictions on the veridicality of their antecedent.

Like in the discussion of propositional anaphora, the pattern discussed here only applies to discourses where the discourse segment containing the antecedent and the one containing the anaphor are in a veridical discourse relation. This section illustrates these patterns with some simple examples from English. The demonstration for veridical discourse relations, and anaphora in veridical, non-veridical, and anti-veridical contexts, respectively, is followed by a brief discussion of some exceptions to these generalizations that arise under non-veridical discourse relations.

This discussion aims to illustrate how the generalization about anaphoric accessibility can be reduced to the basic principles that **(i)** the referent of an individual-denoting phrase exists in its local context and that **(ii)** discourses should be consistent, just like we saw for propositional anaphora. While both come with their respective presuppositions that the referent exist/be true, the veridicality of the embedding context determines how these presuppositions are interpreted relative to the belief state of the speaker. This generalization is the reason why relativizing individual drefs to their local context allows for an account of the observed patterns.

7.2.1 Anaphora in Veridical Contexts

The Basic Pattern

If an anaphor is in a veridical context, it can only have an antecedent that is compatible with a veridical interpretation. This is illustrated in (6).

(6) a. *I know that Mary has* [*a car*]$^{\upsilon_1}$ *and I was surprised that it*$_{\upsilon_1}$*'s red.*

✓it$_{\upsilon}$: the car that Mary has

b. *John said that Mary has* [*a car*]$^{\upsilon_2}$ *and I was surprised that it*$_{?\upsilon_2}$*'s red.*

? it$_{\upsilon}$: the car that Mary has

c. *It's not the case that Mary has* [*a car*]$^{\upsilon_3}$ *and I was surprised that it*$_{\#\upsilon_3}$*'s red.*

✗it$_{\upsilon}$: the car that Mary has

This pattern is expected under the assumption that an individual anaphor requires that the referent of its antecedent exists in its local context. In (6), the anaphoric pronoun (*it*) is embedded under the emotive factive *surprised*. As such, *surprised* presupposes the truth of its prejacent and introduces a veridical local context for its interpretation. Because the anaphor is in a veridical embedding, the speaker is committed to local context of the anaphor, and the requirements on this contexts apply to their global commitments. This commits the speaker to the existence of a referent for *it*, which in turn requires a veridical antecedent.

The first clause in (6-a) provides a possible antecedent—the indefinite DP *a car.* Since it is embedded in a factive—and therefore veridical—context, an utterance of the first clause commits the speaker to the proposition that *Mary has a car*, and therefore also the existence of *a car* (in particular Mary's car). Because the interpretation of the pronoun requires that the speaker is committed to the existence of a referent, this initial assertion is consistent with anaphorically referring to the car in a veridical context. As a result, the indefinite is an accessible antecedent for the anaphoric pronoun.

In (6-c), the potential antecedent is embedded in an anti-veridical context, and its non-existence is asserted. That is in direct contradiction to the interpretation of the veridical anaphor, which requires that the referent exist (according to the speaker). Since this requirement applies directly to the speaker's commitments here, the interpretation of the second utterance is inconsistent in relation to the first.

In (6-b), the potential antecedent is embedded in the non-veridical context *'John said'*. Based on the semantic content alone, the speaker does not signal a particular commitment wrt the truth of the proposition local to the antecedent *(Mary has a car)*, as well as the existence of said car. Since the interpretation of an anaphor in a veridical context use requires commitment to the existence of the referent, the antecedent in the non-veridical context is not immediately accessible for the veridical anaphor *it*.

Licensing vs. Interpretation

Similar to the accommodation cases discussed in **Chapter 2**, as well as the patterns for propositional anaphora discussed in **Chapter 5**, however, non-veridical embeddings could be pragmatically strengthened to a veridical or anti-veridical interpretation (see e.g. de Marneffe et al., 2012).

Based on whether the interlocutors take John to be an epistemic authority on Mary in the context, *John said* in (6-b) might be interpreted as a veridical embedding in (6-b) (in which case we do expect an anaphoric relation to be possible). On the other hand, the interlocutors might also take John to be an unreliable source with regards to Mary. Then we might get an inference that the speaker is biased against the proposition that *Mary has a car*, and subsequently requiring the existence of said car would lead to inconsistency.

This claim is supported in contrast with non-veridical embeddings under *might*, which are less amenable to pragmatic strengthening (cf. similar contrasts for propositional anaphora discussed in **Chapter 5**).

(7) ✗*Mary might have* [*a car*]$^{\upsilon}$. *It*$_{\#\upsilon}$ *is red.*

The contrast between (6-b) and (7) is similar to what we saw for propositional anaphora.

On the one hand, the relative acceptability of (6-b) (as well as the accommodation-cases in **Chapter 2**) suggests that antecedents in semantically non-veridical contexts can be antecedents for anaphora in veridical contexts. On the other hand, cases like (7) suggest that this is only the case, if the interpretation of the antecedent segment can be strengthened to a (pragmatically derived) veridical interpretation.

Like for propositional anaphora, we can distinguish between conditions on licensing and interpretation, where the former characterizes the semantic conditions on the embedding contexts, and the latter characterizes the actual epistemic status, which may in addition be influenced by pragmatic factors.

An overview of the patterns for veridical individual anaphora, taking into account the difference between licensing and interpretation, is given in **Table 7.2**.

Epistemic context of anaphor	Epistemic context of antecedent	Individuals	
		Licensing	Interpretation
Veridical	Veridical	✓	✓
	Non-veridical	✓	✗
	Anti-veridical	✗	✗

Table 7.2: Overview of (Im-)possible Combinations for Veridical Individual Anaphora

Upshot

The interpretation of an individual anaphor requires that its referent exists. If the pronoun is in a veridical context, this condition applies on a global level. Therefore, a possible antecedent is available if its existence has been asserted on a global level, i.e., in case it was introduced in a veridical context. An antecedent may also be accessible if the previous discourse is compatible with the accommodation of the requirement of the existence of its referent, like with certain non-veridical contexts. If the existence of the referent has been negated, co-reference will result in contradictory requirements on the intepretation and an inconsistent discourse.

7.2.2 Anaphora in Anti-Veridical Contexts

The Basic Pattern

An individual anaphor in an anti-veridical context does not place any restrictions on the veridicality of its antecedent.

(8) a. *I know that Mary has* [*a car*]$^{\upsilon}$ *but it's not the case that it*$_{\upsilon}$*'s parked outside.*

✓it$_{\upsilon}$: the car that Mary has

b. *John said that Mary has* [*a car*]$^{\upsilon}$ *but it's not the case that it*$_{\upsilon}$*'s parked outside.*

✓it$_{\upsilon}$: the car that John said Mary has

c. *It's not the case that Mary has* [*a car*]$^{\upsilon}$ *and it's not the case that it*$_{\upsilon}$*'s parked outside.*

✓it$_{\upsilon}$: the car that Mary doesn't have

The local context of the anaphor is the proposition characterizing a set of worlds in which the conjunction of the local existence requirement (p : υ exists) and the expressed content (q : υ is parked outside) is true. This comes about due to the existence-requirement of it$_\upsilon$), and the overt content of the embedded clause. Because it is in an anti-veridical embedding, the speaker's commitments are characterized by a set of worlds in which the this conjunction is false ($\neg(p \wedge q)$). As a result, the speaker's commitments do not entail the truth or falsity of the existence requirement itself. That is ($(\neg(p \wedge q) \wedge p \not\to 0$, as well as ($\neg(p \wedge q) \wedge \neg p \not\to 0$).

Therefore, an utterance of *it's not the case that it$_\upsilon$'s parked outside* does not commit to the speaker to the existence of said car, or its opposite. Not committing the speaker about the existence of a referent, the interperetation of a non-veridical anaphor is consistent with an antecedent of any veridicality.

(8-c) is not associated with a global inference that Mary has a car or that said car exists. In this case, the existence-requirement of the anaphor is satisfied in the local, anti-veridical context. As a result, the antecedent may be in an anti-veridical context as well. On the other hand, (8-a), does come with a global inference that Mary has a car and that said car exists. In this case, the existence requirement is satisfied in the global context (in addition to the local one). (8-b) is compatible with the existence-requirement being satisfied locally or globally, and this may be disambiguated by pragmatic information—for example depending on whether we might take John to be a reliable source, or whether we might think that the fact that Mary's car is not parked outside is sufficient evidence against the existence of said car.

Individual vs. Propositional Anaphora

The pattern for anti-veridical individual anaphora is different from what we saw for propositional ones in **Chapter 5**. While individual anaphora may have veridical antecedents, propositional ones cannot. This is shown with the contrast between (8-a), repeated from above, and (9), repeated from **Chapter 5**.

(8-a) *I know that Mary has* [*a car*]$^{\upsilon}$ *but it's not the case that it*$_{\upsilon}$*'s parked outside.*

✓it$_{\upsilon}$:the car that Mary has

(9) [$^{\phi_1}$ *I know* [$^{\phi_2}$ *that Mary is sick*]] *and/but that*$_{\phi_1/\phi_2}$ *is not the case.*

a. ✗that$_{\phi_1}$: that I know that Mary is sick

b. ✗that$_{\phi_2}$: that Mary is sick

I have been arguing that propositional and individual anaphora interact with veridicality and commitment in largely parallel ways. I suggest that the above contrast does not provide an argument against this assumption. Instead, this difference can be attributed to the different ways in which propositional and individual drefs relate to their local contexts.

Since the relationship between an individual anaphor and its context is more indirect than for propositional anaphora, a speaker may be committed to the existence of an individual referent without being committed to the propositional embedding context. Because of that, an anti-veridical individual anaphor does not require any specific commitment about the existence of its referent.

While individuals are veridical iff they exist in all the worlds that constitute their local contexts, a propositional anaphor is veridical iff it is true in all the worlds that constitute their local context. In a sense, a proposition *is* its own local context, whereas the relationship for individual anaphora is more indirect.

The pattern for anti-veridical individual anaphora (or lack thereof) is indicated in **Table 7.3**.

Epistemic context of anaphor	Epistemic context of antecedent	Individuals	
		Licensing	Interpretation
Anti-veridical	Veridical	✓	✓
	Non-veridical	✓	✓
	Anti-veridical	✓	✓

Table 7.3: Overview of (Im-)possible Combinations for Anti-veridical Individual Anaphora

7.2.3 Anaphora in Non-Veridical Contexts

A pronoun in a non-veridical context does not place any restrictions on the veridicality of its antecedent in terms of licensing either. The requirement of the anaphor is evaluated in relation to a non-veridical intensional context. The non-veridical embedding is compatible with the embedded proposition being true, false, or neither. So, even if the interpretation of the anaphor requires that its antecedent exist in its local context, there will be no inconsistency on a global level. As a result, any possible antecedent is accessible here.

(10) a. *I know that Mary has* [*a car*]$^{\upsilon}$ *and John believes that it*$_{\upsilon}$*'s parked outside.*

✓it$_{\upsilon}$: the car that Mary has

b. *John said that Mary has* [*a car*]$^{\upsilon}$ *and Sue believes that it*$_{\upsilon}$*'s parked outside.*

✓it$_{\upsilon}$: the car that John said Mary has

c. *It's not the case that Mary has* [*a car*]$^{\upsilon}$ *even though John believes that it*$_{\upsilon}$*'s parked outside.*

✓it$_{\upsilon}$: the car that Mary doesn't have

The most straightforward case for non-veridical individual anaphora is (10-b), where the antecedent is in a non-veridical context as well. Neither the antecedent segment nor the segment containing the anaphor commits the speaker to the existence or non-existence of a referent. It exists in a local context, a proposition that may or may not be true, according to the speaker.

That is different with veridical and anti-veridical antecedents. The antecedent segments already commits the speaker to the (non)existence of the referent, and the segment containing the anaphor does not specify whether the speaker should be committed to the local proposition or not. Therefore, the anaphor inherits its veridicality from the antecedent. In (10-a), the speaker may or may not believe that Mary's car is parked outside, but they are committed to its existence. In (10-c), the speaker has a previous commitment that *It's not the case that Mary has a car*. Because of that and the requirement that the local context of the anaphor supports the existence of the referent, the local context of the anaphor (i.e., the proposition that *Mary's car is parked outside*) receives a non-veridical interpretation due to pragmatic enrichment.

Again, we compare the above cases to non-veridical embeddings under *might*, which are not easily strengthened.

(11) a. *I know that Mary has* [*a car*]$^{\upsilon}$ *and it*$_{\upsilon}$ *might be parked outside.*

✓it$_{\upsilon}$: the car that Mary has

b. *John said that Mary has* [*a car*]$^{\upsilon}$ *and it*$_{\upsilon}$ *might be parked outside.*

✓it$_{\upsilon}$: the car that John said Mary has

c. *# It's not the case that Mary has* [*a car*]$^{\upsilon}$ *even though it*$_{\upsilon}$ *might be parked outside.*

✗it$_{\upsilon}$: the car that Mary doesn't have

Notably, (11-c) is unacceptable. Again, we see the property of *might* being incompatible with a false prejacent. The local existence requirement states that υ exists in its local context. As a result, we expect it to be an epistemic possibility. That is not the case with the anti-veridical antecedent in (11-c). This shows that a non-veridical context which cannot be strengthed towards an anti-veridical context cannot host anaphora to anti-veridical antecedents.

Based on the pattern for propositional anaphora, we might expect a similar thing for veridical antecedents: That they are incompatible with anaphora in strictly non-veridical contexts. However, (11) is perfectly acceptable. This difference between the patterns for individual and propositional anaphora can be explained in the same way as what we said above for anaphora in anti-veridical contexts: The existence requirement of the pronoun applies to the local context. However, it may also be satisfied in the global context in addition to that. As a result, it is possible that the referent (Mary's car) exists, while it is not possible, but not certain, whether it is parked outside.

The pattern for non-veridical individual anaphora is indicated in **Table 7.4**.

Epistemic context of anaphor	Epistemic context of antecedent	Individuals	
		Licensing	Interpretation
Non-veridical	Veridical	✓	✓
	Non-veridical	✓	✓
	Anti-veridical	✓	✗

Table 7.4: Overview of (Im-)possible Combinations for Non-veridical Individual Anaphora

7.2.4 Interim Summary

This section presented an argument that the possibilities for individual anaphora systematically pattern along the lines of veridicality. It further suggested that this is due to requirements that the interpretation of an anaphor places on its local context and the more general one that a discourse cannot be inconsistent. These claims are supported by the empirical empirical patterns that the accessibility constraints give rise to:

Veridical Anaphora require antecedents that are compatible with a veridical interpretation. That includes veridical antecedent, or non-veridical antecedents that can be pragmatically strengthened to a veridical interpretation, but not anti-veridical ones.

Non-verdical Anaphora do not commit the speaker about the existence of their referent. They are compatible with antecedents of any veridicality.

Anti-veridical Anaphora also do not commitment the speaker about the existence of a referent, at least in the case of individual anaphora. This is different from what we observed for anti-veridical propositional anaphora in **Section 5.2**. For propositional anaphora, we saw a reversal of the pattern for veridical anaphora: An anti-veridical propositional anaphor requires an antecedent that is compatible with an anti-veridical interpretation.

An overview of the (im-)possibilities for individual anaphora along the lines of veridicality is given in **Table 7.5**.

Epistemic context	Epistemic context	Individuals	
of anaphor	of antecedent	Licensing	Interpretation
Veridical	Veridical	✓	✓
	Non-veridical	✓	✗
	Anti-veridical	✗	✗
Non-veridical	Veridical	✓	✓
	Non-veridical	✓	✓
	Anti-veridical	✓	✗
Anti-veridical	Veridical	✓	✓
	Non-veridical	✓	✓
	Anti-veridical	✓	✓

Table 7.5: Overview of (Im-)possible Combinations of Veridicality of Antecedent and Anaphor in Licensing and Interpretation for Individual Anaphora

7.2.5 Discourse Consistency

Like discussed for propositional anaphora in **Chapter 5**, the generalizations also apply on the level of discourse relations, and veridicality distinctions for discourse relations (in the sense of Asher and Lascarides (2003)) are relevant here. While the patterns above are found for veridical discourse relations (like conjunction), the same restrictions do not apply in the case of non-veridical discourse relations like alternation (/disjunction) or anti-veridical ones like disagreement. In these cases, the two discourse segments containing the antecedent and anaphor are not required to be consistent with each other (or even required to be inconsistent). This contrast is illustrated for conjunction vs. disjunction in (12).

(12) Veridical individual anaphor with anti-veridical antecedent:

a. *Either Mary doesn't have [a dog]$^{\upsilon}$, or it$_{\upsilon}$ is at the sitter.*

b. *# Mary doesn't have [a dog]$^{\upsilon}$ and it$_{\upsilon}$ is at the sitter.*

The same contrasts can be observed for agreement vs. disagreement:

(13) Veridical individual anaphor with anti-veridical antecedent:

A: *Mary doesn't have [a dog]$^{\upsilon}$,*

a. B: *(That's not true.) It$_{\upsilon}$'s just at the sitter right now.*

b. B: *# (Indeed.) It$_{\upsilon}$'s at the sitter.*

Similarly to what we have seen for propositional anaphora (**Chapter 5**), the discourse relations connecting the discourse segments including the antecedent and anaphor are non-veridical. As such, these cases are not counterexamples but rather an extension of the generalization. The extistence-requirement of the anaphor is satisfied in its local context. The difference in cases with non-veridical discourse relations is that there are fewer constraints on the relationship between the local context of the anaphor and the local context of the pronoun.

7.3 Intensionalizing Individual Drefs

7.3.1 Individual Drefs and Possible Worlds

To capture the relationship between discourse referents and their intensional contexts, we want an intensional representation of drefs. Propositions, of course, are intensional already and we don't need to devise a special mechanism to allow them to relate to other propositions.

To distinguish veridical, non-veridical, and anti-veridical individual drefs, we need an intensional representation of drefs that allows them to be interpreted in relation to possible worlds. A veridical dref, as defined in **Chapter 2**, is one to whose existence the speaker is committed. This means that all the worlds that are members of the speaker's commitment set are ones where the dref in question has an existing referent.

Conceptually, we think of a dref as the mapping between a discourse variable and a referent. To capture the generalizations about anaphora and negation, we need to capture the way in which an individual dref depends on its local context. Specifically, we need to allow the mapping from discourse variables to their referents to be sensitive to possible worlds from their local intensional context.

Some Possibilities

There are a few different possible ways of relativizing the mapping to a possible world:

- **As part of the representation of the variable:**

  We can represent discourse variables as variables for individual concepts (i.e. functions of type $\langle w, e \rangle$) (Stone, 1999; Stone and Hardt, 1999, see also Aloni 2000, 2001 for a use of drefs for individual concepts for a different purpose). Given a world, the individual concept maps to an entity.

- **As part of the representation of the members of the context:**

  The context consists of a set of world assignment pairs (Heim, 1983, et seq.). Each element of the context specifies a variable assignment relative to a possible world.

- **As part of the representation of a plural context:**

  The context consists of a set of variable assignments (Van den Berg, 1996) and the world which the mapping is relativized to is stored in the same assignment (Brasoveanu, 2006, 2010a).

  Regarding the kinds of data and generalizations considered here, this is similar to representing the context as a set of world-assignment pairs. In general, this framework is more flexible, though, and could be used to model the discourse effect of dependencies between drefs and multiple kinds of embedding operators simultanously, including individual quantifiers (see also Brasoveanu, 2010b).

Our Requirements and Choice

When interpreting a variable in its local intensional context, we want to be able to retrieve the relativized mappings by world, and the three options outlined above could all provide a way of doing that. Any of them could be used as a framework for the analysis of the interaction of anaphora and negation presented here, as long as they also include a way of representing the components I argued are necessary for a successful account:

- **It is intensional**:

  Sentence meaning is understood as characterizing sets of worlds in which they are true (propositions).

- **It is epistemic**:

  Utterance meaning is understood in relation to a representation of what the speaker considers to be true.

- **It satisfies the following requirements for making an intensional system dynamic**:

  – Intensional operators are externally dynamic, while drefs in introduced in their scope store information about their intensional context.

– Discourse variables are mapped to a referent in relation to their intensional context.

Satisfying these requirements, any of the three mentioned kinds of frameworks would be able to frame our analysis. Therefore, the choice between them is made for concreteness rather than analytical considerations.

Since we have been working with context states consisting of singleton variable assignments (type s), implementing relativized individual drefs by using drefs for individual concepts seems like minimal addition to the present system. Therefore,we will adopt the strategy proposed in Stone (1999); Stone and Hardt (1999). This involves **(i)** formalizing individual drefs as drefs for individual concepts, and **(ii)** encapsulating quantification over possible worlds in the static definition of atomic DRS-conditions.

Propostional Drefs

The framework set up in **Chapters 3 & 4** captures veridicality for propositional drefs and their relationship to their local context, and propositional drefs will stay the same in this chapter. A propositional dref ϕ is a function of type $s(wt)$ from assignments i_s to sets of worlds (wt). The proposition $\phi_{s(wt)}(i_s)$ is the set of worlds that i assigns to ϕ.

Individual Drefs

The idea of an individual concept goes back to Carnap (1947), who writes:

> *I think that individual variables in modal sentences [...] must be interpreted as referring, not to individuals, but to individual concepts.*
>
> — (Carnap, 1947, p. 180)

Individual concepts are functions of type (we), mapping possible worlds to individual referents. An individual dref υ is a function of type $s(we)$ from assignments i_s and worlds w_w to individuals x_e. The individual $\upsilon_{s(we)}(i_s)(w_w)$ is the individual that the assignment i assigns to υ in w.

The system from **Chapters 3 & 4** makes use of indeterminate individuals $\#_e$ to define (pseudo-)partial variable assignments. Here, we will make use of them to define (pseudo-)partial individual concepts as well.

7.3.2 Interpreting Individual Drefs in Relation to Propositional Drefs

Relative Variable Update

Using individual concepts allows for the introduction of individual drefs along with the information about the set of worlds in which their referent exists (and does not). Individual drefs are introduced relative to the set of worlds in which their referent exists (adapted from Stone (1999); Stone and Hardt (1999)).

(14) Variable Update with υ relative to ϕ:

The relation $i[\phi : \upsilon]j$ holds (for assignments i_s, j_s, and the drefs $\phi \in \textbf{Term}_{s(wt)}$,

$\upsilon \in \textbf{Term}_{s(we)}$) iff the conjunction of the following holds:

$\bullet i[\upsilon]j$

$\bullet \forall w_w.(\phi(j)(w) \rightarrow \upsilon'(i)(w) \neq \#$

(an individual referent of υ is determined in each world in $\phi(j)$)

$\bullet \forall w_w.(\neg\phi(j)(w) \rightarrow \upsilon(j)(w) = \#)$

(in each world not in ϕ, υ points to the indeterminate value)

(14) states that j is an update of i that differs at most wrt the value assigned to υ. The second conjunct requires that for each world w in $\phi(j)$, $\upsilon(j)(w)$ doesn't map to # (but an actual individual). The third conjunct states that for each world w not in $\phi(j)$, $\upsilon(j)(w)$ maps to #, ensuring that the referent of υ exists only in the worlds in $\phi(j)$.

The definition of a relative variable update makes reference to the simple variable update ($i[\upsilon]j$), which is understood in the usual way, as in the previous chapters.

Quantificational Encapsulation and the Local Context

(15) illustrates how a dynamic predicate is defined as an abbreviation for a condition involving the corresponding static predicate and the referents of the argument drefs given i and ϕ.

(15) Dynamic predicates with their arguments as DRS conditions:

$R_\phi\{\upsilon\} := \lambda i_s.\forall w_w.\phi(i)(w) \rightarrow R(\upsilon(i)(w))(w)$,

for $R \in \textbf{Term}_{e(wt)}$, $\upsilon \in \textbf{Term}_{s(we)}$, $\phi \in \textbf{Term}_{s(wt)}$

Like before, conditions abbreviate properties of discourse states. (15) states that all worlds that are members of the proposition supplied as the local context are worlds in which the static relation corresponding to the condition holds. (Like the previous version of the interpretation of DRS-conditions defined in **Chapter 3**), the definition of conditions, therefore, encapsulates universal quantification over possible worlds.

As a reminder—following Stone (1999), quantification over possible worlds is encapsulated because it happens statically, as part of the interpretation of DRS conditions, and separately from the dynamic quantification using random introduction of drefs.

This also applies to the above definition of relative variable update (14). It also encapsulates universal quantification over the possible worlds of the local context, introducing a new dref in relation to each possible world in the local context.

To contextualize this in relation to the discussion from **Section 4.4** (Asher and McCready, 2007, see also): Our system allows us to look at collective sets of possible worlds in relation to each other. That is how propositional drefs and the semantics of propositional operators are captured. The treatment of indefinites and anaphora in intensional contexts is captured by distributively relating to each possible world in the local context.

Anaphoric 'Accessibility' in Relation to the Local Context

Relativizing individual drefs to sets of worlds in which they refer gives rise to an accessibility condition based on an existence condition for pronouns and the notion of a local context (Karttunen (1973); Heim (1992)). The latter involves the idea that expressions are interpreted wrt an intensional context set that is pragmatically constrained on a global level by the speaker's commitments (Stalnaker (1978, 2002); Gunlogson (2004)) and locally constrained by the semantics of sentential operators. Based on this, a dref υ is accessible for reference by a pronoun υ', only if the referent of υ is determined in the local context of υ', i.e., if υ maps to an individual (other than #), for each world in the local context.

Now, individual drefs can be interpreted in relation to (each of the worlds in) their local context. As a result, the availability of a dref to specify the mapping of an anaphoric variable can be derived from a general requirement that the referent of a variable exists in its local context.

In this system, the local context is formally defined in reference to the evaluation of DRS conditions ((15), repeated here) and consists of an assignment i_s, of which the condition is predicated, and a dref for a set of possible worlds ϕ, a compositionally supplied intensional argument.

(15) Dynamic predicates with their arguments as DRS conditions:

$R_{\phi}\{\upsilon\} := \lambda i_s.\forall w_w.\phi(i)(w) \rightarrow R(\upsilon(i)(w))(w)$,

for $R \in \mathbf{Term}_{e(wt)}$, $\upsilon \in \mathbf{Term}_{s(we)}$, $\phi \in \mathbf{Term}_{s(wt)}$

Given i (and a model), the interpretation of conditions, as defined in (15) checks for each world w in $\phi(i)$, if the relation holds in w of the individual $\upsilon(i)(w)$. The evaluation of a DRS condition in the context of i_s and $\phi_{s(wt)}$ is defined (mapped to a truth-value other than #), only if the argument dref maps to an individual other than # for each world in $\phi(i)$, i.e., if the referent $\upsilon(i)$ exists in each world in $\phi(i)$. This general requirement that drefs have an existing referent in each world of their local context follows from the definition of DRS conditions.

The existence condition on reference has consequences for anaphora: As a special case of this general requirement, the condition on anaphoric accessibility also follows from the definition of DRS-conditions:

(16) Available Antecedents:

Given a relation $R_{e(wt)}$, an individual dref $\upsilon_{s(we)}$, a propositional dref $\phi_{s(wt)}$, and a discourse state i_s:

a. A DRS-condition predicated of a discourse state $R_\phi\{\upsilon\}(i)$ is defined as $\forall w_w.\phi(i)(w) \rightarrow R(\upsilon(i)(w))(w)$

b. $\forall w_w.\phi(i)(w) \rightarrow R(\upsilon(i)(w))(w)$ can be determined only if $\upsilon(i)(w)$ is determined for all worlds w s.t. $\phi(i)(w)$, therefore $\forall w_w.\phi(i)(w) \rightarrow \upsilon'(i)(w) \neq$ #

c. As a result, a pronoun in the context of i and ϕ can make reference to the dref υ only if the value of $\upsilon(i)(w)$ is determined for all worlds w, s.t. $\phi(i)(w)$

(16) is a direct consequence of the definition of atomic conditions. No designated theory of accessibility is needed, follows from general requirement that 1. the referent of a dref exists in its local context and 2. a discourse be consistent. A dref can provide a referent

for an anaphor, only if its referent exists in the local context of the anaphor. Because any dref exists only in the set of worlds of the local context where it was introduced, another result of this is that the intensional context of the anaphor is a subset of the intensional context of the antecedent.

7.3.3 Modelling the Interpretation of Natural Language

The Semantics of DPs

Indefinites. Based on the definition of relative variable update, we assume that the anaphoric potential of indefinite DPs is captured by allowing them to introduce a dref relative to their local intensional context. For a non-specific indefinite in the scope of negation (17), an individual dref υ_1 is introduced in relation to the propositional dref corresponding to the prejacent of negation ϕ_2.

(17) S: *There isn't a bathroom.* $\rightsquigarrow$

$$\boxed{\begin{array}{c}\phi_1\\ \hline \phi_{DC_S} \Subset \phi_1\end{array}}\,;\mathbf{max}_{\phi_1}\left(\boxed{\begin{array}{c}\phi_2\\ \hline \phi_1 \equiv \overline{\overline{\phi_2}}\end{array}}\,;\mathbf{max}_{\phi_2}\left(\boxed{\begin{array}{c}\phi_2 : \upsilon_1\\ \hline \mathrm{Bathroom}_{\phi_2}\{\upsilon_1\}\end{array}}\right)\right)$$

With υ_1 being introduced in relation to ϕ_2 (in the update $[\phi_2 : \upsilon_1]$), the referent of υ_1 exists in all and only the ϕ_2 worlds. Because ϕ_1 is the complement of ϕ_2, the referent of υ_1 does not exist in any ϕ_1-world. Since S is committed to ϕ_1, υ_1 is an anti-veridical dref wrt S in any context that is the result of a successful update with (17).

We formally defined (anti-)veridicality for individual drefs in the following way:

(18) Veridicality for Individual Drefs:

a. An individual dref υ is veridical in context $C = \langle M, \textsc{Int}, i\}$ wrt $A \in \textsc{Int}$ iff $\forall w(w \in \phi_{DC_A}(i) \rightarrow \upsilon(i)(w) \neq \#)$, and non-veridical otherwise.

b. An individual dref υ is anti-veridical in context $C = \langle M, \textsc{Int}, i\}$ wrt $A \in \textsc{Int}$ iff $\forall w(w \in \phi_{DC_A}(i) \rightarrow \upsilon(i)(w) = \#)$.

The indefinite in (17) is interpreted as an unspecific indefinite, relative to the prejacent of negation, because it is interpreted in the scope of negation. Specific indefinites in similar configurations could be derived compositionally by assuming they take scope over negation. i.e. based on a compositional translation of indefinite DPs as quantifiers given in (19), and some compositional mechanism for scope-taking.

(19) a bathroom $\rightsquigarrow \lambda P_{\mathbf{e(wt)}}.\lambda\phi_{\mathbf{w}}.[\phi : \upsilon]; [Bathroom_{\phi}\{\upsilon\}], P(\upsilon)(\phi)$

Proper Names are taken to be discourse constants, as before. They refer to the same object across assignments. In addition, they will be treated as rigid designators. An individual concept corresponding to a propoer name will be a constant function assigning the same individual to the name across worlds. As a result, drefs for proper names are not introduced in relation to their local intensional context.

(20) S: *Mary slept.* $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_1 \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \hline \end{array}; \mathbf{max}_{\phi_1}\left(\begin{array}{|c|}\hline \upsilon \\ \hline \upsilon \equiv Mary_{\mathbf{e}} \\ \hline \end{array}\right)$$

(where $Mary_{\mathbf{e}} := \lambda i_s.\lambda w_w.mary_e$)

(20) states that υ specifies the same mappings as the discourse constant $\mathtt{Mary_e}$, which refers to the individual $\mathtt{mary_e}$ across all assignments and worlds.

This, of course, has strong predictions. The fact that the referent of a name is assumed to exist in any possible world precludes us from giving an appropriate semantics for counterfactuals like *'if Monica Lewinsky had never existed'* (Stone (1999), p. 17). In addition, the use of a proper name (or even the presence of a name in the object language) will presuppose that any interlocutor is committed to their existence.

However, since our focus lies on the interpretation of indefinite antecedents, an appropriate counterfactual semantics for proper names, or an empirical investigation of the soundness of the strong contextual conditions associated with them, will not be further pursued here.

Discourse-Initial States

Given that we changed the type for individual anaphora, we need to slightly adjust our definition of a discourse-initial state in a context C. Recall from **Section 3.3** that a discourse-initial state contains the contextual information about which interlocutors are associated with which drefs for commitment states, but no information about any other discourse variables. We construct a discourse-initial state in two steps: First, we assume an assignment i_0 that contains no information about any discourse variables (21).

(21) Properties of i_0 in C:

a. No individual drefs:

For any individual discourse Variable $\upsilon \in \mathbf{Term}_{s(we)}$:

$[\![\upsilon(i_0)]\!]^{M^C, i_0} = \{\langle w, \#_e \rangle \mid w \in D_w\}$

b. No propositional drefs:

For any propositional discourse Variable $\phi \in \mathbf{Term}_{s(wt)}$:

$[\![\phi(i_0)]\!]^{M^C, i_0} = \{\langle w, \#_t \rangle \mid w \in D_w\}$

Now, a discourse state i_0 in context C is one that maps any individual discourse variable to # in M^C, given assignment i_0, for any world $w \in D_w$. Put differently, it maps any individual discourse variable to an individual concept that is undefined for any world. Since the type for propositional drefs stays the same as in the previous chapters, so does the axiom for i_0 not containing any information about propositional discourse variables: For any propositional discourse variable, i_0 maps to a set of worlds for which membership is undefined for any world.

A discourse-initial state i in C is one that is an output after introducing drefs for the interlocutors's commitment states in i_0.

(22) For all members $x, x', x'', \ldots$ of $\textsc{Int}^C$:

$i_0[\phi_{DC_x}, \phi_{DC_{x'}}, \phi_{DC_{x''}}, \ldots]i$

Like before, the initial state i is an assignment where discourse referents for the commitment states of the interlocutors are available, but none of them contain any information about their commitments, and possible values for these drefs include any non-empty, well defined subset of D_w.

Interpretation of Individual Anaphora

With these components in place, we illustrate the interpretation of individual anaphora in the discourse in (8-a).

(8-a) *Mary has [a car]$^{\upsilon_2}$. It$_{\upsilon_2}$ is parked outside.*

Step 1: Antecedent Sentence The semantic representation of the antecedent sentence is given in (23).

(23) S: *Mary has a car.* $\rightsquigarrow$

$$\boxed{\begin{array}{c}\phi_1\\\hline \phi_{DC_S} \Subset \phi_1\end{array}}\,;\mathbf{max}_{\phi_1}\left(\boxed{\begin{array}{c}\upsilon_1, \phi_1 : \upsilon_2\\\hline \upsilon_1 \equiv \mathrm{Mary}_{\mathbf{e}}\\ \mathrm{Car}_{\phi_1}\{\upsilon_2\}\\ \mathrm{Have}_{\phi_1}\{\upsilon_1, \upsilon_2\}\end{array}}\right)$$

υ_2 is introduced relative to ϕ_1, the set of worlds where *Mary has a car*. Its referent exists in all (and only the) worlds in ϕ_1.

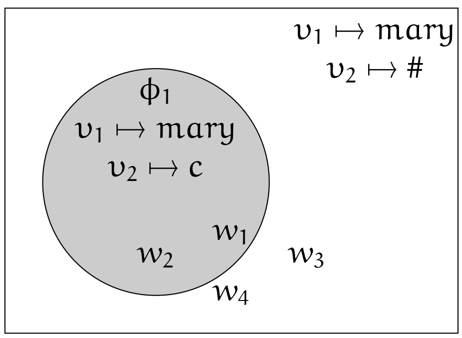
To illustrate the effect of the update, consider a context $C_1 = \langle M_1, \{S\}, i\rangle$, where i is a discourse-initial state, and M_1, is s.t. $D_e = \{mary, c, S, \#\}$, $D_w = \{w_1, w_2, w_3, w_4\}$, and $[\![car]\!]^{M_1}$, $[\![have]\!]^{M_1}$, $[\![parked]\!]^{M_1}$, $[\![outside]\!]^{M_1}$ characterize the propositions shown in **Table 7.6**.

In M_1, c is a car in all worlds, $mary$ has c in w_1, w_2. c is parked and outside in w_1 and w_3.

	$\lambda w.\mathtt{car}(c)(w)$	$\lambda w.\mathtt{have}(\mathtt{mary}, c)(w)$	$\lambda w.\mathtt{parked}(c)(w)$	$\lambda w.\mathtt{outside}(c)(w)$
w_1	1	1	1	1
w_2	1	1	0	0
w_3	1	0	1	1
w_4	1	0	0	0

Table 7.6: The propositions $\lambda w.\mathtt{car}(c)(w)$, $\lambda w.\mathtt{have}(m, c)(w)$, $\lambda w.\mathtt{parked}(c)(w)$, $\lambda w.\mathtt{outside}(c)(w)$ in M_1.

Updating with (23) in C_1 results in an output assignment j that assigns static referents to drefs as in **Figure 7.1**.



$$\begin{aligned} \phi_{DC_S}(j) &= \{w_1, w_2\} \\ \phi_1(j) &= \{w_1, w_2\} \\ \upsilon_1(j) &= \lambda w.\mathtt{mary}_e \\ \upsilon_2(j) &= \begin{bmatrix} w_1 \mapsto c \\ w_2 \mapsto c \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix} \end{aligned}$$

Figure 7.1: Drefs after updating with *'Mary has a car'* in C_1

Figure 7.1 illustrates that υ_2 refers to c in all ϕ_1-worlds, i.e. the worlds corresponding to S's assertion. υ_2 maps to # in all other worlds.

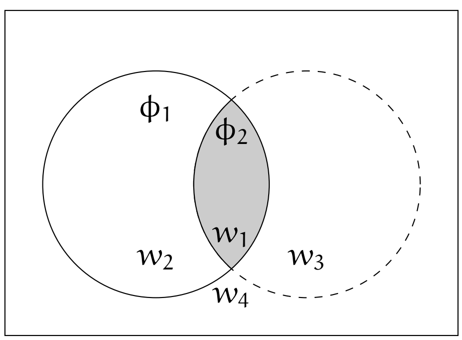
Step 2: Sentence Containing the Anaphor. The representation of the antecedent sentence is given in (24).

(24) S: *It is parked outside.* $\rightsquigarrow$ $\begin{array}{|c|}\hline \phi_2 \\ \hline \phi_{DC_S} \Subset \phi_2 \\ \hline \end{array}$; $\mathbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \mathtt{Parked}_{\phi_2}\{\upsilon_2\} \\ \mathtt{Outside}_{\phi_2}\{\upsilon_2\} \\ \hline \end{array}\right)$

Here, υ_2 is picked up anaphorically in the context of ϕ_2, the set of worlds where *It is parked outside*. Because the interpretation of DRS-condition is based on encapsulated quantification over the possible worlds in the intensional context, an update with (24) requires of its output that a referent of υ_2 exists in all worlds in ϕ_2.

Given this local existence-requirement, and the fact that the antecedent referent exists in all and only the worlds of its own local intensional context ϕ_1, ϕ_2 can only contain worlds that are also in ϕ_1. The consequence of this is that an anaphoric relation is possible, only if the local context of the anaphor is a subset of the local context of the pronoun.

Consider C_1', a context that is the result of updating with the antecedent in C_1, i.e. $C_1' = \langle M_1, \{S\}, j \rangle$, where $[\![(23)(i^{C_1})(j)]\!]^{M_1} = 1$. An update with (24) in C_1' results in an output k, which assigns drefs as shown in **Figure 7.2**.



$$
\begin{aligned}
\phi_{DC_S}(k) &= \{w_1\} \\
\phi_1(k) &= \{w_1, w_2\} \\
\phi_2(k) &= \{w_1\} \\
\upsilon_1(j) &= \lambda w.\mathtt{mary}_e \\
\upsilon_2(j) &= \begin{bmatrix} w_1 \mapsto c \\ w_2 \mapsto c \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix}
\end{aligned}
$$

Figure 7.2: Drefs after updating with *'Mary has a car'*; *'It is parked outside'* in C_1

The content of the asserted proposition $\phi_2(k)$ can semi-formally be characterized as $\lambda w.\upsilon_2(k)(w)$ is parked outside in w. For w_1, this is true, as $[\![\upsilon_2(k)(w_1)]\!]^{M_1} = c$, and $[\![\mathtt{parked}(c)(w_1) \wedge \mathtt{outside}(c)(w_1)]\!]^{M_1} = 1$, so $w_1 \in \phi_2(k)$. For w_2, this is false, as $[\![\upsilon_2(k)(w_2)]\!]^{M_1} = c$, and $[\![\mathtt{parked}(c)(w_2) \wedge \mathtt{outside}(c)(w_2)]\!]^{M_1} = 0$, so $w_2 \notin \phi_2(k)$. For w_3 and w_4, this is undefined (#), because $[\![\upsilon_2(k)(w_3)]\!]^{M_1} = \#$ and

$[\![\upsilon_2(k)(w_4)]\!]^{M_1} = \#$.

The mechanisms borrowed from Stone (1999), i.e. encapsulating quantification over possible worlds in the definition of DRS-conditions, and relativizing individual drefs to possible worlds, allows for anaphoric accessibility to fall out naturally from the requirement of a consistent discourse. An update containing an anaphor (like (24)) will be successful, only if there is an output under which the local context of the anaphor is a subset of the local context of the antecedent.

7.4 Analysis: Individual Anaphora and Negation

This section illustrates how the analysis accounts for the unavailablity of anti-veridical antecedents for veridical anaphora (2-b), modal subordination with negative antecedents (3-a), bathroom-configurations under disjunction (5-a) and inter-speaker disagreement (5-b), repeated here.

(2-b) *Mary doesn't have [a car]$^{\upsilon_2}$. #It$_{\upsilon_2}$ is parked outside.*

(3-a) Modal subordination:

There isn't [a bathroom]$^{\upsilon_1}$ in this house. It$_{\upsilon_1}$ would be easier to find.

(5-a) Disjunction:

Either there isn't [a bathroom]$^{\upsilon_1}$ in this house, or it$_{\upsilon_1}$ is in a weird place.

(5-b)Disagreement:

A: *There isn't [a bathroom]*$^{\upsilon_2}$ *in this house.*

B: *(What are you talking about?) It*$_{\upsilon_2}$ *is just in a weird place.*

7.4.1 The Counterfactual Bathroom

First, let us see how the veridical use of a pronoun with a counterfactual antecedent where both discourse segments are uttered by the same speaker is ruled out. The representations for an utterance by speaker S of the antecedent sentence, and anaphoric sentence are given in (17) and (25), respectively.

(17) S: *There isn't [a bathroom]*$^{\upsilon_1}$. $\rightsquigarrow$

$$\boxed{\begin{array}{c}\phi_1\\ \hline \phi_{DC_S} \Subset \phi_1\end{array}}\,;\mathbf{max}_{\phi_1}\left(\boxed{\begin{array}{c}\phi_2\\ \hline \phi_1 \equiv \overline{\overline{\phi_2}}\end{array}}\,;\mathbf{max}_{\phi_2}\left(\boxed{\begin{array}{c}\phi_2 : \upsilon_1\\ \hline \mathrm{Bathroom}_{\phi_2}\{\upsilon_1\}\end{array}}\right)\right)$$

(25) # S: *It*$_{\upsilon_1}$ *is in a weird place.* $\rightsquigarrow$ $\boxed{\begin{array}{c}\phi_3\\ \hline \phi_{DC_S} \Subset \phi_3\end{array}}\,;\mathbf{max}_{\phi_3}\left(\boxed{\begin{array}{c}\phi_3 : \upsilon_2\\ \hline \mathrm{Place}_{\phi_3}\{\upsilon_2\}\\ \mathrm{Weird}_{\phi_3}\{\upsilon_2\}\\ \mathrm{In}_{\phi_3}\{\upsilon_1, \upsilon_2\}\end{array}}\right)$

Consider a context $C_2 = \langle M_2, \{S\}, i\rangle$, where i is a discourse-initial state, and M_2 is s.t. $D_e = \{b, p, S, \#\}.D_w = \{w_1, w_2, w_3, w_4\}$, and the interpretation of the predicates $\mathrm{bathroom}$, place, weird, and in characterizes the propositions in **Table 7.7**.

	$\lambda w.\mathtt{bathroom}(b)(w)$	$\lambda w.\mathtt{place}(p)(w)$	$\lambda w.\mathtt{weird}(p)(w)$	$\lambda w.\mathtt{in}(b,p)(w)$
w_1	1	1	1	1
w_2	1	0	0	0
w_3	0	1	1	1
w_4	0	0	0	0

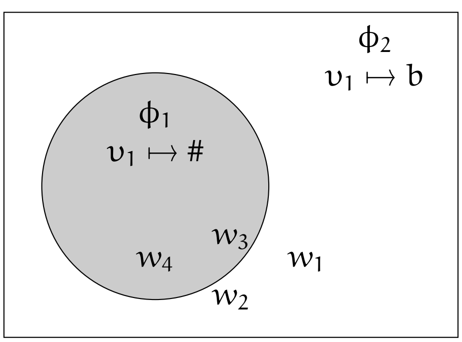
Table 7.7: Truth-values of $[\![\lambda w.\mathtt{bathroom}(b)(w)]\!]^{M_2}$, $[\![\lambda w.\mathtt{place}(p)(w)]\!]^{M_2}$, $[\![\lambda w.\mathtt{weird}(p)(w)]\!]^{M_2}$, $[\![\lambda w.\mathtt{in}(b,p)(w)]\!]^{M_2}$ for the worlds in D_w.

Step 1: Update with the Antecedent Segment

(17) S: *There isn't [a bathroom]$^{\upsilon_1}$.* $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_1 \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \hline \end{array}\,;\mathbf{max}_{\phi_1}\left(\begin{array}{|c|}\hline \phi_2 \\ \hline \phi_1 \equiv \overline{\overline{\phi_2}} \\ \hline \end{array}\,;\mathbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \phi_2 : \upsilon_1 \\ \hline \mathtt{Bathroom}_{\phi_2}\{\upsilon_1\} \\ \hline \end{array}\right)\right)$$

An update with (17) introduces three drefs: A proposition ϕ_1 (for the worlds where there is no bathroom), restricting the set of discourse commitments held by S, a proposition ϕ_2 (for the worlds where there is a bathroom), the complement of ϕ_1, and an individual υ_1, (for a bathroom in ϕ_2). In C_2, the update results in an output assigment j, which maps drefs as illustrated in **Figure 7.3**.



$$\begin{aligned}
\phi_{DC_S}(j) &= \{w_3, w_4\} \\
\phi_1(j) &= \{w_3, w_4\} \\
\phi_2(j) &= \{w_1, w_2\} \\
\upsilon_1(j) &= \begin{bmatrix} w_1 \mapsto b \\ w_2 \mapsto b \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix}
\end{aligned}$$

Figure 7.3: Drefs after updating with *'There isn't [a bathroom]$^{\upsilon_1}$'* in C_2

The resulting assignment j maps υ to b for each world in $\phi_2(j)$, and maps it to # for each world not in $\phi_2(j)$. $\phi_2(j)$ is the maximal set of worlds in which υ_1 is a bathroom

($\mathtt{Bathroom}_{\phi_2}\{\upsilon_1\}$). The speaker is committed to ϕ_1, the opposite of ϕ_2. Since there are no ϕ_1-worlds in ϕ_2, $\upsilon_1(j)$ maps to # for every world in ϕ_1, and υ_1 is an anti-veridical dref.

Step 2: Update with the Segment Containing the Anaphor

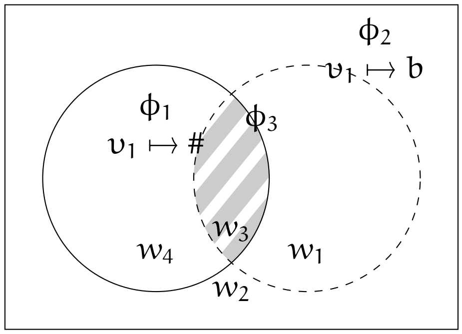
(25) #S: *It*$_{\upsilon_1}$ *is in a weird place.* $\rightsquigarrow \boxed{\begin{array}{c}\phi_3\\ \hline \phi_{DC_S} \Subset \phi_3\end{array}}; \mathbf{max}_{\phi_3}\left(\boxed{\begin{array}{c}\phi_3 : \upsilon_2\\ \hline \mathtt{Place}_{\phi_3}\{\upsilon_2\}\\ \mathtt{Weird}_{\phi_3}\{\upsilon_2\}\\ \mathtt{In}_{\phi_3}\{\upsilon_1, \upsilon_2\}\end{array}}\right)$

An update with (25) in $C_2' = \langle M_2, \{S\}, j\rangle$ (where $[\![(17)(i_{C_2})(j)]\!]^{M_2} = 1$) fails due to the interpretation of the condition $\mathtt{In}_{\phi_3}\{\upsilon_1, \upsilon_2\}$. The anaphor υ_1 cannot be interpreted in its local context ϕ_3. To see why, we consider an update that is like (25), minus the problematic condition:

(26) $\boxed{\begin{array}{c}\phi_3\\ \hline \phi_{DC_S} \Subset \phi_3\end{array}}; \mathbf{max}_{\phi_3}\left(\boxed{\begin{array}{c}\phi_3 : \upsilon_2\\ \hline \mathtt{Place}_{\phi_3}\{\upsilon_2\}\\ \mathtt{Weird}_{\phi_3}\{\upsilon_2\}\end{array}}\right)$

An update with (26) in C_2' yields an output k, which is illustrated in **Figure 7.4**.

Because S asserts ϕ_1 and ϕ_3, their commitments should be a non-empty subset of both. Discourse consistency requires that ϕ_1 and ϕ_3 overlap. For the assertion of our antecedent sentence and subsequently our impoverished update (26), S's commitments would be characterized by the worlds contained in the gray hatched area. A further update including a the condition $\mathtt{In}_{\phi_3}\{\upsilon_1, \upsilon_2\}$ imposes a conflicting requirement. The anaphor υ_1 is interpreted in the intensional context of ϕ_3 and has previously been intro-



$$
\begin{aligned}
\phi_{DC_S}(k) &= \{w_3\} \\
\phi_1(k) &= \{w_3, w_4\} \\
\phi_2(k) &= \{w_1, w_2\} \\
\phi_3(k) &= \{w_1, w_3\} \\
\upsilon_1(k) &= \begin{bmatrix} w_1 \mapsto b \\ w_2 \mapsto b \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix} \\
\upsilon_2(k) &= \begin{bmatrix} w_1 \mapsto p \\ w_2 \mapsto \# \\ w_3 \mapsto p \\ w_4 \mapsto \# \end{bmatrix}
\end{aligned}
$$

Figure 7.4: Drefs after updating with *'There isn't [a bathroom]*$^{\upsilon_1}$*'*;*'It*$_{\upsilon_1}$ *is in a weird place'* in C_2

duced relative to ϕ_2. Therefore, the existence-requirement demands that ϕ_3 is a subset of ϕ_2 (i.e. it contains only worlds where $\upsilon_1 \neq \#$).

There is no way that both discourse-consistency and the existence-requirement are satisfied at the same time. If consistency is satisfied, ϕ_1 and ϕ_3 overlap, and there are worlds in ϕ_3 in which υ_1 does not exist, so the interpretation fails. If the existence-requirement is satisfied, ϕ_3 is a subset of ϕ_2, and the intersection of ϕ_1 and ϕ_3 is empty. The speaker would be committed to a contradiction. Either way, there is no successful update with (17);(25).

7.4.2 The Hypothetical Bathroom

Now, let us see how an update with the same first utterance (17) can be followed by an utterance with a non-veridical pronoun (27).

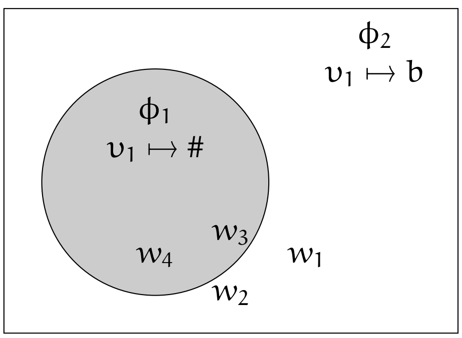
(17) S: *There isn't [a bathroom]*$^{\upsilon_1}$. $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_1 \\ \hline \phi_{DC_S} \Subset \phi_1 \\ \hline\end{array} ; \mathbf{max}_{\phi_1}\left(\begin{array}{|c|}\hline \phi_2 \\ \hline \phi_1 \equiv \overline{\overline{\phi_2}} \\ \hline\end{array} ; \mathbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \phi_2 : \upsilon_1 \\ \hline \mathrm{Bathroom}_{\phi_2}\{\upsilon_1\} \\ \hline\end{array}\right)\right)$$

(27) S: *It would be in a weird place.*

$$\rightsquigarrow \begin{array}{|c|}\hline \phi_3 \\ \hline \phi_{DC_S} \Subset \phi_3 \\ \hline\end{array} ; \mathbf{max}_{\phi_3}\left(\boxed{\mathrm{would}_{\phi_3}\{\phi_4\}} ; \mathbf{max}_{\phi_4}\left(\begin{array}{|c|}\hline \phi_4 : \upsilon_2 \\ \hline \mathrm{Place}_{\phi_4}\{\upsilon_2\} \\ \mathrm{Weird}_{\phi_4}\{\upsilon_2\} \\ \mathrm{In}_{\phi_4}\{\upsilon_1, \upsilon_2\} \\ \hline\end{array}\right)\right)$$

First, we assume the same update as for the first utterance of the previous example, with context C_2 and output j, again illustrated here in **Table 7.5**.



$$\begin{aligned} \phi_{DC_S}(j) &= \{w_3, w_4\} \\ \phi_1(j) &= \{w_3, w_4\} \\ \phi_2(j) &= \{w_1, w_2\} \\ \upsilon_1(j) &= \begin{bmatrix} w_1 \mapsto b \\ w_2 \mapsto b \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix} \end{aligned}$$

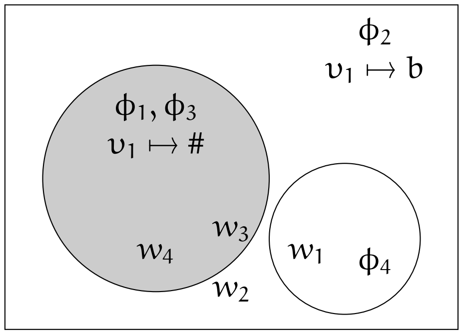
Figure 7.5: Drefs after updating with *'There isn't [a bathroom]*$^{\upsilon_1}$*'* in C_2

Again, υ_1, a bathroom, exists in all and only the counterfactual ϕ_2-worlds. Now, consider what happens when we update with (27). The output is illustrated in **Figure 7.6**.

(27) S: *It would be in a weird place.*

$$\rightsquigarrow \begin{array}{|c|}\hline \phi_3 \\ \hline \phi_{DC_S} \Subset \phi_3 \\ \hline\end{array} ; \mathbf{max}_{\phi_3}\left(\boxed{\mathrm{would}_{\phi_3}\{\phi_4\}} ; \mathbf{max}_{\phi_4}\left(\begin{array}{|c|}\hline \phi_4 : \upsilon_2 \\ \hline \mathrm{Place}_{\phi_4}\{\upsilon_2\} \\ \mathrm{Weird}_{\phi_4}\{\upsilon_2\} \\ \mathrm{In}_{\phi_4}\{\upsilon_1, \upsilon_2\} \\ \hline\end{array}\right)\right)$$

Because the anaphor υ_1 is interpreted in the condition $\mathrm{In}_{\phi_4}\{\upsilon_1, \upsilon_2\}$, it will still be required to have an existent referent in all of the ϕ_4-worlds. This set of worlds, however,



$$
\begin{aligned}
\phi_{DC_S}(k) &= \{w_3, w_4\}\\
\phi_1(k) &= \{w_3, w_4\}\\
\phi_2(k) &= \{w_1, w_2\}\\
\upsilon_1(k) &= \begin{bmatrix} w_1 \mapsto b\\ w_2 \mapsto b\\ w_3 \mapsto \#\\ w_4 \mapsto \# \end{bmatrix}\\
\phi_3(k) &= \{w_3, w_4\}\\
\phi_4(k) &= \{w_1\}\\
\upsilon_2(k) &= \begin{bmatrix} w_1 \mapsto p\\ w_2 \mapsto \#\\ w_3 \mapsto \#\\ w_4 \mapsto \# \end{bmatrix}
\end{aligned}
$$

Figure 7.6: Drefs after updating with *'There isn't [a bathroom]$^{\upsilon_1}$'*; *'It$_{\upsilon_1}$ would be in a weird place'* in C_2

is the local context corresponding to the prejacent of the non-veridical embeddor *would*. Crucially, this is constrained to be a set of worlds that is not entailed by ϕ_3.[1]

Because ϕ_3 is in turn entailed by the discourse commitments of the speaker, the conditions in ((27)) state that ϕ_4 is not entailed by the speaker's commitments, i.e. a non-veridical proposition. As a consequence, the antecedent for *it*$_{\upsilon_1}$ can be provided by an anti-veridically introduced dref.

7.4.3 The Optional Bathroom

In contrast to propositions expressed by two assertions of the same speaker, the propositions expressed by the disjuncts of a disjunction are not required to be compatible with each other.

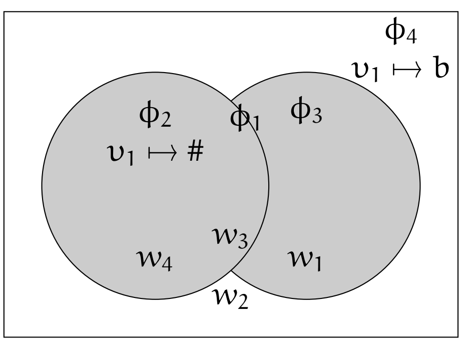
[1]This is glossing over the specifics of modal semantics of *would* specifying the relationship between ϕ_3 and ϕ_4. See Stone (1999); Stone and Hardt (1999) for some more detailed discussion.

(28) S: *There is no bathroom or it is in a weird place.* $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_1 \\ \hline \phi_{\mathrm{DC}_S} \Subset \phi_1 \\ \hline\end{array}; \mathbf{max}_{\phi_1}\left(\begin{array}{|c|}\hline \phi_2, \phi_3 \\ \hline \phi_1 \equiv \phi_2 \Cup \phi_3 \\ \hline\end{array}; \begin{array}{l} \mathbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \phi_4 \\ \hline \phi_2 \equiv \overline{\overline{\phi_4}} \\ \hline\end{array}; \mathbf{max}_{\phi_4}\left(\begin{array}{|c|}\hline \phi_4 : \upsilon_1 \\ \hline \mathtt{Bathroom}_{\phi_4}\{\upsilon_1\} \\ \hline\end{array}\right)\right); \\ \mathbf{max}_{\phi_3}\left(\begin{array}{|c|}\hline \phi_3 : \upsilon_2 \\ \hline \mathtt{Place}_{\phi_3}\{\upsilon_2\} \\ \mathtt{Weird}_{\phi_3}\{\upsilon_2\} \\ \mathtt{In}_{\phi_3}\{\underline{\upsilon_3}, \upsilon_2\} \\ \hline\end{array}\right) \end{array}\right)$$

The two sequential assertions in the previous example were interpreted in conjunction, leading to a requirement that the intersection between the propositions expressed by the two clauses couldn't empty. The same doesn't apply under disjunction.

An update with (28) in C_2 yields an output j, which is illustrated in **Figure 7.7**.



$$\begin{array}{ll}
\phi_{\mathrm{DC}_S}(j) & = \{w_1, w_3, w_4\} \\
\phi_1(j) & = \{w_1, w_3, w_4\} \\
\phi_2(j) & = \{w_2, w_4\} \\
\phi_3(j) & = \{w_1\} \\
\phi_4(j) & = \{w_1, w_2\} \\
\upsilon_1(j) & = \begin{bmatrix} w_1 \mapsto b \\ w_2 \mapsto b \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix} \\
\upsilon_2(j) & = \begin{bmatrix} w_1 \mapsto p \\ w_2 \mapsto \# \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix}
\end{array}$$

Figure 7.7: Drefs after updating with *'(Either) there is no bathroom or it is in a weird place'* in C_2

$\phi_4(j)$ is the maximal set of worlds in which there is a bathroom. The referent of $\upsilon_1(k)$ (the bathroom) exists in all and only the ϕ_4-worlds. $\phi_2(k)$ is the opposite of $\phi_4(k)$, and the propositions expressed by the first disjunct. The referent of $\upsilon_1(k)$ (the bathroom) does not exist in any of the worlds in $\phi_2(k)$. $\phi_3(k)$, the proposition expressed by the second disjunct, is the maximal set of worlds where $\upsilon_1(k)$ is in a weird place. $\phi_1(k)$ is the union of ϕ_2 and ϕ_3, and the proposition expressed by the assertion.

Because the anaphor υ_1 is interpreted in the intensional context of ϕ_3, $\upsilon_1(k)$ exists in every ϕ_3-world, and $\phi_3(k)$ is a subset of $\phi_4(k)$. Because of that, $\phi_3(k)$ does not overlap with $\phi_2(k)$, where υ_1 does not exist. However, the relation expressed by disjunction, i.e. set union, does not require the propositions expressed by the two disjuncts to overlap, and the update with (28) is successful.

7.4.4 The Contested Bathroom

The case of inter-speaker disagreement is similar to the case of disjunction because the discourse segment containing the antecedent and the one containing the anaphor are not required to be compatible with each other.

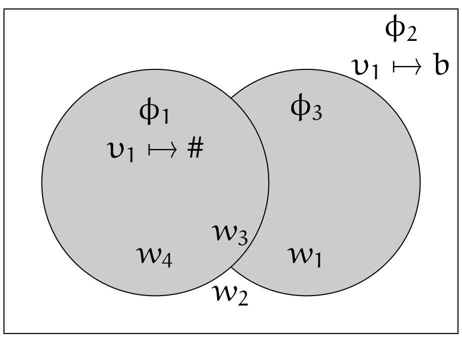
(17) A: *There isn't [a bathroom]*$^{\upsilon_1}$. $\rightsquigarrow$

$$\begin{array}{|c|}\hline \phi_1 \\ \hline \phi_{DC_A} \Subset \phi_1 \\ \hline \end{array} ; \mathbf{max}_{\phi_1}\left(\begin{array}{|c|}\hline \phi_2 \\ \hline \phi_1 \equiv \overline{\overline{\phi_2}} \\ \hline \end{array} ; \mathbf{max}_{\phi_2}\left(\begin{array}{|c|}\hline \phi_2 : \upsilon_1 \\ \hline \mathsf{Bathroom}_{\phi_2}\{\upsilon_1\} \\ \hline \end{array}\right)\right)$$

(29) B: *It*$_{\upsilon_1}$ *is in a weird place.* $\rightsquigarrow \begin{array}{|c|}\hline \phi_3 \\ \hline \phi_{DC_B} \Subset \phi_3 \\ \hline \end{array}; \mathbf{max}_{\phi_3}\left(\begin{array}{|c|}\hline \phi_3 : \upsilon_2 \\ \hline \mathsf{Place}_{\phi_3}\{\upsilon_2\} \\ \mathsf{Weird}_{\phi_3}\{\upsilon_2\} \\ \mathsf{In}_{\phi_3}\{\upsilon_1, \upsilon_2\} \\ \hline \end{array}\right)$

Consider a context $C_4 = \langle M_4, \textsc{Int}, i\rangle$. This context is just like C_2, except that $\textsc{Int} = \{A, B\}$, and $D_e = \{b, p, A, B, \#\}$ in M_4.

An update with (17); (29) in C_4 yields an output j, which is illustrated in **Figure 7.8**.



$$
\begin{array}{ll}
\phi_{DC_A}(j) & = \{w_3, w_4\} \\
\phi_{DC_B}(j) & = \{w_1\} \\
\phi_1(j) & = \{w_3, w_4\} \\
\phi_2(j) & = \{w_1, w_2\} \\
\phi_3(j) & = \{w_1\} \\
\upsilon_1(j) & = \begin{bmatrix} w_1 \mapsto b \\ w_2 \mapsto b \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix} \\
\upsilon_2(j) & = \begin{bmatrix} w_1 \mapsto p \\ w_2 \mapsto \# \\ w_3 \mapsto \# \\ w_4 \mapsto \# \end{bmatrix}
\end{array}
$$

Figure 7.8: Drefs after updating with A's utterance of *'There isn't [a bathroom]*$^{\upsilon_1}$*'*; then B's utterance of *'It*$_{\upsilon_1}$ *is (just) in a weird place'* in C_4

$\phi_2(j)$ is the maximal proposition where there is a bathroom and υ_1 is introduced relative to ϕ_2. A is committed to ϕ_1, the opposite of ϕ_2, and therefore A's commitment state contains only worlds in which υ_1 has no referent, i.e. it is an anti-veridical dref. B, however, is not bound by A's commitments and asserts ϕ_3. $\phi_3(j)$ is the maximal proposition where υ_1 is in a weird place. The anaphor *it*$_{\upsilon_1}$ is interpreted in the intensional context of ϕ_3. Therefore, $\phi_3(j)$ is a subset of $\phi_2(j)$, the set of worlds in which the referent of υ_1 exists. Since B is not committed to ϕ_2 being false, the output j characterizes a consistent discourse. Althouhg A and B are in disagreement, their respective commitments are non-contradictors.

7.5 Comparison to Other Approaches

Quantificational encapsulation over sets of worlds with relativized drefs allows us to have a dynamic system, but also allows us to specify relations between drefs for static propositions. Here, sentential operators specify relations over propositional drefs, which satisfy the inferential rules familiar from static propositional logic.

Interpreting individual drefs in relation to these static propositions allows for these inferences to influence the availability of drefs in certain contexts. Double negation elimination and other truth-conditional inferences will affect the veridicality of drefs, as well as the accessibility of potential antecedents.

This section compares the account presented here to alternative approaches in some more detail, expanding on the discussion in **Chapter 2**. The discussion there involved a broader-level comparison of three classes of accounts **(i)** flat-update dynamic systems with intensional drefs, like the one presented here, as well as Stone (1999); Stone and Hardt (1999); Brasoveanu (2006, 2010a,b); **(ii)** conditional variable update (Krahmer and Muskens, 1995; Gotham, 2019; Mandelkern, 2022), and **(iii)** accommodation-based accounts (e.g Roberts, 1989; Frank and Kamp, 1997; Geurts, 1999; Nouwen, 2003; Lewis, 2021).

The main point about approaches with conditional variable update there were that they are limited by a lookahead-problem, because they give rise to cases where the embedding context of the antecedent fully determines its subsequent (non-)accessibility, precluding any influence of the context of the anaphor.

Concerning accommodation-based approaches, I argued that they also need to make reference to veridicality and speaker commitments on some level to capture the relevant generalizations. Therefore, this class of possible approaches is not necessarily empirically distinguishable from intensional flat-update accounts based on judgment data. In the absence of robust behavioral or neurophysiological evidence for one or the other kind of analysis as a model of mental representations of linguistic meaning, it does not seem possible to present a convincing empirical argument. Accordingly, we made our decision for a flat-update system based on a conceptual, theoretical argument. I argued that the presented account is more explanatory, because it reduces anaphoric accessibility to more fundamental principles of discourse interpretation.

While the discussion in **Chapter 2** was based on a very broad classification of possible approaches, the present section provides a more nuanced discussion of specific accounts in some detail, which shows more specifically how they compare to the presented account.

7.5.1 Flat Update and Intensional Drefs

As noted in the above deliberation about how we might intensionalize individual drefs (**Section 7.3**), other dynamic systems would be suitable to formalize the analysis developed here if they have the following features: **(i)** The context can store information about the relationship between dref mappings and their local intensional context, as well as **(ii)** the relationship between local intensional contexts and the speaker's commitments, and **(iii)** this relationship may be non-overlapping (i.e., it allows for anti-veridical propositional drefs).

Various intensional dynamic systems could be devised, which satisfy these properties, and conceive of drefs as mappings from discourse variables to entities in our model domain relative to each of the possible worlds in the local context (i.e., distributively over worlds). In our current system, this works because of relativized individual drefs and encapsulated universal quantification over the elements of the local context as part of individual variable update and the interpretation of conditions.

This is illustrated in **Table 7.8**, which shows the output for an update with an utterance of *Mary*$_{\upsilon_1}$ *doesn't have [a car]*$_{\upsilon_2}$ in a toy context where $D_w = \{w_1, w_2, w_3, w_4\}$, the assertion is true in w_1, w_2, and false in w_3, w_4 (cf. above **Section 7.4.1**).

ϕ_{DC_S}	ϕ_1	ϕ_2	υ_1	υ_2
$\{w_1, w_2\}$	$\{w_1, w_2\}$	$\{w_3, w_4\}$	$\lambda w.mary$	$\{w_1 \mapsto \#,$ $w_2, \mapsto \#,$ $w_3, \mapsto b,$ $w_4, \mapsto b\}$

Table 7.8: Output for update with 'Mary doesn't have a car'in ICDRT

Table 7.8 illustrates the mechanism of relativizing individual drefs to their intensional context, which works by understanding them as functions from possible worlds to entities that may be interpreted in relation to a propositional dref.

Brasoveanu (2006): Plural ICDRT

The framework of Plural intensional CDRT (Brasoveanu, 2006, 2010a,b) was developed as an analysis of different cases of discourse subordination, such as quantificational subordination and modal subordination. The analysis of modal subordination also allows for individual drefs to be interpreted intensionally, in relation to the possible worlds of the local context. The result of the same update as above, with *Mary doesn't have a car*, in this system is shown in **Table 7.9**.

I	ϕ_u	ϕ_{DC_S}	ϕ_1	ϕ_2	υ_1	υ_2
i_1	w_1	w_1	w_1	#	$mary$	#
i_2	w_2	w_2	w_2	#	$mary$	#
i_3	w_3	w_3	#	w_3	$mary$	b
i_4	w_4	#	#	w_4	$mary$	b
i_5	#	#	#	#	#	#

Table 7.9: Output for update with 'Mary doesn't have a car'in IPCDRT

In this system, propositional drefs are drefs for possible worlds (instead of sets of worlds), but due to the plural nature of the system, a dref for possible worlds may store a set. Each member of the set of variable assignments that constitutes the discourse context stores a possible world that is a member of the proposition or a dummy value #. E.g. in **Table 7.9**, ϕ_1 is the set of all worlds that some assignment in the context assigns to ϕ_1.

Brasoveanu's (2006) system treats propositional operators as quantifiers over possible worlds. The proposed semantics for quantification involves a notion of structured inclusion. Roughly speaking, this makes sure that in **Table 7.9**, each row corresponds to a possible world. Since ϕ_1 is a (structured) subset of ϕ_{DC_S}, each world that some assignment in the context assigns to ϕ_1 is also a world that some assignment assigns to ϕ_{DC_S} (subset), but also no row contains different values for the two unless ϕ_1 is mapped to # (structured subset).

The mechanism of structured inclusion achieves that each row corresponds to a possible world. That allows for a unification of a DRT-style dynamic semantics with e-type approaches of anaphora because each element of the context set can also be conceived of as a partial situation.

One feature added here in relation to Brasoveanu's (2006) system is a dref for the universe ϕ_U, which is a structured superset of any other propositional dref. That is necessary to preserve the correspondence between rows and possible worlds when including anti-veridical propositional drefs / counterfactual propositions.

The relativization of individual drefs is achieved by introducing new drefs distributively over possible worlds. For the example above, the dref for Mary's car is introduced by a selectively distributive update which introduces υ_2. The result is that the same update is made for each non-# ϕ_2-equivalence class of the context, i.e., each subset of assignments with the same value for ϕ_2. In this case, b is the only possible value for υ_2, and so the mapping $\upsilon_2 \mapsto b$ is introduced for each row where $\phi_2 \neq \#$. The interpretation of anaphoric pronouns will also take into account the worlds of the local context, similar to the analysis outlined in this work. In IPCDRT, this is also done by invoking selective distributivity over the worlds in the local context: An update containing a discourse variable is interpreted in relation to each world of the local context. As a result, the local context may only contain worlds s.t. an assignment can be found in the context set, for which the variable can be mapped to a suitable referent in that world.

Without giving all of the formal details here, the update in **Table 7.9** can be derived in IPCDRT by extending the system in (Brasoveanu, 2006, 2010a,b) with the following features:

- A representation of the universe in order to include anti-veridical worlds
- A representation of speaker commitments in order to distinguish between (non-/anti-)veridical propositions
- An externally dynamic semantics of negation which introduces a local context, similarly to the modal operators in Brasoveanu (2006, 2010a) (see also the externally dynamic semantics for negation in Brasoveanu and Farkas (2011))

Since this kind of analysis has a lot of the same properties, it would work very similarly to the kind of analysis proposed here. With the added features suggested here, it is my understanding that an account along these lines would be equivalent for the empirical cases discussed here. However, it is a more general kind of account, extending to other empirical cases in addtion. It also allows for an account of discourse subordination with other types (like quantificational subordination), or multiple dependencies.

7.5.2 Conditional Dref Introduction

Krahmer and Muskens (1995)

Focusing on double negation and bathroom sentences, Krahmer and Muskens (1995) address the issue of anaphora to negated indefinites by providing a semantics which satisfies the equivalence of double negation elimination on a dynamic level, in addition to the truth-conditional level.

Krahmer and Muskens's (1995) account for the disjunction and double negation cases within the framework of Double Negation DRT. The analysis involves a semantics for negation that switches from the extension of an expression to its anti-extension, and a semantics for disjunction that analogizes it to conditionals, both in terms of their truth conditions and their dynamic potential. Accordingly, the sentences in (30), respectively, are taken to be equivalent to each other both dynamically and truth-conditionally.

(30) a. Either there is no bathroom in this house, or it is in a weird place.

b. If it is not true that there is no bathroom in this house, it is in a weird place.

c. If there is a bathroom in this house, it is in a weird place.

Their analysis relies on the conventional semantics associated with negation and disjunction and therefore does not extend to cases without overt disjunction or double negation, like the modal subordination and inter-speaker disagreement cases.

This account, like others which conditionally introduce a dref for double negation utteraces while simple negative utterances are considered externally static, could not address the lookahead problem described in **Section 2.3.3**. Based on the embedding context of an indefinite, either a dref is introduced, or not. Since there is no relativization, this account could not be extended to discourse subordination.

Introducing Drefs via Projective Content

Some other accounts addressing the interaction of negation and anaphora are ones which allow indefinites to introduce a drefs by means of some sort of projective content. This content is often disjunctive (Gotham, 2019) or conditional (Mandelkern, 2022).

Its introduction associated with the contribution of negation (Gotham, 2019) or the indefinite (Mandelkern, 2022). All of the accounts noted here are presented in works which empirically focus on accounting for double negation and bathroom disjunctions. In my understanding, these accounts work well for these cases, which they were designed for. Due to their conditional or disjunctive nature, these projective variable updates might also be extended to account for modal subordination. However, I believe that they do not work for modal subordination with negative antecedents, or any other cases in which an individual dref that was introduced anti-veridically is picked up later. The following outlines the reasoning for this in some detail for the two accounts.

Gotham (2019) accounts for the possibility of anaphora to indefinites under double negation and in bathroom sentences by assuming that the semantics of negation is associated with an excluded middle inference, which is treated as a kind of projective content associated with negation. While negation is still treated as an externally static operator, material in the scope of negation may introduce new drefs via the excluded middle inference.

Gotham's account is provided in Dynamic Predicate Logic (Groenendijk and Stokhof (1991)), which has the same syntax as first-order predicate logic, and is interpreted semantically as shown in (31).

(31) DPL semantics (from Gotham (2019): Fig. 1)

a. $[\![P(t_1, \ldots, t_n)]\!]^{M,f} = \{g \mid f = g \ \& \ \langle [\![t_1]\!]^{M,g}, \ldots, [\![t_n]\!]^{M,g} \rangle \in \mathcal{I}^M(P)\}$

b. $[\![t_1, = t_2]\!]^{M,f} = \{g \mid f = g \ \& \ [\![t_1]\!]^{M,g} = [\![t_2]\!]^{M,g}\}$

c. $[\![\neg\phi]\!]^{M,f} = \{g \mid f = g \ \& \ [\![\phi]\!]^{M,g} = \varnothing\}$

d. $[\![\phi \wedge \psi]\!]^{M,f} = \{h \mid \exists g.g \in [\![\phi]\!]^{M,g} \& \ h \in [\![\psi]\!]^{M,g}\}$

e. $[\![\phi \vee \psi]\!]^{M,f} = \{g \mid f = g \ \& \ [\![\phi]\!]^{M,g} \cup [\![\psi]\!]^{M,g} \neq \varnothing\}$

f. $[\![\phi \rightarrow \psi]\!]^{M,f} = \{g \mid f = g \ \& \ [\![\phi]\!]^{M,g} \subseteq \{h \mid [\![\psi]\!]^{M,h} \neq \varnothing\}\}$

g. $[\![\exists x\phi]\!]^{M,f} = \{h \mid \exists g.f[x]g \ \& \ h \in [\![\phi]\!]^{M,g}\}$

h. $[\![\forall x\phi]\!]^{M,f} = \{g \mid f = g \ \& \ \{h \mid g[x]h\} \subseteq \{h \mid [\![\phi]\!]^{M,h} \neq \varnothing\}\}$

The excluded middle inference makes use of program disjunction (Groenendijk and Stokhof (1991): 88), a binary operator $\cup$ with the following semantics:

(32) $[\![\phi \cup \psi]\!]^{M,f} = [\![\phi]\!]^{M,f} \cup [\![\psi]\!]^{M,f}$

An example of anaphora to indefinites under double negation (33) is translated into DPL as in (34).

(33) a. It's not true that Mary doesn't own a car.

b. It's parked outside.

(34) a. $\neg\neg\exists x.(car(c) \wedge own(m, x))$

b. $parked(x)$

Gotham suggests that (34) is interpreted in the context of the excluded middle inference associated with the negation of $\neg\exists x.(car(c) \wedge own(m, x)$, shown in (35).

(35) $\exists x.car(c) \wedge own(m, x) \cup \neg\exists x.car(c) \wedge own(m, x)$

(35) could be paraphrased as 'Either Mary owns a car or she doesn't', and is formulated using program disjunction. An update with (35) results in result the union of assignments where either the first update is made (and a dref for Mary's car is introduced), or the second (and only assignments survive in which no such update can be made). If (35) is evaluated with an assignment f, in a model where Mary doesn't own a car, the output is $\{g \mid g = f\}$, and no new dref is introduced. In a model where Mary does own a car, $\{g \mid f[x]g \wedge g \in [\![car(x) \wedge own(m, x)]\!]^{M,g}\}$, i.e. a dref for Mary's car is introduced.

A subsequent update with the negative antecedent ($\neg\neg\exists x(car(x) \wedge own(m, x)$) still works as a test, but eliminates all assignments that cannot be updated with $\exists x(car(x) \wedge own(m, x)$, leaving only the ones where a dref for Mary's car has been introduced. In a model where Mary does not own a car, the test fails and the output will be an empty set of assignments ($\neg\neg\exists x(car(x) \wedge own(m, x)$ is false). In a model where Mary does have a car, the test succeeds, but no new drefs are introduced ($\{h \mid h = g\}$). Because of the previous update with the excluded middle, discourse contains only assignments where x maps to Mary's car. As a result, the dref x for Mary's car is available for subsequent reference. Based on the same mechanism, Gotham accounts for bathroom-anaphora under disjunction.

This account can still rule out anaphora to indefinites under single negation like (36-a), which is translated into DPL as in (36-b).

(36) a. Mary doesn't own a car.

b. $\neg\exists x(car(x) \wedge own(m, x))$

(36-b) is interpreted in the context of an update with the excluded middle, again yielding the union of all assignments where a dref for Mary's car is introduced, or where no such update can be made. A subsequent update with $\neg\exists x(car(x) \wedge own(m, x))$ works as a test, and eliminates all assignments that can be updated with $\exists x(car(x) \wedge own(m, x))$, leaving only the ones where no dref is introduced. As a result, no dref for Mary's car is made available.

This rules out any subsequent anaphora to indefinites under single negation. While this accounts for the unavailability of veridical anaphora in these cases, it also rules out modal subordination with anti-veridical antecedents, or anaphora in other non-veridical contexts.

Mandelkern's (2022)) bounded theory gives a two-dimensional semantics for natural language utterances that offers a treatment of indefinites as existential quantifiers and a treatment of connectives in the classical way while addressing their dynamic behavior.

Natural language meaning is characterized on two dimensions: One static dimension, which characterizes contents as sets of indices, where connectives are classical, indefinites are treated as existential quantifiers, and definites are interpreted as involving a conjunction of their nuclear scope and restrictor. The second dimension of meaning captures the dynamic behavior of indefinites, as well as anaphoric definites and pronouns.

In Mandelkern's bounded theory, natural language utterances are translated to a predicative language which is closed under a two-place definite operator ιx, and a two-place indefinite operator ϵx. The semantic interpretation on the static dimension characterizes the truth-conditions of a formula p, relative to an assignment g and a world w. On the dynamic dimension, the interpretation characterizes the conditions under which the bounds of p are satisfied (i.e. p is satt). This concept is similar to, but possibly distinct from, presupposition satisfaction. This is done relative to a context, an assignment and a world, where a context c is a set of assignment-world pairs.

A formula is interpreted as an update in a context c. Updating c with p, yields context where p is true and has its bounds satisfied $\{\langle g, w\rangle \in c \mid p \text{ is true and satt at } \langle c, g, w\rangle\}$.

The interpretation on the static, truth-conditional, dimension is given in (37).

(37) a. Atoms: $[\![A(x_1, \ldots, x_n)]\!]^{g,w} = 1$ iff $\langle g(x_1), \ldots, g(x_n)\rangle \in \mathcal{I}(A, w)$

b. Conjunction: $[\![p\&q]\!]^{g,w} = 1$ iff $[\![p]\!]^{g,w} = [\![q]\!]^{g,w} = 1$

c. Disjunction: $[\![p \vee q]\!]^{g,w} = 1$ iff $[\![p]\!]^{g,w} = 1$ or $[\![q]\!]^{g,w} = 1$

d. Negation: $[\![\neg p]\!]^{g,w} = 1$ iff $[\![p]\!]^{g,w} = 0$

e. Indefinites: $[\![\text{з}x(p, q)]\!]^{g,w} = 1$ iff $\exists a : [\![p\&q]\!]^{g[x\mapsto a],w} = 1$

f. Definites: $[\![\iota x(p, q)]\!]^{g,w} = [\![p\&q]\!]^{g,w}$

On the level of bounds, indefinites are interpreted with a witness bound, which requires that, if the indefinite is true, the variable associated with the indefinite points to the witness of its truth.

(38) Witness bound for indefinites:

$\text{з}x(p, q)$ is satt at $\langle c, g, w\rangle$ only if:

If $[\![\exists x(p\&q)]\!]^{g,w} = 1$, then $p\&q$ is true and satt at $\langle c, g, w\rangle$

Crucially, this bound projects across negation.

This is used to characterize the contrast between indefinites under negation (39-a) and under double negation (39-b):

(39) a. We don't have [a cat]^x.

$\neg\text{з}x(cat(x), exists(x))$

b. It's not the case that we don't have [a cat]^x.

$\neg\neg\text{з}x(cat(x), exists(x))$

An update with (39-a) in a context c will yield an output context c' containing all pairs

$\langle g, w \rangle \in c$, s.t. (39-a) is true and satt at $\langle c, g, w \rangle$. For a pair $\langle g, w \rangle \in c$, (39-a) is true just in case there is no cat in w. (39-a) is satt at $\langle c, g, w \rangle$ just in case, if there is a cat in w, $g(x)$ is a cat in w. Based on the truth-conditions, c' contains only pairs where there is no cat in w, and the witness-bound is trivially satisfied in these cases. As a result, no condition on $g(x)$ is enforced, and c' may contain assignments mapping x to any value, or none at all.

On the other hand, an update with (39-b) in c has different results. For a pair $\langle g, w \rangle \in c$, (39-b) is true just in case there a cat in w. (39-b) is satt at $\langle c, g, w \rangle$ just in case, if there is a cat in w, $g(x)$ is a cat in w. Based on the truth-conditions, the output context c' contains only pairs where there is a cat in w, and the witness-bound requires that it only contains pairs, where in addition $g(x)$ is a cat that exists in w. As a result any assignment in the output maps x to a cat.

Anaphoric definites are interpreted with a familiarity bound, which requires that the content of their scope is true and satt throughout the definite's local context.

(40) Familiarity bound for definites:

$\iota x(p, q)$ is satt at $\langle c, g, w \rangle$ only if:

$\forall \langle g', w' \rangle \in c : p$ is true and satt at $\langle c, g', w' \rangle$.

Based on this, an account of the licensing conditions for anaphoric definites like (41) works in the following way.

(41) [The cat]$_x$ is tabby.

$$\iota x(\mathtt{cat}(x), \mathtt{tabby}(x))$$

An update with (41) in a context c will yield an output context c' containing all pairs $\langle g, w\rangle \in c$, s.t. (41) is true and satt at $\langle c, g, w\rangle$. For a pair $\langle g, w\rangle \in c$, (41) is true just in case $g(x)$ is a cat and tabby w. (41) is satt at $\langle c, g, w\rangle$, just in case all pairs $\langle g', w'\rangle \in c$ are s.t. $g'(x)$ is defined and $g'(x)$ is a cat.

While the witness-bound on indefinites applies pointwise to the elements of the input contexts, and imposes a condition on which elements may also be a member of the output, the familiarity bound applies by universally quantifying over the input context: The update is satt only if any assignment in the input maps x to a cat. As discussed above, this is the case for an input that was previously updated with a double-negation sentence (39-b), but not one that was updated with a negated sentence (39-a).

The analysis extends to anaphoric pronouns, which are interpreted as definites with a tautological restrictor. It is also extended to bathroom sentences, by assuming the following bound for disjunction:

(42) $p \vee q$ is satt at $\langle c, g, q\rangle$ iff p is satt at $\langle c^{\neg q}, g, q\rangle$ and q is satt at $\langle c^{\neg p}, g, q\rangle$.

However, this account rules out modal subordination, or any anaphora to singly negated indefinites, similar to Gotham's (2019) account.

Elliott (2020): Positive Closure

The dynamic semantics developed in Elliott (2020, 2021) is designed to account for double negation cases and bathroom-disjunctions. The more recent version in Elliott (2021) is called Externally-dynamic dynamic semantics. It is based on Charlow's (2017) monadic grammar and uses a trivalent logic. Elements of the context are (ts) pairs of truth-values and variable assignments.

The meaning of an utterance is a dynamic proposition, here formalized as a mapping from an input assignment to a set of truth-value/assignment pairs (type $D(t) := s((ts)t)$).

Elliott (2020) suggests that an indefinite introduces an 'indeterminate [dref] with respect to both the positive and negative information associated with a sentence' (p. 7). Instead of relativizing variable assignments to possible worlds, they are relativized to possible truth values for the containing sentence (true, false, or neither).

(43) Random variable assignment relative to a restrictor r_{et}:

a. $\varepsilon^n(P_{et})(\mathcal{P}_{e,D(t)}) := \lambda g_s. \bigcup_{P(x_e)} \mathcal{P}(x)(g^{[n\mapsto x]})$

b. E.g. for *Mary has a car*:

$\varepsilon^n(car_{et})(\lambda y_e.\lambda g.have_{e(et)}(y)(m), g) =$

$\lambda g_s. \bigcup_{car(y_e)}\{\langle have_{e(et)}(y)(m), g^{[n\mapsto y]}\rangle\}) =$

$\lambda g.(\{\langle 1, g^{[n\mapsto y]}\rangle \mid car(y) \wedge have(y)(m)\} \cup \{\langle 0, g^{[n\mapsto y]}\rangle \mid car(y) \wedge \neg have(y)(m)\})$

Given a restrictor (and a dynamic proposition), random variable assignment maps an input assignment g to a set of truth-value/assignment pairs. According to Elliott, we extend the input g to map n to a car, and tag those assignments extended to a car that Mary has with 1 (true), and those extended to a car that Mary does not have with 0 (false).

The semantics of indefinites is given in terms of a positive closure operator, which makes sure that 'if its prejacent isn't true with respect to the input assignment, referential information is filtered out'(Elliott, 2021, p. 8).

We do not need to consider the exact formalization of the positive closure operator, or the semantics given for negation. Instead, we will discuss the implications of the analysis on the basis of the result of the update associated with positive and negative utterances provided by Elliott.

An affirmative utterance with an indefinite like our example *Mary has a car* is interpreted as in (44).

(44) $$[\![\text{Mary has [a car]}^n]\!] = \lambda g.\ \ \{\langle 1, g^{[n \mapsto x]}\rangle \mid \mathit{car}(x) \wedge \mathit{have}(x)(m)\} \cup \{\langle 0, g\rangle \mid \neg\exists x(\mathit{car}(x) \wedge \mathit{have}(x)(m))\}$$

The update with *Mary has a car* maps an input assignment g to a set that contains $\langle 1, g^{[n \mapsto x]}\rangle$, for each $x \in D_e$ of which it can be truthfully said that x is a car that Mary has, and contains $\langle 0, g\rangle$ in case no such x exists in D_e.

Negation in this system works as a simple (strong Kleene) truth-functional operator, which applies to the truth-value component of the elements of the context:

(45) $$[\![\text{Mary doesn't have [a car]}^n]\!] = \lambda g.\ \{\langle \neg 1, g^{[n\mapsto x]}\rangle \mid car(x) \wedge have(x)(m)\} \cup \{\langle \neg 0, g\rangle \mid \neg\exists x(car(x) \wedge have(x)(m))\}$$
$$= \lambda g.\ \{\langle 0, g^{[n\mapsto x]}\rangle \mid car(x) \wedge have(x)(m)\} \cup \{\langle 1, g\rangle \mid \neg\exists x(car(x) \wedge have(x)(m))\}$$

Elliott describes the effect of negation as follows: 'if the negative sentence is true, no referential information is introduced. This captures the external staticity of negation.'The way in which negation is applied here does allow for an account of double negation cases, and is sucessfully extended to disjunction. In fact, Elliott provides a lot of insightful discussion and arguments of anaphora under disjunction. However, this account has the same issue as the other accounts that were specifically developed for double negation: It conditions the introduction of a dref on the context of an indefinite without taking into account the context of the anaphor.

The account is more general than classic dynamic accounts, because the introduction of a dref is not conditional on the presence of negation. It is conditional on the relationship between the local context of the anaphor and the overall truth-value of the sentence (i.e. what the speaker will be committed to). I have argued that this is a clear advantage.

However, the account will not be able to account for cases of modal subordination with negative antecedents, negative subordination, or inter-speaker disagreement cases.

Similarly, it would not be able to account for modal subordination.

Of course, it was not designed to do so. Like any conditional variable update account, it could be supplemented with a mechanism of accommodation for cases of discourse subordination and inter-speaker disagreement.

7.5.3 Accommodation and Bridging

Another way of dealing with anaphora to antecedents which are in otherwise inaccessible parts of the discourse is finding ways of making them accessible again. These include mechanisms, such as accommodation (Roberts, 1987, 1989), anaphoric reference to parts of the referential context itself Geurts (1999); Frank (1996); Frank and Kamp (1997), or bridging (e.g. like suggested by Geurts (1999) for anaphora to indefinites under negation, and developed by Nouwen (2003) for complement anaphora).

Roberts' DRS-copying

These kinds of analyses—like the study of modal subordination itself—go back to Roberts (1987, 1989), who developed a DRT-copying mechanism to model an analysis of modal subordination based on the idea that an antecedent can be accommodated. The mechanism allows for a previously subordinated modal DRS to be reused by accommodation to provide the modal base for a subsequent modal. Roberts also suggests that a similar mechanism of accommodation may also provide analysis for bathroom sentences.

Roberts suggests that this mechanism of accommodation should be constrained in the following ways: (see also the discussion of these constraints in Geurts (1999)).

1. Modal subordination requires non-factual mood

2. It must be plausible that the modally subordinate utterance has a hypothetical common ground suggested by the immediately preceding context

3. Modal subordination may not make available to anaphoric expressions antecedents that have no explicitly representation in the given DRS

Based on the basic mechanism and these constraints, Roberts' account allows anaphoric reference to drefs which have previously introduced into a subordinate context.

The most convincing arguments that Roberts's ideas needed to be built upon further are ones that address the issue that accommodation is captured representationally by means of a DRS-copying mechanism (Kibble, 1994). In particular, the copying mechanism assumed by Roberts allows for copying the DRS introduced by the antecedent of a conditional, but not its consequent. Kibble (1994) notes this restriction and provides the following counterexample:

(46) Kibble (1994):

If John's at home, he'll be reading a book. It'll be a murder mystery.

While this is an argument against the DRS-copying mechanism devised by Roberts, this is not an argument against accommodation per se.

Guerts / Frank & Kamp: Anaphoric Retrieval of Context

Geurts (1999) argues that an accommodation mechanism, as suggested by Roberts, is not constrained enough. Instead of accommodating the immediately preceding context, Geurts (1999) suggests an an analysis which is in the spirit of Roberts, but more constrained because it makes use of anaphora. Like in Roberts' analysis, a previously subordinated part of the discourse context is reintroduced for the interpretation of an anaphor in modal subordination (where contexts are treated as world-assignment pairs). This mechanism is restricted by the assumption that modals presuppose their domain, and that it may be retrieved anaphorically. Building on this, Frank (1996); Frank and Kamp (1997) develop a similar anaphoric account by incorporating an analysis of relative modality. This allows them to extend the empirical coverage to cases of modal subordination with antecedents that involve graded modality or counterfactual contexts (such as negation).

Stone (1999) provides an argument against accounting for modal subordination by use of anaphora to dynamic local contexts (i.e. ones including variable assignments). That is because an embedded context may only store all drefs that have been introduced globally up to that point, plus locally introduced ones. It could not store any global drefs that are introduced subsequently. This becomes a problem for cases like (47).

(47) Stone (1999)

a. *[A wolf]*$^{\upsilon_1}$ *might walk in.*

b. *We would be safe because John has [a gun]*$_{\upsilon_2}$.

c. *He would use it*$_{\upsilon_2}$ *to shoot it*$_{\upsilon_1}$.

Under accounts which make dynamic embedded contexts available anaphorically, the modal *would* in (47-c) anaphorically retrieves the local context introduced by *might* in (47-a). Then, the anaphora *it*$_{\upsilon_2}$ and $_{\upsilon_1}$ in (47-c) are interpreted in relation this local context. However, the local context introduced by *might* in (47-a) is not predicted to store a dref υ_2 for John's gun. Mixed subordinated / global anaphora in this configuration are expected to be unacceptable, but that is not the case. Using only one global referential context (i.e. variable assignment / set of assignments / set of world-assignment pairs), where drefs are understood in relation to possible worlds, and allowing individual denoting expressions to influence the relationship between assignments, worlds and the individuals they map to addresses this problem.

Bridging

The main arguments against accommodation-based accounts were either specific arguments against the representational nature Robert's DRS-copying mechanism, or arguments about specific accounts being too unconstrained.

Even Geurts (1999), who argued that Roberts' accommodation-account is too unconstrained to provide an analysis for modal subordination suggests that bathroom-sentences may receive an analysis in terms of bridging inferences. Geurts writes: *'a speaker utters [a bathroom sentence], the speaker may infer that according to the speaker the house has a bathroom'* (p. 129), and suggests that more generally *'under certain cir-*

cumstances, such an inferential process may give rise to new [drefs]' (*Ibid.*)

One account that relies on bridging, offering a more constrained mechanism which is similar to the idea of accommodating an antecedent, and at the same time makes explicit the assumed constraints is Nouwen's (2003) analysis of complement anaphora. Nouwen does not directly address (singular) anaphora to antecedents in the scope of negation. However, we might characterize the contexts in which complement anaphora are licensed as somehow negative in a sense that is relevant for anaphora (see discussion in **Chapter 6**). Even independently of the empirical connections between complement anaphora and anaphora to negated content, Nouwen's (2003) account of accommodation might possibly be extended to analyze anaphora to indefinites under negation and other non-veridical operators.

Nouwen suggests that an anaphoric pronoun may introduce its own dref by way of bridging, just in case the following conditions are satisfied:

1. **Inferability:**

   The referent of the pronoun is inferable. (Specifically, I interpret this to mean that the previous discourse needs to support the inference that the referent exists in the context of the pronoun.)

2. **Uniqueness:**

   '[T]here should be no other similarly inferable entities with the same semantic features.' (p. 105)

3. **Use of semantically available information:**

   'Whenever the inference involved is [...] in need of information which is not semantically present, pronominal reference is no longer an option' (p. 106) Specifically, Nouwen points out that in case of bridging definites, the inference could not be supported by implicature.

4. **Support of discourse coherence by the anaphoric link:**

   The anaphoric resolution supports discourse coherence; This is discussed more specifically in relation to a constraint to avoid contradiction: (For complement anaphora) Nouwen suggests that a referent *'(once inferable) is only used as a resolution for pronominal reference once there is a semantic reason to do so.'*

Nouwen's analysis works for the cases of bridging and complement anaphora discussed in Nouwen (2003), however, could not account for the famous marble-contrast discussed in **Chapter 2** (repeated here as (48)), where a unique semantically inferable referent could be retrieved to establish an explanatory discourse relation, but no overt antecedent is available.

(48) Roberts (1987), attributed to Barbara Partee:

a. *I dropped ten marbles and found all but one of them. It's probably under the couch.*

b. *I dropped ten marbles and found nine of them. # It's probably under the couch.*

If we wanted to extend an account along these lines to modal subordination and anaphora to negated content, we would need a principled way of ruling this out. The account presented here can rule out the infelicitous marble-cases. In (29-a), a dref is introduced and can be picked up subsequently. Even if a dref is introduced under negation, it will be with an overt expression, while its epistemic status, the question of whether or not it exists in the global context (i.e. the speaker's commitments) can be inferred based on additional (semantic or pragmatic) information.

7.5.4 Lewis (2021): Hybrid Account

I have argued that accommodation- and bridging-type accounts, while designed for discourse subordination, would need to say more in order to address double negation cases and bathroom sentences. On the other hand, conditional variable introduction accounts run into problems with discourse subordination.

The argument presented here relies on the systematically analogous ways in which the epistemic status of the antecedent and anaphor come into play, and in which both interact with their local context. It claims that this suggests that the phenomena under consideration here should receive a unified account.

However, we might also consider a hybrid account—one where double negation and bathroom-disjunction are analyzed by conditioning the introduction of drefs, and discourse subordination and inter-speaker disagreement are addressed by means of accommodation. An interesting account of this kind is Lewis's (2021) dynamic pragmatic account.

In her own words, the paper presents an argument 'in favor of adopting a dynamic pragmatics, a theory that explains context change through general Gricean principles, and combining it with a static, d-type theory of anaphora, in which pronouns go proxy for definite descriptions' (p. 1403).

Lewis also notes that negation is not special in often preventing anaphora to the material in its scope, nor in exhibiting some notable counterexamples. She characterizes the connection between veridicality and speaker commitment on a pragmatic level:

According to Lewis, the introduction of a dref is a pragmatic matter, and discourse referents are 'objects under discussion', in a pretty general manner. A discourse referent is added, in case the speaker signals a plan to talk about the referent again in subsequent discourse. Lewis suggests that the certain uses are default strategies to signal this. In particular, using a Russellian denoting phrase (which I interpret to mean a DP) to assert the existence of an individual would be a prototypical strategy to signal such a plan.

This account ties the introduction of a dref to the speaker's commitments about the existence of a referent. As discussed above, this is an advantage in addressing double negation and disjunction cases, as well as anaphora to any antecedents in veridical contexts which might be ruled out in classic dynamic semantics based on the involved operators (e.g. any conditional, modal, or negative contexts).

The static d-type semantics for pronouns then assumes that pronouns require discourse-uniqueness, i.e. the availability of 'a unique discourse referent in the context that satisfies the [implicit] descriptive material'.

Lewis discusses a variety of cases where anaphora to negated content is possible, including cases of discourse subordination. These cases are explained by appealing to a mechanism of accommodation. Rather than characterizing this kind of accommodation as an ad hoc kind of mechanism, Lewis suggests that it is expected that reasoning about the speaker's discourse plans may be updated and adapted'—and discourse referents are accommodated.' (p. 1425).

Lewis addresses the infelicitous marble examples ((48-b) above, repeated here) by suggesting that they are improved if a pause is added between the utterances.

(48-b)*I dropped ten marbles and found nine of them. # It's probably under the couch.*

I do believe that this example raises a lot of questions about what is accommodated here. Are we picturing a situation in which we are searching for the marble? And if so, how different is this pronoun use from a deictic use? What can we learn from deictic uses, and cases of bridging? As Lewis suggests, a dynamic pragmatic account may also address some questions about anaphora which have been noted as counterexamples to (see discussion in Lewis, 2021, Section 4.4).

Lewis's approach addresses all of the different cases of anaphora to negated content discussed here in an insightful and thought-provoking way. Nevertheless, I would like to point out two potential issues for the account:

First, it does not seem constrained enough. For example, we might suggest that the above constraints that Nouwen (2003) discusses for complement anaphora should also apply to the accommodation mechanism proposed here in order to rule out licensing of the unacceptable complement anaphora cases.

Further, Lewis's account does not directly apply to the inter-speaker disagreement cases discussed here. I argued that these show that one speaker can make reference to an entity or proposition as veridical, even if another introduced the antecedent non-veridically. This could not be explained on the basis of the first speaker changing their intentions, or the second adapting their understanding of what they might be. However, again, an idea from Nouwen's account might help here, who suggested that pronouns may introduce drefs under the special conditions outlined by the given constraints. In fact, pronouns could be characterized as Russellian denoting phrases, but if anaphoric pronouns can license themselves, we return to the question whether the licensing conditions are constrained enough.

An account which incorporating insights from Lewis's and Nouwen's accounts may very well provide a suitable account of the interaction of anaphora and negation.

7.5.5 Summary and Conclusions

The discussion of the various accounts builds on the assumption that anaphora to indefinites under negation in cases with double-negation, bathroom-disjunctions, modal subordination with non-/anti-veridical antecedents, and inter-speaker disagreement should receive a unified analysis. This view, represented in this work, is supported by the fact that both individual and propositional anaphora very systematically follow patterns along the lines of veridicality-based distinctions.

Within the understanding laid out here, accounts of anaphora and negation which rely on conditional introduction of drefs run into a lookahead problem: They can account for double negation and bathroom disjunctions, but exhibit limitations when considering modal subordination with negative antecedents.

Accommodation-based accounts and flat-update dynamic systems do not run into this kind of problem, and both of these classes of accounts seem in principle viable analyses for all of the kinds of data under consideration.

The present work is an exploration of the latter kind of account, its empirical coverage and limitations. As discussed here, an account based on intensionalized drefs and their relations to the interlocutors' commitments can capture many generalizations about how individual anaphora, propositional anaphora, and negativity-tags interact with negation, and other non-veridical operators in a principled way.

As discussed in **Chapter 2**, the strongest arguments I have for preferring a flat-update account over an accommodation-type one are conceptual in nature: I suggest that interpreting all DPs in relation to their local context, and keeping track of their epistemic provides a principled explanation. It allows us to understand the different kinds of counterexamples to the classic generalizations on the same terms, and explains the constraints in terms of basic principles of discourse interpretation. Further, Stone (1999) suggested that an accommodation-based account might fail to provide an understanding of the parallel ways in which anaphora interact with modality/negation and tense.

It is also worth noting that most accommodation based accounts do not necessarily account for double negation, although the Lewis's (2021) dynamic pragmatic account does.

For an empirically motivated argument for flat update or accommodation as a suitable model of the parts of linguistic competence that have to do with the interaction of anaphora and negation (or more broadly: veridicality), we may want to look for other behavioral or neurophysiological evidence for cognitive representations or processes.

For unacceptable veridical anaphora with non-veridical antecedents, Dwivedi et al. (2006, 2010) have found EEG measurements of neurophysiological responses which they interpret to be part of the P600 family (*John is considering writing a novel. It ends quite abruptly.*, at the verb of sentence 2). The the neurophysiological response thus shows a different kind of event-related potential than would be expected in purely semantic conceptual violations (like *John ate # democracy*). This evidence might be interpreted as evidence for a discourse-level structural kind of violation Dwivedi et al. (2006), or contradictory information Dwivedi et al. (2010). More research is certainly needed in this area.

7.6 Conclusion

The presented analysis explains the constraints on individual anaphora in terms of a fundamental requirement that the referent of a DP exist in its local context, and that a discourse be consistent. It is formalized in a flat-update dynamic semantics where all non-veridical operators are externally dynamic. Indefinites in their scope update the variable assignment globally, along with the information about the sets of worlds in which their referents exist. In addition, the discourse needs to store the information about the relationship to the speaker's commitments. The resulting understanding of discourse referents is therefore both intensional and epistemic.

The analysis provides an understanding of when the surrounding context allows for an anaphoric relation between expressions introducing anaphora and potential antecedents, and represents a unified account of anaphora to indefinites under double negation, bathroom-disjunctions, modal subordination with negative antecedents, and inter-speaker disagreement.

Chapter 8

Conclusion

8.1 Summary

8.1.1 Flat Update and Veridicality-Constraints

I have provided arguments for the claim that anaphora are sensitive to negation and polarity, because they are senstivity to the veridicality of their antecedents, and the veridicality of their own intensional context.

In particular, **Chapter** 2, developed the generalization that propositional anaphora require that their embedding context is compatible with an interpretaion under which their antecedent is true in the local intensional context, and the parallel generalization for individual anaphora, which can have antecedents which exist in their local context. Chapter three also showed the effects that this generalization has on the possible combinations of (non-/anti-)veridical anaphora and antedents.

Chapter 5 provided further evidence for the claim that the interaction of anaphora and negation can be analyzed in terms of veridicality and the epistemic status of drefs in their local context. **Chapters 3 & 4**, provided a formal framework which was used to emply the suggested generalizations as a model of the interpretation and anaphoric availability for propositional anaphora. If we assume that a local truth-requirement constrains propositional anaphora, and find a way to represent how propositions relate to each other, the observed patterns observed can be explained in terms of the basic constraint of discourse consistency.

Chapter 7 also provided evidence for this claim, by showing that the interaction between individual anaphora and negation could also be explained in this way. Assuming that a local existence-requirement on individual variables constrains individual anaphora can provide an analysis of anaphora to indefinites under double negation, in bathroom-sentences, in cases of modal subordination with negative antecedents, and anaphora to indefinites under negation in inter-speaker disagreement. It can explain why they are systematically available, and how they are interpreted.

The presented analysis results in a dynamic semantics where all non-veridical operators are externally dynamic. Indefinites in their scope update the variable assignment globally, along with the information about the sets of worlds in which their referents exist. The analysis provides an understanding of when the surrounding context allows for an anaphoric relation between expressions introducing anaphora and potential antecedents.

8.1.2 Discourse-Negativity

The generalizations and analyses developed in **Chapters 2 & 5** characterize the interaction of anaphora and polarity in terms of a sensitivity to anti-veridical propositional discourse referents available in discourse. This suggests that the kind of polarity that anaphora are sensitive to should be characterized on the level of discourse. **Chapter 6** provided further evidence for this, showing that **(i)** as long as a propositional dref is introduced, an utterance can be rendered discourse-negativity due to a pragmatic inference that this dref is anti-veridical; and **(ii)** scope-relationships which may point towards a more structural generalization could be addressed by assuming certain independently motivated inferences as well. Further, the suggested discourse-level kind of representation, i.e. capturing discourse-negativity in terms of anti-veridical propositional drefs could capture the variety of discourse-negative surveyed here. The chapter provided further evidence for a veridicality-based generalization of discourse-negativity by reporting experimental data which suggest that anti-veridical embedded clauses without negation may introduce discourse-negativity as well, at least to some extent.

There are two avenues for further research here: First, the generalizations and results about a characterization of discourse-negativity on the level of discourse representation should have consequences for the analysis of polarity particles and negativity-tags. Since these anaphoric expressions are analyzed as sensitive to discourse-polarity, the question of their polarity-sensivity should be analyzed represents the flipside of the question how discourse-polarity as a property of the antecedent should be analyzed.

Second, the categorical characterization of discourse-negativity based on a representation on the level of discourse provided here could not address the gradient kind of data that we saw in Chapter 7. We have found a way to characterize all of the various kinds of utterances containing different kinds of negative expressions as discourse-negative, by assuming that all of them can introduce an anti-veridical propositional dref. The question of why weakly negative expressions or quantifiers in object position might introduce discourse-negativity less robustly than overt negative marking has not been addressed here. I have been working with the assumption that this question can be addressed separately. I briefly discuss ways of addressing both kinds of questions, how to draw conclusions for the analysis of PolPs and negativity-tags, and how to address gradient negativity here.

8.2 Future Research

8.2.1 Polarity Particles and Negativity-Tags

The generalizations about arguments about the nature of discourse-negativity presented here will have some consequences for an analysis of the polarity-sensitivity of polarity particles and negativity-tags. Future research could target the question how compatible the results and generalizations are with different kinds of analyses of PolPs present in the literature.

Prominent accounts are the feature-based account (Farkas and Bruce, 2010; Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019), ellipsis-based accounts (Kramer and Rawlins, 2009; Holmberg, 2013), and the kind of accounts that Krifka (2013); Meijer et al. (2015); Claus et al. (2017) call saliency-accounts. The representational assumptions of the latter are compatible with the kind of analysis argued for here, as are the assumptions about the interpretation of PolPs in relation to antecedent polarity. At the same time, we showed that certain constraints on anaphoric resolution, which Krifka (2013); Meijer et al. (2015); Claus et al. (2017) assume to be modulated by salience could be explained in terms of local truth-requirements and discourse consistency. I suggest that further questions of how to address gradient data about antecedent choice in the case of referential ambiguities and PolP choice in case of contextual synonimity should be addressed in terms of processing, rather than representation (see discussion in the following subsection). While Meijer et al. (2015); Claus et al. (2017) do not find an effect of what they describe as 'salience', on PolP choice, as might be predicted by the analysis in Krifka (2013), they suggest that *'It may turn out that the most valid account is one that succeeds in integrating processing insights in addition to theoretical insights from different approaches.'(p. 46)*

That future research should in fact aim to combine insights from various approaches to PolPs becomes apparent when we compare the saliency-account to the achievements of other accounts. The feature-based account (Farkas and Bruce, 2010; Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019), as well as the particular implementations of ellipsis-based accounts in (Kramer and Rawlins, 2009; Holmberg, 2013), may require antecedent polarity as a part of the overt representation of the sentence introducing the propositional antecedent. As such, their representational assumptions are not immediately compatible with the generalizations and findings discussed here. However, especially the feature-based account discussed in (Farkas and Bruce, 2010; Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019) has made significant contributions to the understanding of the pragmatic function of PolPs in responses to polar questions and assertions, and in turn of the pragmatic function of assertions and polar questions as well. Also, the relative and absolute polarity-features of the feature-based accounts are based on Pope's (1972) typological work on PolP answering systems.

A veridicality-based account of the interaction of anaphora and negation may offer a way to combine insights from Krifka (2013) and (Farkas and Bruce, 2010; Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019). It would incorporate insight from Krifka on a representational level: An anti-veridical propositional dref is one that is introduced for the prejacent proposition of negation. It would incorporate insight from Farkas and Bruce (2010); Roelofsen and Farkas (2015); Farkas and Roelofsen (2019) on the level of explaining which kinds of interpretions for PolPs are possible and how they make up answering systems which cover a certain set of pragmatic functions. This might work by reinterpreting absolute and relative polarity features in terms of veridicality: Absolute polarity might be characterized in terms of a veridical vs. anti-veridical propositional anaphor (e.g. *yes* vs. *no*), whereas relative reatures might be characterized in terms of the discourse relation between the discourse segments containting the antecedent and the anaphor, in terms of a veridical vs. anti-veridical discourse relation (such as e.g. agreement vs. disagreement). Following (Farkas and Bruce, 2010; Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019), the typology of answering systems could be described by assuming that response particles may lexicalize either of these kinds of meanings, or a mixture of both.

Future work should investigate in more detail which kinds of representational commitments the varioud accounts of the polarity-sensitivity of PolPs make wrt the representation of polarity, and to what extend they might be compatible with the arguments presented here. To the extent that any account relies on the presence of negation in a morphosyntactic or semantic representation of the antecedent clause, they would not be entirely compatible with the claims made here. In addition, future research should address and explicate whether a veridicality-based account of PolPs allows for combining insights from various accounts of PolPs. A third question that would be interesting in this context would be the question whether the account of discourse-polarity and anaphoric polarity-sensitivity may be extended to analyze other kinds of negativity-tags, such as *why not* questions or polar additive tags, and whether their analysis can provide further theoretical insights or arguments that may help support or reject some of the possible analyses.

8.2.2 Weakly Negative Expressions and Gradient Negativity

Based on the experimental results in Brasoveanu et al. (2013, 2014), we know that downward-entailing quantifiers and negative proximatives contribute discourse-negativivity less robustly than anti-additives. Similarly, sentential objects provide negativity less strongly than subjects and sentence adverbs, and anti-veridical attitudes less strongly than negation.

We see gradience wrt the proportions of negativity-tags a certain kind of negative antecedent gives rise to. However, we have assumed categorical representational characterization of discourse-negativity: An utterance is discourse-negative just in case it introduces an anti-veridical propositional dref, and discourse-positive if it doesn't. This means that the gradience found in the experiments is still left unexplained. Future research might address the question of how this kind of gradience might be addressed based on the generalizations put forward here.

We can draw some inspiration from the discussion of closely related gradient data from similar experiments. The preference patterns for PolPs, particularly with negative antecedents have been discussed in Meijer et al. (2015); Claus et al. (2017); Farkas and Roelofsen (2019). Recall that in the context of a negative antecedent, in German either *nein* ('no') or *ja* ('yes') can be used to agree. Given this contexual synonymity, Meijer et al. (2015); Claus et al. (2017) investigate which of the PolPs is preferentially used with this function.

Their acceptability-ratings study showed that with negative antecedents, *ja* was preferred over *nein* in agreeing uses (unlike what Brasoveanu et al. (2013) have found for English). However, they found some inter-speaker variation here as well. While most speakers followed a *ja* > *nein* pattern, some participants consistently rated *nein* higher than *ja* in these contexts. For disagreeing with a negative antecedent, the preference order was *doch* > *ja* > *nein*.

The goal of this research is to provide compare the feature-based account (Farkas and Bruce, 2010; Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019) and Krifka's (2013) saliency-account, based on how well they might explain the preference-patterns. In their evaluation of the two accounts against the experimental findings, Claus et al. (2017) suggest that both of them are suited to address the gradient data.

In particular, Roelofsen and Farkas (2015) appeal to the relative morphological markedness of agreeing uses of *ja/nein*, based on generalizations over different kinds of polarity-features, and a pressure to realize the morphologically more marked, as well as pragmatic blocking mechanisms. Claus et al. (2017) appeal to resolution of propositional anaphora, the relative salience of their different possible antecedents, and pragmatic blocking as well.

The latter suggest that the gradience in PolP preference patterns is part of the grammar, in turn apparently assuming that salience is part of the grammar. I believe that these arguments could be stronger, for the following reasons:

There certainly are theories of anaphora, which attempt to representationalize salience (e.g. centering theory, references). But the often invoked, but ill-defined notion of salience could similarly be conceptualized in terms of competence, or more general cognitive principles, like attention allocation or activation in working memory.

Further, the experimental results led Claus et al. (2017) to revise those particular assumptions from Krifka (2013) that were related to salience. The assumption that for a negative utterance, the prejacent of negation is the more salient proposition was rejected, and the opposite was assumed. Additionally, the assumption that the context will modulate the relative salience the two propositional drefs introduced by a negated utterance was discarded.

It is arguable if the relative salience of the possible propositional antecedents modulates the choice of PolP in this case, and even if so, whether this generalization should lead to an explanation in the realm of linguistic representations or processing mechanisms. Claus et al. suggest that future research might come up with a more comprehensive analysis of the data discussed in their paper, by incorporating insights from processing, as well as the various accounts of PolPs. I am not eager to make claims about the best explanation for PolP preference patterns in for agreeing responses to negative antecedents. Regarding the question of why some utterances behave more or less like a negative antecedent, I would like to suggest a line of research addressing the question whether the the gradience of discourse-negativity could be addressed in terms of processing.

The question of whether the gradience should be addressed in terms of linguistic representations or language processing is a hard one to address directly. However, the question should be addressed in future research. The arguments and generalizations made here rely on the assumption that the linguistic generalizations on the level of discourse-representation are categorial in nature. In turn, they make an assumption that utterances can be more or less robustly negative based on processing factors, purposely leaving questions of gradience to processing. To investigate the validity of these assumptions, future research may develop and investigate some concrete processing-based hypotheses.

As such, the representational claims made here are also meant to provide a basis for research addressing this kind of question.We might ask whether we can we theoretically explain the gradient effects in relation to processing mechanisms, such as parsing, ambiguity resolution, or anaphor resolution. We might address these kinds of hypotheses by investigating whether we find the behavioral effects that we might expect on the basis of these kinds of hypotheses in online experiments. Here, I would like to begin outlining some of these kinds of hypotheses for weakly negatives expressions, which give rise to the question of why they license negativity-tags less robustly than their anti-additive counterparts.

Why are DE less negative than AA? — Dynamic Ambiguity Hypothesis

One possible hypothesis for addressing the question of why downward-entailing/downward-asserting expressions introduce discourse-negativity less robustly than anti-additive expressions draws from literature on complement anaphora. In an analysis of complement anaphora, Kibble (1997) suggests that negative downward-entailing quantifiers (such as *few*) are dynamically ambiguous: They have two possible interpretations, which are truth-conditionally equivalent but differ in the licensing of subsequent anaphora.

For negativity-tags, this hypothesis can be related to the fact that, while utterances with downward-entailing expressions (1-b) may pattern with discourse-negative utterances like (1-a), they may also pattern with discourse-positive utterances like (1-c). That is, they may be followed by negativity-tags or positivity-tags.

(1) a. *None of the boys ate their cake.*
   (i) ✗So did I.
   (ii) ✓Neither did I.

 b. *Few of the boys ate their cake.*
   (i) ✓So did I.
   (ii) ✓Neither did I.

 c. *All of the boys ate their cake.*
   (i) ✗So did I.
   (ii) ✓Neither did I.

Complement anaphora (2-b-ii) following an utterance with a *few*-quantifier exhibit a pattern that is similar to that of a negativity-tag in the following ways: **(i)** The follow-up *They threw it away* in (2-b-ii), is also available with a more strongly negative antecedent (2-a) under the same interpretation; and **(ii)** the same weakly negative antecedent (2-b) could alternatively be followed by *They really enjoyed it*, which is compatible with a positive antecedent under the same interpretaition (2-c).

(2) a. *None of the boys ate their cake.*

(i) *#They really enjoyed it.*

(ii) ✓*They threw it away.* (they ≈ the boys who didn't eat their cake)

b. *Few of the boys ate their cake.*

(i) ✓*They really enjoyed it.* (they ≈ the boys who ate their cake)

(ii) ✓*They threw it away.* (they ≈ the boys who didn't eat their cake)

c. *All of the boys ate their cake.*

(i) ✓*They really enjoyed it.* (they ≈ the boys who ate their cake)

(ii) *#They threw it away.*

In comparison to discourse-polarity tags, we might therefore describe complement anaphora as a way in which downward-entailing utterances pattern with negative utterances in discourse (2-b-ii), although optionally so: They could also pattern with positive utterances (2-b-i).

The parallels between complement anaphora and negtivity tags in these cases suggest that the literature on complement anaphora might help with understanding weak discourse-negativity.[1] Here, in particular, we take the parallels to suggest that the li-

[1]There are different kinds of analyses of complement anaphora. Bridging: Nouwen, quasi-generic: references, expectation-denial and focus: moxey and sanford.
The question of which is the most appropriate analysis for complement anaphora is a complex question, which I wouldn't want to make any claims about here. It is imaginable that other accounts of complement anaphora might also help us come up with a different hypothesis for why DE/DA are weakly discourse-negative. For now, we go with the dynamic ambiguity hypothesis, for a few reasons: 1. It is the one for which I most straightforwardly see how it could relate to a processing hypothesis. 2. It could be supported by the low confidence-ratings found for negative proximatives in Experiment 1 from Chapter 8. 3. It also provides a hypothesis for why negative objects are less robustly discourse-negative (see below).

censing of both negativity-tags and complement anaphora is sensitive to *few* possibly being interpreted as discourse-negative, while there still is a possible discourse-positive interpretation.

This could be addressed within Kibble's (1997) analysis of complement anaphora. Kibble suggests, following Zwarts (1996), that sentences involving negative downward-entailing quantifiers like *Few boys ate their cake* can be paraphrased in two ways which are truth-conditionally equivalent. This is based on an understanding of DE quantifiers as negations of upward monotonic proportional quantifiers with a meaning similar to something like *many*. Zwarts notes that the complement (*not many*) and the contradual (*many …don't*) are truth-conditionally equivalent, and might both be considered as interpretation for *few* The two readings could be paraphrased as:

(3) a. It's not the case that many boys ate their cake.

b. Many boys are such that they didn't eat their cake.

Based on van den Berg's (1996) account of dynamic quantification, the two interpretations, while truth-conditionally equivalent, lead to two different anaphoric potentials. Kibble defines the complement and contradual as derived dynamic quantifiers. If we assume a decompositional analysis of *few*, this would be understood as a quantifier scope ambiguity.

This possible explanation of how the anaphoric possibilities for downward-entailing quantifiers like *few* arise could also provide a basis for a processing-based explanation of their weak discourse-negativity by appealing to semantic parsing and resolution of quantifier scope ambiguity. The antecedent is dynamically ambiguous due to a quantifier scope ambiguity. In order to use a subsequent anaphor, the ambiguity needs to be resolved in a way that is compatible with the interpretation of the anaphor.

This kind of explanation could help us understand why weak negatives in forced-choice tasks can sometimes be followed up by negativity-tags, and in other cases won't be. They can be interpreted as discourse-negative or discousre positive. Similarly, it might help us understand why we saw lower confidence ratings for negative proximatives in Experiment 1 in Chapter 8. Anti-additive negative antecedents would allow for either agreeing follow-ups with either *no* or *yes*, and therefore are associated with lower confidence compared to the positive baseline. However, antecedents with *hardly* led to even lower confidence ratings. This could be explained if the antecedent was ambiguous and introduced a further source of uncertainty.

The dynamic ambiguity hypothesis assumes that the antecedent utterance is scopally ambiguous due to the interaction of two scope-taking elements decompositionally introduced by *few*: Negation and a proportional quantifier which could roughly be paraphrased as *many*. We would expect the processing profile to be similar to other cases in which quantifier scope ambiguities are disambiguated by anaphora. In particular, we might expect to see an effect of reanalysis for either negativity-tags or positivity-tags. This is because evidence suggests that quantifier scope relations have a preferred interpretation, and are associated with a processing difficulty if a subsequent disambiguating anaphoric definite is incompatible with this preferred interpretation. If weak negatives are similar, we expect an asymmetry between different possible disambiguiating follow-ups in online processing measures.

Were we indeed able to see evidence of reanalysis and therefore ambiguity for weakly discourse-negative utterances, this would help us to understand why they are weakly negative. It would also provide some evidence for the assumption that the gradient nature of discourse-negativity could find a competence-based explanation.

Final Words

I hope that this work can contribute to an understanding of how anaphora interact with negation, and non-veridical operators. I have provided an argument that an understanding of anaphoric polarity-sensitivity falls into place once we assume an intensional representations of drefs and allow certain expressions to be interpreted in relation to their local context.

This understanding of anaphoric polarity-senstivity can give a unified analysis for individual anaphora in non-veridical contexts, propositional anaphora across veridicality distinctions, and discourse-negativity. I have also indicated in some places how the presented analysis may provide further insight by drawing connections to polarity particles, negativity-tags and complement anaphora. Other ways in which the questions and themes discussed here will lead to further research questions involve the question of how the suggested analysis of individual anaphora compares to a bridging analysis, as well as how certain gradient data can be explained in terms of semantic processing.

Appendix A

Formal Appendix to Chapter 3

This appendix contains the formal definitions for the system developed in **Chapter 3**.

A.1 Static Logic Definitions

(1) The sets of types

a. Basic static types:

BasSTyp $:= \{t, w, e\}$ (truth values, worlds, individuals)

b. Static types:

STyp: smallest set including **BasSTyp** and all elements (σ, τ), s.t. $\sigma, \tau \in$ **STyp**

c. Basic dynamic types:

BasDynTyp $:= \{s\}$ (variable assignments)

d. Dref types:

DRTyp: smallest set s.t. if $\tau \in$ **STyp**, then $(s\tau) \in$ **DRType**

e. Basic types:

BasTyp := **BasSTyp** $\cup$ **BasDynTyp**

f. Types:

Typ: Smallest set including **BasTyp** and all elements $(\sigma\tau)$, s.t. $\sigma, \tau \in$ **Typ**

(2) Basic expressions

For any type $\tau \in$ **Typ**, there is a denumerable set of τ-constants $\mathbf{Con}_\tau$ and a denumerably infinite set of τ-variables $\mathbf{Var}_\tau = \{v_{\tau 0}, v_{\tau 1}, \dots\}$. Some relevant sets of constants that are lexical translations of NL expressions include the following:

a. $\mathbf{Con}_e = \{mary_e, sue_e, \dots\}$

b. $\mathbf{Con}_{e(wt)} = \{sleep_{e(wt)}, tired_{e(wt)}, dance_{e(wt)}, car_{e(wt)}\}$

c. $\mathbf{Con}_{e(e(wt))} = \{own_{e(e(wt))}\}$

d. $\mathbf{Con}_s = \{i, i', \dots, j, j', \dots, k, k', \dots\}$

(3) Terms

a. For any type $\tau \in$ **Typ**, the set of τ-terms is the smallest set, s.t.:

(i) $\mathbf{Con}_\tau \cup \mathbf{Var}_\tau \subseteq \mathbf{Term}_\tau$

(ii) $\alpha(\beta) \in \mathbf{Term}_\tau$ if $\alpha \in \mathbf{Term}_{(\sigma,\tau)}$ and $\beta \in \mathbf{Term}_\tau$ for any $\sigma \in$ **Typ**

(iii) $(\lambda v.\alpha) \in \mathbf{Term}_\tau$ if $\tau = (\sigma, \rho), v \in \mathbf{Var}_\sigma$ and $\alpha \in \mathbf{Term}_\rho$ for any $\sigma, \rho \in$ **Typ**

b. For formulas (terms of type t), the following apply in addition:

(i) $(\alpha = \beta) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha, \beta \in \mathbf{Term}_\tau$

(ii) $(i[\delta]j) \in \mathbf{Term}_\tau$ if $\tau = t$ and $i, j \in \mathbf{Var}_s$ and $\delta \in \mathbf{Term}_\sigma$, for any $\sigma \in \mathbf{DRTyp}$

(iii) $(\neg\alpha) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha \in \mathbf{Term}_t$

(iv) $(\alpha \wedge \beta) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha, \beta \in \mathbf{Term}_t$

(v) $(\alpha \vee \beta) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha, \beta \in \mathbf{Term}_t$

(vi) $(\alpha \rightarrow \beta) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha, \beta \in \mathbf{Term}_t$

(vii) $(\alpha \leftrightarrow \beta) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha, \beta \in \mathbf{Term}_t$

(viii) $\exists\nu(\alpha) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\nu \in \mathbf{Var}_\sigma$ for any $\sigma \in \mathbf{Typ}$ and $\alpha \in \mathbf{Term}_t$

(ix) $\forall\nu(\alpha) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\nu \in \mathbf{Var}_\sigma$ for any $\sigma \in \mathbf{Typ}$ and $\alpha \in \mathbf{Term}_t$

(x) $(\alpha \subseteq \beta) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha, \beta \in \mathbf{Term}_{(\sigma,t)}$

(xi) $(\alpha \in \beta) \in \mathbf{Term}_\tau$ if $\tau = t$ and $\alpha \in \mathbf{Term}_\sigma$ and $\beta \in \mathbf{Term}_{(\sigma,t)}$

c. For sets (terms of type (σ, t)), the following apply in addition:

(i) $\overline{\alpha} \in \mathbf{Term}_\tau$, if $\tau = (\sigma, t)$ and $\alpha \in \mathbf{Term}_\tau$

(ii) $(\alpha \cap \beta) \in \mathbf{Term}_\tau$, if $\tau = (\sigma, t)$ and $\alpha, \beta \in \mathbf{Term}_\tau$

(iii) $(\alpha \cup \beta) \in \mathbf{Term}_\tau$, if $\tau = (\sigma, t)$ and $\alpha, \beta \in \mathbf{Term}_\tau$

(4) Frames, models, assignments, interpretation function

a. A *standard frame* $\boldsymbol{F}$ is a set $\{D_\tau : \tau \in \mathbf{Typ}\}$ s.t. D_e, D_t, D_w, and D_s are pairwise disjoint sets and $D_{\sigma\tau} = \{f : f$ is a total function from D_σ to $D_\tau\}$, for any $\sigma, \tau \in$ **Typ**. D_e and D_t each contain an indeterminate value # in any frame.

b. A *model* ***M*** is a pair $\langle F^M, [\![\cdot]\!]^{M,g}\rangle$, s.t.:

(i) F^M is a standard frame

(ii) An ***M***-assignment g is a function that assigns to each variable $\upsilon \in$ **Var**$_\tau$ an element $g(\upsilon) \in D^M_\tau$ for any type $\tau \in$ **Typ**. Given an ***M***-assignment g, if $\upsilon \in$ **Var**$_\tau$ and $d \in D^M_\tau$, then $\theta^{\upsilon/d}$ is the ***M***-assignment identical to θ except that it assigns υ to d.

(iii) $[\![\cdot]\!]^{M,g}$ is the interpretation function assigning an object in the domain of the model as the semantic value of a term.

c. ***M*** satisfies the following axioms:

1.**Logical constants:**

$[\![\cdot]\!]^M$ assigns an object $[\![\alpha]\!]^M \in D^M_\tau$ to each $\alpha \in$ **Con**$_\tau$ for any $\tau \in$ **Typ**

2.**Unspecific drefs: dVar**(δ)

for any discourse variable name δ of any type $\tau \in$ **DRType**, **dVar** is a non-logical constant intuitively identifying the non-constant functions of type $s\tau$ (for any $\tau \in$ **STyp**)

3.**Unique dref names:**

dVar$(\delta_\tau) \wedge$ **dVar**$(\delta'_\tau) \rightarrow \delta \neq \delta'$, for any distinct dref names δ, δ', for any type $\tau \in$ **DRTyp**

4.**Identity of assignments:**

$\forall i_s, j_s.[i[]j](i = j)$

5.**Enough assignments:**

$\forall i_s, \upsilon_{s\tau}, f_\tau[\mathbf{dVar}(\upsilon)](\exists j_s.[i[\upsilon]j](\upsilon(j) = f))$, for any type $\tau \in$ **STyp**

d. The interpretation function $[\![\cdot]\!]^{M,g}$ is defined as follows:

•Constants: If $\alpha \in \mathbf{Con}_\tau$ for any $\tau \in$ **Typ**:

–$[\![\alpha]\!]^{M,g} = [\![\alpha]\!]^M$;

•Variables: If $\nu \in \mathbf{Var}_\tau$ for any $\tau \in$ **Typ**:

–$[\![\nu]\!]^{M,g} = g(\nu)$;

•Function application: If $\alpha \in \mathbf{Term}_{(\tau,\sigma)}$, $\beta \in \mathbf{Term}_\tau$, for any $\tau, \sigma \in$ **Typ**:

–$[\![\alpha(\beta)]\!]^{M,g} = [\![\alpha]\!]^{M,g}([\![\beta]\!]^{M,g})$;

•Set membership: If $\alpha \in \mathbf{Term}_\sigma$, $\beta \in \mathbf{Term}_{(\sigma,t)}$, for any $\sigma \in$ **Typ** :

–$[\![\alpha \in \beta]\!]^{M,g} = [\![\beta(\alpha)]\!]^{M,g}$

•Lambda-abstraction: If $\alpha \in \mathbf{Term}_\tau, \nu \in \mathbf{Var}_\sigma$, for any $\sigma, \tau \in$ **Typ**:

–$[\![\lambda\nu.\alpha]\!]^{M,g} = \{\langle d, [\![\alpha]\!]^{M,g(\nu/d)}\rangle \mid d \in D^M_\tau\}$;

•Equality: If $\alpha, \beta \in \mathbf{Term}_\tau$, for any $\tau \in$ **Typ**:

–$[\![\alpha = \beta]\!]^{M,g} =$

1 iff $[\![\alpha]\!]^{M,g} = [\![\beta]\!]^{M,g}$,

0 otherwise;

•Variable update: If $\delta \in \mathbf{Term}_\sigma$, for any $\sigma \in \mathbf{DRTyp}$, $i, j \in \mathbf{Term}_s$:

–$[\![i[\delta]j]\!]^{M,g} =$

1 iff

·$[\![\forall\delta'_\sigma[\mathbf{dVar}(\delta') \wedge \delta' \neq \delta](\delta'(i) = \upsilon(j))]\!]^{M,g} = 1$ and

·$[\![\forall\delta'_\tau[\mathbf{dVar}(\delta')](\delta'(i) = \delta'(j))]\!]^{M,g} = 1$, for all $\tau \neq \sigma, \tau \in$ **DRTyp**;

0 otherwise;

•Unary connective(s): If $\alpha \in \mathbf{Term}_t$:

–Negation: $[\![\neg\alpha]\!]^{M,g} =$

iff $[\![\alpha]\!]^{M,g} = \#$,

1 iff $[\![\alpha]\!]^{M,g} = 0$,

0 otherwise;

•Binary connectives: If $\alpha, \beta \in \mathbf{Term}_t$:

–$[\![\alpha \wedge \beta]\!]^{M,g} =$

iff $[\![\alpha]\!]^{M,g} = \#$ or $[\![\beta]\!]^{M,g} = \#$,

otherwise: 1 iff $[\![\alpha]\!]^{M,g} = 1$ and $[\![\beta]\!]^{M,g} = 1$,

0 otherwise;

–$[\![\alpha \vee \beta]\!]^{M,g} =$

iff $[\![\alpha]\!]^{M,g} = \#$ or $[\![\beta]\!]^{M,g} = \#$,

otherwise: 1 iff $[\![\alpha]\!]^{M,g} = 1$ or $[\![\beta]\!]^{M,g} = 1$, 0 otherwise;

–$[\![\alpha \rightarrow \beta]\!]^{M,g} =$

iff $[\![\alpha]\!]^{M,g} = \#$ or $[\![\beta]\!]^{M,g} = \#$,

otherwise: 1 iff $[\![\alpha]\!]^{M,g} = 0$ or $[\![\beta]\!]^{M,g} = 1$,

0 otherwise;

–$[\![\alpha \leftrightarrow \beta]\!]^{M,g} =$

iff $[\![\alpha]\!]^{M,g} = \#$ or $[\![\beta]\!]^{M,g} = \#$,

otherwise: 1 iff $[\![\alpha]\!]^{M,g} = 1$ and $[\![\beta]\!]^{M,g} = 1$, or $[\![\alpha]\!]^{M,g} = 0$ and $[\![\beta]\!]^{M,g} = 0$

0 otherwise;

•Quantifiers: If $\alpha \in$ **Term**$_t$, $\nu \in$ **Var**$_\tau$, for some $\tau \in$ **Typ**:

–$[\![\exists\nu(\alpha)]\!]^{M,g} =$

iff $\langle d, \# \rangle \in [\![\lambda\nu.\alpha]\!]$, for some $d \in D_\tau^M$,

otherwise: 1 iff $\langle d, 1 \rangle \in [\![\lambda\nu.\alpha]\!]$, for some $d \in D_\tau^M$,

0 otherwise;

–$[\![\forall\nu(\alpha)]\!]^{M,g} =$

iff $\langle d, \# \rangle \in [\![\lambda\nu.\alpha]\!]$, for some $d \in D_\tau^M$,

otherwise: 1 iff $\langle d, 1 \rangle \in [\![\lambda\nu.\alpha]\!]$, for all $d \in D_\tau^M$,

0 otherwise;

•Set relation(s) and operators: If $\alpha, \beta \in$ **Term**$_{(\tau,t)}$, for some $\tau \in$ **Typ**

–$[\![\alpha \subseteq \beta]\!]^{M,g} = [\![\forall x_\tau(\alpha(x) \rightarrow \beta(x))]\!]^{M,g}$

–$[\![\overline{\alpha}]\!]^{M,g} = \{d \mid d \in D^M_\sigma \wedge [\![\alpha]\!]^{M,g}(d) = 0\}$

–$[\![\alpha \cap \beta]\!]^{M,g} = \{d \mid d \in D^M_\sigma \wedge [\![\alpha]\!]^{M,g}(d) = 1 \wedge [\![\beta]\!]^{M,g}(d) = 1\}$

–$[\![\alpha \cup \beta]\!]^{M,g} = \{d \mid d \in D^M_\sigma \wedge ([\![\alpha]\!]^{M,g}(d) = 1 \vee [\![\beta]\!]^{M,g}(d) = 1)\}$

e. Truth

•A formula $\varphi \in$ **Term**$_t$ is *true* in M relative to g iff $[\![\phi]\!]^{M,g} = 1$.

•A formula $\varphi \in$ **Term**$_t$ is *true* in M iff it is true in M relative to any assignment g.

f. Discourse-truth and discourse-acceptability:

•A dynamic formula $D \in$ **Term**$_{s(st)}$ is discourse-true relative to an input assignment i_s and model M, iff $D(i)(j)$ is *true* in M for some $j \in$ **Term**$_s$.

•A dynamic formula $D \in$ **Term**$_{s(st)}$ is discourse-acceptable relative to an input assignment i_s and model M, iff $[\![D(i)(j)]\!]^M \neq \#$ for some $j \in$ **Term**$_s$.

A.2 Dynamic Semantics Abbreviations

(5) Conditions (type st):

a. Dynamic predicates with their arguments as DRS conditions:

$R_\phi\{v_1, \ldots, v_n\} :=$

$\lambda i_s.\forall w(w \in \phi(i) \rightarrow R(v_1(i), \ldots, v_n(i))(w))$

for any non-logical constant R of type $e^n t$,

where $e^n t$ is defined as follows: $e^0 t := t$ and $e^{m+1} t := e(e^m t)$;

for $\upsilon_1, \ldots, \upsilon_n \in \mathbf{Term}_{se}$, $\phi \in \mathbf{Term}_{st}$

b. Relations over drefs:

(i) Dynamic equality:

$\alpha \equiv \beta :=$

$\lambda i_s.\alpha(i) = \beta(i)$

for $\alpha, \beta \in \mathbf{Term}_\tau$, and $\tau \in \mathbf{DrTyp}$

(ii) Dynamic inclusion:

$\phi_1 \Subset \phi_2 :=$

$\lambda i_s.\phi_1(i) \subseteq \phi_2(i)$

for $\phi_1, \phi_2 \in \mathbf{Term}_{s(wt)}$

(6) Discourse constants for proper names:

a. $\mathcal{M}ary_{\mathbf{e}} := \lambda i_s.mary_e$

b. $\mathcal{S}ue_{\mathbf{e}} := \lambda i_s.sue_e$

(7) Special drefs for propositions

a. $\phi_{DC_A\mathbf{w}}$: The set of worlds that refers is compatible with the discourse commitments of A

(8) DRSs (type $s(st)$):

a. Variable Update:

$[\delta] := \lambda i_s.\lambda j_s.i[\delta]j$, for $\delta \in$ **Term**$_{s\tau}$ and $\tau \in$ **STyp**

b. Conditions:

$[C] := \lambda i_s.\lambda j_s.i = j \wedge C(j)$ for $C \in$ **Term**$_{st}$

(9) Relations over DRSs:

a. Dynamic conjunction as function composition:

$\mathcal{D}_1;\mathcal{D}_2 := \lambda i_s.\lambda j_s.\exists h_s[\mathcal{D}_1(i)(h)](\mathcal{D}_2(h)(j)),$

where $\mathcal{D}_1, \mathcal{D}_2 \in$ **Term**$_{s(st)}$

(10) Multiple drefs and multiple conditions:

a. $[C_1, \ldots, C_n] := \lambda i_s.\lambda j_s.i = j \wedge C_1(j) \wedge \cdots \wedge C_2(j),$

where $C_1, \ldots, C_n \in$ **Term**$_{st}$

b. $[\delta_1, \ldots, \delta_n] := [\delta_1]; \ldots; [\delta_n],$

where $\delta \in$ **Term**$_\sigma$, for any $\sigma \in$ **DRTyp**

(11) DRS notation: $[\delta_1, \ldots, \delta_n \mid C_1, \ldots, C_n] := \lambda i_s.\lambda j_s.([\delta_1, \ldots, \delta_n]; [C_1, \ldots, C_n])(i)(j)$

…where $\delta \in$ **Term**$_\sigma$, for any $\sigma \in$ **DRTyp** and $C_1, \ldots, C_n \in$ **Term**$_{st}$

A.3 Compositional Translations

(12) Defining dynamic ‘Meta-types’ for a intensional dynamic system

a. **e** is short for $s(we)$ (drefs referring to entities)

b. **w** is short for $s(wt)$ (drefs referring to sets of worlds/truth-values)

c. **t** is short for $s(st)$ (dynamic correspondent of truth-values)

(13) Proper Names:

a. Mary $\rightsquigarrow \lambda P_{\mathbf{e(wt)}}.\lambda\phi_{\mathbf{w}}.[\upsilon \mid \upsilon \equiv Mary_{\mathbf{e}}]; P(\upsilon)(\phi)$

(where $Mary = \lambda i_s.mary_e$)

b. Sue $\rightsquigarrow \lambda P_{\mathbf{e(wt)}}.\lambda\phi_{\mathbf{w}}.[\upsilon \mid \upsilon \equiv Sue_{\mathbf{e}}]; P(\upsilon)(\phi)$

(where $Sue := \lambda i_s.sue_e$)

(14) DP-Pronouns:

$\text{it}_\upsilon \rightsquigarrow \lambda P_{\mathbf{e(wt)}}.\lambda\phi_{\mathbf{w}}.P(\upsilon)(\phi)$

(15) Unary Predicates:

a. $\text{sleep}_{\mathbf{e(wt)}} \rightsquigarrow \lambda\upsilon_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[Sleep_\phi\{\upsilon\}]$

b. $\text{tired}_{\mathbf{e(wt)}} \rightsquigarrow \lambda\upsilon_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[Tired_\phi\{\upsilon\}]$

(16) Declarative mood:

$\text{DEC}_S := \lambda\mathcal{P}_{\mathbf{wt}}.[\phi_{1\mathbf{w}} \mid \phi_{DC_S} \Subset \phi_1]; \mathcal{P}(\phi_1)$

Appendix B

Formal Appendix to Chapters 4 & 5

This appendix contains additional and/or updated definitions for the analysis of propositions and propositional anaphora in **Chapters 4 & 5** (supplementing/updating the definitions in **Appendix A.**)

B.1 Static Definitions

(1) a. Assertive Information States:

$[\![\textsc{Assertive}_x(w)]\!]^{M,g} =$

$\{w' \in D_w : w'$ conforms to the (communicative) commitments of $[\![x]\!]^{M,g}$ in $[\![w]\!]^{M,g}\}$

b. 'Prove' as veridical assertive attitude:

$[\![\mathrm{prove}(x_e)(p_{wt})(w_w)]\!]^{M,g} = [\![\textsc{Assertive}_x(w) \subseteq p \land p(w)]\!]^{M,g}$

c. 'True' as vacuous (veridical) propositional embedder:

$[\![\mathtt{true}((p_{wt})(w_w))]\!]^{M,g} = [\![p(w)]\!]^{M,g}$

d. 'Mistaken' as anti-veridical assertive attitude:

$[\![\mathit{mistaken}(x_e)(p_{wt})(w_w)]\!]^{M,g} = [\![\textsc{Assertive}_x \subseteq p \wedge \neg p(w)]\!]^{M,g}$

e. Belief states:

$[\![\mathrm{Dox}_x(w)]\!]^{M,g} = \{w' \in D_w : w'$ conforms to what $[\![x]\!]^{M,g}$ believes in $[\![w]\!]^{M,g}\}$

f. 'Believe' as non-veridical doxastic attitude:

$[\![\mathit{believe}(x_e)(p_{wt})(w_w)]\!]^{M,g} =$

$[\![\mathrm{Dox}_x(w) \subseteq p]\!]^{M,g}$

B.2 Dynamic Semantic Abbreviations

(2) Conditions (type st):

a. Relations over drefs:

(i) Dynamic complementation:

$\overline{\overline{\phi_{\mathbf{w}}}} := \lambda i_s.\overline{\phi(i)}$

for $\phi \in \mathbf{Term}_{s(wt)}$

(ii) Dynamic intersection:

$\phi_1 \Cap \phi_2 := \lambda i_s.\phi_1(i) \cap \phi_2(i)$

for $\phi_1, \phi_2 \in \mathbf{Term}_{s(wt)}$

(iii) Dynamic union:

$\phi_1 \Cup \phi_2 := \lambda i_s.\phi_1(i) \cup \phi_2(i)$

for $\phi_1, \phi_2 \in \mathbf{Term}_{s(wt)}$

(3) Relations over DRSs:

a. Maximization over propositional drefs:

$\mathbf{max}_{\phi}(D) :=$

$\lambda i.\lambda j.D(i)(j) \wedge \forall k[D(i)(k)](\neg(\phi(j) \subset \phi(k)))$

for $\phi \in \mathbf{Term}_{s(wt)}, D \in \mathbf{Term}_{s(st)}$

B.3 Compositional Translations

(4) Declarative mood:

$\text{DEC}_S : \lambda \mathcal{P}_{\mathbf{wt}}.[\phi \mid \phi_{DC_S} \Subset \phi]; \mathbf{max}_{\phi}(\mathcal{P}(\phi))$

(5) Propositional Operators:

a. Negation:

$\text{NOT} \rightsquigarrow \lambda \mathcal{P}_{(\mathbf{wt})}.\lambda \phi_{\mathbf{w}}.[\phi'_{\mathbf{w}} \mid \phi \equiv \overline{\phi'}]; \mathbf{max}_{\phi'}(\mathcal{P}(\phi'))$

b. Conjunction:

$\text{AND} \rightsquigarrow$

$\lambda \mathcal{P}_{1\mathbf{wt}}.\lambda \mathcal{P}_{2\mathbf{wt}}.\lambda \phi_{\mathbf{w}}.[\phi_1, \phi_2 \mid \phi \equiv \phi_1 \Cap \phi_2]; \mathbf{max}_{\phi_1}(P_1(\phi_1)); \mathbf{max}_{\phi_2}(P_2(\phi_2))$

c. Disjunction:

$\text{OR} \rightsquigarrow$

$\lambda \mathcal{P}_{1\mathbf{wt}}.\lambda \mathcal{P}_{2\mathbf{wt}}.\lambda \phi_{\mathbf{w}}.[\phi_1, \phi_2 \mid \phi \equiv \phi_1 \Cup \phi_2]; \mathbf{max}_{\phi_1}(P_1(\phi_1)); \mathbf{max}_{\phi_2}(P_2(\phi_2))$

(6) Unary Predicates:

a. $\text{dance}_{\mathbf{e}(\mathbf{wt})} \rightsquigarrow \lambda \upsilon_{\mathbf{e}}.\lambda \phi_{\mathbf{w}}.[Dance_{\phi}\{\upsilon\}]$

(7) Binary predicates:

a. own $\rightsquigarrow \lambda\upsilon_{\mathbf{e}}.\lambda\upsilon'_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[\mathit{Own}_{\phi}\{\upsilon', \upsilon\}]$

(8) Propositional Attitudes:

a. proved $\rightsquigarrow \lambda\upsilon_{\mathbf{e}}.\lambda\mathcal{P}_{\mathbf{wt}}.\lambda\phi_{\mathbf{w}}.[\phi'_{\mathbf{w}} \mid \mathit{Prove}_{\phi}\{\upsilon, \phi'\}]; \mathbf{max}_{\phi'}(\mathcal{P}(\phi'))$

Appendix C

Formal Appendix to Chapter 7

This appendix contains additional and/or updated definitions for the analysis of indefinites and individual anaphora in **Chapter 7** (supplementing/updating the definitions in **Appendices A & B.**)

C.1 Dynamic Semantics Abbreviations

(1) DRS conditions (type st):

a. Dynamic predicates with their arguments as DRS conditions:

$\mathrm{R}_{\phi}\{\upsilon\} := \lambda i_s.\forall w_w.\phi(i)(w) \rightarrow \mathrm{R}(\upsilon(i)(w))(w)$,

for $\mathrm{R} \in \mathbf{Term}_{e(wt)}$, $\upsilon \in \mathbf{Term}_{s(we)}$, $\phi \in \mathbf{Term}_{s(wt)}$

(2) DRSs (type $s(st)$):

a. Relative Variable Update:

The relation $i[\phi : \upsilon]j$ holds (for assignments i_s, j_s, and the drefs $\phi \in$ $\mathbf{Term}_{s(wt)}$, $\upsilon \in \mathbf{Term}_{s(we)}$) iff the conjunction of the following holds:

•$i[\upsilon]j$

•$\forall w_w.(\phi(j)(w) \rightarrow \upsilon'(i)(w) \neq \#$

(an individual referent of υ is determined in each world in $\phi(j)$)

•$\forall w_w.(\neg\phi(j)(w) \rightarrow \upsilon(j)(w) = \#)$

(in each world not in ϕ, υ points to the indeterminate value)

C.2 Compositional Translations

(3) Unary Predicates:

a. bathroom $\rightsquigarrow \lambda\upsilon_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[Bathroom_{\phi}\{\upsilon\}]$

b. weird $\rightsquigarrow \lambda\upsilon_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[Weird_{\phi}\{\upsilon\}]$

c. place $\rightsquigarrow \lambda\upsilon_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[Place_{\phi}\{\upsilon\}]$

(4) Binary predicates:

a. in $\rightsquigarrow$

(i) $\lambda\upsilon_{\mathbf{e}}.\lambda\upsilon'_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.[In_{\phi}\{\upsilon', \upsilon\}]$

(ii) $\lambda\mathcal{Q}_{((\mathbf{e}(\mathbf{wt}))(\mathbf{wt}))}.\lambda\upsilon_{\mathbf{e}}.\lambda\phi_{\mathbf{w}}.$

$\mathcal{Q}(\lambda\upsilon'_{\mathbf{e}}.\lambda\phi'_{\mathbf{w}}.[in_{\phi}\{\upsilon, \upsilon'\}])(\phi)$

(5) Copular *be* $\rightsquigarrow \lambda \mathcal{P}_{\mathbf{wt}}.\mathcal{P}$

(6) Indefinite Determiner:

$\mathrm{a}^{\upsilon} \rightsquigarrow \lambda \mathsf{P}_{\mathbf{e(wt)}}.\lambda \mathsf{P}'_{\mathbf{e(wt)}}.\lambda \phi_{\mathbf{w}}.[\phi : \upsilon]; \mathsf{P}(\upsilon)(\phi); \mathsf{P}'(\upsilon)(\phi)$

Bibliography

Dorit Abusch. Presupposition triggering from alternatives. *Journal of Semantics*, 27(1): 37–80, 2010. Publisher: Oxford University Press.

Maria Aloni. Conceptual covers in dynamic semantics. *Logic, language and computation*, 3:23–48, 2000. Publisher: CSLI.

Maria D. Aloni. *Quantification under conceptual covers*. Institute for Logic, Language and Computation, 2001.

Patrícia Amaral. Entailment, assertion, and textual coherence: the case of *almost* and *barely*. *Linguistics*, 48(3):525–545, 2010.

Patrícia Matos Amaral. *The meaning of approximative adverbs: Evidence from European Portuguese*. The Ohio State University, 2007.

Pranav Anand and Valentine Hacquard. Epistemics with attitude. In *Semantics and linguistic theory*, volume 18, pages 37–54, 2008.

Pranav Anand and Valentine Hacquard. Factivity, belief and discourse. *The art and craft of semantics: A festschrift for Irene Heim*, 1:69–90, 2014. Publisher: MIT Working Papers in Linguistics Cambridge, MA.

Pranav Anand, Daniel Hardt, and James McCloskey. The Santa Cruz sluicing data set. *Language*, 97(1):e68–e88, 2021. Publisher: Linguistic Society of America.

Nicholas Asher. *Reference to Abstract Objects in Discourse*, volume 50 of *Studies in Linguistics and Philosophy,*. Springer Netherlands, Dordrecht, 1993.

Nicholas Asher and Alex Lascarides. *Logics of conversation*. Cambridge University Press, 2003.

Nicholas Asher and Elin McCready. Were, would, might and a compositional account of counterfactuals. *Journal of semantics*, 24(2):93–129, 2007. Publisher: Oxford University Press.

Dale J. Barr, Roger Levy, Christoph Scheepers, and Harry J. Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3):255–278, 2013. Publisher: Elsevier.

Douglas Bates, Martin Maechler, Ben Bolker, Steven Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, Bin Dai, Fabian Scheipl, Gabor Grothendieck, Peter Green, John Fox, Alexander Bauer, and Pavel N. Krivitsky (shared copyright on simulate.formula). lme4: Linear Mixed-Effects Models using 'Eigen' and S4, June 2021. URL https://CRAN.R-project.org/package=lme4.

Maria Bittner. Tense, mood, and centering. 2009.

Maria Bittner. Time and modality without tenses or modals. In Renate Musan and Monika Rathert, editors, *Tense across languages*, pages 147–188. Niemeyer, 2011.

Adrian Brasoveanu. *Structured Nominal and Modal Reference*. Doctoral dissertation, Rutgers University, 2006.

Adrian Brasoveanu. Donkey pluralities: Plural Information States Versus Non-Atomic Individuals. *Linguistics and philosophy*, 31(2):129–209, 2008. Publisher: Springer.

Adrian Brasoveanu. Decomposing Modal Quantification. *Journal of Semantics*, 27(4): 437–527, 2010a.

Adrian Brasoveanu. Structured anaphora to quantifier domains. *Information and Computation*, 208(5):450–473, 2010b.

Adrian Brasoveanu and Donka F. Farkas. How indefinites choose their scope. *Linguistics and philosophy*, 34(1):1–55, 2011. Publisher: Springer.

Adrian Brasoveanu, Donka Farkas, and Floris Roelofsen. N-words and sentential negation: Evidence from polarity particles and VP ellipsis. *Semantics & Pragmatics*, 6: 1–33, 2013.

Adrian Brasoveanu, Karen De Clercq, Donka Farkas, and Floris Roelofsen. Question Tags and Sentential Negativity. *Lingua*, 145:173–193, 2014.

Sylvain Bromberger. *On What We Know We Don't Know: Explanation, Theory, Linguistics, and How Questions Shape Them*. The University of Chicago Press: Chicago and London, 1992.

Paul-Christian Bürkner. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80(1):1–28, 2017.

Rudolf Carnap. *Meaning and necessity: a study in semantics and modal logic*. University of Chicago Press, Chicago, 1947.

Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76(1): 1–32, 2017.

Simon Charlow. A modular theory of pronouns and binding. *Logic and engineering of natural language semantics (LENLS)*, 14, 2017.

Herbert H. Clark. *Semantics and Comprehension*. De Gruyter Mouton, 1976.

Berry Claus, A. Marlijn Meijer, Sophie Repp, and Manfred Krifka. Puzzling response particles: An experimental study on the German answering system. *Semantics & Pragmatics*, 10(19), 2017.

Chris Collins and Paul M. Postal. Disentangling two distinct notions of NEG raising. *Semantics and Pragmatics*, 11:5, 2018.

Robin Cooper and Jonathan Ginzburg. Negation in dialogue. *Proceedings of SemDial*, 15:130–139, 2011. Publisher: Citeseer.

Robin Cooper and Jonathan Ginzburg. Negative inquisitiveness and alternatives-based negation. In *Logic, Language and Meaning*, pages 32–41. Springer, 2012.

Mark Davies. The Corpus of Contemporary American English: 520 million words, 1990-present, 2008. Available online at http://corpus.byu.edu/coca/.

Marie-Catherine de Marneffe, Christopher D. Manning, and Christopher Potts. Did It Happen? The Pragmatic Complexity of Veridicality Assessment. *Computational Linguistics*, 38(2):301–333, February 2012. Publisher: MIT Press.

Henriëtte De Swart. Scope Ambiguities with Negative Quantifiers. In Klaus von Heusinger and Urs Egli, editors, *Reference and Anaphoric Relations*, Studies in Linguistics and Philosophy, pages 109–132. Springer Netherlands, Dordrecht, 2000. doi: 10.1007/978-94-011-3947-2_6.

Judith Degen and Judith Tonhauser. Prior beliefs modulate projection. 2021. Publisher: PsyArXiv.

P. J. E. Dekker. Transsentential meditations: Ups and downs in dynamic semantics. *Linguistics*, 14:65–117, 1993.

Veena D. Dwivedi, Natalie A. Phillips, Maude Laguë-Beauvais, and Shari R. Baum. An electrophysiological study of mood, modal context, and anaphora. *Brain Research*, 1117(1):135–153, October 2006. ISSN 0006-8993.

Veena D. Dwivedi, John E. Drury, Monika Molnar, Natalie A. Phillips, Shari Baum, and Karsten Steinhauer. ERPs reveal sensitivity to hypothetical contexts in spoken discourse. *Neuroreport*, 21(11):791–795, 2010. Publisher: LWW.

Patrick D. Elliott. Towards a principled logic of anaphora. December 2020.

Patrick D Elliott. Externally-Dynamic Dynamic Semantics, 2021. URL `https://patrl.keybase.pub/handouts/ens.pdf`.

Donka Farkas and Kim Bruce. On reacting to assertions and polar questions. *Journal of Semantics*, 27(1):81–118, 2010.

Donka F. Farkas. Mood choice in complement clauses. *Approaches to Hungarian*, 4: 207–225, 1992.

Donka F. Farkas. Specificity and Scope. In *in L. Nash and G. Tsoulas (eds), Langues et Grammaire 1*, pages 119–137, 1994.

Donka F. Farkas. Dependent indefinites. In *Empirical issues in formal syntax and semantics*. Citeseer, 1997.

Donka F. Farkas. Specificity distinctions. *Journal of semantics*, 19(3):213–243, 2002. Publisher: Oxford University Press.

Donka F. Farkas and Adrian Brasoveanu. Kinds of (Non) Specificity. *The Wiley Blackwell Companion to Semantics*, pages 1–26, 2020. Publisher: Wiley Online Library.

Donka F. Farkas and Floris Roelofsen. Polarity particles revisited. *Semantics and Pragmatics*, 12(0):15, November 2019. Number: 0.

Justin Fitzpatrick. The whys and how comes of presupposition and NPI licensing in questions. In *Proceedings of the 24th west coast conference on formal linguistics*, volume 24, pages 138–145. Citeseer, 2005.

Anette Frank. *Context Dependence in Modal Constructions*. Doctoral dissertation, University of Stuttgart, 1996.

Anette Frank and Hans Kamp. On context dependence in modal constructions. In *Semantics and linguistic theory*, volume 7, pages 151–168, 1997.

Jon Gajewski. Neg-Raising and Polarity. *Linguistics and Philosophy*, 30:289–328, 2007.

Daniel Gallin. *Intensional and Higher-Order Modal Logic*, volume 19 of *North-Holland Mathematics Studies*. North-Holland, Amsterdam, 1975.

Bart Geurts. On No. *Journal of Semantics*, 13(1):67–86, January 1996.

Bart Geurts. *Presuppositions and pronouns*. Brill, 1999.

Anastasia Giannakidou. *Polarity Sensitivity as (Non)Veridical Dependency*. John Benjamins Publishing Company, Amsterdam, 1998.

Jonathan Ginzburg and Ivan A. Sag. *Interrogative investigations: The Form, Meaning, and Use of English Interrogatives*. Number 123 in CSLI lecture notes. CSLI Publications, Stanford, Calif, 2000.

Ben Goodrich, Jonah Gabry, Imad Ali, and Sam Brilleman. rstanarm: Bayesian Applied Regression Modeling via Stan, 2020. URL https://mc-stan.org/rstanarm/.

Matthew Gotham. Double Negation, Excluded Middle and Accessibility in Dynamic Semantics. In *Proceedings of the 22ˆ{nd} Amsterdam Colloquium*, Amsterdam, 2019.

Vera Gribanova. Head movement and ellipsis in the expression of Russian polarity focus. *Natural Language and Linguistic Theory*, pages 1–43, 2015.

Paul Grice. Logic and Conversation. In *Studies in the Way of Words*. Harvard university press, Cambridge, 1975.

Jane Grimshaw. Locality and Extended Projection. In Everaert Martin Coopmans, Peter and Jane Grimshaw, editors, *Lexical Specification and Insertion*, number 197 in Current Issues in Linguistic Theory. John Benjamins Publishing Company, Philadephia, PA, 2000.

Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100, February 1991.

Jeroen Antonius Gerardus Groenendijk and Martin Johan Bastiaan Stokhof. *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD Thesis, Univ. Amsterdam, 1984.

Christine Gunlogson. *True to Form : Rising and Falling Declaratives as Questions in English*. Doctoral dissertation, University of California, Santa Cruz, June 2004.

Javier Gutiérrez-Rexach. The semantic basis of NPI licensing in questions. *MIT Working Papers in Linguistics*, 31:359–376, 1997.

Liliane Haegeman and Raffaella Zanuttini. Negative heads and the Neg Criterion. 1991. Publisher: Walter de Gruyter, Berlin/New York Berlin, New York.

Charles L. Hamblin. Questions in montague english. In *Montague grammar*, pages 247–259. Elsevier, 1976.

Jorge Hankamer and Ivan Sag. Deep and Surface Anaphors. *Linguistic Inquiry*, 7: 391–426, 1976.

Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation, University of Massachusetts Amherst, 1982.

Irene Heim. File Change Semantics and the Familiarity Theory of Definiteness. In A. von Stechow, R. Bauerle, and C. Schwarz, editors, *Meaning, Use and Interpretation*. Berlin: W. de Gruyter, 1983.

Irene Heim. Presupposition Projection and the Semantics of Attitude Verbs. *Journal of Semantics*, 9(3):183–221, 1992.

Lisa Hofmann. Why not? – Polarity ellipsis in *why*-questions. 2018.

Lisa Hofmann. The anaphoric potential of indefinites under negation and disjunction. In *Proceedings of the 22nd Amsterdam Colloquium*, pages 181–190, 2019a.

Lisa Hofmann. Sentential negativity and polarity-sensitive anaphora. In *ESSLLI 2019 Student Session*, 2019b.

Anders Holmberg. The syntax of answers to polar questions in English and Swedish. *Lingua*, 128:31–50, 2013.

Vincent Homer. Neg-raising and positive polarity: The view from modals. *Semantics and Pragmatics*, 8:4–1, 2015.

Laurence R. Horn. *A Natural History of Negation*. University of Chicago Press), 1989.

Laurence R. Horn. Assertoric inertia and NPI licensing. In *CLS*, volume 38, 2002.

Laurence R. Horn. *Almost* forever. In *Pragmatics and autolexical grammar: In honor of Jerry Sadock*, pages 3–21. 2011.

Laurence R. Horn. Licensing NPIs: Some Negative (and Positive) Results. In P. Larrivée and C. Lee, editors, *Negation and Polarity: Experimental Perspectives*, volume 1 of *Language, Cognition and Mind*. Springer International Publishing Switzerland, 2016.

Sabine Iatridou and Ivy Sichel. Negative DPs, A-Movement, and Scope Diminishment. *Linguistic Inquiry*, 42(4):595–629, 2011.

Ray Jackendoff. An Interpretive Theory of Negation. *Foundations of Language*, 5(2): 218–241, 1969.

Hans Kamp. A theory of truth and semantic representation. *Formal semantics-the essential readings*, pages 189–222, 1981.

Hans Kamp and Uwe Reyle. *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers, 1993.

Lauri Karttunen. Presuppositions of Compound Sentences. *Linguistic Inquiry*, 4:169–93, 1973.

Lauri Karttunen. Discourse Referents. In *Notes from the linguistic underground*, pages 363–385. Brill, 1976.

Lauri Karttunen. Syntax and semantics of questions. *Linguistics and philosophy*, 1(1): 3–44, 1977. Publisher: Springer.

Edward L. Keenan and Robert D. Hull. The logical presuppositions of questions and answers. 1973.

Rodger Kibble. Dynamics of epistemic modality and anaphora. page 11, Tilburg, 1994. ITK.

Rodger Kibble. Complement anaphora and dynamic binding. In *Semantics and Linguistic Theory*, volume 7, pages 258–275, 1997.

Jeffrey C. King and Karen S. Lewis. Anaphora. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018. URL https://plato.stanford.edu/archives/fall2018/entries/anaphora/.

Edward Klima. Negation in English. In J. Fodor and J. Katz, editors, *The structure of language*, pages 246–32. Prentice Hall, New Jersey, 1964.

Emiel Krahmer and Reinhard Muskens. Negation and Disjunction in Discourse Representation Theory. *Journal of Semantics*, 12(4):357–376, 1995.

Ruth Kramer and Kyle Rawlins. Polarity particles: An Ellipsis Account. In Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman, editors, *Proceedings of NELS*, 2009.

Manfred Krifka. Response Particles as Propositional Anaphors. *Semantics and Linguistic Theory (SALT)*, 23:1–18, 2013.

Margaret Kroll. Polarity Reversals under Sluicing. *Semantics and Pragmatics*, 12(0): 18, November 2019.

Margaret Kroll and Deniz Rudin. Identity and interpretation: Syntactic and pragmatic constraints on the acceptability of sluicing. In Andrew Lamont and Katerina Tetzloff, editors, *Proceedings of the 47th annual meeting of the North East Linguistics Society. Amherst, MA: Graduate Linguistics Student Association*, pages 177–190, 2017.

Margaret Ilona Kroll. *Comprehending Ellipsis*. Doctoral dissertation, UC Santa Cruz, 2020.

William Ladusaw. *Polarity Sensitivity as Inherent Scope Relations, Outstanding Dissertations in Linguistics*. Garland Publishing, New York, 1980.

William A. Ladusaw. Expressing negation. In *Semantics and linguistic theory*, volume 2, pages 237–260, 1992.

Itziar Laka. *Negation in Syntax: On the nature of functional categories and projections*. PhD Thesis, Massachusetts Institute of Technology, 1990.

Fred Landman. *Towards a Theory of Information: The Status of Partial Objects in Semantics*. Foris Publications, 1986.

John M. Lawler. *any Questions?* 1971.

David Lewis. Scorekeeping in a language game. In *Semantics from different points of view*, pages 172–187. Springer, 1979.

Karen S. Lewis. Anaphora and negation. *Philosophical Studies*, 178(5):1403–1440, May 2021.

Sebastian Löbner. Polarity in Natural Language: Predication, Quantification and Negation in Particular and Characterizing Sentences. *Linguistics and Philosophy*, 23(3): 213–308, June 2000.

Matthew Mandelkern. Witnesses. *Linguistics and Philosophy*, pages 1–27, 2022. Publisher: Springer.

Jim McCloskey. Ellipsis, Polarity, and the Cartography of Verb-Initial Orders in Irish. In Enoch Aboh, Eric Haeberli, Genoveva Puskás, and Manuela Schönenberger, editors, *Elements of Comparative Syntax: Theory and Description*, pages 99–151. De Gruyter Mouton, 2017a.

Jim McCloskey. The Expression of Polarity Lecture Notes. 2017b.

A. Marlijn Meijer, Berry Claus, Sophie Repp, and Manfred Krifka. Particle responses to negated assertions: Preference patterns for German ja and nein. In *Proceedings of the 20th amsterdam colloquium*, pages 286–295. Univ. Amsterdam Amsterdam, 2015.

Jason Merchant. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press, 2001.

Jason Merchant. Fragments and Ellipsis. *Linguistics and Philosophy*, 27:661–738, 2004.

George Edward Moore. A reply to my critics. In P. A. Schlipp, editor, *The philosophy of GE Moore*, pages 660–667. Northwestern University, 1942.

Linda M. Moxey. Effects of what is expected on the focussing properties of quantifiers: A test of the presupposition-denial account. *Journal of Memory and Language*, 55 (3):422–439, October 2006.

Linda M. Moxey, Anthony J. Sanford, and Eugene J. Dawydiak. Denials as Controllers of Negative Quantifier Focus. *Journal of Memory and Language*, 44(3):427–442, April 2001.

Linda M. Moxey, Ruth Filik, and Kevin B. Paterson. On-line effects of what is expected on the resolution of plural pronouns. *Language and Cognitive Processes*, 24(6):843–875, 2009. Publisher: Taylor & Francis.

Sarah E. Murray. Varieties of update. *Semantics and Pragmatics*, 7(0):2–1–53, 2014.

Sarah E. Murray. Evidentiality and Illocutionary Mood in Cheyenne. *International Journal of American Linguistics*, 82(4):487–517, 2016.

Reinhard Muskens. Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy*, 19(2):143–186, 1996.

Rick Nouwen. Complement Anaphora and Interpretation. *Journal of Semantics*, 20(1): 73–113, February 2003.

Doris Penka. Uninterpretable negative features on negative indefinites. In *Proceedings of the 16th Amsterdam Colloquium*, pages 19–22. Citeseer, 2007.

Doris Penka. *Negative Indefinites*. Oxford Studies in Theoretical Linguistics. Oxford University Press, Oxford, New York, 2010.

Emily Pope. *Questions and answers in English*. PhD Thesis, Diss. Massachusetts Institute of Technology, 1972.

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

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

Floris Roelofsen and Donka Farkas. Polarity particle responses as a window onto the interpretation of questions and assertions. *Language*, 91(2):359–414, 2015.

Deniz Rudin. *Rising above commitment*. Doctoral dissertation, University of California, Santa Cruz, 2018.

Anthony J. Sanford, Eugene J. Dawydiak, and Linda M. Moxey. A Unified Account of Quantifer Perspective Effects in Discourse. *Discourse Processes*, 44(1):1–32, June 2007.

Scott A. Schwenter. Pragmatic variation between negatives: Evidence from Romance. *Selected Papers from NWAV 30*, 8(3):17, 2002.

C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948. Conference Name: The Bell System Technical Journal.

Todd Nathaniel Snider. *Anaphoric Reference to Propositions*. Doctoral dissertation, Cornell University, 2017.

Robert Stalnaker. Assertion. In P. Cole, editor, *Pragmatics*, number 9 in Syntax and Semantics, pages 315–332. London: Academic Press, 1978.

Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25:701–721, 2002.

Robert Stalnaker, Milton K. Munitz, and Peter Unger. Pragmatic presuppositions. In *Proceedings of the Texas conference on performatives, presuppositions, and implicatures. Arlington, VA: Center for Applied Linguistics*, pages 135–148, 1977.

Mark Steedman. Information Structure and the Syntax-Phonology Interface. *Linguistic Inquiry*, 31(4):649–689, 2000.

Erik Stenius. Mood and Language-Game. *Synthese*, 17(3):254–274, 1967.

Matthew Stone. The anaphoric parallel between modality and tense. *Technical Reports (CIS)*, page 177, 1997.

Matthew Stone. Reference to possible worlds. *RuCCS Report*, 46, 1999.

Matthew Stone and Daniel Hardt. Dynamic discourse referents for tense and modals. In *Computing meaning*, pages 301–319. Springer, 1999.

R Core Team. R: A Language and Environment for Statistical Computing, 2014. URL https://www.r-project.org/.

Stan Development Team. RStan: the R Interface to Stan, 2020. URL https://mc.stan.org/rstan/.

Judith Tonhauser. Prosodic cues to presupposition projection. In *Semantics and Linguistic Theory*, volume 26, pages 934–960, 2016.

Judith Tonhauser, David I. Beaver, and Judith Degen. How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics*, 35(3): 495–542, 2018. Publisher: Oxford University Press.

Martin van den Berg. *Some Aspects of the Internal Structure of Discourse: The Dynamics of Nominal Anaphora*. PhD Thesis, dissertation, Universiteit van Amsterdam, 1996.

Frank Veltman. Logics for conditionals. *Studia Logica*, 46(2), 1987.

Frank Veltman. Defaults in update semantics. *Journal of philosophical logic*, 25(3): 221–261, 1996. Publisher: Springer.

P. C. Wason. The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, 4(1):7–11, February 1965.

Chris Wilder. Small clauses and related objects. *GAGL: Groninger Arbeiten zur germanistischen Linguistik*, (34):215–236, 1992.

Ludwig Wittgenstein. *Philosophische Untersuchungen*. 1952.

Juhani Yli-Vakkuri and James McGilvray. Reference and Extension. 2010.

Rafaella Zanuttini. *Negation and clausal structure*. New York and Oxford: Oxford University Press, 1997.

Hedde Zeijlstra. On the syntactically complex status of negative indefinites. *The Journal of Comparative Germanic Linguistics*, 14(2):111–138, 2011. Publisher: Springer.

Hedde Zeijlstra. Negation and Negative Dependencies. *Annual Review of Linguistics*, 2:233–254, 2016.

Frans Zwarts. Nonveridical Contexts. *Linguistic Analysis*, 25(3-4):286–312, 1995.

Frans Zwarts. Facets of negation. *Quantifiers, logic and language*, pages 385–421, 1996. Publisher: CSLI Publications.