# assertion
# &
# rejection

julian j schlöder

# Assertion and Rejection

Julian J. Schlöder

# Assertion and Rejection

# Assertion and Rejection

**Promotiecommisie**

| | | |
|---|---|---|
| Promotor: | Prof. dr. ing. R.A.M. van Rooij | Universiteit van Amsterdam |
| Co-promotor: | Dr. R. Fernández Rovira | Universiteit van Amsterdam |
| Co-promotor: | Prof. dr. A. Lascarides | University of Edinburgh |
| | | |
| Overige leden: | Prof. dr. F.J.M.M. Veltman | Universiteit van Amsterdam |
| | Prof. dr. A. Betti | Universiteit van Amsterdam |
| | Dr. P.J.E. Dekker | Universiteit van Amsterdam |
| | Prof. dr. M. Steedman | University of Edinburgh |
| | Prof. dr. M. Stone | Rutgers University |
| | Prof. dr. Y. Venema | Universiteit van Amsterdam |
| | | |
| Faculteit: | Faculteit der Natuurwetenschappen, Wiskunde en Informatica | |

# Contents

## Part One: Logics for Rejection

# Part Two: Linguistic Aspects

# Appendix

# Preface

I distinctly remember a session of *Computational Semantics and Pragmatics* (Winter 13/14) where I presented an idea for a term paper. My proposal was to conduct a small study on the different ways in which speakers, in context, express agreement and disgreement. Immediately, another student opined that this is simple: aren't people just saying *yes* and *no*? As I recall it, I didn't have much to respond to that. After interjecting and saving my face, Raquel Fernández took me under her wings. Over the last four and a half years, she gave me the impetus and the analytic tools I needed to develop my ideas. Without her constant insight and encouragement, nothing I present here would have come to fruition.

As I know now, the matters of *assertion and rejection* are far from simple. In my Master's thesis, I attempted to find dependable definitions of what it *means* to agree and disagree in a speech act theory framework. This initial analysis, as well as the gathering of further data to illuminate the rejection phenomenon, forced a number of observations on me: that *anything can be rejected*; (hence) rejection is *not about truth*; (hence) *dissent is not negative assent*; and *prosody matters*. Uncomfortable as these claims may be, I have learned to welcome them. To carve out a stance that embraces them without reservation, I have written this thesis.

I hope to have succeeded in presenting this stance in a way that can be discerned independently of the formal theories I develop. While large parts of this thesis *are* about the construction of particular formalisms, my linguistic analyses can, or so I hope, be appreciated *a priori*. I have tried to avoid formalisation for formalisation's sake, gratuitous use of symbols, and other uses of mathematics as staffage. After all, English, for all its failings, can be surprisingly precise when suitably employed. Even though some odd bits of facetiousness made it into the final manuscript, I would hope that these will be taken not as a loss of clarity, but a gain of enjoyment.

Starting out, I took for granted that among the speech acts there is a particular one called *rejection* that lives among its brethren in harmony and with equal rights. This assumption is somewhat highly contentious—its denial, in fact, entrenched. I defend my view in Chapter 2. In this connection, I have greatly benefitted from

"Are you looking forward to the day
when you emerge from these cloistered walls
into what some call the world?"

John Williams, *Stoner*

# Chapter 1

# Introduction

Au contraire mon capitaine!

—Q

What is rejection, and how does it relate to assertion?

An arcane question, for sure, but an important one. Someone engaged in a systematic inquiry will assert some premises, make some hypotheses and in the process may come to reject some hypothesis or other. What is rejected will affect the further course of inquiry. But how?

In the less than systematic natural language, too, proposals are made and some of them are rejected—this is a plain fact, taking a rightful place in JL Austin's early observations about speech acts. To understand how speakers coordinate on their proposals, on what they take for granted, and on what they expect of the future conversation, we need to understand how they reject.

Yet, linguists and philosophers are wont to let assertion take centre stage. The first of them is Gottlob Frege. In his *Die Verneinung* he argues that rejection is no more than negative assertion. Rejection, then, is a chimera, a shadow thrown by assertion, obstructing our clear view. Timothy Smiley in his *Rejection* gives a memorable voice to the opposition: *assertion and rejection [are] distinct activities on all fours with one another.*

Let me nail my colours to the mast. Rejection is a speech act, distinct from and irreducible to assertion, that cancels or precludes the effects of other speech acts. It is principally performed by using the word *no*, though any number of other options are available—including particular inflections of one's voice. I defend that it is *sensible and useful* to countenance rejection as a distinct speech act, describe logical inferences involving rejected alongside asserted content, address puzzles and challenges to classical logic that arise, and model how intonation affects the interpretation of an utterance. The major upshots of these investigations are a new defence of classical logic, an axiomatic treatment of the notion of public commitment, an inferentialist semantics for epistemic modals, a novel theory of intonation, and a defence of classical truth conditions for taste predicates.

## 1.1   Overview

**Outset.**   This thesis is concerned with the speech acts of assertion and rejection and, more broadly, with agreement and disagreement in dialogue. These topics have recently seen increased attention in the literature on semantics, pragmatics and philosophy of language. Core points of interest for this thesis are disagreement about modals or taste (MacFarlane, 2014) disagreements about non-propositional content (Huvenes, forthcoming), logics of non-asserted content (Rumfitt, 2000; Murzi & Carrara, 2015), the influence of rejection moves on dialogue structure (Krifka, 2013), and the influence of intonation (Goodhue & Wagner, 2015). Many of these results appear to indicate that the data on disagreement in dialogue is somewhat at odds with orthodox semantic theory or classical logic. In this thesis, I argue that this appearance is misleading.

I start from the ground up and investigate the *foundations* of what it means to disagree. My discussion starts with an old point by Frege (1919): that it is superfluous to talk of negative judgements (roughly, rejections), since one can just as well speak of positive judgements (roughly, assertions) about negated propositions. Much of the current discussion is very much within that paradigm, as, e.g. disagreement may be defined as one party believing a proposition and another believing its negation (Kölbel, 2004). However, there are some detractors who see a second kind of disagreement that is not reducible to negative assertion (Khoo, 2015).

My allegiance is with this latter camp. I defend, *contra* Frege, that there are disagreement moves in dialogue that are not reducible to negative assertion. I take this to be good grounds for considering a basic speech act of *weak rejection* that covers the breadth of data on disagreement. Such a speech act has been claimed unsystemisable (Dickie, 2010), but I demonstrate—by way of giving a systematisation—that this is not so. This speech act can be defined both in terms of common ground (Stalnaker, 1978) and in terms of public commitment (Brandom, 1983). I prove that these definitions are equivalent in cooperative dialogue.

In the commitment framework I can additionally formulate a *general rejection* that makes sense for non-cooperative settings and moreover accounts for rejections of non-propositional contents. Additionally, Brandom observed that committing means being able to vindicate. I analyse *Why*-questions as one particular way of prompting vindication, embedded in a more general treatment of *giving and eliciting reasons* in Chapter 7. This explains how *Why*-questions can *resist* a prior assertion without outright disagreeing with it (Bledin & Rawlins, 2016).

These insights on rejection can be leveraged and expanded to inferentially explain the epistemic operator *might*. The resulting *epistemic multilateral logic* shows that classical logic can be inferentially motivated as the logic of asserted content. This is unexpected, since a body of evidence seems to indicate that *might* is incompatible with classical inference (Yalcin, 2007; MacFarlane, 2014).

Moreover, in this logic one can follow an old suggestion by Terence Parsons (1984) and *reject* the semantic paradoxes. Another challenge to the orthodoxy of logic is the *faultless disagreement* argument that purports to show that truth can be relative (Kölbel, 2004; MacFarlane, 2014). I apply a theory of commitment to defuse this argument and show that one can remain with classical truth conditions here as well.

Additionally, during my research, I have found that sometimes it is only intonation that makes the difference between agreement and disagreement. This point has sometimes been acknowledged in prior literature (Ladd, 1980; Goodhue & Wagner, 2015), but with little consequence for the formal study of neither disagreement nor intonation. Indeed, some have argued that such intonational contributions are resistant to formal treatment altogether (Bolinger, 1985). I demonstrate that this is not so by developing a novel, formal, general theory of intonation. This theory is motivated by broad and independent data, but in particular accounts for the disagreement related cases.

**Parts and Chapters.** This thesis is separated into two parts. Part One (Chapters 2–4) develops in detail a *logic for rejected content* and defends that classical logic is the logic of asserted content. This part should be primarily interesting to logicians and philosophers, but also to linguists. Part Two (Chapters 5–7) deals with some linguistic aspects surrounding disagreement and should be primarily interesting to linguists, but also to philosophers. It addresses taste disagreements, intonational aspects and *Why*-questions used to doubt an assertion. In detail:

- **Chapter 2** argues that it is useful and possible to develop a logic that includes rejected content. It in particular deals with the arguments of Imogen Dickie (2010) and Gottlob Frege (1919), who doubt that my project is a sensible one. In this chapter I develop *weak bilateral logic*, a logic that countenances rejected content.

- **Chapter 3** re-casts the results of Chapter 2 in greater generality by translating them to a commitment framework. In this chapter, I present a partial axiomatisation of what it means to commit to something in dialogue. These axioms make intuitively compelling predictions about the *projections* that some speech acts make in both cooperative and non-cooperative settings. Based on these predictions, they allow me to formalise *general rejection*. Then, I propose some further axioms that demarcate the line between cooperativity and non-cooperativity. A completeness result demonstrates that the weak bilateral logic of Chapter 2 is equivalent to the *logic that preserves cooperative commitment*. This moreover allows me to explain how weak bilateral logic can also account for strongly rejected content.

- **Chapter 4** extends weak bilateral logic to *epistemic multilateral logic* by adding content that has been *weakly asserted* (the dual speech act to weak rejection). A linguistic study of the adverb *perhaps* is used to motivate a speech act of

*weak assertion* that moreover fills a natural gap in the account of Chapter 2. I then show that epistemic multilateral logic is an inferentialist defence of classical logic: it derives the *classicality of the logic of asserted content* from pure, simple and harmonious inference rules. Moreover, it has the potential to avoid the classical Liar paradox.

- **Chapter 5** addresses the phenomenon of *taste disagreement*. Data on this phenomenon is typically taken to indicate that the classical notion of truth needs to be revised. I argue in this chapter that this would be a too hasty conclusion. By defining a novel linguistic-semantic analysis of taste predicates as *reframable predicates*, I am able to give a formal semantics that explains the phenomenon and is compatible with non-revisionary realism about taste.

- **Chapter 6** presents a new theory of intonation in dialogue. This theory was originally intended to explain cases of disagreement in dialogue that are indistinguishable from agreement moves except for their intonation. However, the analysis in this chapter goes well beyond this specific phenomenon, and instead presents a novel and *general* theory of intonation from the ground up. My main argument shows that one cannot model focus without also considering tune and that prior accounts of focus suffer from not doing so. My own proposal hence models focus and tune *jointly*. One particularly striking case where intonation can make the difference between agreement and disagreement is intonationally marked *verbal irony*. I show how my model accounts for this.

- **Chapter 7** gives a linguistically motivated formal analysis of *Why*-questions in dialogue. On one hand, this expands on Chapter 3, as a comprehensive analysis of commitment in dialogue must also include the commitment to *vindicate* what one has committed to by answering *Why* questions. On the other hand, this also means that *Why*-questions are *resistance moves* that can be a (soft) form of disagreement. I present a comprehensive analysis of *Why* questions in dialogue as *enthymeme elicitors*. I moreover build on the results of Chapter 6 to explain the phenomenon of focus in *Why*-questions.

## 1.2   Remarks about Linguistic Examples

Many of my motivating examples consist of two participants engaged in dialogue. When I construct such an example, I usually have a woman interacting with a man. This is a pragmatic choice to increase readability: it allows me to use gendered pronouns like *he* and *she* to non-ambiguously refer to a speaker. The specific names I give to these fictitious interlocutors are the result of my whimsy or derived from various works of popular culture—sometimes the work that inspired me to construct a particular example. When I use an example from a natural language corpus, however, I repeat the monikers used in that corpus, even if they are just individual letters.

When annotating linguistic examples with acceptability judgements, I use the following (loose) convention, as usual in the linguistics literature: an asterisk (\*) denotes ungrammaticality; a pound sign (#) denotes pragmatic unacceptability; single question marks (?) and double question marks (??) denote weak/strong judgements of relative pragmatic unacceptability (i.e. that an utterance sounds odd or unidiomatic, but not outright infelicitous). This categorisation is, of course, rather vague and already theory-laden. In Chapter 6, for instance, I annotate utterances with misplaced focal stresses as pragmatically unacceptable (#) and then give a pragmatic theory to explain them—a syntactician pursuing a syntactic explanation may well claim ungrammaticality (\*). In addition, these judgements may vary from individual to individual and many arguments, I am sure, can be had about whether ?? or # is the correct annotation for some examples.

I cannot solve these methodological issues here. The judgements stated in this thesis are a synthesis of a literature consensus (where available), my own intuitions, these of my advisors and co-authors, and native speaker informants. Similar remarks apply to my use of the arrow $\rightsquigarrow$ as a catch-all indicator for pragmatic entailment. Again, judgements and theories may differ on whether some utterance entails what I claim it entails—and on whether this entailment is *pragmatic.* In Chapter 6, I offer a potential explanation for why some particular intuitions about focal stresses may differ. There I also introduce my annotation conventions for prosodic tunes.

## 1.3  List of Included Papers

A number of published, soon-to-be published and hopefully-to-be-published papers form the very basis of this dissertation, some of which are the result of joint work.

- Schlöder, JJ & Lascarides, A (2015). Interpreting English pitch contours in context. In: Howes, C & Larsson, S (eds.), *Proceedings of the 19th Workshop on the Semantics and Pragmatics of Dialogue*, 131–139.
  Order of authors follows relative contribution.

- Schlöder, JJ, Breitholtz, E & Fernández, R (2016). Why? In: Hunter, J, Simons, M & Stone, M (eds.), *Proceedings of the 20th Workshop on the Semantics and Pragmatics of Dialogue*, 5–14.
  Order of authors follows relative contribution.

- Schlöder, JJ, Venant, A & Asher, N (2017). Aligning intentions: Acceptance and rejection in dialogue. In: Truswell, R (ed.), *Proceedings of Sinn und Bedeutung 21*, to appear.
  Order of authors follows relative contribution.

- Schlöder, JJ (2017). Towards a formal semantics of verbal irony. In: Howes, C & Rieser, H (eds.), *Proceedings of the Workshop on Formal Approaches to the Dynamics of Linguistic Interaction*, 55–59.

- Incurvati, L & Schlöder, JJ (2017). Weak rejection. *Australasian Journal of Philosophy*, 95(4):741–760.
  Alphabetical order of authors.
- Incurvati, L & Schlöder, JJ (under review). Weak assertion.
  Alphabetical order of authors.
- Schlöder, JJ & Lascarides, A (under review). Understanding focus: Tune, placement and coherence.
  Order of authors follows relative contribution.
- Schlöder, JJ (under review). How to be a realist about taste.

By chapter:

- Chapter 2 is primarily based on *Weak rejection* and partly on *Weak assertion.*
- Chapter 3 is based on *Aligning intentions: Acceptance and rejection in dialogue* and *Weak rejection.*
- Chapter 4 is based primarily on *Weak assertion* and partly on my unpublished work.
- Chapter 5 is based on *How to be a realist about taste.*
- Chapter 6 is based on *Understanding focus: Tune, placement and coherence*, *Interpreting English pitch contours in context* and *Towards a formal semantics of verbal irony.*
- Chapter 7 is based on *Why?* and additional unpublished work with Ellen Breitholtz and Raquel Fernández.

Additionally, I reference *Pragmatic rejection* (Schlöder & Fernández, 2015b), *Clarifying intentions in dialogue: A corpus study* (Schlöder & Fernández, 2015a), *A formal semantics of the final rise* (Schlöder, 2015) and my Master's thesis *Uptake, Clarification and Argumentation* (Schlöder, 2014). The results of these works, however, are not included in this thesis.

# Part One

# Logics for Rejection

# Chapter 2

# Weak Rejection

[A] sign for a rejection opposed to Frege's assertion ... is a futile complication.
All we need in logic ... is the assertion sign and a negation.
[W]hatever is more than these ... cometh of evil.

—Peter Geach, *Assertion*

This chapter presents the foundational and conceptual basis of Part One of this thesis: the speech act *weak rejection* and its attendant *weak bilateral logic*. There are some substantial challenges to get out of the way first. Gottlob Frege (1919) is credited with a convincing argument that seems to show that it is useless to countenance a speech act for rejection in one's logic. In addition, Imogen Dickie (2010) argues that it is not just useless, but indeed impossible to model satisfactorily inferences that involve both asserted and rejected content. I address both of their arguments, following Timothy Smiley's (1996) seminal response to Frege.

I demonstrate that it is possible to give a formal logic in which weakly rejected sentences can appear as premisses and conclusions by constructing such a logic. In this *weak bilateral logic*, the logic of asserted content is precisely classical logic. This logic is hence already a partial defence of classical logic, but some of its inference rules might appear *ad hoc*. However, in Chapter 4 I build on the work of this chapter to present a fully principled defence of classical logic.

This chapter is based on, and draws from, *Weak Rejection* (co-authored with Luca Incurvati), *Weak Assertion* (co-authored with Luca Incurvati).

## Chapter Contents

## 2.1   Rejection and Logic

**Frege on Rejection.**   In his essay *Die Verneinung* (1919), Gottlob Frege argues that—at least where logic is concerned—it is entirely useless to distinguish *affirmative judgements* (*bejahende Urteile*, lit. *yes-ing judgements*) from *negative judgements* (*verneinende Urteile*, lit. *no-ing judgements*). After all, so the story goes, a negative judgement reduces to an affirmative judgement about a negated proposition. And, Frege continues, since a logician *must* reduce further wherever possible, the useless notion of *negative judgement* ought to be cut. Accordingly, if we want to analyse a discourse, we need not separate *force for negation* from *force for affirmation*, since negating a proposition (as an act or judgement) is the same as affirming the negated proposition (1919, p154).

In modern parlance, one can understand Frege as follows. Assertion and rejection are *speech acts* that express *attitudes* towards sentences: assertions express *assent* and rejections express *dissent*.[1] Then, Frege's argument says that it is useless to include primitive logical notions for the speech act rejection or the attitude dissent, since these are reducible to *assertion of a negative* or *assent to a negative*.[2]

A close reading of Frege's essay reveals an ambiguity in this argument. The German term *Verneinung* can mean both *negation* in a logical sense and *negative answer* (to a polar question) in a linguistic sense. Generally, Frege seems to talk about negation. However, on page 153, he explicitly proposes to realise judgements as answers to polar questions; that is, a negative judgement (if there is any) would be a expressed as a *negative answer to a polar question*. Thus, Frege shows awareness of this distinction, but nonetheless seems to presuppose that the different kinds of *Verneinung* must be analysed the same way. That is, Frege's argument tacitly rests on the claim that if *one* of the following *Verneinungen* is a negative judgement, then *all* of them are.

(1)   a.  It is not the case that *p*.

     b.  Is it the case that *p*? No.

The form (1b) is never explicated in *Die Verneinung*; this leaves open that by *negative answer*, Frege also means (1a) (since (1a) is indeed a well-formed answer to a polar question). However, in his later *Gedankengefüge* (Frege, 1923) he does mention bare yes/no answers like (1b) as indicating affirmative/negative judgements.

---

[1]Some authors use the term *rejection* for the attitude corresponding to the speech act *denial* (Restall, 2005; Ripley, 2011); this difference is merely in terminology.

[2]An early response to this argument is due to Kent Bendall (1979). Frege burdens the defender of a distinct speech act for rejection with *three* basic operations—rejection, assertion and an embeddable operator *not*—whereas only the latter two are required. But Bendall shows that one can give a classical propositional calculus in which there are no embedded negations; hence one only needs to assume rejection and assertion.

In his widely influential paper *Rejection* (1996), Timothy Smiley presses Frege on this. All that Frege can show, or so Smiley argues, is that *it is not the case* is not a marker for rejection, but fails to show that there is *no* marker for rejection. In fact, so Smiley continues, (1b) *is* a linguistic form of a rejecting speech act and there is no need to demand that (1b) must be analysed in the same way as (1a).

Then, if putting a polar question to oneself and replying *No* is a rejection, we can take replying *Yes* to have assertive force. Thus, on Smiley's account, (2) is a rejection of $p$ and (3) is an assertion of *not p*.

(2)    Is (it the case that) $p$? No.

(3)    Is (it the case that) *not p*? Yes.

Frege's reductive strategy would reduce (2) to (3). Smiley, however, contends that (2) and (3) are distinct linguistic phenomena that should not be conflated.

**Frege's Argument.**    I take for granted that Frege succeeds in showing that embeddable negation (i.e. *not* as in (1a)) is not a sign for rejection. Thus, I present here how his argument would unfold for negative answers to polar questions. Frege takes the inference (3) to be a valid instance of natural language reasoning and takes (4) and (5) to be possible analyses.

(3)  a.  If the accused was not in Berlin, he did not commit the murder.

    b.  Was he in Berlin? No.

---

    c.  Did he commit the murder? No.

(4)  a. Assert: If not $p$, then not $q$.

    b. Assert: Not $p$.

---

    c. Assert: Not $q$.

(5)  a. Assert: If not $p$, then not $q$.

    b. Reject: $p$.

---

    c. Reject: $q$.

In (4), the word 'no' in response to 'was he in Berlin?' is taken to *modify* the propositional clause, i.e. (4b) expresses *assent* to the proposition that he *was not* in Berlin. In (5), 'no' is interpreted as expressing an attitude the speaker has *towards* the propositional clause, i.e. (5b) expresses *dissent* from the proposition that he *was* there.

Now, Frege argues, (4) is validated by *modus ponens*: the antecedent of the conditional in (4a) is asserted in (4b). However, the propositional clause in the antecedent of (5a) is not the same as the one in (5b), so *modus ponens* cannot apply there; Frege also emphasises that rejections, as speech acts, cannot embed under conditionals, so the *not* in the antecedent of (3a) must indeed modify the propositional clause as in (4a) and (5a).

Thus, to preserve the validity of (5), the rejection in (5b) must be analysed as the assertion in (4b)—and then we might as well forego the distinction. The following is a possible reconstruction of Frege's argument:

  i. (5) is a candidate analysis of (3).

 ii. (3) is valid.

iii. The inference from (3ab) to (3c) is an instance of *modus ponens.*

 iv. Hence, *modus ponens* must be applied to (5ab).

  v. Thus, (5b) is (4b).

Now, Smiley (1996, p4) points out that (iii) need not be true. He draws attention to the rule in (6).

(6)  a. Assert: If $p$, then $q$.

   b. Reject: $q$.
   _____

   c. Reject: $p$.

This appears to be valid: If I assert a conditional but reject its consequent, I ought to reject the antecedent as well, lest I contradict myself. If we identify rejection with negative assertion, we understand (6) as *modus tollens*, but Smiley's point is that we do not *need* to do so. The inference in (6) is valid *prima facie* and does not require theorising about the nature of rejection to be appreciated. Likewise, Smiley continues, we do not *need* to understand (3) as a *modus ponens* inference.

Thus, Smiley concludes, all that Frege shows is that *modus ponens* does not apply to (5) if rejection is distinct from assertion—not that (5) is invalid. Frege seems to agree with this; he also concludes that the inference from (5ab) to (5c) cannot be performed 'in the same way as before' (Frege, 1919, p154). Thus, Frege's argument must be one of parsimony: maintaining (5b) as distinct from (4b) is not a logical mistake, but requires us to add a new primitive for rejection as well as novel inference rules to validate (5) without appealing to *modus ponens.*

However, Smiley continues, there is nothing unparsimonious in doing so if this explains more data. Smiley takes inferences like (3) to be new data, unaccounted for by the Fregean approach. But Frege may still insist that his reductive strategy can *analyse* just as many examples as Smiley can. To break the impasse, I now present an instance of (5) that resists analysis as (4).

**New Data.**  Consider the following variant of (3).

(7)  a.  If the chair of logic is not here, the chair of metaphysics is not here.

   b.  Is the chair of logic here? No, the position is still open.

   _____

   c.  Is the chair of metaphysics here? No.

This also is a valid inference: it is natural in situations in which, for instance, the chairs of logic and metaphysics are in personal union, or the speaker knows that the chair of metaphysics would only come to meet the new chair of logic, once appointed.

However, the analysis as (4) does not apply here: saying 'no, the position is still open' is not equivalent to asserting that *the chair of logic is not here.* Following Frege's analysis of (3), this means that this is not a *modus ponens* inference, since (7b) does not coincide with the antecedent of (7a). One could analyse (7b) as a negative assertion by giving the negation *external scope*, i.e. read (7b) as *it is not the case that the chair of logic is here.* However, this does not coincide with the antecedent of (7a) and so does not license *modus ponens* either. To say otherwise, moreover, would obscure the difference between (7b) and the alternative rejection 'No, she is not.'

This also means that it is wrong to consider (7b) as *equivalent* to a negative assertion. Thus, Frege's reductive strategy makes the wrong predictions here. However, given a suitable conception of what it means to reject a sentence, it seems plausible that (7) can be analysed as (5). It is my goal to develop such a conception. First, however, I note that (7) also spells trouble for Smiley's proposal.

**Smiley's Bilateralism.**  Smiley wants to give an account of inferences that can use rejections *as premisses and conclusions*. Thus he conceives a *bilateral* logic that countenances rejected sentences alongside asserted ones. Bilateral logics have been given an elegant formulation by Incurvati & Smith (2009). They ask us to 'imagine a speech community for whom any sentence is explicitly structured into a propositional content clause and a force-indicator' (p3). In a formal language, we can represent propositional content clauses in the usual language of atoms and Boolean connectives, and use $+$ as a marker for assertive force and $-$ as a marker for rejective force. That is, (2) and (3) can be analysed as follows.

(2)    Is (it the case that) $p$? No.

   *is* Reject: $p$

   *is* $-p$

(3)    Is (it the case that) *not p*? Yes.

   *is* Assert: $\neg p$

   *is* $+\neg p$

To make this analysis useful, a bilateral logic must include so-called *coordination principles* that specify the relationship between rejected and asserted content. The bilateralist logics of Smiley (1996), Rumfitt (2000) and Incurvati & Smith (2009) use the following *Smilean reductio* principle for that purpose.

$$(\bot) \ \frac{+A \quad -A}{\bot} \qquad (\text{SR}_1) \ \frac{\begin{array}{c} [+A] \\ \vdots \\ \bot \end{array}}{-A} \qquad (\text{SR}_2) \ \frac{\begin{array}{c} [-A] \\ \vdots \\ \bot \end{array}}{+A}$$

The absurdity sign $\bot$ is considered a punctuation mark and is therefore not force–marked (Tennant, 1999; Rumfitt, 2000).[3] With these principles, a bilateralist can validate (6) as follows.

1. $+(p \to q)$ is a premiss.

2. $-q$ is a premiss.

3. Towards a Smilean *reductio*$_1$ assume $+p$.

4. From 1, 3 and *modus ponens*, infer $+q$.

5. From 2 and 4 infer $\bot$.

6. From 5 and $(\text{SR}_1)$, infer $-p$.

To explain (5), however, the bilateralist needs to say something about the meaning of negation in their framework. Ian Rumfitt (2000) offers a very complete account of this.

**Rumfitt's Inferentialism.**   In general, *logical inferentialism* is the view that the meaning of a logical constant is given by a pair of inference rules that specifies how to *introduce* and how to *eliminate* the constant in a natural deduction system. This is in direct opposition to *referentialism* which seeks to specify the meaning of a logical (or linguistic) constant through how it contributes to the *truth conditions* of a complex sentence.

   Logical inferentialism was first proposed by Gerhard Gentzen (1935) and further developed by Dag Prawitz (1965) and later Michael Dummett (1991). One major problem for inferentialism is that it appears that not every pair of introduction and elimination rules confers a coherent meaning on the constant involved (Prior, 1960).[4] This issue can be addressed by demanding that there be a certain balance

---

[3]Smiley himself does not use $\bot$ at all—which those who find usage of $\bot$ inelegant might consider an advantage—but specifies the antecedents of the Smilean *reductios* as parallel inferences of $+B$ and $-B$ for any $B$.

[4]Prior's example is the connective *tonk*, which is explained as follows: a premiss $p$ allows one to introduce $p \ tonk \ q$ for arbitrary $q$, and eliminating $p \ tonk \ q$ allows one to infer $q$. As this trivialises the logic, *tonk* is usually taken to be incoherent.

between introduction and elimination rules, known as *harmony* (Prawitz, 1965; Dummett, 1991; Tennant, 1997). Opinions differ on what the correct notion of harmony is (Steinberger, 2009). However, one criterion is universally regarded as a sufficient condition for harmony: if the elimination rule is precisely the inverse of the corresponding introduction rule, then a coherent meaning is conferred.

As Dummett (1991) points out, however, while the harmony requirement rules out the known problematic cases, it also appears to sanction intuitionistic logic: double negation elimination does not appear to be a harmonious rule in a standard natural deduction system for propositional logic. According to Rumfitt (2000), this appearance is deceiving. It is the restriction of standard natural deduction systems to *asserted* content that prevents one from giving an inferentialist explanation of classical negation, or so he argues. Once *rejected* content is also countenanced, as it should, one can formulate harmonious rules for classical logic. Rumfitt (2000) and Incurvati & Smith (2009) suggest the following rules for negation.

$$(+\neg \text{I.})\ \frac{-A}{+\neg A} \quad (+\neg \text{E.})\ \frac{+\neg A}{-A} \quad (-\neg \text{I.})\ \frac{+A}{-\neg A} \quad (-\neg \text{E.})\ \frac{-\neg A}{+A}$$

These rules are clearly harmonious, because they are inverses of each other. Also, they immediately entail classical negation elimination, as $+\neg\neg A \vdash\!\dashv -\neg A \vdash\!\dashv +A$. They then allow a bilateralist to analyse (5) as follows.

1. $+(\neg p \to \neg q)$ is a premiss.
2. $-p$ is a premiss.
3. From 2 and $(+\neg \text{I.})$, infer $+\neg p$.
4. From 1, 3 and *modus ponens*, infer $+\neg q$.
5. From 4 and $(+\neg \text{E.})$, infer $-q$.

This, however, is a troublesome explanation. While this works for (3), it is a bad analysis of (7). As I demonstrated above, the inference in (7) is *not* reducible to *modus ponens* on asserted premisses. In particular, step 3 in the above derivation is a mistake for (7): as I argued, the premiss $-p$ should *not* be analysed as equivalent to the assertion of a negative (i.e. as $+\neg p$).

Therefore, I want to explore an alternative theoretical option that, to my knowledge, has not received substantial attention in the bilateralist literature: to account for (5) by an inference that does *not* appeal to *modus ponens* on asserted premisses. This option is linked to another challenge that has been posed to bilateralism: that there are certain, *feeble* rejections that are allegedly unsuitable for inference.

**The Feeble Rejection Challenge.** It has been argued that bilateralists cannot account for the phenomenon of rejection as a whole because there are cases like (7b) where a rejection appears to be *less informative* than a negative assertion. In

such cases it would be a mistake to apply rule ($+\neg$I.). I will call rejections *strong* if ($+\neg$I.) is correct for them and *feeble* otherwise.[5]

Dickie (2010) calls the existence of feeble rejections alongside strong rejections the 'messiness' of rejection. She argues that it puts bilateralists in a bind: either bilateralism is explicitly stated for strong rejections only or it also covers feeble rejections. Neither would be satisfying:

- If $-A$ exclusively denotes the strong rejection of $A$, then the crucial Smilean *reductio*$_1$ principle $+A \vdash \bot \Rightarrow\ \vdash\! -\!-A$ is invalid: if it is absurd to assert $A$ then $A$ may be rejected, but not necessarily *strongly* rejected, or so Dickie argues.[6] In particular, then, the bilateralist cannot verify the inference pattern (6) as its proof rests on Smilean *reductio*$_1$.

- If $-A$ may also denote a feeble rejection, then, says Dickie, the messiness of rejection precludes one from giving an evidence–preserving proof theory. That is, the challenge to the bilateralist is to verify the inference pattern (5)—as well as other desiderata—*without* appealing to ($+\neg$I.).

Therefore, Dickie concludes, rejection is too unspecific to be useful in inferences.[7] The feeble rejection challenge to bilateralism is twofold: (i) weak rejections are, allegedly, unsuitable for inference; and (ii) if a logic for weak rejection can be found, it must deny ($+\neg$I.) and thus, allegedly, cannot derive classical logic in a harmonious manner. I will return to (ii) in Chapter 4. This chapter addresses (i).

To assess Dickie's argument, it is useful to take a closer look at the rejection phenomenon. Consider the following examples:

(8)　　Did Homer write the *Republic*?

　　　– No, Homer (a guy who actually existed and wrote the *Odyssey* and the *Iliad*) did not also write the *Republic*.　　　　　　　　　(Dickie, 2010)

(9)　　Did Homer write the *Iliad*?

　　　– No, in fact, Homer did not exist.　　　　　　　　　　　　(Dickie, 2010)

(10)　　Was Homer a unicorn?

　　　– No, there is no such property as the property of being a unicorn.

　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　(Dickie, 2010)

---

[5]Usually, feeble rejections are called *weak*. I use *feeble* to avoid confusion with my definition of weak rejection below.

[6]In fact, Rumfitt (1997) already makes the argument against Smilean reductio for strong rejections. He also considers rejections that do not amount to a negative assertion. He calls them *external* rejections (and my strong rejections *internal*); compare with my analysis of example (7). I will make sense of this terminology in Chapter 3.

[7]Such claims have a long history, see (Horn, 1989, chs1–2). The long standing view can be succinctly put as follows: assertions tell us how the world is—rejections tell us much less.

(11)    Is it the case that X or Y will win the election?

    – No, X or Y or Z will win.                                    (adapted from Grice (1991))

(12)    Is Franz here?

    – No, not as far as I know.                                    (my construction)

Here, the speaker marks the sentence in question as *unassertible* in one way or another. There are many grounds for finding a sentence unassertible, falsity as in (8) being just one of them—hence 'messiness.' The same speaker would not necessarily be comfortable with asserting the negated sentence: in many cases, the negated sentence would be unassertible too. Thus (9)–(12) appear to be feeble.

A rejected sentence need not be unassertible *in principle.* In examples (11) and (12), the speaker *ascribes* unassertibility to the sentence in question, but the sentence does not suffer from faults such as malformedness or presupposition failure. Also, the speaker does not reject the sentences there *as false.* If the speaker of (11) were interpreted as saying that it is *false* that X or Y will win, then they would be interpreted as asserting that Z *is* going to win, but this would be an overstatement. Similarly, the epistemic state the speaker self–reports in (12) prevents us from interpreting them as asserting that Franz is *not* here.

Thus, the linguistically acceptable rejection of a sentence cannot always be judged by considering the sentence in isolation.[8] This is reason enough for me to, like Rumfitt, pursue an inferentialist semantics of rejection: if rejections (can) figure in inferences, then an inferentialist can specify *how*—but talk of *truth conditions* seems problematic here. Indeed, despite their alleged messiness, rejections like (11) and (12) still plausibly satisfy the inference schemes (5) and (6), as the following instantiations show.

(13)  a.  If the election will not be won by X or Y, then we will not have a socialist president.
      b.  Is it the case that X or Y will win the election?

                          – No, X or Y or Z will win.
      _____
      c.  Is it the case that we will have a socialist president? – No.

(14)  a.  If there is a seminar today, Franz is here.
      b.  Is Franz here? – No, not as far as I know.
      _____
      c.  Will there be a seminar talk? – No.

An appropriate context for (13) is one in which X and Y are the only socialists on the ballot; one for (14) is one in which Franz is chairing the seminar. Thus, as

_____

[8]To my knowledge, this point has been mostly overlooked in the literature on rejection so far. A notable exception is Grice (1991), who raises attention to example (11).

these examples show, rejections—including feeble ones—may be used as premisses and conclusions in inferences.

Thus, it seems overly hasty to say that these feeble rejections mean that rejection is too messy for inference. Plainly, feeble rejections can appear in natural language inferences. Thus, one should not rule out that there is a formal theory regimenting these inferences. I provide such a theory here: a *weak bilateral logic* for inferences involving assertion and rejection. However, to face the feeble rejection challenge, the bilateralist theory of Smiley and his colleagues must be revised; this in particular requires a refined understanding of what a rejection *is*.

## 2.2   Rejection and Discourse

I now argue that it is natural to assume that the speech act of *assertion* has a foil that I call *weak rejection*. After stating the purpose of such an act and giving it a precise definition, I show how it applies to the cases discussed in the previous section. Afterwards, I compare this weak rejection to similar proposals by Ian Rumfitt and Greg Restall, and demonstrate how one can speak of *inferences* involving weak rejections.

**Purpose.**   Let me for the moment assume that Stalnaker's (1978; 1988) account of assertion is broadly correct. In Chapter 3, I relate the discussion here to the alternative *public commitment account* of assertion (Brandom, 1983; Lascarides & Asher, 2009).

According to Stalnaker, interlocutors in dialogue keep track of a shared *common ground*: the set of propositions the interlocutors assume to be mutually accepted. Stalnaker claims that the *essential effect* of an assertion is to *expand* common ground. That is, the essential effect of an assertion that $p$ is to make it so that $p$ is common ground. However, it would be an overstatement to conclude that every assertion always enacts a common ground update. To make a proposition mutually accepted requires unanimity. Plainly, not every proposition raised in a dialogue is acceptable to all interlocutors. Thus, there must be a mechanism by which expansion of the common ground can be *blocked*. I claim that a speech act for rejection is such a mechanism. Stalnaker appears sympathetic to this.

> It should be made clear that to reject an assertion is ... to refuse to accept the assertion. If an assertion is rejected, the context [JJS: common ground] remains the same as it was.                (Stalnaker, 1978, p87, fn9)

Thus, an assertion does not expand the common ground immediately, but does so only in the *absence of rejection.* Put differently, asserting that $p$ *proposes* to make $p$ common ground and *making* it common ground is a further process that needs to be negotiated by the interlocutors (Clark, 1996). One can take this to be the

*purpose* of rejection moves in dialogue: to manage the common ground. In that sense, assertion and rejections are foils of each other. Assertion serves to expand common ground, and rejection serves to curtail that process.

**Definition.** Rejection moves have a clear purpose, but one may still think that such moves can be reduced to assertions. But this is not so. The purpose of rejection is not reducible to an operator (like 'not' or 'it is unassertible that') that allows one to *assert* that one is not allowing a sentence into common ground. In asserting, one proposes to expand the common ground, but there must be a way to *not* expand it. That is, there must be something that operates dually to assertion, on the same level as assertion.

Returning to the parlance of acts and attitudes, this means the following. An assertion expresses assent to a proposition *p* and *mutual assent* to *p* entails that the common ground is expanded with *p*. As argued above, a rejection's purpose is then to communicate that one does not assent. I define *weak rejection* to be the *minimal* act that fulfills this purpose.

---

(I) The **weak rejection** of *p* is the speech act that expresses the weakest attitude that is incompatible with assent. I call this attitude *dissent* or *finding a sentence unassertible*.

---

Of course, one might attempt to assert the negation of a sentence and thereby prevent that sentence from being added to the common ground. But then something *more* than just blocking a context expansion has happened. I contend that weak rejection, as defined in (I), is more basic. In fact, weak rejection is just as basic as assertion. Assertion and weak rejection are co-determined by their joint influence on the common ground in the following sense: if we understand one of them, we understand the other as its foil. That is, assertion expresses the weakest attitude that rules out weak rejection; and weak rejection expresses the weakest attitude that rules out assertion.

**Uniqueness.** It remains to explain why the broad variety of feeble rejections seen in the previous section should all be considered instances of a *single* speech act. The motivation for the speech act of weak rejection is grounded in the *mechanisms* surrounding the common ground. Thus, one should abstract the relevant data to the level on which these mechanisms operate. Then it becomes clear that the seemingly unrelated instances of rejection turn out to be all cases in which one indicates that one *does not assent* to the proposition under discussion. The differences between the examples are to be located only in the *grounds* that a speaker has for doing so.

With this observation, we can also abstract all examples seen so far to the same *form*: a core content (presented as a polar question), an occurrence of the

negative polarity particle *no*, and a sentence that *explains* why the speaker is
rejecting.[9] Thus, the most parsimonious interpretation seems to be this: *no* is a
force–marker for weak rejection, but the grounds for rejection can vary and are
reported in a separate propositional clause.

However, a methodological clarification is in order. In the previous section,
I followed Frege and Smiley in considering answers to polar questions; Smiley
explicitly says that these are questions that one puts to *oneself.* But in this
section, I follow Stalnaker in considering assertions and rejections *in dialogue.*
This untangles as follows. I maintain that one way of *linguistically realising* a
rejecting speech act is by putting the question to oneself and answering it. Thus
a *dialogue* in which one speaker asserts *Homer is a unicorn* and another rejects it
would unfold as follows.

(15)  a. Amy: Homer is a unicorn.

   b. Bob:  Is Homer a unicorn? No, there is no such property as the property
            of being a unicorn

This is clumsy, and it is hardly surprising that speakers usually choose different
ways of expressing rejection. Some models of dialogue predict that in the con-
text after (15a), the question *Is Homer a unicorn?* is saliently under discussion
(Ginzburg, 2012); so it makes sense that we usually forego the part where we put a
question to ourselves. In any case, it is certainly *possible* to express oneself like this
and we may take this as the *genuine* form of performing rejection and assertion.
This point is made, with great wit, in the *Stellingen* of Jeroen Groenendijk.[10]

> Wittgenstein, Philosophische Untersuchungen, par. 22:
>
>> "Wir könnten sehr gut auch jede Behauptung in der Form einer
>> Frage mit nachgesetzter Bejahung schreiben; etwa: 'Regnet es? Ja!'.
>> Würde das zeigen, dass in jeder Behauptung einer [sic] Frage steckt?"
>
> Ja!                                              (Groenendijk & Stokhof, 1984, p571)

**Comparison to Rumfitt.**   Rumfitt (1997) develops an account of rejection that
seems rather similar to weak rejection as defined above. He contrasts answering
'No' with 'refus[ing] to give the answer Yes' and claims that the latter expresses
dissent 'in another perfectly good sense' (p226). If one were to stipulate that there
is a separate speech act that *expresses* the refusal to answer *yes*, this would seem
similar to weak rejection.

---

[9]In my Master's thesis, I argue that a cooperative speaker is expected to have (and provide)
reasons for rejecting a proposal (Schlöder, 2014). I expand on that point in Chapters 3 and 7.

[10]Translation (mine):

*Wittgenstein, Philosophical Investigations, par. 22:*

> *"We could very well write each statement in the form of a question with an
> affirmative placed after it; for example 'Is it raining? Yes!'. Would this show that
> every assertion contains a question?"*

*Yes!*

However, Rumfitt (1997) argues further that such refusals to answer *yes* call for a three–valued logic (true/false/neither). In a later paper (Rumfitt, 2000), however, he takes a rejection to mark a sentence as 'other than true' and says that this notion is broader than his earlier one (p799–800).[11] Therein lies a difference between his account and mine. While Rumfitt gives truth-conditional explanations of his speech acts, my definitions of (the essential effects of) assertion and rejection are *discursive.*

That is, a weak rejection need not mark a sentence as having or not having any particular truth-value. The difference can be appreciated by recalling (11).

(11)   Is it the case that X or Y will win the election?

– No, X or Y or Z will win.

On my account, one can say that the speaker is performing a weak rejection: the speaker is ruling out assent to the proposition *X or Y will be elected.* However, this proposition might well be true! In fact, another speaker with different opinions on candidate Z's chances might even agree with it. So it is not clear whether (11) would be a rejection in Rumfitt's sense.

**Comparison to Restall.**   It is also useful to compare this notion of weak rejection with Restall's (2005) account of assertion and denial. Note that he uses different terminology. To him, acceptance and rejection are the *attitudes* corresponding to the *speech acts* of assertion and denial.

Restall says that to reject is 'to *refuse* to accept [assent]' (p197, fn6, emphasis his). Clearly, the refusal to assent to *p* rules out assent to *p*; thus Restall's rejection entails my dissent. But there is a weaker attitude than *refusal to assent* that is incompatible with assent: the *refraining from* assenting. It seems that refraining from doing something is less forceful than refusing it, so Restall's refusal is not the *weakest* attitude one can display to reject—thus it is stronger than a weak rejection. To see this, note that example (12) is a weak rejection, but does not seem to fall under Restall's account.

(12)   Is Franz here?

– No, not as far as I know.

In Restall's view, 'a statement is rejected [dissented from] if [. . . ] to accept [assent to] it would be a change of mind, and not merely a supplementation with new information' (p196, fn6). However, the speaker of (12) could come to assent to the

---

[11]In a yet later paper, Rumfitt retracts this statement (Rumfitt, 2014): 'In that paper [Rumfitt 2000] ... −*A* signified that *A* was rejected as false.' It seems that this is a reaction to Dickie's (2010) critique of the logical calculus given in Rumfitt's 2000 paper. Dickie shows that Rumfitt's rules of inference commit him to the claim that rejections are not feeble. Thus, Rumfitts informal claim to rejection-as-other-than-true (which would include feeble rejections) cannot be maintained.

rejected sentence simply due to getting *additional* information instead of *revising* any. So it seems that Restall would not include (12) as a rejection.

Moreover, Restall has different goals. He discusses logics whose formulae are not marked as asserted or rejected, but my aim is to explain how rejections can appear *as premisses and conclusions* in inferences. Restall discards this project because he finds it implausible to close attitudes under inference. His reason for this is a version of Harman's mind cluttering objection (Harman, 1986): Since assenting to something requires awareness of that something, and since moreover infinitely many propositions follow from any given proposition, nobody can be said to have assented to them all. My response is as follows.

**Inference.**   The immediate explanation of bilateralist inference would indeed be to say that inference preserves *attitudes*. That is, if someone has some attitudes towards some propositions, then that someone *also* has certain further attitudes to further propositions. This, while intuitive, requires some additional conceptual clarification.

Say we validate an inference of the form $+A \vdash +B$, i.e. that the assertion of $A$ entails the assertion of $B$. Surely this does not mean that someone who has uttered a assertion of $A$ must have also uttered an assertion of $B$ (Dutilh Novaes, 2015). Similarly, we cannot say that someone who assents to $A$ also assents to $B$. As Restall (2005) points out, this is implausible. This problem can be avoided as follows. Say that $+A \vdash +B$ means that someone who has asserted $A$ is *taken to assent* to $B$. To be taken to assent to something does not necessarily entail being aware of it. Rather, to be taken to assent to $B$ is to be obliged to assent to $B$ once $B$ becomes salient.

Such obligations are external to the individual, so no-one's mind is being cluttered here. Instead, they are instituted by the conventions of proper language use. To take *modus ponens* as an example, if one assents to *if A, then B* and moreover assents to $A$, they are taken to assent to $B$—if they were to refuse to assent to $B$, they would be using the words *if ... then* wrongly. However, as long as $B$ is not salient, this requirement is not relevant. One may take this obligation to be normatively constitutive of cooperative dialogue: not conceding what you are (implicitly) taken to assent to would be uncooperative or even deceptive; I return to this point in Chapter 3. This also makes sense of the passive voice in 'taken to': if a speaker is taken to assent to $B$ then anyone who attributes the attitude *assent to B* to that speaker (i.e. takes her to assent) is normatively correct.

This explanation of inference is similar to the proposal of Catarina Dutilh Novaes (2015): that if $A \models B$ and you have granted $A$ you are required to grant $B$. To be required to grant $B$ does not mean to *have* granted $B$—rather to do so once $B$ is put forward. I prefer to speak of *being taken to have certain attitudes*. The notion of *granting* seems to be married to assertion, whereas my aim is to validate inferences involving other speech acts as well. Then, the same considerations as above apply, *mutatis mutandis*, to rejections and *being taken to dissent* from some

propositions. For instance, a typical bilateralist inference has it that if one asserts that *if A then B* and dissents from *B*, then they are taken to dissent from *A* as well (i.e. the inference in (6)). In the formal logic that I develop now, the inference rules are justified in that they preserve the attitudes one is taken to have.

## 2.3 Weak Bilateral Logic

My goal is to show by demonstration that it is possible to find a bilateral logic that meets the feeble rejection challenge. To this end, I develop a *weak bilateral logic* where the expression $-A$ denotes the weak rejection of *A*; this in particular makes room for feeble rejections. This logic is based on Rumfitt's (2000) bilateral natural deduction calculus. I evaluate which of his rules are valid for $-$ being a sign for weak rejection and propose *restrictions* of those that not. Whenever I do so, I mark the rule with an asterisk.

I will use lowercase Latin letters to denote propositional atoms, uppercase Latin letters to denote propositional logic formulae, and lowercase Greek letters to denote force–marked formulae or absurdity, i.e. $\varphi$ can be $+A$, $-A$ or $\bot$. As before, the absurdity sign $\bot$ is considered a punctuation mark and is therefore not force–marked. Following Rumfitt, I situate weak bilateral logic in a Gentzen–style natural deduction calculus.

**Coordination principles.** In Section 2.1, I discussed the bilateralist coordination principle of Smilean *reductio*. My notion of $-$ as denoting weak rejection differs from previous bilateral logics, so the validity of these rules must be established again.

$$(\bot) \frac{+A \qquad -A}{\bot} \qquad (\mathrm{SR}_1) \frac{\begin{array}{c}[+A]\\ \vdots\\ \bot\end{array}}{-A} \qquad (\mathrm{SR}_2) \frac{\begin{array}{c}[-A]\\ \vdots\\ \bot\end{array}}{+A}$$

($\bot$) expresses that it is absurd to both strongly assert and (weakly) reject the same proposition; that is $+A$ and $-A$ express incompatible attitudes. This follows immediately from the definition of weak rejection in the previous section. The justification for the two halves of Smilean *reductio* goes as follows.

The antecedent of $(\mathrm{SR}_1)$ says that it is absurd to assert *A*. This means that the active premises indicate that the speaker already is taken to have an attitude towards *A* that is incompatible with assent to *A*. By definition, dissent is the weakest attitude one can have that rules out assent; thus who cannot assent to *A* is taken to have an attitude that is *at least* as strong as dissent from *A*. That is, the antecedent of $(\mathrm{SR}_1)$ entails dissent from *A*. Conversely, the antecedent of $(\mathrm{SR}_2)$ says that the active premises are incompatible with dissent from *A*. This entails assent to *A* since anything weaker than that would not be incompatible with dissent from *A* (by definition).

**Negation.**   A bilateralist sees a connection between rejection and negation. Rumfitt (2000) expresses this connection in a way that commits him to all rejections being strong (Dickie, 2010; Rumfitt, 2014). In contrast, a strong rejection is in particular a weak rejection, but not every weak rejection is strong (it might be feeble). Thus, assent to a negative entails dissent, but dissent does not entail assent to a negative. Thus, I maintain the bilateralist principle $(+\neg\text{E.})$ but the converse $(+\neg\text{I.})$ must be dropped.

$$(+\neg\text{E.}) \ \frac{+\neg A}{-A}$$

This rule has a dual $(-\neg\text{I.})$, which is derivable from $(+\neg\text{E.})$.

$$(-\neg\text{I.}) \ \frac{+A}{-\neg A}$$

The derivation of $(-\neg\text{I.})$ is as follows.

$$\frac{+A \quad \dfrac{\dfrac{[+\neg A]^1}{-A}\ (+\neg\text{E.})}{\dfrac{\bot}{} }\ (\bot)}{\dfrac{\bot}{-\neg A}\ (\text{SR}_1)^1}$$

A paraconsistent logician would disagree that assent to a negative entails dissent. One would hardly deny, however, that *sometimes* one rejects a proposition because one judges it false. For now, I restrict my attention to just such cases. However, generally, the paraconsistentist worry can only be about alleged true contradictions like Liars: a paraconsistent reasoner would assent to the Liar and its negation, but not dissent from either (Murzi & Carrara, 2015). I offer an alternative explanation of the Liar paradox in bilateral logic in Chapter 4.

For now, I leave the paraconsistent worries aside. To find further rules for negated premisses, I investigate further the issue raised by Dickie (2010).

**Deduction and the Specificity Argument.**   A logic where $-$ denotes rejections that can be feeble cannot validate the following principle of classical negation introduction (Dickie, 2010):

$$\begin{array}{c} [+A] \\ \vdots \\ (\text{CNI}) \ \dfrac{\bot}{+\neg A} \end{array}$$

Since $-A, +A \vdash\!\!\vdash \bot$, (CNI) would validate $-A \vdash\!\!\vdash +\neg A$. But then $-$ would be a sign for strong rejection. Intuitively, it is clear why (CNI) must fail: inferring absurdity from an assertion only means that $A$ is rejected, but $+\neg A$ says *specifically* that $A$ is absurd to assert *because* it is judged false.

However, the same problem does not occur when we consider instead the rule (CR), which is classically equivalent to (CNI).

$$(\text{CR}) \ \frac{+(A \to (B \land \neg B))}{+\neg A}$$

Accepting (CR) does not allow you to draw the inference $-A \vDash +\neg A$, since $-A$ need not entail that $+(A \to (B \land \neg B))$. This can be understood as follows. Asserting that *A implies a contradiction* in the antecedent of (CR) is *specific*: it tells you a definite inferential property of *A*. Observing that *asserting A is absurd* in (CNI) is *unspecific*: this is true whenever $-A$ is a premiss, but there are many potential grounds for having this premiss.

Thus, the antecedent of (CNI) is *strictly weaker* than the antecedent of (CR). That means that (CNI) is a stronger principle than (CR). Thus, while Dickie succeeds in showing that a bilateralist cannot have (CNI), the bilateralist can still retreat to (CR). However, this retreat comes at a price. If we reject (CNI), but find (CR) acceptable, then we must reject the Deduction principle (Ded).

$$(\text{Ded}) \ \frac{\begin{matrix}[+A]\\\vdots\\+B\end{matrix}}{+(A \to B)}$$

Since $\bot \vDash +(B \land \neg B)$ by Smilean *reductio* and $+(B \land \neg B) \vDash \bot$ by $(+\neg E.)$ and $(\bot)$, accepting (Ded) means accepting that (CNI) and (CR) are equivalent. In particular, then, under (CR)+(Ded), $-A \vDash +\neg A$. Hence (CR) excludes (Ded).

Note that to appreciate the difference between (CNI) and (CR), we need not accept (CR) as valid: even if (CR) fails, one can still see that it is a strictly weaker principle than (CNI)—and (Ded) wrongly makes them equivalent. Thus, (Ded) fails in logics where $-$ denotes weak rejection.

Moreover, nothing hinges on the choice of cashing out this argument using (CR): any other rule that spells out a specific way in which one might judge a proposition false would do (e.g., $+(A \to \neg A) \vDash +\neg A$). Thus it seems that the problem with feeble rejections is merely reflected in (CNI), but properly located within (Ded). I call this the *specificity argument against Deduction.*

---

(II) **The specificity argument against Deduction** goes as follows. In a bilateral logic where the sign $-$ for rejection can denote feeble rejections, there are at least two different grounds for rejection (assent to a negative and at least one other, lest there be only strong rejections). These grounds are not specified in the notational device $-$. Thus, a *derivation* using rejected premisses is unspecific with regards to these grounds. The *assertion of a conditional* makes no use of $-$ and is thus not unspecific in this sense; hence it is strictly more informative.

---

**The Restriction Strategy.**   One might be inclined to draw the conclusion that this dooms the bilateralist project. But this would be too hasty. To solve the specificity problem, one can *restrict* the Deduction principle to derivations with only asserted premisses. Consider the following definition:

**2.3.1.** DEFINITION (Subderivations). Write (a) for the assumption that there is a proof of $\psi$ from $\varphi$ and (b) for the assumption that there is a proof of $\psi$ from $\varphi$ where there are no undischarged rejected assumptions and all premisses used in the proof are asserted.

$$(a)\quad \begin{array}{c} \varphi \\ \vdots \\ \psi \end{array} \qquad (b)\quad +\begin{array}{c} \varphi \\ \vdots \\ \psi \end{array}$$

Note: (b) is still valid if there *appears* a formula of the form $-B$ in the proof, as long as it has been derived from asserted premisses or is a discharged assumption.

Then, one can phrase the rules for asserted conditionals as follows.

$$(+\rightarrow\text{I.})^* \; \frac{\begin{array}{c}[+A]\\ +\vdots\\ +B\end{array}}{+(A \rightarrow B)} \qquad (+\rightarrow\text{E.}) \; \frac{+(A \rightarrow B) \qquad +A}{+B}$$

There is no apparent issue with $(+\rightarrow\text{E.})$. Now, $(+\rightarrow\text{I.})^*$ simply eliminates the source of unspecificity: rejected premisses. This solution rests on the expressive poverty of the language of propositional logic. In this language, we cannot model assent to higher-order judgements about sentences; we cannot phrase, e.g. a predicate that expresses unassertibility (but see Chapter 4).

Thus, the only way to express dissent from *A in an assertion* is by assent to $\neg A$. One might argue that one can also assertively express dissent from *A* by assenting to $A \rightarrow (B \wedge \neg B)$ for some *B* and that this need not entail assent to $\neg A$. But I would not know how this could be motivated and, in any case, ignoring it seems to be innocent: that $(A \rightarrow (B \wedge \neg B)) \rightarrow \neg A$ is true even in Minimal Logic (Johansson, 1937). This then also explains a weakened version of $(+\neg\text{I.})$.

$$(+\neg\text{I.})^* \; \frac{\begin{array}{c}+\vdots\\ -A\end{array}}{+\neg A}$$

That is, if we derive $-A$ from asserted premisses only, then because we *already* derived $+\neg A$ or something stronger than $+\neg A$.

The restriction strategy confers a particular meaning onto the symbol $\neg$. It can be paraphrased as follows: $\neg A$ is the inferentially weakest propositional logic formula such that assent to $\neg A$ entails dissent from *A*; i.e. for all *B* with $+B \mathrel{\vdash\!\!\!-} -A$, $+B \mathrel{\vdash\!\!\!-} +\neg A$. Again, potential paraconsistentist worries may arise, but I need to delay addressing them until Chapter 4.

This restriction strategy of excluding non-asserted premisses only works in the present context of assertions and rejections of *propositional logic* formulae. In particular, there is a clear counterexample to (+¬I.)* if we consider *first order* formulae. Recall that presupposition failures are grounds for rejection.

(9)  Did Homer write the *Iliad*?

  – No, in fact, Homer did not exist.

If one asserts *Homer did not exist*, also asserting *Homer wrote the Iliad* would be absurd; by Smilean *reductio* this means that assent to *Homer did not exist* entails dissent from *Homer wrote the Iliad*. But it would be a mistake to apply (+¬I.)* to conclude *Homer did not write the Iliad*.

However, talk about what exists requires the expressive power of a first order language. Thus, this is not a counterexample in *propositional* bilateral logic. In the present logic, one would model (9) simply as $-p$ where $p$ is the propositional atom corresponding to *Homer wrote the Iliad*. This abstraction is blind to the grounds for rejection, and in this case these grounds cannot even be *phrased*.

Thus, in general, increasing the expressive power of the underlying language also increases the severity of the specificity problem. In a more expressive language we have ways of assertively dissenting that do not reduce to assent to a negative, e.g. by asserting that some presuppositions fails. Hence, extending the language requires further restrictions on the Deduction principle and on the meaning-conferring rule (+¬I.)*. In Chapter 4, I first show how this works for a simple modal language, and then for a language including a disquotational predicate. Extending weak bilateral logic to a full first order language, however, is beyond the scope of this dissertation.

**Conjunction and Disjunction.**  The following rules specify the behaviour of conjunction and disjunction on asserted premisses.

$$(+\wedge\mathrm{I}.)\ \frac{+A \quad +B}{+(A \wedge B)} \quad (+\wedge\mathrm{E}._1)\ \frac{+(A \wedge B)}{+A} \quad (+\wedge\mathrm{E}._2)\ \frac{+(A \wedge B)}{+B}$$

$$(+\vee\mathrm{I}._1)\ \frac{+A}{+(A \vee B)} \quad (+\vee\mathrm{I}._2)\ \frac{+B}{+(A \vee B)} \quad (+\vee\mathrm{E}.)^*\ \frac{+(A \vee B) \quad \begin{matrix}[+A] & [+B]\\ +\vdots & +\vdots\\ \varphi & \varphi\end{matrix}}{\varphi}$$

The rules on $\wedge$ appear immediately reasonable. Regarding the $\vee$-introduction rules, one should take care not to be lead astray by pragmatics. Under some readings, an assertion of $A \vee B$ pragmatically entails that the speaker does not assent to either $A$ or $B$, so moving from assent to $A$ to assent to $A \vee B$ to *not* assenting to $A$ might seem like a slip. However, the *derived* assent is merely implicit in the dialogue and not itself subject to pragmatics: a speaker who asserts

*A* is *taken to* assent to $A \lor B$, but never *said* 'A or B'. Hence, implicatures should not be computed here.

The minor premisses of $(+\lor E.)^*$ are again restricted due to a specificity problem. If *A* and *B* are rejected feebly, then one may still assent to their disjunction. For example, the following expressions of attitudes seem reasonable (also recall example (11)).

(16)    Is it the case that A or B will win? Yes, nobody else is running.

        Is it the case that A will win? No, B might win.

        Is it the case that B will win? No, A might win.

Thus $-A, -B, +(A \lor B) \not\vdash \bot$ in general. The following derivation is admitted by unrestricted $(+\lor E.)$ but not by $(+\lor E.)^*$ because there are weak rejections in the minor premisses.

$$\cfrac{+(A \lor B) \qquad \cfrac{-A \qquad [+A]^1}{\bot}\,(\bot) \qquad \cfrac{-B \qquad [+B]^1}{\bot}\,(\bot)}{\bot}\,(+\lor E.)^1$$

This problem does not arise if feeble rejections of *A* and *B* are excluded from the minor premisses. As above, restricting the subderivations does just that.

**Smilean Inference.**    Recall the inference pattern (5).

(5)  a. Assert: If not *p*, then not *q*.

    b. Reject: *p*.

                     

    c. Reject: *q*.

In Section 2.1, I presented some valid instantiations of (5) that involve feeble rejections. I argued there that these examples show that (5) is an inference pattern in its own right, not reducible to a *modus ponens* inference on asserted premisses. Smiley, in his rebuttal of Frege, argues that it would be reasonable to introduce novel inference rules that validate (5). So, instead of reducing this pattern to rules on asserted premisses (as Frege would do), I take Smiley at his word and regard it as a new mode of inference altogether. Let me christen it *Smilean inference.*

$$(\text{SI}) \quad \cfrac{\begin{array}{c}[+\neg A]\\ +\vdots\\ +\neg B\end{array} \qquad -A}{-B}$$

The subderivation states that in the restricted language of only asserted content, if one strongly rejects *A*, then one also strongly rejects *B*. (SI) says that in that case the weak rejection of *A* also entails the weak rejection of *B*. I take Smilean

inference to be a basic rule in a bilateral calculus. It is the final addition to the calculus of *weak bilateral logic.*

---

(III) **Weak bilateral logic (WBL)** is a natural deduction calculus over the following rules of inference:

  – $(\bot)$, $(\mathrm{SR}_1)$ and $(\mathrm{SR}_2)$.
  – $(+\rightarrow\mathrm{E.})$ and $(+\rightarrow\mathrm{I.})^*$.
  – $(+\wedge\mathrm{E.}_1)$, $(+\wedge\mathrm{E.}_2)$, $(+\wedge\mathrm{I.})$, $(+\vee\mathrm{E.})^*$, $(+\vee\mathrm{I.}_1)$ and $(+\vee\mathrm{I.}_2)$.
  – $(+\neg\mathrm{E.})$ and $(+\neg\mathrm{I.})^*$.
  – $(\mathrm{SI})$.

---

Smilean inference might appear *ad hoc* at first glance, but its merits can be appreciated by comparing it to Rumfitt's inference rule $(-\neg\mathrm{E.})$.

$$(-\neg\mathrm{E.})\ \frac{-\neg A}{+A}$$

Clearly, $(-\neg\mathrm{E.})$ can only be valid for strong rejections and thus invalid in weak bilateral logic. However, Smilean inference can be understood as the appropriate weakening of $(-\neg\mathrm{E.})$ to weak rejections. That is, $(\mathrm{SI})$ is provably equivalent to the restriction of $(-\neg\mathrm{E.})$ to asserted premises (proof in Appendix A).

$$(-\neg\mathrm{E.})^*\ \frac{\overset{+\vdots}{-\neg A}}{+A}$$

In this sense, Smilean inference completes the weak bilateral picture by licensing the dual principle to $(+\neg\mathrm{I.})^*$. The principle $(-\neg\mathrm{E.})$ has been criticised by Dummett (2002) for presupposing the classical principle of double negation elimination: the consequent of $(-\neg\mathrm{E.})$ should properly be $+\neg\neg A$, or so Dummett argues. However, while stipulating $(-\neg\mathrm{E.})^*$ might be seen as presupposing a classical semantics, Smilean inference is grounded solely in the linguistic data and does not presuppose any prior semantics.

**Rules for Rejected Premises.** Bilateralist logics verify inference rules that involve rejected sentences as premises and conclusions. In particular, recall the inference pattern (6).

(6)  a. Assert: If $p$, then $q$.
    b. Reject: $q$.
    _____
    c. Reject: $p$.

This pattern is valid in weak bilateral logic for the same reasons that it is valid in other bilateral logics; the proof given in Section 2.1 still works. Moreover, weak

bilateralist logic validates the following introduction/elimination rules for rejected premisses.

$$(-\wedge\mathrm{I.}_1)\ \frac{-A}{-(A \wedge B)} \qquad (-\wedge\mathrm{I.}_2)\ \frac{-B}{-(A \wedge B)} \qquad (-\wedge\mathrm{E.})\ \frac{-(A \wedge B) \quad \overset{[-A]}{\overset{\vdots}{\varphi}} \quad \overset{[-B]}{\overset{\vdots}{\varphi}}}{\varphi}$$

$$(-\vee\mathrm{I.})^*\ \frac{\overset{+\vdots}{-A} \quad \overset{+\vdots}{-B}}{-(A \vee B)} \qquad (-\vee\mathrm{E.}_1)\ \frac{-(A \vee B)}{-A} \qquad (-\vee\mathrm{E.}_2)\ \frac{-(A \vee B)}{-B}$$

$$(-\rightarrow\mathrm{I.})\ \frac{+A \quad -B}{-(A \rightarrow B)} \qquad (-\rightarrow\mathrm{E.}_1)\ \frac{-(A \rightarrow B)}{-B} \qquad (-\rightarrow\mathrm{E.}_2)^*\ \frac{\overset{+\vdots}{-(A \rightarrow B)}}{+A}$$

The proofs that these are valid are in Appendix A.

**Classicality.**   Rumfitt's (2000) bilateral logic is a defence of classical logic: his inference rules for rejection and negation give an elegant and motivated explanation of the classical rules for negation. However, as Dickie (2010) shows, the existence of feeble rejections rules out (CNI) and so Rumfitt's defence of classical logic fails.

My solution to the specificity problem addresses Dickie's criticism. I show that, in a defined sense, the *asserted fragment* of weak bilateral logic is classical logic. Intuitively, this is because on this fragment, the restrictions to only asserted premisses are irrelevant and thus it is just Rumfitt's calculus again. The proofs of the following results are in Appendix A.

**2.3.2.** LEMMA (Contraposition).  $+(\neg A \rightarrow \neg B), +B \vdash +A$.

**2.3.3.** PROPOSITION (Classical Negation).  *The following can be derived:*

$\vdash +(A \rightarrow \neg\neg A).$

$\vdash +(\neg\neg A \rightarrow A).$

$\vdash +\neg(A \wedge \neg A).$

$\vdash +((A \rightarrow (B \wedge \neg B)) \rightarrow \neg A).$

In particular, then, negation behaves classically in weak bilateral logic. This delivers a Classicality theorem.

**2.3.4.** THEOREM (Classicality). *Let $\vdash^{CPL}$ be the derivability relation of classical propositional logic and $\vdash$ be the derivability relation of WBL. It holds that $A \vdash^{CPL} B$ iff $+A \vdash +B$.*

The Classicality theorem follows because the axioms for Classical Negation together with the introduction/elimination rules for asserted $\vee, \wedge$ and $\rightarrow$ give a standard axiomatisation of propositional logic and, conversely, all weak bilateral rules on asserted premises are classically valid. In particular, if $A$ is a classical tautology, then $+A$ is a WBL-tautology. But while some nonclassical logics also verify all classical tautologies (e.g. the Logic of Paradox (Priest, 1979)), the Classicality theorem says that the asserted fragment of WBL also has classical *inference.*

This result might seem odd when considering Liar sentences. A paracomplete logician might want to simultaneously reject the Liar sentence, its negation and their disjunction (Field, 2008). But doing so is inconsistent in weak bilateral logic, since $+(l \vee \neg l)$ is a theorem. Note, however, that there are treatments of the semantic paradoxes on which both the Liar sentence and its negation should be (weakly) rejected, while their disjunction should be asserted (see, e.g. McGee 1990). Thus, the semantic paradoxes are not immediate counterexamples to weak bilateral logic.

However, if we enrich the propositional language to one in which Liar sentences can be phrased, we introduce a new source for the specificity problem. Thus, the restrictions on the subderivations in *-ed rules will need to be tightened. I will investigate this issue in Chapter 4.

**Semantics for Weak Bilateral Logic.**   To conclude this exposition of weak bilateral logic, I now give a semantics for which weak bilateral logic is sound and complete. Smiley gives semantics to his bilateral logic by defining which classical valuations are *correct* for a set of signed formulae. To wit, a classical valuation $V$ is correct for $+A$ if $V(A) = 1$, and for $-A$ if $V(A) = 0$.

Correctness can be generalised to weak bilateral logic as follows. Let an $\omega$–*pointed model* be a sequence of $\omega$ many classical valuations (*points*). It is correct for $+A$ if $A$ holds at *every* point, and for $-A$ if $A$ fails at *some* point. This formalises the intuition that there are many ways for a formula to be unassertible, and any of them is grounds for rejecting it.

**2.3.5.** DEFINITION. A $\omega$–*pointed model* is a mapping $V$ from $\omega$ to models of propositional logic.

- For any $x \in \omega$, $V \models_x A$ iff $V(x) \models A$ (as a model of propositional logic).

- $V \not\models \bot$.

- $V \models +A$ iff $\forall x \in \omega : V \models_x A$.

- $V \models -A$ iff $\exists x \in \omega : V \not\models_x A$.

**2.3.6.** THEOREM (Soundness). *WBL is sound on $\omega$–pointed models.*

**Proof sketch:**
This is a standard induction on the length of derivations; the full proof is in

Appendix A. Only the rules involving restricted subderivations require more than standard methods; I present here the inductive step for one of them, $(+\neg \text{I.})^*$:

Assume $\Gamma \vdash^D +\neg A$ by an application of $(+\neg \text{I.})^*$, i.e., $\Gamma \vdash^{D'} -A$ where $D'$ uses only asserted premises from $\Gamma$. Let $\Gamma'$ be the asserted formulae in $\Gamma$. Then $\Gamma' \vdash^{D'} -A$. Assume that $\Gamma' \not\models +\neg A$. Then there is a model $V$ of $\Gamma'$ and a point $y \in \omega$ such that $V \models_y A$. Construct an $\omega$–pointed model $V'$ where *every* point is $y$, i.e., for any $x$ and atom $p$, $V' \models_x p$ iff $V \models_y p$. Because $\Gamma'$ contains only asserted formulae, $V' \models \Gamma'$. Also, because $V \models_y A$, $V' \models +A$. But by induction, $V' \models -A$. Contradiction. $\qquad\square$

To prove completeness, one first needs to derive Lemma 2.3.7 from the classicality theorem. Then, one can prove a model existence result as follows: Given a WBL-consistent set of formulae, construct an $\omega$-pointed model by constructing for each rejected formula one point where it fails, while letting all asserted formulae hold at *all* points.

**2.3.7.** LEMMA. *Let $\Gamma$ be a consistent set of only asserted formulae. Assume $\Gamma \cup \{-A\}$ is consistent. Then the set $\{\neg A\} \cup \{B \mid +B \in \Gamma\}$ is satisfiable in classical propositional logic.*

**2.3.8.** THEOREM (Completeness). $\Gamma \models \varphi$ *implies* $\Gamma \vdash \varphi$.

The full proofs are in Appendix A.

## 2.4 Conclusion

In this chapter, I have addressed a worry from the prior literature on rejection: that rejections should not or cannot be premises in inferences. First, there is Frege's classical argument, and second Dickie's feeble rejection challenge. I have upheld Smiley's response to Frege and added a case where Frege's reductive strategy does not apply. To contradict Dickie, I have presented some examples where even a feeble rejection appears in a natural inference. As weak bilateral logic shows, such inferences can be formalised in an appropriately articulated logic. This logic is sound and complete for $\omega$-pointed models. Thus, rejections—weak and strong—can appear as premises and conclusions in inferences.

On this account, rejection is a primarily dialogical phenomenon, characterised by its interplay with assertion in managing common ground. Hence, I do not appeal to the truth or falsity of the asserted or rejected sentences. This allows my logic to account for cases in which a sentence rejected even if (possibly) true and not unassertible in principle. Moreover, the discursive view directly explains the—in the prior literature merely stipulative—Smilean *reductio* principles.

My account generalises on both the Fregean view and Rumfitt's bilateralism. The logic WBL reduces to Rumfitt's logic if weak rejections are not present in

the dialogue, but accounts for their occurrence by only including the bilateralist rules from the prior literature that are still valid for weak rejections. The singular addition I make to these rules is the *Smilean inference*: an inference pattern endorsed by Smiley and entertained by Frege; the latter at least insofar as he recognises some natural language instances for the purposes of his argument.

Weak bilateralist logic shows that the logic of asserted content is classical. But there are still two issues left open. First, the rules of WBL are not—or at least not obviously—in harmony. In particular, Smilean inference has no dual and it is not clear how it can be understood as an introduction or elimination rule. Second, WBL obscures the difference between strong and weak rejection. This appears odd, since *prima facie* the meaning of many *no*'s in natural language seems to be a strong rejection. I address the first point in Chapter 4 and the second now in Chapter 3.

# Chapter 3

# Commitment

The previous chapter motivated the notion of *weak rejection* from a Stalnakerian perspective on assertion. In this chapter I elaborate how weak rejections can be understood in a *public commitment* account of assertion (Brandom, 1983; Lascarides & Asher, 2009). Doing so has certain explanatory advantages. A commitment account is more transparent with regards to cooperativity (Asher & Lascarides, 2013) and dialogue dynamics (Farkas & Bruce, 2010).

I begin by considering a challenge that has been put to me by Nicholas Asher: to explain that certain speech acts make certain projections, e.g. that an assertion projects acceptance, *even in non-cooperative settings*. My solution to the puzzle includes an *axiomatic* account that describes (partially) what it means to commit to something in dialogue. In this account, I can formalise weak rejection in greater generality than I did in Chapter 2. Then, by adding further axioms, I can separate cooperative commitment from non-cooperative commitment. A completeness result shows that WBL inference is in fact the preservation of cooperative commitment, giving due to my arguments about cooperativity in Chapter 2.

This chapter is based on, and draws from, *Weak rejection* (co-authored with Luca Incurvati), and *Aligning intentions: Acceptance and rejection in dialogue* (co-authored with Antoine Venant and Nicholas Asher).

## Chapter Contents

## 3.1    Projected Continuations and Commitment

**Backward and Forward.**    The interpretation of an utterance is to a substantial degree *backward-looking.* That is, to interpret the current utterance of a dialogue, one needs to look at the *prior dialogue* to know, *inter alia*, what is being responded to (Asher, 2008), which referents are available for anaphora resolution (Kamp & Reyle, 1993), or what to do with presuppositions (van der Sandt, 1992). Moreover, the literature consensus is that this prior dialogue is in some way *structured* to facilitate this process, though opinions differ on *what* that structure *is* (Asher & Lascarides, 2003; Ginzburg, 2012; Roberts, 2012).

This also means that an interpreted utterance is *forward-projecting*: it constrains the space of possible interpretations for future utterances and, thereby, also constrains the future structure of the dialogue. This has been taken to mean that an utterance *projects* particular *continuations* of the dialogue (Clark, 1996; Farkas & Bruce, 2010; Krifka, 2015). For instance, an assertion projects acceptance, but disprefers rejection, while a polar question projects two possible continuations: positive or negative answer, where any further preference is marked (Krifka, 2015).

One can take these projections to be part of the *meaning* of particular speech acts, and directly encode, say, that assertion projects acceptance in the dialogue-update rule for assertion (Farkas & Bruce, 2010).[1] One can, however, also attempt to derive these projections from more general principles. Nicholas Asher and Alex Lascarides (2003; 2009; 2013) do this as follows.

**Intentions from Observable Behaviour.**    The *basic* function of a speech act is, to Asher & Lascarides, the undertaking of a commitment to the speech act's content.[2] They then stipulate axioms for the *typical expectations* that interlocutors have of one another, based on what they have committed to. These axioms interface with the dialogue commitments to validate inferences towards the *interlocutors' intentions.* Such intentions, by and large, correspond to the preferred continuations of the dialogue. That is, if you ask a question, you *intend* to get an answer and, thus, an answer would be your preferred continuation. Standard assumptions of cooperativity entail that a cooperative interlocutor follows your overt preferences (unless they cannot); thus, your preferred continuation is *projected.* These inferred intentions projections are, however, *derivative* and not part of a speech act's meaning.

Let me introduce some notation. Here and henceforth, $>$ denotes a default conditional ($p > q$ expressing roughly that 'if $p$, then typically $q$'), $C_A\varphi$ means that the speaker $A$ is publicly *committed* to the content $\varphi$ and $I_A\varphi$ means that the speaker $A$ *intends* to establish a state that brings about $\varphi$. Then, write $\mathrel{\vert\!\sim}$ for a

---

[1] Arguably, this would make it an *empirical* problem to find out which act projects what.

[2] They derive this from the writings of Charles Hamblin (1970) and Robert Brandom (1983). These two authors were only concerned with assertions, but Asher & Lascarides apply a commitment account to *any* speech act.

defeasible proof theory satisfying *defeasible modus ponens* on $>$, i.e. $\Gamma, p > q, p \mathrel{|\!\sim} q$ unless $\Gamma, p > q, p, q \mathrel{|\!\sim} \bot$. This in particular means that if $p$ and $r$ are logically independent, then $p > q, r > \neg q, p, r \mathrel{|\!\not\sim} q$ and $p > q, r > \neg q, p, r \mathrel{|\!\not\sim} \neg q$, i.e. two contradictory default conditionals cancel each other out. If $p > q$ and $p$ are in the context, but $q$ is not inferred, I say that the default inference is *defeated*.

With this notation, the relevant axioms for analysing assertion are as follows.

**(Intent to Share Commitment)**   $C_A \varphi > C_A I_A C_B \varphi$.

**(Cooperativity)**   $C_A I_A \varphi > I_B \varphi$.

The first axiom expresses that, typically, commitments are intended to be shared. The second axiom expresses that, typically, publicised intentions are shared by cooperative interlocutors. Both axioms are taken to be properties of conversation in general, not just of assertion (following Clark 1996). Now, under the assumption that a speaker $A$ who asserts that $p$ undertakes the commitment $C_A p$, the derivation that *assertion projects agreement* goes as follows.

| | | |
|---|---|---|
| | $C_A p$ | $A$ asserts that $p$ (i.e. commits to $p$). |
| $\mathrel{|\!\sim}$ | $C_A I_A C_B p$ | by (Intent to Share Commitment). |
| $\mathrel{|\!\sim}$ | $I_B C_B p$ | by (Cooperativity). |
| $\rightsquigarrow$ | $B$ will agree. | |

This seems appropriate. However, Nicholas Asher posed the following three problems to me. First, it is unclear what the *meaning* of the operators $I_A/I_B$ is. While Asher & Lascarides (2003, 2008) give a modal semantics to these operators, it is unclear on what basis one would judge a statement like $I_B C_B p$ true. Second, in non-cooperative settings both axioms above plausibly fail, e.g. in debates where the interlocutors' intentions are not to convince their opponents, but to score points with an electorate (Asher *et al.*, 2017). Yet, in such settings, it seems, an assertion still affects the future dialogue in broadly the same way. Third, the axioms above are stipulative and need to be justified.

I address these problems as follows. In line with the approach of the previous chapter, I give a partial, *proof-theoretic* definition of what it *means to commit*. It is then straightforward to model *and derive* projected continuations in a *temporal* logic. More generally, my account then derives appropriate analogues of *Intent to Share Commitment* and *Cooperativity* from first principles.

**The Temporal Dimension.**   Talk of projected continuations is talk of *future* speech acts. Thus, my first step is to amend the formal language of commitment logic with temporal logic operators $\Diamond$ (*eventually*) and $\Box$ (*henceforth*, equivalent to $\neg \Diamond \neg$). For simplicity, I assume that these modal operators obey the axioms of the normal modal logic KT4C. Nicholas Asher suggests to me that more sophisticated temporal logics such as LTL (Pnueli, 1977) would be a more appropriate choice

here. I am inclined to agree, but I remain with KT4C as not to commit myself needlessly to additional assumptions about the structure of time. I do not need anything more expressive, and KT4 seems to express universal truths about temporal reasoning, if we accept that *henceforth* includes *now* (Goranko & Galton, 2015). Then, the convergence axiom (C) expresses that time does not split: KT4 supports that *eventually, it will forever be p* and *eventually, it will forever not be p* are consistent. But this does not correspond to how *eventually* is used in language. So I add (C) to rule out such cases.

It might further be objected that KT4C presupposes a view of time as a sequence of discretised instants. However, for my purposes, we can consider one step in time to be equivalent to one speech event in the dialogue. The, potentially continuous, time between and during speech events is of no consequence here.

**3.1.1.** Definition. **Temporal Logic** is the normal modal logic KT4C.

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$.

(N) $\vdash \varphi$ entails $\vdash \Box\varphi$.

(T) $\Box\varphi \to \varphi$. (If something is true henceforth, it is also true now.)

(4) $\Diamond\Diamond\varphi \to \Diamond\varphi$. (If eventually eventually $\varphi$, then eventually $\varphi$.)

(C) $\Diamond\Box\varphi \to \Box\Diamond\varphi$. (see above)

This extension of the language of commitment logic allows me to speak of *future projections*. Say that a speaker $A$ *projects* a situation $\alpha$ if that speaker is committed to $\alpha$ eventually obtaining; that is, a speech act $a$ projects $\alpha$ if making that speech act entails the projection of $\alpha$.

---

(I) A speech act $a$ **projects** a situation $\alpha$ if in the situation where a speaker $S$ makes $a$ it holds that $C_S\Diamond\alpha$.

---

**Basic Principles of Commitment.**   Part of my goal is to derive these projections from more fundamental principles. Here is one such principle: To undertake a commitment is also to undertake a commitment to *keeping the commitment*. Quite literally, a commitment that does not persist is not a commitment. Thus, I consider the following axioms to be *first principles* of any talk of commitment.

**3.1.2.** Definition (Temporal Basic Principles).  The following axioms are (partly) constitutive of **commitment**:

(a) $C_S\varphi > C_S\Box C_S\varphi$ (for each speaker $S$).

(b) $C_S\Diamond C_S\Diamond\varphi > C_S\Diamond\varphi$ (for each speaker $S$).

That is, speakers are (defeasibly) committed to keeping their commitments.

The principle (a) states that if one makes a commitment, one commits to keeping it. Principle (b) expresses an analogue to the temporal logic Axiom 4. This can be spelled out as follows. If I commit that the future is such that I eventually will be committed to the future being such that eventually $\varphi$, then all futures where I have kept that commitment are such that eventually $\varphi$; thus, I am already committed to the future being such that eventually $\varphi$. These axioms have a privileged status over the principles *Intent to Share Commitment* and *Cooperativity* discussed earlier. The principles in Definition 3.1.2 are *constitutive* of commitment: that one commits to keeping a commitment is part of what it means to commit.

Note that the Basic Principles are formulated using the *nonmonotonic* (i.e. defeasible) conditional introduced above. This is for the following reason. If commitment is the basic operation in dialogue and the Basic Principles are foundational to commitment, they need to be maintained (at least as a pretence) even in non-cooperative dialogue. This is part of what Asher & Lascarides (2013) call *basic cooperative*; the minimal cooperativity required to engage in conversation at all. Plainly, there is the non-cooperative move of making a false commitment: a commitment one does not consider oneself to be bound by (e.g. an unfaithful promise (Searle, 1969, ch3) or a plausibly deniable commitment (Pinker *et al.*, 2008)). In such cases, if the unfaithfulness is discovered later, the inference from commitment to kept commitment must be defeated. This is a special case of the rationale behind using a logic of defeasible inference in pragmatic inference (Asher & Lascarides, 2003): Since new information may cause a revision of a previous interpretation, many interpretative inferences are left defeasible.

One may argue that such insincere commitments are not commitments at all. I am inclined to agree in principle, but the insincere speaker still *feigns* a commitment and one must recognise *in hindsight* that it was insincere. Since my goal is to reduce reasoning about future projections to observable behaviour only, the model needs to deal with such deceit and what it means for the dialogue.

As a case in point, the non-cooperative move of making a false commitment would not *work* if the principle (a) were not active. In fact, this is analogous to the non-cooperative action of lying. But even in non-cooperative settings, speakers do not lie all the time and it is frequently rational to assume that truth is spoken. Thus, all else being equal, it is frequently rational to maintain the inference that speakers voice their beliefs, as well as the inference that commitments are intended to be kept. But in non-cooperative settings one should be ready to revise this. The possibility to be insincere extends to publicised commitments to future commitments as well: one can commit to eventually committing to something without intending to ever doing so; hence (b) is a mere default as well.

There is, of course, more to the undertaking of a commitment than what is recorded in Definition 3.1.2. For one, undertaking a commitment to a proposition entails that one is willing and able to defend that proposition (Brandom, 1983); I return to this in Chapter 7.

**Discursively Relevant Intent.**    Simply put, intentions are private and hence opaque to everyone but the person having them. However, if an intention is to be discursively relevant, it ought to be apparent to the interlocutors (Austin, 1962; Grice, 1975). As Asher & Lascarides (2013) argue, speakers can be purposefully *deceptive* about their true intentions by broadcasting signals that make *other* intentions inferable. Thus, my goal here is not to arrive at a cognitive theory of intent, but to develop a reductive strategy that *replaces* talk of intent (in discourse) by talk of *projections.*

First note that having an intention has an inherently temporal component: an intention relates to some (potential) future. Furthermore, intending an act or state $\alpha$ includes a certain dedication to undertake that act or bring forth that state (Cohen & Levesque, 1990). This is a stronger attitude than merely *preferring* $\alpha$ over not-$\alpha$ or desiring or wanting $\alpha$. For a simple example, consider someone who desires to go out for fine dining, but cannot afford to do so. It is perfectly reasonable to attribute to someone that *they want to go out, but choose not to*, but *they intend to go out, but choose not to* strikes me as more than just a little odd.

Such a dedication to action can be conceptualised as a commitment. Say that an intention to $\alpha$ can be represented as a *commitment* (to oneself) to bring about $\alpha$ and a publicised intention to $\alpha$ is a *public commitment* (to one's interlocutors) to bring about $\alpha$. This reduction of intent compares well with the felicity conditions for speech acts posited by Austin (1962). According to him, a successfully performed speech act must be made with the right intentions and, moreover, these intentions must be apparent to one's addressees and properly *taken up* by them. In the terminology I am developing here, a speech act publicises a commitment to a (potential) future state or act; uptake consists in the addressees sharing the commitment. Only after all participants have thereby fulfilled the intentional felicity conditions does the illocutionary force of the act apply.

That is, what has been called an intention (in discourse) reduces to operators that are already quite well understood: the operators for *public commitment* ($C_A$, $C_B$) and the *temporal modal operators* ($\Diamond$, $\Box$). So one might write $C_A \Diamond \alpha$ instead of Asher & Lascarides' $I_A \alpha$. But to outright identify $I_A$ with $C_A \Diamond$ would be a mistake. The commitment expressed in $C_A$ is *public* commitment, and this notion has nothing to say about (not publicised) commitments to oneself—these are just as opaque as private intentions. But this is not a problem, but part and parcel of my general argument. As said, as far as intentions are required for the understanding of a speech act, they ought to be (implicitly or explicitly) publicised. In fact, the intentions going along with a speech act must be *apparent* to the addressee of that act. And even in non-cooperative settings, where the apparent intentions might be purposefully misleading, a speaker nevertheless needs to display *some* intention (even if it is dishonest) when making a speech act. Making explicit how speakers may reason from displayed behaviour to apparent intentions is one of my main goals here.

However, the form $C_A \Diamond \alpha$ for the intention that $\alpha$ is indistinguishable from the claim that $\alpha$ *will happen.* Surely, claiming that something will happen is much stronger than intending it. The difference is, however, smaller than it appears. As argued above, an intention $\alpha$ incurs a dedication to bringing forth $\alpha$. It is a minor step from there to conclude that the publicised intention to $\alpha$ amounts to the claim that one will cause $\alpha$ unless something unexpected happens. As a case in point, notice that it is absurd to intend something that one saliently cannot do.[3] In this sense, intentions must be understood as *defeasible* commitments: they reduce to the claim that the content of the intention will obtain *unless* circumstances unexpectedly prevent this. Thus, as far as the dialogue is concerned, I am on the safe side when all projections ('intentions') I derive are *defeasibly* inferred.

**The Dialogue Dimension.** Trivially, according to the Basic Principles above, any speech act projects the maintenance of the commitments undertaken by making that act. However, committing to a proposition is rarely if ever a goal in and of itself, so there must be something more to their purpose. To get to the bottom of this, one first must ask what the overall purpose of conversation itself is; since speakers make speech acts to engage in conversation, their purpose can only be understood with respect to the purpose of the conversation.

The broadest characterisation of such a purpose is that conversation serves the *exchange of information* and thereby the establishment of particular information *as shared* (Stalnaker, 1978; Skyrms, 2010). In the context of a commitment-based model, establishing information as shared means having a shared commitment to a particular proposition (Lascarides & Asher, 2009). Therefore, I fix a *discursive* basic principle: that both speakers engaged in a dialogue *project that information will be established as shared.*

**3.1.3.** DEFINITION (Discoursive Basic Principle). The following axiom is (partly) constitutive of conversing (where $p <> q$ abbreviates $p > q \land q > p$):

  (c) $C_S \Diamond \Box (C_A \varphi <> C_B \varphi)$ (for each pair of speakers $A$, $B$).
     Speakers are committed to resolving issues.

The Basic Principle (c) states that both speakers are committed to aligning with each other on all issues $\varphi$ raised in the conversation. This does not mean to agree to $\varphi$, but rather to reach a state where they *share* a single commitment with regards to $\varphi$; that is, either both speakers commit to $\varphi$ or both do not commit to $\varphi$. Of course, is an idealisation. Plainly, dialogues might be aborted at any point and so leave issues open. But this does not affect my general point that it is still

---

[3]There might be situations where a speaker claims to intend something objectively impossible. In such cases, the speaker is mistaken. This seems to be similar phenomenon to someone claiming to *know* something false.

part of a conversation's *purpose* to resolve issues. Cases where speakers *agree to disagree* is one where the defeasible biconditional is defeated, i.e. the interlocutors agree that they both undertook a commitment, but relinquish the commitment to share that commitment. Note that (c) leaves open that both speakers remain *not committed* to some $\varphi$; e.g. for speakers $A$ and $B$ that $\neg C_A\varphi$ and $\neg C_B\varphi$. This is the case for any issue $\varphi$ that is never raised in the dialogue.

Nicholas Asher informed me of an objection. He and his collaborators have argued that there are some, in particular non-cooperative, contexts where a dialogue takes place that is *not* properly analysed as an exchange of information (Asher & Paul, 2016; Asher *et al.*, 2017). For instance, as mentioned earlier, a political debate is not (primarily) about the exchange of information between two politicians, but about convincing *an audience* such as their electorate; in cross-examination, a lawyer might ask a witness questions that are not designed to be answered, but rather to sway the jury or judge in some way.

I agree in principle, but contend that even in these situations the speakers are still operating under *the pretence* of trying to exchange information. To wit, politicians in debate will still exchange arguments *with each other* and *as if trying to convince* each other; a lawyer must still phrase a question that *appears* to be an honest inquiry for information, lest she be reprimanded by the judge. Then note that, in line with the previous argument on intentionality, it does not matter whether the *covert* intent of the interlocutors is something else—their *displayed* intent is what is relevant to the interpretation of conversation.

**Understanding.**   Finally, I make an assumption that is, arguably, a simplifying assumption: that, generally, speakers understand each other. It is not immediately clear what this means on a commitment account. Antoine Venant *et al.* (2014) have argued that *understanding* is also something that needs to be displayed publicly, if it is to be conversationally relevant (arguably, this is the purpose of backchannel utterances). Displaying something publicly is, on a pure commitment account, always the undertaking of a commitment. Thus, so Venant *et al.* argue, an addressee $B$ understanding that $A$ committed to $\varphi$ can be modelled precisely as $B$ undertaking a commitment to that fact; $C_B C_A \varphi$.

**3.1.4.** Definition (Simplifying Assumption). Speakers understand each other:
  (d)  $C_A\varphi > C_B C_A\varphi$ (for each pair of speakers $A$, $B$).
      If one speaker makes a commitment, the other speaker commits to that fact.

This mirrors Stalnaker's (1978) claim that if one makes an assertion, then the common ground records that the assertion has taken place. On the commitment account, the common ground is derivatively defined as *shared commitment* (Asher & Lascarides, 2008; Farkas & Bruce, 2010). Principle (d) expresses that if one speaker undertakes a commitment, then all other speakers undertake a commitment

| | Level | Joint Action | Ex. Clarification |
|---|---|---|---|
| 1 | contact | A and B pay attention to each other | *Are you talking to me?* |
| 2 | perception | A produces a signal and B perceives it | *What did you say?* |
| 3 | understanding | A conveys a meaning and B understands it | *What did you mean?* |
| 4.1 | recognition | A projects a future and B recognises it | *What do you want?* |
| 4.2 | adoption | A proposes an action and B accepts it | *Why? / How?* |

Table 3.1: Grounding hierarchy for speaker A and addressee B, including constructed examples for possible clarification requests indicating failure to ground the corresponding level (adapted from Schlöder & Fernández (2015a)).

to that fact, i.e. there is a shared commitment to the fact that the commitment has been made.

The assumption (d) is phrased as a defeasible conditional to acknowledge the fact that understanding is not always achieved. This, however, does not do justice to a host of issues that some rightly believe to be of central importance to the analysis of dialogue: grounding (Clark, 1996), clarification requests (Ginzburg, 2012), and misunderstandings (Hulstijn, 2000), among others. Here I am concerned with the dialogue dynamics that are relevant *after* these issues have been negotiated.

To be more precise, the simplifying assumption (d) abstracts away from what I have called *lower-level grounding* in previous work (Schlöder & Fernández, 2015a), i.e. from roughly the first three levels in Clark's grounding hierarchy (Clark, 1996); see Table 3.1. The simplifying assumption abstracts away from levels 1–3 and assumes that level 3 is achieved by default. In what follows, I investigate how the principles (a)–(d) relate to the *uptake* levels 4.1 and 4.2. That is, I derive which projections can be derived from assertions and rejections (level 4.1) and that cooperative speakers are expected to take up these projections (level 4.2).

To sum up, I claim that the following four principles are fundamental to the notion of public commitment and to the activity of conversation.

---

(II) **Basic Principles of the Commitment Account**

    (a) $C_S\varphi > C_S\Box C_S\varphi$ (for each speaker $S$).

    (b) $C_S\Diamond C_S\Diamond\varphi > C_S\Diamond\varphi$ (for each speaker $S$).

    (c) $C_S\Diamond\Box(C_A\varphi <> C_B\varphi)$ (for each pair of speakers $A$, $B$).

    (d) $C_A\varphi > C_B C_A\varphi$ (for each pair of speakers $A$, $B$).

---

## 3.2   Assertion and Rejection

**Assertion, Rejection and Commitment.**   As already mentioned, it is clear
how to understand an *assertion that p* in a commitment framework: in asserting,
the speaker undertakes a commitment to some *p*. It is less clear, however, how a
*rejection of p* should be understood. One could follow Frege (1919) and Rumfitt
(2000) in considering this act to be (equivalent to) the assertion that ¬*p*; so a
rejection of *p* would be the undertaking of a commitment to ¬*p*.

   However, as I have argued in Chapter 2, this would not capture the breadth of
linguistic data. There, I characterised the *weak rejection of p* as expressing the
weakest attitude that is incompatible with the assertion of *p*. On the commitment-
based account, there are no attitudes, but only commitments. But we may consider
the weakest *commitment* one can undertake that rules out assertion. This is the
*commitment to not committing*.

---

(III)  **Assertion and Rejection in the Commitment-Based Account**

   In asserting that *p*, a speaker *A* undertakes the commitment that $C_A p$.

   In rejecting *p*, a speaker *A* undertakes the commitment that $C_A \neg C_A p$.

---

These are incompatible in the following sense. If $C_A p$ then by Basic Principle (a),
and temporal axiom (T), $C_A p \mathrel{|\!\sim} C_A \square C_A p \mathrel{|\!\sim} C_A C_A p$. Thus, $C_A p, C_A \neg C_A p \mathrel{|\!\sim} C_A \bot$.
Thus, the commitments of a speaker asserting and rejecting the same proposition
are absurd and she appears as incoherent.

   One may be able to articulate these ideas in a *dynamic* logic where there are
action-operators for speech acts that *update* a commitment model with the new
commitments that a speaker undertakes (see Venant & Asher 2015). However, I
will remain with a static logic. In the model theory I give in Section 3.5, a dialogue
is represented as a sequence of speech acts, and there is one static commitment
model *per initial segment* of that sequence. Updating the dialogue model with a
new speech act is to lengthen the sequence and compute a new commitment model,
based on the prior information. I relegate the model update to an extraneous
mechanism (see the development of Hindsight Logic in Chapter 6 for details).

   There are advantages to having this set-up. A dynamic logic allows only the
*update* of information, but has generally a hard time supporting the *revision*
of previous information. However, in dialogue it is frequently the case that
a new speech act prompts one to revise the interpretation of a former act *in
hindsight* (Asher & Lascarides, 2003; Schlöder & Lascarides, 2015). By maintaining
a *sequence* of static commitment models, prior information is preserved and
accessible. For instance, if a commitment gets retracted, it needs to be removed
when computing the present commitment state. On a standard dynamic account,
one would delete the update that led to the commitment in the first place. But in

the discourse between the commitment and its retraction, the commitment might have been relevant. Thus, deleting the update would render the representation of the discourse incoherent. However, in a sequential static theory it is trivial to record that one *was* committed to something for some time, but now *is* not.

**Projection of Assertion.** Now I am able to derive a particular projection of the speech act of assertion. The following derivation show that an assertion projects agreement. Suppose that speaker $A$ asserts the proposition $p$. Then:

$$
\begin{array}{lll}
& C_A p & A \text{ asserts that } p. \\
\vdash & C_A \Box C_A p & \text{by Basic Principle (a).} \\
\vdash & C_A \Box\Box C_A p & \text{by temporal axiom (4).} \\
\mid\!\sim & C_A \Diamond\Box C_B p & \text{by Basic Principle (c).} \\
\mid\!\sim & C_A \Diamond C_B p & \text{by temporal axiom (T).} \\
\rightsquigarrow & A \text{ projects a future where } B \text{ agrees to } p.
\end{array}
$$

Note in particular that this derivation shows that, in general, $C_A p \mid\!\sim C_A \Diamond C_B p$. Thus, the Basic Principles *derive* an analogue the axiom *Intent to Share Commitment*: that commitments project sharing.

This projection seems to be of particular significance with regards to the purpose of asserting and rejecting in dialogue. So, while an arbitrary number of projections can be derived from my axioms, it appears to be right to single this one out as the *primary* or *salient* one for assertion (though I cannot offer a formal measure of this saliency here). That is, in the terminology of the grounding hierarchy (Table 3.1), this projection corresponds to the grounding level 4.1: one speaker is projecting a future and the addressee is expected (by default) to *recognise* what is being projected. To *share* the projection would correspond to level 4.2.

For a rejection move $C_A \neg C_A p$ this immediately delivers that $C_A \Diamond C_B \neg C_A p$, i.e. that $A$ projects that $B$ accepts that $A$ is not committed to $p$. If we accept that what is *common ground* is what the speakers have *shared a commitment to* (Asher & Lascarides, 2008; Farkas & Bruce, 2010), then the projection of $A$ rejecting $p$ is that $B$ accepts that $p$ will not be common ground. In the special case where $B$ has previously asserted that $p$, this corresponds to a projected future in which $B$ accepts that the essential effect of her assertion has been cancelled (see Chapter 2). However, there is more to rejection than this.

**Rejecting Rejections.** Assertions are not the only speech act that can be rejected. In fact, *any* speech act can have its force *cancelled* (or *aborted*; Austin 1962). Thus, there is a more general sense in which a speech act can be rejected. My definition (III) is about rejecting *propositions*, but speech acts are not propositions.

As a case study, I now address the phenomenon of *rejections of rejections*. Consider the following constructed dialogue and its German translation.

(17)  a. Amy: It is the case that *p*.        (18)  a. Amy: Es ist der Fall, dass *p*.

      b. Bob:  No, it's not.                             b. Bob:  Nein, ist es nicht.

      c. Amy: Yes, it is.                                  c. Amy: Doch, ist es.

It is not obvious—in English—that Amy's utterance in (17c) is a rejection of Bob's rejection in (17b), rather than a re-assertion of Amy's own (17a). However, in the German version, Amy's utterance in (18c) contains the negative polarity item *doch* which requires a negated antecedent (Krifka, 2013). Thus, (18c) must rhetorically relate to (18b) and it is implausible to model it as a re-assertion of *p*. While English lacks a particle like *doch*, there are cases where a rejection syntactically mirrors the shape of a previous rejection. The following examples are from the AMI corpus (Carletta, 2007).

(19)  a. A: That's dependent on the television.

      b. B: No, <u>I don't think so</u>.

      c. A: <u>I do know so</u>.                                    (AMI corpus, dialogue TS3008d)

(20)  a. A: Mushroom <u>is a vegetable</u>.

      b. B: I don't think it is.

      c. A: <u>It's vegetable.</u>                                     (AMI corpus, dialogue IS1003d)

In (19), the form of (19c) mirrors (19b), suggesting that (19c) functions as a rejection of the rejection (19b). In contrast, the form of (20c) suggests that it indeed works as a re-assertion of (20a).

I do not want to argue that rejection-of-rejection and re-assertion have different *effects*; surely, both (19c) and (20c) have the effect of *A* indicating that she stands by her original assertion. Rather, I take (18) and (19) to evince that there *is* a linguistic phenomenon of rejection-of-rejection that needs to be modelled as such. I now show that my approach can be generalised to model these cases without simply defining rejections-of-rejections to be re-assertions.

First note that a naïve generalisation of the idea that a rejection is a *commitment to not commit* will not do. This can be appreciated by considering the logical forms added in (21). That is, (21a) results in *A* undertaking a commitment to *p* and (21b) in *B* undertaking a commitment not to share *A*'s commitment. However, representing the rejection-of-a-rejection in (21c) as *A* committing not to share in *B*'s commitment is odd.

(21)  a. Amy: It is the case that *p*. $C_A p$

      b. Bob:  No, it's not.          $C_B \neg C_B p$

      c. Amy: Yes, it is.            $??C_A \neg C_A \neg C_B p$

The commitment $A$ is predicted to undertake in (21c) misses the target: it expresses, roughly, that $A$ is committed to her commitments being compatible with $B$ being committed to $p$. This is hardly a rejection of (21b). Clearly, the previous characterisation of rejection is only correct for rejections of *propositions*, but speech acts are not propositions and rejecting *them* is different. One can make sense of this as follows.

**Rejecting Speech Acts.** Following Clark (1996) and my own earlier work (Schlöder, 2014), I take a speech act to *propose a joint project* and hence the *rejection* of a speech act to be the act that expresses the unwillingness to participate in the associated project. But if rejections are speech acts and speech acts are joint project proposals, then a rejecting speech act itself proposes a joint project. This is the project of cancelling a previously proposed project. For instance, an assertion proposes the joint project of making some proposition common ground; the rejection of this assertion proposes to cancel this project.

This view has striking consequences. First, that general rejection is *not about truth*, since some speech acts cannot be understood truth-conditionally. Second, that it is *not* the case that a rejection move immediately entails the cancellation of a prior proposal, but the project of cancellation is *itself* up for discussion (Schlöder, 2014). For instance, one can query a rejecting *no* with the question *why not?* to discuss the cancellation of the rejected project; as, for instance, in the following example from the British National Corpus (Burnard, 2000).

(22) a. Anon 2: Just take it into school.

b. Richard: No.

c. Richard: No way!

d. Anon 2: Why not? (BNC, KSV, 3786–3789)

Now, in the model articulated so far, speech acts make *projections*. We can conceptualise these to describe the desired state after the completion of the joint projects associated with an act. One can express cancellation of a project by denying that it will be completed. Thus, it makes sense to say that to reject a speech act is to deny its projection.

---

(IV) **General Rejection**

Suppose $a$ is a speech act that projects $\alpha$. If a speaker $B$ rejects $\alpha$, $B$ is undertaking a commitment to $C_B \neg \Diamond \alpha$. I assume that, all else being equal, there is a *primary* or *salient* projection $\alpha$ that is associated with $a$ and that it is this $\alpha$ that is targeted by unmarked rejections of $a$.

---

This definition generalises on the earlier definition of rejecting a proposition. As the following derivation shows, rejecting an assertion of $p$ (i.e. rejecting its projection) results in rejecting the asserted proposition.

$$C_B \neg \Diamond C_B p \quad B \text{ rejects } \alpha = \Diamond C_B p.$$
$$\vdash \quad C_B \Box \neg C_B p \quad \text{by a modal logic validity.}$$
$$\vdash \quad C_B \neg C_B p \quad \text{by temporal axiom (T).}$$

Thus, generalising the notion of rejection to arbitrary speech acts is conservative over my prior discussion of rejections of propositions.

This yields a more perspicuous derivation of what is projected by a rejection of an assertion. If $B$ rejects $A$'s assertion of $p$ this is $C_B \neg \Diamond C_B p$ according to (IV). Thus, the projection of a rejection of an assertion is computed as follows.

$$C_B \neg \Diamond C_B p \quad B \text{ rejects the assertion that } p.$$
$$\vdash \quad C_B \Box \neg C_B p \quad \text{by a modal logic validity}$$
$$\vdash \quad C_B \Box \Box \neg C_B p \quad \text{by temporal axiom (4).}$$
$$\mathrel{|\!\sim} \quad C_B \Diamond \Box \neg C_A p \quad \text{by Basic Principle (c).}$$
$$\mathrel{|\!\sim} \quad C_B \Diamond \neg C_A p \quad \text{by temporal axiom (T).}$$
$$\rightsquigarrow \quad B \text{ projects a future where } A \text{ retracts her commitment to } p.$$

Now, I take this projection to a retraction to be the discursively salient projection of a rejection of an assertion. Thus, the meaning of the rejection of that rejection (e.g. in (17)) is as follows.

$$C_A \neg \Diamond \neg C_A p \quad A \text{ rejects } \alpha = \Diamond \neg C_A p.$$
$$\vdash \quad C_A \Box C_A p \quad \text{by a modal logic validity.}$$
$$\vdash \quad C_A C_A p \quad \text{by temporal axiom (T).}$$
$$\rightsquigarrow \quad A \text{ is committed to her commitment that } p.$$

The intuitive reading of this derivation is that $A$ is confirming or renewing her commitment to $p$ as made in (21a). This seems correct for rejecting a rejection.

## 3.3 Cooperative Commitments

**Cooperative Axioms.** Everything said so far goes for cooperative as well as non-cooperative settings. One of my initial goals, however, was to reproduce the *Cooperativity* axiom of Asher & Lascarides. This axiom models that cooperative speakers adopt publicised intentions. Put differently, once an addressee has *understood* the projection of an act, she is expected to *take up* that act and participate in the projection. I now demonstrate that this kind of cooperativity varies with assumptions made on the strength of commitments. Consider the following three modal logic axioms.

$$\text{(D) } \neg C_A \bot \qquad \text{(4) } C_A \varphi \to C_A C_A \varphi \qquad \text{(5) } \neg C_A \to C_A \neg C_A \varphi.$$

These axioms express that one is committed to ones current commitment state. This relates to cooperativity as follows. First, (4) and (5) express that one is

committed to one's *implicit* commitments as well. If, say, the commitment $C_A p$ entails that $A$ is *not* committed to $q$ ($\neg C_A q$), then (5) entails that she is committed to that consequence as well ($C_A \neg C_A q$); consider, for example, $p=$ *Homer does not exist* and $q=$ *Homer wrote the Iliad*. Similarly, if $C_A p$ entails $C_A q$, then (4) makes $A$ committed to that consequence ($C_A C_A q$). Thus, (4)+(5) express that cooperative speakers honour their implicit commitments.

Second, these axioms formalise sincerity on commitments: (D)+(5) expresses that one cannot claim to have made a commitment ($C_A C_A p$) that one in fact has not made ($\neg C_A p$); (D)+(4) that a speaker cannot deny a commitment ($C_A \neg C_A p$) they in fact have undertaken ($C_A p$). To see the consequences of this, note how (D)+(5) rule out the non-cooperative act in (23a) and (D)+(4) the one in (23a′).[4]

(23)  a.  Amy: I have always maintained that $p$.    $C_A C_A p$

    *(when, in fact, A never asserted p, i.e. $\neg C_A p$).*

  a.′ Amy: I never claimed that $p$.    $C_A \neg C_A p$

    *(when, in fact, A asserted p previously, i.e. $C_A p$.)*

Claims similar to (23) can be found, e.g., in political debates. This demonstrates that in non-cooperative settings one cannot trust one's interlocutor to be sincere about their own commitments. Thus, I refer to KD45 operators $C_A$ as *cooperative commitments*. Note that, in fact, a defeasible version of Axiom (4) is already entailed by the Basic Principle (a) and the temporal axiom (T). However, the monotonic version and the addition of Axiom (D) are required to make (23a′) impossible (contradictory), not merely incoherent.

It might appear counterintuitive at this point that Axiom (5) allows one to infer from a *non*-commitment ($\neg C_S p$) that the speaker $S$ is committed to that non-commitment ($C_S \neg C_S p$). The concern can be appreciated in the following derivation: if $\neg C_S p$ then $C_S \neg C_S p$, (by (5)) so $C_S \square C_S p \neg C_S p$ (by (a)) and hence $C_S \square \neg C_S p$ (by (D)+(4)). This says that $S$ is committed to never change her mind, which seems odd if she is neutral towards $p$, which may be expressed as $\neg C_S p \wedge \neg C_S \neg p$. However, recall that the inferences licensed by in (a) are defeasible. Indeed, (a) should not be applied in this case. This can be seen as follows. The fact that a speaker is *neutral* to a proposition $p$ recorded only *partially* in the fact that $\neg C_S p \wedge \neg C_S \neg p$. To be neutral also means to, in principle, be open to accepting it (given sufficient convincing, say). Thus, to represent a speaker being

---

[4]I contend that it is perfectly natural to model claims about past commitments as the undertaking of commitments about one's (present) commitments, since past commitments subsist (unless they are retracted). However, given the present framework with a temporal dimension, one may think it more principled to model these cases with backward-looking temporal operators, say $\blacksquare/\blacklozenge$. Then, the content of (23a) is $C_A \blacksquare C_A p$ and the fact that $A$ never asserted $p$ is $\neg \blacklozenge C_A p$. Note that these two are consistent, since $C_A$ does not have Axiom T. To derive a contradiction, we need to eliminate the temporal operators with the temporal Axiom T: $C_A \blacksquare C_A p \vdash C_A C_A$ and $\neg \blacklozenge C_A p \vdash \neg C_A p$. Then, (D)+(5) on $C_A$ derive a contradiction. So it seems that the backwards temporal operators do not add anything here.

neutral also means that $\neg C_S \square \neg C_S p$. Adding this to the premises we use to reason about the discourse defeats the unwanted inference.

So, the logical form $C_S \neg C_S p$ does double duty: it is the commitment one *undertakes* when rejecting $p$, but it is *also* a commitment that one *has* as a part of cooperativity. These are conflated in the *static* representation of one point in the conversation, but in the *dynamics* of model update, the difference matters. This means that a commitment structure must not only record commitments, but also needs to make sure that the correct default inferences are drawn. In the present proof theoretic approach this is realised by stipulating defeaters. Accordingly, in the model theory I give below, the semantics of $>$ is sensitive to time.

**Cooperativity.** I now show that if the operators $C_A$ and $C_B$ are assumed to be $\square$-like KD45 modals, then the Basic Principles entail the cooperative principle that *understanding entails acceptance*; this inference is of course only made defeasibly, since rejections can occur for innocuous reasons in cooperative settings as well. This precisely formalises a claim made in my earlier work: that a cooperative addressee is expected to take up a proposal unless there is a reason not to (Schlöder & Fernández, 2014). Sometimes, speakers demonstrate awareness of this default expectation, as evinced by the following example from the British National Corpus.

(24)  a. Anon 2: Do you agree with that?

　　　b. George: I have no reason to disagree. Yes.　　　(BNC, FMN, 492–493)

The following theorem shows that under cooperative commitments (i.e. KD45), *understanding entails uptake.* That is, if an addressee understood what their interlocutor's projection is (level 4.1 in Table 3.1), then that addressee is expected to take up that projection (level 4.2 in Table 3.1). Unless, that is, the addressee has a reason not to, which would defeat the defeasible inference to uptake.

**3.3.1.** THEOREM (Cooperative Commitment). *Assume that $C_A$ and $C_B$ satisfy the KD45 axioms. Then $C_B C_A \Diamond \varphi \mathrel{\vert\!\sim} C_B \Diamond \varphi$.*

**Proof:**

$$
\begin{array}{lll}
& C_B C_A \Diamond \varphi & \text{(assumption).} \\
\vdash & C_B \square C_B C_A \Diamond \varphi & \text{by Basic Principle (a).} \\
\vdash & C_B \square \square C_B C_A \Diamond \varphi & \text{by temporal axiom (4).} \\
\mathrel{\vert\!\sim} & C_B \Diamond \square C_A C_A \Diamond \varphi & \text{by Basic Principle (c).} \\
\mathrel{\vert\!\sim} & C_B \Diamond \square C_A \Diamond \varphi & \text{by Axioms (D)+(5).} \\
\mathrel{\vert\!\sim} & C_B \square \Diamond C_A \Diamond \varphi & \text{by temporal axiom (C).} \\
\mathrel{\vert\!\sim} & C_B \Diamond \Diamond C_B \Diamond \varphi & \text{by Basic Principle (c).} \\
\mathrel{\vert\!\sim} & C_B \Diamond C_B \Diamond \varphi & \text{by temporal axiom (4).} \\
\mathrel{\vert\!\sim} & C_B \Diamond \varphi & \text{by Basic Principle (b)} \qquad\qquad \square
\end{array}
$$

Note that this result reproduces the *Cooperativity* axiom of Asher & Lascarides.

## 3.4 Relation to Weak Bilateral Logic

In this section, I relate the commitment-based account of assertion and rejection to the weak bilateral logic of Chapter 2. First, I observe that the two models similarly relate to cooperativity. This observation is formally vindicated by a completeness result. Then, the combination of the two theories allows me to give due to the data on *strong rejections* in a principled way.

**Cooperativity and Inference.** In Chapter 2 I describe the inference relation of weak bilateral logic to preserve the attitudes that speakers are *taken to* have expressed. To be taken to assent to (dissent from) a proposition $p$ is to be required to assent to (dissent from) $p$ once $p$ becomes salient. As mentioned there, it would be uncooperative or deceptive to ignore this requirement. This relates to the cooperativity described here as follows.

First, translate the language of weak bilateral logic into the language of commitment logic by mapping the attitudes expressed by the speech acts assertion and rejection to their respective commitments; that is, I identify assent with commitment and dissent with the commitment not to commit. Note that since the signs $+$ and $-$ of WBL do not embed, this translation is only partial onto the language of modal logic. Moreover, since weak bilateral logic is about a single speaker's attitudes, I fix here a speaker $S$.

**3.4.1.** DEFINITION. Let $\varphi$ be a formula in WBL. The *modalisation* of $\varphi$ is:

$$\varphi^m = \begin{cases} \bot, & \text{if } \varphi = \bot \\ C_S A, & \text{if } \varphi = +A \\ C_S \neg C_S A, & \text{if } \varphi = -A. \end{cases}$$

Let *modalised WBL* be the inference rules of WBL where premises and conclusion have been modalised.

**3.4.2.** DEFINITION. Call a formula $\varphi$ of modal logic *bilateral* if there is a $\psi$ in WBL with $\psi^m = \varphi$. For the inverse mapping write $\varphi^b = \psi$.

Now, a speaker $S$ has two options to disregard the preservation of attitudes:

- To be taken to assent to $p$, but refusing to assent to $p$.
- To be taken to dissent from $p$, but refusing to dissent from $p$.

These are similar to the non-cooperative moves in (23), where a speaker fails to acknowledge a commitment they have (or failed to have) undertaken earlier. Thus, it seems suggestive that the cooperativity expressed in weak bilateral logic is related to the cooperativity expressed by the KD45 axioms.

Indeed, given the embedding above, it can be shown that weak bilateral logic is *the logic of cooperative commitment.* That is, weak bilateral inference preserves

cooperative commitments. Thus the notion of *taken to have a certain attitude* towards some proposition (see Chapter 2) is equivalent to having undertaken a cooperative commitment towards that proposition.

**Soundness and Completeness.**   Here, I outline a proof that, given the above embedding, the proof theory $\vdash^{\mathrm{KD45}}$ coincides with the proof theory $\vdash$ of weak bilateral logic. The full proofs are in Appendix A.

**3.4.3.** THEOREM (Soundness). *If $\Gamma$ is a set of formulae in WBL such that $\Gamma \vdash \varphi$, then $\{\psi^m \mid \psi \in \Gamma\} \vdash^{KD45} \varphi^m$.*

**3.4.4.** THEOREM (Completeness). *(i) Modalised WBL axiomatises the class of KD45 frames. Moreover, (ii) if $\Gamma$ is a collection of bilateral modal formulae and $\Gamma \vdash^{KD45} \varphi$ for a bilateral $\varphi$, then there is a proof of $\varphi^b$ in WBL from the premisses $\{\psi^b \mid \psi \in \Gamma\}$.*

The following Lemma is the only challenge in the Soundness proof.

**3.4.5.** LEMMA. *Let $\Gamma$ be a set of modal formulae such that for any $\varphi \in \Gamma$ there is an $X$ such that $\varphi = C_S X$ where $X$ does not contain further modal operators. Let $A$ and $B$ be formulae not containing modal operators. If $\Gamma, C_S A \vdash^{KD45} C_S B$ then $\Gamma \vdash^{KD45} C_S(A \to B)$.*

**Proof sketch:**
By contraposition. If there is a model of $\Gamma$ with a witness for $\neg C_S(A \to B)$, one can construct a model of $\Gamma$, $C_S A$ and $\neg C_S B$ by duplicating the witness as in the proof of Theorem 2.3.6. $\qquad\qquad\square$

To show Completeness, one can take the completeness result on $\omega$–pointed models from Chapter 2 and map $\omega$–pointed models to Kripke-style KD45 models as follows: take $\omega$ to be the set of worlds and $\omega \times \omega$ to be the accessibility relation. Then, the translation between the truth definitions interacts with Definition 3.4.1 as expected.

**Recovering Strong Rejection.**   My characterisation of the speech act of rejection, both here and in Chapter 2, seems to obscure that there *are* strong rejections, i.e. rejections that are best interpreted as equivalent to the assertion of a negative. In fact, it seems that many, if not most, rejections are strong in this sense. This can be explained as follows.

Recall that on the commitment-based account, an assertion of $p$ (by a speaker $S$) is the undertaking of the commitment $C_S A$, a (weak) rejection is $C_S \neg C_S p$, and a strong rejection can be put as the undertaking of $C_S \neg p$. Now observe that under the KD45, $C_S A$ and $C_S \neg A$ are *contrary* whereas $C_S A$ and $C_S \neg C_S A$ are *contradictory* in the classical square of opposition.

Linguistic evidence supports the claim that natural language has a general preference for contrary negation over contradictory negation (Horn, 1989, chs4–5). The most familiar case is the *neg–raising* phenomenon: when a speaker is stating certain attitudes (e.g. belief) towards a proposition, a linearly wide–scoped negation can be interpreted to take narrow scope. For example, in 'I don't believe *A*,' a literal interpretation could be $\neg\mathcal{B}A$ (with a belief modality $\mathcal{B}$), i.e. *the speaker lacks the belief that A.* However, the typical reading of such an utterance is *the speaker believes that not-A,* i.e. $\mathcal{B}\neg A$. I argue that this preference also applies to strong *vs.* weak rejections. Consider the following rejections of *Homer wrote the Iliad.*

(25)    Did Homer write the *Iliad*? No.

(26)    Did Homer write the *Iliad*? No, he did not.

(27)    Did Homer write the *Iliad*? No, Homer did not exist.

Example (26) is a strong rejection, committing the speaker to *Homer did not write the Iliad,* whereas (27) is a weak rejection that does not entail any such commitment. The bare *no* in (25) might appear to be ambiguous between the two, but this is not quite so. All else being equal, (25) gets the reading of of (26) and the weak reading in (27) arises only *because* of the added sentence *Homer did not exist.* In that sense, strong rejections are the default readings of bare *no.* That is, a bare *no* is read as a strong rejection by default, but this reading is cancellable: (27) is a continuation of (25) that cancels the default reading.

Thus, in sum, the commitment account can give a straightforward account of rejection in natural language. All rejections are *a priori* weak rejections, but there is a preference to read (contradictory) commitment-to-not-commit as (contrary) commitment to a negative.[5]

---

(V) By default, a rejection is given a **strong reading**. That is, for any speaker $S$ and proposition $p$, $C_S\neg C_S p > C_S\neg p$.

---

Note that in the case where a speaker rejects both $p$ and $\neg p$, the two instances of (V) would simply cancel each other out. That is, both rejections would be predicted to be feeble.

As the above Completeness result shows, (V) is equivalent to the claim that in a cooperative setting the default reading of dissent is assent to a negative and the default ground for unassertibility is falsity (in the parlance of Chapter 2). Since weak bilateral logic does not have a default conditional, this is not expressed there. However, if this inference is encoded as a *monotonic* rule of inference, weak bilateral logic narrows to Rumfitt's bilateral logic (Rumfitt, 2000).

---

[5]There are competing proposals for the *mechanism* that realises this preference (Horn, 1989, ch5). I do not take a stance, as any mechanism accounting for the linguistic data would do.

**3.4.6.** THEOREM. *WBL plus the following inference rule is equivalent to bilater-alist logic.*

$$(+\neg\text{I.})\ \frac{-A}{+\neg A}.$$

**Proof:**
Trivial, since $(+\neg\text{I.})$ allows us to conduct *any* derivation from only asserted pre-misses; thus the restrictions placed on Rumfitt's logic are lifted.             □

Strictly speaking, this move is unsound: $(+\neg\text{I.})$ encodes the default reading as non–cancellable. Nonetheless, there is an upshot. This result shows that one can now conceive of Rumfitt's bilateralism as being about the default interpretation of bare *no*. As such, it can reclaim linguistic plausibility *vis-à-vis* the data showing that there are weak rejections. Moreover, this distinction aligns in terminology with Rumfitt's earlier (1997) claim that strong rejections are akin to *internal negation* and weak rejections akin to *external negations*: a weak rejection is an external negation scoped on a commitment operator and a strong rejection an internal one.

## 3.5   Model Theory

In this section I demonstrate that the proof-theoretic principled I have outlined in this chapter can also be regimented in a sufficiently general semantics. In principle, one could just consider the Basic Principles to be axioms for a Kripkean multimodal logic. However, I want to make a more principled suggestion.

**Tree Models.**   A model theory can be obtained by understanding speech acts as moves in an unbounded conversational game (Asher & Paul, 2013). The start of the game records both participants initial commitments, as represented in a Kripke model satisfying the appropriate axioms. Computing a speech act leads to a model *transition* (Venant & Asher, 2015). I abstract away from all particulars of a conversation and model it as a *sequence of commitment models* and treat two speech acts as equivalent if they effect the same transition. I refer to individual models as *states*.

**3.5.1.** DEFINITION (State). Let $W$ be a set of worlds and $V : \text{At} \to \mathcal{P}(W)$ be a valuation on these worlds. A *state* on $W, V$ is a model for two speaker public commitment modal logic, i.e. a structure $t = \langle W, V, R_A^t, R_B^t, \rangle$ where $R_A^t$ and $R_B^t$ are accessibility relations.

Then, the space of all possible conversations, given a starting state, is a tree in which each vertex is a Kripke model for commitment. The temporal operators $\Diamond$ and $\Box$ can be understood as quantifying over future commitment states. To

define defeasible inferences involving temporal reasoning, the semantics of $>$ need to be modified. Typically, the interpretation of $>$ selects a set of *normal* worlds (Asher & Morreau, 1990); I modify this to select a set of normal world-time pairs.

**3.5.2.** DEFINITION (Tree Model). A *tree model* is a tuple $M = \langle W, V, P, S, T, * \rangle$ where:

- $W$ is a set of worlds and $V$ is a valuation on $W$.

- $P$ is a set of states on $W, V$.

- $S \subseteq P^{<\omega}$ is a tree. That is, $\emptyset \in S$ and $S$ is a set of finite sequences on $P$ such that if $x \in S$ and $y$ is an initial segment of $x$, then also $y \in S$.

  If $x \in P^{<\omega}$, $x = \langle t_1, t_2, ..., t_n \rangle$, write $x \oplus t = \langle t_1, t_2, ..., t_n, t \rangle$ for the extension of $x$ with $t$.

- $T$ is a function that maps an $x \in S$ and a $w \in W$ to a maximal branch in $S$ extending $x$. That is, $T(x, w) \subseteq S$ is a set of finite sequences such that:
  - $x \in T(x, w)$ and for all $y \in T(x, w)$, $x$ is an initial segment of $y$.
  - for all $y, y' \in T(x, w)$, either $y$ is an initial segment of $y'$, or $y'$ is an initial segment of $y$, or $y = y'$.
  - $\bigcup T(x, w)$ is infinite or there is no $t \in P$ such that $(\bigcup T(x, w) \oplus t) \in S$.
  - If $x \oplus t \in T(x, w)$, then $T(x \oplus t, w) = T(x, w) \setminus \{x\}$.

- $*$ is a function $* : S \times W \times \mathcal{P}(S \times W) \to \mathcal{P}(S \times W)$ with:
  - for all $x, w \in S \times W$ and $X \subseteq S \times W$, $*(x, w, X) \subseteq X$.
  - for all $x, w \in S \times W$ and $X, Y \subseteq S \times W$:
  $$*(x, w, X \cup Y) \subseteq *(x, w, X) \cup *(x, w, Y).$$

For $x \in S$ write $M^x = \langle W, V, R_A^x, R_B^x \rangle$ for the last state in $x$.

Intuitively, the set $S$ represents the possible conversations along transitions between the commitment models $P$. A finite sequence represents a finite initial segment of a conversation and the last state in the sequence is the current commitment state in that conversation. This means that the utterances or dialogue moves in these conversations are not explicitly recorded, but abstracted to the *effect* they have on a commitment model. In particular, then, all utterances that effect the same model transition are considered equivalent.

**Validity in Tree Models.** The assignment $T$ is used to give truth-conditions to the temporal operators. Intuitively, $T$ assigns *a future timeline* to every world in that state. Note that this time at a world is linear, and the temporal modals are evaluated with respect to the same world at different points in time.

Finally, $*$ is used to give truth-conditions to defeasible conditionals. The function $*$ assigns to a sequence-world pair (i.e. a point in time) and a set of sequence-world pairs (i.e. the extension of a proposition, computed globally on the entire tree) an smaller set of sequence-world pairs (thought to be the points in time where the input proposition holds in a typical ('normal') manner). Note that this means that what is 'normal' is sensitive to time: a proposition can hold normally at some world at some time, but hold abnormally at the same world at a different time.

Then, the following semantic clauses are straightforward.

**3.5.3.** DEFINITION (Semantic Validity). Validity is defined in a model $M = \langle W, V, P, S, T, * \rangle$ relative to some $x \in S$ and $w \in W$.

- $M, x, w \models p$ iff $w \in V(p)$,

- $M, x, w \models \neg\varphi$ iff $M, x, w \not\models \varphi$,

- $M, x, w \models \varphi \wedge \psi$ iff $M, x, w \models \varphi$ and $M, x, w \models \psi$,

- $M, x, w \models C_S\varphi$ iff for all $v \in R_S^x(w)$, $M, x, v \models \varphi$,

- $M, x, w \models \Diamond\varphi$ iff there is a $y \in T(x, w)$ such that $M, y, w \models \varphi$.

- $M, x, w \models \varphi > \psi$ iff $\forall(y, v) \in *(x, w, \{(y, v) \in S \times W \mid M, y, v \models \varphi\})$, $M, y, v \models \psi$.

Write $\llbracket \varphi \rrbracket = \{(y, v) \in S \times W \mid M, y, v \models \varphi\}$ for the global extension of $\varphi$.

Note that this definition of $\Diamond$ satisfies KT4C—and quite a lot more. The temporal logic axioms and the Basic Principles I developed in the preceding discussion were intentionally chosen to be minimal, *basic* assumptions. They are not *complete* for this model theory, as this would require some stronger assumptions. My goal merely is to demonstrate that there is a motivated model theory that is general enough to include the Basic Principles.

Now, in this model theory, it is possible to define *structural* admissibility conditions on tree models such that the Basic Principles are sound on admissible tree models. The formal forms of the admissibility conditions are complex and not particularly elucidating, so they—and their attendant Soundness proof—are given in Appendix A.

The following intuition suffices here. In a defined sense, the Basic Principles express that *not too much* can change in the commitment structures when traversing the branches given by $T$, since speakers are committed to keeping their commitments. But the function $*$ that is responsible for realising the Basic Principles is sensitive to time. Thus, while it is predicted that a *normal* conversation proceeds along the restrictions the Basic Principles encode, the model acknowledges that a dialogue can take a different turn at any time. This is in particular the case when the speakers agree to disagree, in which case principle (c) is defeated.

## 3.6  Conclusion

The theory presented in this chapter is intended to be foundational: it offers a partial account of *commitment* through principles that are partly constitutive of the concept. By conceptualising speech acts as the undertaking of particular commitments, the account then explains from first principles why assertions and rejections make particular projections. It also explains what *uptake* is, namely the sharing of these projections. Then, I can define *general rejection* as the dual of uptake: cancellation. To reject an arbitrary speech act is to deny that its projection will obtain. I have demonstrated how this works for the particular case of rejecting a rejection.

The *Basic Principles* of commitment apply in full generality to cooperative and non-cooperative situations. This allows the model to modulate cooperativity by adding additional constraints on what it means to commit: by strengthening how tightly speakers are bound by their commitments, the model excludes particular non-cooperative moves. Ultimately, this derives the (Cooperativity) axiom of Asher & Lascarides: that cooperative speakers *take up what they understand* (if they can).

In addition, a completeness result shows that the weak bilateral logic from Chapter 2 is in fact the logic of cooperative commitment. That is, when the force-indicators + and − from weak bilateral logic are mapped to their commitment analogues, the axioms of WBL are sound and complete for the independently derived logic of cooperative commitment. This means that WBL inference *preserves cooperative commitment.* This moreover yields an explanation of what *strong rejection* is: the strengthening of a weak rejection by raising a contradictory negation to a contrary negation. This strengthening is also independently motivated.

One immediate avenue to expand the account is to extend it to further speech acts, in particular questions. As mentioned early on, questions project answers and this is also something that should be derivable from first principles. This project, however, faces some challenges. One would first need to include suitable semantics for questions into the commitment structures. Then, it is well possible that one needs to phrase additional principles that constitute commitment with respect to *questions*; i.e. my Basic Principles potentially underspecify what it means to be committed to a question.

Finally, there is an aspect of what it means to commit that is not addressed by the model in this chapter. If someone is committed to a proposition, one is usually also committed to *defending* that proposition. I consider this idea again in Chapter 7.

# Chapter 4

# Weak Assertion

The results of Chapters 2 and 3 give an account of *weak* and *strong* rejection: the former is the minimal foil to assertion and the latter is a default strengthening of this foil according to independent linguistic preferences. To a theorist with a certain desire for symmetry—as logicians are wont to be—this suggests that there might be a *weak assertion* as a foil to strong rejection.

In this chapter, I argue that there are *linguistic* grounds for there indeed being a speech act of *weak assertion* and that alongside the force-indicators *yes* and *no*, the adverb *perhaps* serves to indicate this speech act. Then, just as weak rejection serves to explain the meaning of *not*, weak assertion serves to explain the meaning of *might*. I show how this explanation solves a number of epistemic puzzles that appear to challenge classical logic.

Moreover, while weak bilateral logic did not fully succeed in defending classical logic against arguments from harmony, the *epistemic multilateral logic* that extends weak bilateral logic with weak assertion resolves this complaint: extending weak bilateral logic with weak assertions yields a fully harmonious definition of the negation operator.

Sections 4.1 – 4.4 are based on, and draw from, *Weak assertion* (co-authored with Luca Incurvati). Section 4.5 is based on my unpublished work.

## Chapter Contents

# 4.1   Linguistic Evidence

In this section I argue that *perhaps* should be modelled as a force-indicator similar to how Smiley and Rumfitt model *yes* and *no* (see Chapter 2). A number of arguments demonstrate that *might* is not a force-indicator (von Fintel & Gillies, 2007; Swanson, 2010; MacFarlane, 2014). The primary piece of evidence is that *might* felicitously embeds in places where a force-indicator should not (see Rumfitt 2000), e.g. antecedents of conditionals, quantification, or embedded clauses.[1] This is correct, but I argue that it would be wrong to conclude that therefore *no* epistemic particle may be analysed as indicating force. In particular, I argue that *perhaps* is indeed a force-indicator.

***perhaps* and *might*.**   At first glance, *perhaps* and *might* are rather similar, aside from some syntactic differences.

(28)   a. Perhaps it is raining.

      b. It might be raining.

(28a) and (28b) can be justifiably uttered in precisely the same circumstances. I take this to mean that they have the same inferential content (i.e. in a fixed context, the same inferences can be drawn from either utterance). However, the fact that (28a) and (28b) are equivalent should not be taken to mean that *perhaps* and *might* make the same *compositional* contribution to an utterance (cf. Rabern 2012). In particular, this is not sufficient to claim that *might* and *perhaps* have identical embedding behaviour. For instance, while (29a) is infelicitous, (29b) appears to be fine.[2]

(29)   *a. Suppose that perhaps it is raining.

      b. Suppose that it might be raining.

These judgements are not uncontroversial. For instance, Joshua Crabill (2013) claims that (29b) is infelicitous. However, he identifies *might* and *perhaps* from the outset and does not stop to consider ways in which they might differ. Thus, his claim that (29b) is infelicitous might be an oversight.

To see the felicity of (29b), consider the following two natural cases for embeddings of *might* under *suppose*.

---

[1]Similar arguments from embedding behaviour are made about *discourse particles* like German *ja*, *doch* and *wohl* (Zimmermann, 2004). Indeed, force-indicators and discourse particles are similar in that they do not *modify* content, but indicate something *about* the content (this is incompatible with certain embeddings). However, while discourse particles indicate something about how the content relates to the state of the discourse, force-indicators indicate the force with which the content is to be interpreted.

[2]One may object that in the ungrammatical sentences with *perhaps*, I have simply put *perhaps* in the wrong position. This is not so; I follow the principle of testing the indicative forms from (28) under different embeddings. This usually works for *might*, but not for *perhaps*. A sceptical reader can verify that, say, 'Suppose that it is perhaps raining' is just as bad as (29a).

(30)    Biologists supposed it might be a gene like the one causing (...)

> (The New York Times, *Geneticists' Target: Immortal Cells*, Dec 22, 1992)

(31)    (...) he supposes it might even have dwindled to around five thousand.

> (The New Yorker, *On the Road with "Little Failure"*, Oct 16, 2014)

This is *prima facie* evidence for (29b) being generally acceptable. Given that these excerpts are taken from newspapers that are well regarded for their attention to style and grammar, it is unlikely that these are instances of 'loose talk'.

There are further differences between *might* and *perhaps*. The next case shows that *perhaps* also does not embed in the antecedent of conditional statements, whereas *might* does.

(32)    *a.  If perhaps it is raining, I'd better take an umbrella.

      b.  If it might be raining, I'd better take an umbrella.

Again (32b) strikes me as felicitous. In particular, it seems to express a perfectly reasonable thought. It appears entirely plausible that one takes an umbrella not only if it *is* raining, but also if it *might* be. Moreover, Eric Swanson (2010) argues against treating *might* as a force-indicator by observing that it embeds under quantification. Again, *perhaps* differs from *might* in that it does not embed.

(33)    a.  Every day it might be raining.

     *b.  Every day perhaps it is raining.

I now argue that these differences are not mere matters of grammar, but rather a reflection of the fact that *perhaps* and *might* have different functions in discourse.

**perhaps as a Force-Indicator.**    The embedding behaviour of *perhaps* is sometimes taken to evince that it belongs (at least syntactically) to the category of *speaker-oriented adverbs* like *frankly*, *fortunately* or *evidently*, which speakers use to comment on their own utterances (Jackendoff, 1972; Mittwoch, 1977; Bellert, 1977; Ernst, 2009). Such adverbs do not embed under phrases that take propositional arguments, such as *it is the case that*.

(34)    *a.  It is the case that frankly it is raining.

     *b.  It is the case that fortunately it is raining.

     *c.  It is the case that evidently it is raining.

     *d.  It is the case that perhaps it is raining.

This is strong evidence for the claim that these adverbs do not modify propositional content.

However, it is not quite right to say that *perhaps* is used to *comment* on the performance of a speech act. If one says *frankly p*, *fortunately p* or *evidently p*, all

the effects of asserting *p* still obtain, but this is not the case for *perhaps p*. For instance, on a commitment account of assertion (Chapter 3), uttering *frankly p* commits one to *p*, and on a knowledge norm account (Williamson, 2000), uttering *frankly p* requires one to know that *p*. But if one says *perhaps p*, one is *not* thereby committing to *p* or required to know that *p*. Thus, *perhaps p* is not the assertion that *p* plus some comment on this act.

Moreover, note that speaker-oriented adverbs cannot co-occur: e.g. *frankly fortunately it is raining* or *fortunately evidently it is raining* sound bad.[3] Apparently, one can only make one such comment per speech act. However, they can co-occur with *perhaps*. Here is a natural example.

(35)    Frankly, perhaps Route 4 isn't what Ms. Milby needs to investigate.

    (The Washington Post, *Commuter Advice From Several Directions*, Dec 6, 2001)

Thus, the embedding behaviour of *perhaps* fits neither with the speaker-oriented adverbs nor with the compositional operators. This might be taken as evidence that either of these classes needs to be broadened, but I want to raise attention to a third option. Whereas *frankly* is used to *comment* on the performance of a speech act (but does not modify the force of that act), *perhaps* is used to *indicate which* speech act is being performed. Thus, in (35), the function of *frankly* is to comment on the performance of the speech act indicated by *perhaps*. The content of that act remains *Route 4 isn't what Ms. Milby needs to investigate.*

That *perhaps* does not modify content, but indicates which speech act is being performed, can also—and arguably more clearly—be observed when *perhaps* is applied to non-propositional contents. First consider a case where *perhaps* appears in imperative mood and contrast it with the same utterance without *perhaps*.

(36)    Perhaps try to get a theatre to support it (...)

    (The Guardian, *Stage musicals in 2013: how to make your own*, Feb 26, 2013)

(37)    Try to get a theatre to support it (...)

(36) and (37) seem to have the same *content* but different *forces*. (37) has the force of a command, whereas (36) is more of a suggestion. Evidently, *perhaps* in (36) indicates force.

Now, consider (38) as an example for the use of *perhaps* in a polar question.

(38)    Is it perhaps [made of] resin?                                   (BNC, KCV, 4908)

Considering the positive answers to *is it perhaps resin?* reveals that it puts *is it resin?* under discussion, i.e. *perhaps* does not modify the question's core proposition.

---

[3]Certain paraphrases, however, are acceptable. E.g. *frankly, it is fortunate that it is raining.* This does not affect the argument, since here *fortunate* is used as an embeddable adjective, not a speaker-oriented adverb.

(39)   a.  Is it perhaps resin?

      b.  Yes, it is.

  ?c.  Yes, perhaps it is.

  #d.  Yes, but perhaps it is something else.

If *perhaps* were to modify the core proposition of the question in (38a), we would expect (39c) to be the appropriate positive answer to (39a). But, evidently, the proper answer is (39b).[4] Moreover, (39d) indicates that the polarity particle *yes* here means *it is resin*, since this is incompatible with *perhaps it is something else*, whereas *perhaps it is resin* would be compatible.[5] Thus, *perhaps* in (39a) affects the force of the question, but not its content: it seems to indicate a biased question instead of a neutral polar question (Bellert, 1977).

Note that *might* can be used in polar questions too. If *might* is a compositional operator, then it is expected to modify the core proposition of the question. This appears to be the case.

(40)   a.  Might it be resin?

  ?b.  Yes, it is.

      c.  Yes, it might be.

Here, it seems that the preferred answer to the polar question in (40a) includes *might*.[6] This indicates that the core proposition of (40a) is *modified* by *might*.[7]

In sum, the evidence for *perhaps* being a force-indicator is as follows: (i) *perhaps* exhibits the embedding behaviour that is expected of a force-indicator; (ii) the role of *perhaps* cannot be reduced merely to that of commenting on one's performance of a speech act; (iii) in imperative mood, *perhaps* evidently indicates force; (iv) in polar questions, *perhaps* evidently does not modify the core proposition. Thus, or so I claim, *perhaps* does not modify content, but indicates force.

Now recall Smiley's suggestion that assertion and rejection can be realised, respectively, by *yes* and *no* answers to polar questions posed to oneself (see Chapter 2). But *perhaps* too can be used as an answer to a question as in (41a) to express something inferentially equivalent to (41b).

(41)  a.  Is it raining? Perhaps.

     b.  Perhaps it is raining.

---

[4]While (39c) does not seem to be downright infelicitous, it appears to be mockery by repetition. The appropriate answer using *perhaps* appears to be *I don't know—perhaps it is.*

[5]I thank Maria Aloni for pointing out examples like (39d).

[6]Again, (40) is not downright infelicitous, but it seems to *overanswer* the question.

[7]It is also possible to use *might* in highly polite requests, e.g. *might you have time tomorrow?*. Here, it seems, answering *yes* is taken to mean that one *does* have time tomorrow (and not that one *might have*). This appears to be a conventionalised form of politeness. In any case, this does not affect the present argument. All I intend to show is that *perhaps* functions as a force-indicator. Some overlap with certain uses of *might* is compatible with this claim.

Following the argument that *perhaps* indicates force, I propose to read (41a) and (41b) as the speaker performing a speech act that expresses the *perhaps*-attitude towards the proposition *it is raining*.[8] Henceforth, I refer to this speech act as *weak assertion* and to what is usually called assertion as *strong assertion.*

There is one immediate objection. It seems that *perhaps* can embed in the *consequents* of conditionals, which, at first glance, appears odd for a force-indicator.

(42)  a.  If it is going to rain, then perhaps we should stay in.

b.  If it is going to rain, then should we stay in? Perhaps.

However, such an embedding is perfectly compatible with *perhaps* not being a compositional operator. For instance, *frankly* (which is usually agreed not to be an operator) also embeds like this.

(43)     If it is going to rain, then frankly we should stay in.

Nobody would take (43) to evince that *frankly* is not a speech act adverb. Instead, conditionals of this kind are to be analysed as *conditional performances* of a particular speech act (Edgington, 1995; Schnieder, 2010). This is compatible with a force-modifying expression in the consequent of (42): *perhaps* modifies the speech act that is being conditionally performed.

Note moreover that *yes* and *no* too can appear—in their force-marking function—in the consequents of conditionals.

(44)  a.  It looks like rain outside. Should we stay in?

b.  If it is going to rain, then should we stay in? Yes.

(45)  a.  It looks like rain outside. Should we still go out?

b.  If it is going to rain, then should we still go out? No.

The conditionals in (44) and (45) appear perfectly reasonable in their respective contexts. Given an understanding of conditional speech acts, *yes* and *no* here indicate the speech act that is being *conditionally* performed.

**might and the Frege–Geach Problem.**   Accepting that *perhaps* is a force-indicator, but that *might* embeds in the antecedent of conditionals, yields a revenge version of the Frege–Geach embedding argument (named for Peter Geach (1965)) who observed that Frege's argument generalises; see Chapter 2). If my claim of the inferential equivalence of *perhaps* and *might* in non-embedded indicative mood clauses (see (28)) is accurate, then (46) should be valid.

(46)  a.  If it might be raining, then the streets might be wet.

b.  Perhaps it is raining.

c.  Perhaps the streets are wet.

_____

[8]I leave open here what the function of *perhaps* is in imperative and interrogative mood.

I have argued that *perhaps* in (46b) expresses an attitude and does not modify the propositional clause. But expressions of attitude do not embed and the *might* in (46) must necessarily *modify* the proposition *it is raining* (Schroeder, 2008). Thus to apply *modus ponens* to (46ab), *perhaps* in (46b) must be understood to be modifying *it is raining* after all—for otherwise (46b) does not coincide with the antecedent of (46a).

However, this argument rests on the assumption that (46) *must* be a direct application of *modus ponens*. In Chapter 2, I gave Smiley's (1996) analogous argument for embeddings with *not*. His crucial observation was that there are genuine inferences involving rejection—and no negation at all—that we can recognise as valid without theorising about rejection and negation. Here is an analogous case of an inference involving *perhaps* but no *might*.

(47)  a.  If it is raining, the streets are wet.

   b.  Perhaps it is raining.

   ---

   c.  Perhaps the streets are wet.

This is *prima facie* valid and we need not theorise about *perhaps* to recognise it as such. Also, (47) is clearly not a *modus ponens* inference. So why should (46)?

Indeed, to validate (46) it suffices that (46b) is *equivalent* to the antecedent of (46a)—there might be more than one step in the proof of (46c). If the weak assertion of *it is raining* is equivalent to the strong assertion of *it might be raining*, then (46) remains valid without impugning on the arguments that *perhaps* is a force-indicator.

In Chapter 2 I explained the *operator* negation through the *force-indicator no*, sidestepping Frege's Argument. Here, I explain the operator *might* through the force-indicator *perhaps*, solving a variant of the same problem. The inference (47) is validated below.

**What is Weak Assertion?** I have presented some evidence for the existence of a specific speech act I call *weak assertion* and which can be realised using *perhaps*. However, it is left to explain what that speech act *is*. In Chapter 2, I motivate the definition of weak rejection through the mechanisms of common ground management: if assertion is a mechanism to update common ground, then, or so I argue, there must be a mechanism to prevent such update. There, I identify this mechanism to be the speech act of *weak rejection*. Then, in Chapter 3, I motivate that weak rejections are by default read as strong assertions.

This, now, leaves a gap. What is the mechanism to express that a weak rejection is *not* strong? Recall the following example from Chapter 2.

(48)   Is it the case that X or Y will win the election?

   – No, X or Y or Z will win.

Here, the speaker does not assent to *X or Y will win*, but also not to *neither X nor*

*Y will win.* According to my analysis in Chapter 2, the particle *no* here expresses that the speaker does not assent to *X or Y will win.* But that the speaker does not assent to *neither X nor Y will win* is not explicit in (48). I take this to be a pragmatic implicature: if the speaker had meant that *Z will win*, she would have said so.[9] Indeed, this implicature can be cancelled.

(49)    Is it the case that X or Y will win the election?

    – No, X or Y or Z will win. In fact, it's going to be Z.

By the commonsense knowledge that only one person will win the election, the speaker in (49) *does* express that *neither X nor Y will win.*

    Now, I argue that *weak assertion* is the mechanism by which one can *explicate* in a non-cancellable manner that one does not assent to *neither X nor Y will win.* Consider (50).

(50)    Is it the case that X or Y will win the election?

    – Perhaps. #In fact, it's going to be Z who will win.

Here, the speaker does not assent to *neither X nor Y will win* and this is non-cancellable. Thus, I propose that a *weak assertion of p* rules out the *strong rejection of p.*

    That is, my analysis of the situation is as follows: Weakly asserting *X or Y will win*, as in (50), excludes assent to the negative of *X or Y will win* (and nothing more). This yields four speech acts with the following essential effects:

- The strong assertion of $p$ proposes to add $p$ to common ground.

- The strong rejection of $p$ proposes to add $\neg p$ to common ground.

- The weak rejection of $p$ excludes that $p$ be added to common ground.

- The weak assertion of $p$ excludes that $\neg p$ be added to common ground.

Jointly, these speech acts allow precise management of what is accepted—and not accepted—into the common ground. Thus, just as the purpose of weak rejection is to exclude strong assertion (expressing assent), weak assertion works to exclude strong rejection (expressing assent to a negative).

---

(I)  A **weak assertion** of $p$ expresses the weakest attitude that is incompatible with a strong rejection of $p$ (i.e. with assent to $\neg p$).

---

This attitude is then easily recognised as equivalent to *dissent from $\neg p$.* This is in line with the common suggestion that *might p* is in some sense incompatible with *not p* (e.g. Yalcin 2007). Indeed, it has also been previously observed that

---

[9]Pragmatically, (48) may then be analysed as an epistemic free choice disjunction (Kamp, 1973; Zimmermann, 2000).

*might* and *perhaps* can be used to reject a negative (Khoo, 2015; Bledin & Rawlins, 2016).

(51)  a.  It is not raining. – Perhaps it is. $\rightsquigarrow$ *rejection*

    b.  It is not raining. – No, it might be raining.

The rejection moves in (51) can be explained as follows: the weak assertion of *it is raining* excludes the strong assertion of *it is not raining*, thus (51a) implies a weak rejection move. Then, (51b) can be explained through an inferential relation between *perhaps* and *might*. At first glance, it might seem that one can also reject the positive *it is raining* like this.

(52)  It is raining. – No, it *might* be raining.

However, the rejection in (52) is a case of a *pragmatic* rejection by implicature; i.e. *might* implicates *not surely* in the same way that *some* implicates *not all* in (53) (Khoo, 2015; Schlöder & Fernández, 2015b).

(53)  She ate all the cookies. – No, she ate *some* cookies.

This is evinced by the fact that the implicature of (52) (without the *no*) can be cancelled as in (54).

(54)  It might be raining—in fact, it *is* raining!

Thus, (52) should not be mistaken for evidence regarding the *semantic* contribution of *might*. Also note that for the examples in (52) and (53) a prosodic focus on *might* and *some* is mandatory, as is a particular intonational tune. I return to such *rejections by implicature* and how they interface with prosody in Chapter 6.

In sum, just as the purpose of weak rejection is to exclude strong assertion (i.e. assent), weak assertion works to exclude strong rejection (i.e. assent to a negative).

## 4.2  Epistemic Multilateral Logic

The analysis in the previous section can be understood as follows. In Chapter 2, I characterised bilateralism as follows: assertion and rejection express *attitudes* towards propositions. Weak bilateral logic models the speech acts *strong assertion* and *weak rejection*, which express *assent* and *dissent* respectively.

I now add to this *weak assertion* and explain this speech act by saying that a weak assertion of *p* excludes the strong assertion of *not p* (and vice versa). One option would be to say that, thanks to weak bilateral logic, we understand *not* already and thus the attitude expressed by a weak assertion is *dissent from a negative*. Just as well, however, one could say that say that we understand weak assertion already and that this understanding contributes to our understanding of

*not.* This latter route appears to be more principled: it is to consider the speech acts as more basic than any embeddable operator.

Hence, I take the three speech acts to be *sui generis* and coordinate them *jointly* in a *multilateral* logic. This logic forms the basis of an inferentialist explanation of the operators *not* and *might*, according to which their meaning is given by conditions on weak assertion, strong assertion and weak rejection.

Note that, while it would be satisfyingly symmetric, I do not include strong rejection as a primitive alongside weak rejection, weak assertion and strong assertion. Strong rejection reduces to an understanding of assent and negation; but as I am about to show, the three other speech acts suffice to explain these notions and thus strong rejection is derivative.

**Bilateral Coordination Principles.** In Chapter 2, I described the *coordination principles* principles coordinating assertion and rejection. Since I only *added* a new speech act to the framework, the speech act of (strong) assertion and (weak) rejection coordinate as in weak bilateral logic. In particular, the notion of *inference* remains the same as before: the preservation of being taken to have a particular attitude towards some content.

$$\text{(Rejection)} \ \frac{+A \qquad -A}{\bot}$$

$$\text{(SR}_1) \ \frac{\begin{matrix}[+A]\\\vdots\\\bot\end{matrix}}{-A} \qquad \text{(SR}_2) \ \frac{\begin{matrix}[-A]\\\vdots\\\bot\end{matrix}}{+A}$$

***not* and *might*.** I will symbolically express the (linguistic) operator *might* as the logical operator $\Diamond$ and represent weakly asserted content with the prefix $\oplus$. Then, the following rules state under what conditions one is taken to assent to a statement having an epistemic modal as its main operator and what consequences can be drawn from an assertion of such a statement.

$$(+\Diamond\text{I.}) \ \frac{\oplus A}{+\Diamond A} \qquad (+\Diamond\text{E.}) \ \frac{+\Diamond A}{\oplus A}$$

As argued in the linguistic analysis, when *might* and *perhaps* take scope over the same non-embedded clause, they can be interchanged without affecting the inferential meaning of the sentence. Thus, $\Diamond$s can be introduced by moving from weak assertion to strong assertion, and they can be eliminated symmetrically.

I add two additional rules for $\Diamond$ to account for the fact that *perhaps it might be raining* is equivalent to *perhaps it is raining.* This is not uncontentious. Seth Yalcin (2007) and Malte Willer (2013) agree that nesting *might* makes no difference, but Sarah Moss (2015) defends that nested epistemic vocabulary is meaningful.

However, all of Moss's examples involve probabilistic adverbs (*likely*, *probably*) which are not immediately relevant here.

$$(\oplus \Diamond \text{I.}) \; \frac{\oplus A}{\oplus \Diamond A} \qquad (\oplus \Diamond \text{E.}) \; \frac{\oplus \Diamond A}{\oplus A}$$

Then, the following rules give the meaning of negation. Earlier, I argued that the purpose of a weak assertion of $A$ is to rule out assent to $\neg A$; hence the attitude expressed by $\oplus$ is equivalent to dissent from a negative. Accordingly, $\oplus A$ and $-\neg A$ interchange. Analogous arguments show that weakly asserting a negative is equivalent to dissent: observe, e.g. that one can reject *it is raining* with *perhaps it is not*. Thus $\oplus \neg A$ and $-A$ also interchange.

$$(-\neg \text{I.}) \; \frac{\oplus A}{-\neg A} \qquad (-\neg \text{E.}) \; \frac{-\neg A}{\oplus A}$$

$$(\oplus \neg \text{I.}) \; \frac{-A}{\oplus \neg A} \qquad (\oplus \neg \text{E.}) \; \frac{\oplus \neg A}{-A}$$

The two quartets of introduction/elimination rules are rather satisfying from an inferentialist perspective. First of all, they are obviously in harmony as the elimination rules are the direct inverse of the introduction rules. Second, they are *pure* and *simple* in the sense of Dummett (1991); i.e. each quartet of introduction/elimination rules mentions precisely one operator and that operator is the principal operator in the propositions it occurs in.

One might object that these rules fail to *completely* specify the meaning of *might* and *not*, as I have neglected to give rules for negation in asserted propositions and for *might* in rejected sentences. It is, however, not necessary to give such rules, since the rules above fix the meaning of $\neg$ and $\Diamond$ already. To see this, note that the following mixed rules are derivable from the ones above and the coordination principles on strong assertion and weak rejection.[10]

$$(-\Diamond \text{I.}/+\neg \text{E.}) \; \frac{+\neg A}{-\Diamond A} \qquad (-\Diamond \text{E.}/+\neg \text{I.}) \; \frac{-\Diamond A}{+\neg A}$$

This means that the introduction/elimination rules above, together with the coordination principles, indeed explain the inferential meaning of $\neg$ and $\Diamond$: they specify how to introduce and eliminate these operators in every speech act considered here.

---

[10]The proofs of $(-\Diamond \text{I.}/+\neg \text{E.})$ and $(-\Diamond \text{E.}/+\neg \text{I.})$ are in the Appendix.

**Conditionals and the Specificity Problem.**   As in weak bilateral logic, the obvious way to phrase the Deduction principle fails in epistemic multilateral logic. Again consider the classical rule (CNI).

$$(\text{CNI})\ \frac{\begin{array}{c}[+A]\\ \vdots\\ \bot\end{array}}{+\neg A}$$

In the presence of epistemic vocabulary, it is particularly obvious why (CNI) must fail. With the epistemic $\Diamond$, one can state at two least two reasons why strongly asserting $A$ would be absurd: because one has assented to $\neg A$ or because one has assented to $\Diamond\neg A$. Admitting (CNI) would conflate this distinction. Again, the problem is the *unspecificity of derivation*: the antecedent of (CNI) could indicate *either* $+\neg A$ or $\oplus\neg A$. However, the antecedent of (CR) is not similarly underspecified.

$$(\text{CR})\ \frac{+((A \to (B \wedge \neg B))}{+\neg A}$$

The antecedent of (CR) expresses a definite, *specific* inferential property of $A$—and the antecedent of (CNI) does not. Thus, as already argued in Chapter 2, (CR) is a strictly weaker principle than (CNI), but admitting an unrestricted Deduction principle would make them equivalent.

   In weak bilateral logic, I could separate (CR) from (CNI) by *restricting* the Deduction principle to derivations that only use strongly asserted premises. Since WBL only includes the standard Boolean connectives, this solves the specificity problem: the only way to express dissent from a classical proposition $p$ by strong assertion is to strongly assert $\neg p$ or something stronger than $\neg p$. However, this solution will not do here, since epistemic language is richer: $+\Diamond\neg A$ is a strong assertion that is sufficient to infer dissent from $A$, but it is *weaker* than $+\neg A$.

   It is, however, possible to restrict the Deduction principle *further* to exclude the new source of the specificity problem, namely $+\Diamond\neg A$. Any derivation of $-A$ from $+\Diamond\neg A$ needs to eliminate the $\Diamond$. Hence, the specificity problem here relates to the elimination of $\Diamond$s. Thus, I exclude such derivations from the Deduction principle. Again, write $\overset{+}{:}$ for a subderivation that only uses premises marked by $+$ and in which there are no undischarged assumptions that are not marked with $+$. Then the appropriate restriction of the Deduction principle can be put as follows:

$$(\text{Ded*})\ \frac{\begin{array}{c}[+A]\\ \overset{+}{:}\\ +B\end{array}}{+(A \to B)}\ \text{if } (+\Diamond\text{E.}) \text{ and } (\oplus\Diamond\text{E.}) \text{ were not used to derive } +B.$$

The obvious way to state *modus ponens* harmonises with (Ded*) according to the *levelling peaks* criterion for harmony (Dummett, 1991).

$$(MP) \frac{+(A \to B) \qquad +A}{+B}$$

A *peak* is an application of an introduction rule followed immediately by the corresponding elimination rule. To *level* a peak is to eliminate the successive introduction/elimination. Thus, a peak for $\to$ is produced by applying (Ded*) to obtain $+(A \to B)$ and immediately applying (MP) to $+A$ to arrive at $+B$. Such peaks can be levelled by applying the proof of $+B$ from $+A$ (from the antecedent of (Ded*)) to the minor premiss $+A$ of (MP). The restrictions on (Ded*) do not interfere with this.

**Multilateral Coordination Principles.** I maintain the *bi*lateral coordination principles (Rejection) and Smilean *reductio*. But I also add another force-indicator $\oplus$. So there must be additional coordination principles to integrate weak assertion into the bilateralist picture. Smilean *reductio* and (Rejection) express the thought that assent and dissent are incompatible attitudes. Now, I add new coordination principles to encode a relationship between assent and the attitude expressed by weak assertion. First consider the inference rule (Assertion).

$$(\text{Assertion}) \frac{+A}{\oplus A}$$

This principle expresses the thought that one can reject *not A* by asserting that $A$. That is, if one assents to $A$ then one rules out assent to $\neg A$. A paraconsistent logician would object here. But I take it as a plain fact that *sometimes* we can reject *by* asserting a negative, so the paraconsistentist demand can only be to exclude (Assertion) in the select cases where we do not intend a rejection by asserting $\neg A$: allegedly, the semantic paradoxes and their ilk. I show in section 4.5 how EML can coherently account for the paradoxes without restricting (Assertion).

Then, note that weakly asserted content factors in inference. As noted above, *perhaps it is raining; so I'd better take an umbrella* expresses perfectly plausible reasoning between weakly asserted and strongly asserted contents. Similarly, one may reason with weakly asserted premises under hypothetical strong assertions as follows.

(55) a. Assume I'm in a situation where (I assent to) *it is raining*.

    b. In that scenario, (I am taken to assent to) *the streets are wet*.

    c. Perhaps *it is raining*.

    d. Thus, perhaps *the streets are wet*.

Following my definition of weak assertion (I), (55c) states that I can not assert that I am *not* in the scenario described in (55a). Since every such scenario is also

one that is described by (55b), I also cannot assert that I'm not in a scenario described by (55b); this is expressed as the weak assertion in (55d).

Clearly, this is a generally valid form of inference. Thus, hypothetical reasoning about strongly asserted content interacts with weakly asserted content. This hints at another coordination principle; I call it *Weak Inference.*

$$\text{(Weak Inference)} \quad \frac{\begin{array}{c} [+A] \\ +\vdots \\ +B \end{array} \quad \oplus A}{\oplus B} \text{ if } (+\Diamond E.) \text{ and } (\oplus\Diamond E.) \text{ were not used to derive } +B.$$

This states precisely that inferences of the form of (55) are valid. Note that the subderivation is restricted analogous to the Deduction principle. This is because only *one* weakly asserted proposition can interact with hypothetically asserted content at the same time. To see this, note that even if $A$ and $C$ are propositions that cannot be mutually asserted, one would be able to say *perhaps A, perhaps C* when one has no reason to accept or reject either one. For instance, the coordination principle (Assertion) makes it so that one cannot assent at the same time to $A$ and *not A*, but one can perfectly well utter *perhaps it is raining, perhaps it is not.*

This affects (Weak Inference) as follows. Let $A$ and $C$ be such propositions that cannot be strongly asserted together, but can be weakly asserted together. Suppose I have weakly asserted both $A$ and $C$ and I have found a derivation $[+A]\cdots+B$. Now, I want to apply (Weak Inference). If in my derivation of $+B$ from $+A$ I have made use of the premiss $\oplus C$, it would be wrong to conclude $\oplus B$. The conclusion that $\oplus B$ is justified in that the premiss $\oplus A$ indicates that the hypothetical scenario where I assent to $A$ is not excluded; but a hypothetical scenario where I assent to $A$ *and C is* ruled out.

The restrictions placed on the subderivation avoid this problem. By restricting the permissible premisses to only asserted ones, contradictory weakly asserted propositions are avoided. They also cannot be smuggled in by eliminating a $\Diamond$.

Again, the paraconsistent logician might register a complaint here. One may argue that one can perfectly well conceive of hypothetical scenarios where one assents to a proposition and its negation, e.g. in Graham Priest's *Sylvan's Box* scenario (Priest, 1997). But while I indeed assume that one cannot do so (per the rule (Assertion)), my argument for the restrictions on (Weak Inference) does not rest on this—the argument does not refer to negation at all, but only requires that there are *some* propositions $A$ and $C$ that one cannot conceive together. I take it to be *prima facie* plausible that there are such $A$ and $C$—that some things rule out some other things.

Together with (Ded*) and the rules for $\Diamond$, (Weak Inference) immediately entails the following rule of *weak modus ponens.*

$$\text{(wMP)} \quad \frac{+(A \to B) \qquad +\Diamond A}{+\Diamond B}$$

**Conjunction and Disjunction.** Aside from applying the appropriate restriction of subderivations to the rule for disjunction elimination $(+\vee\text{E.})$ (see Chapter 2), the bilateral rules for conjunction and disjunction can be maintained without further modification.

$$(+\wedge\text{I.}) \ \frac{+A \qquad +B}{+(A \wedge B)} \qquad (+\wedge\text{E.}_1) \ \frac{+(A \wedge B)}{+A} \qquad (+\wedge\text{E.}_2) \ \frac{+(A \wedge B)}{+B}$$

$$(+\vee\text{I.}_1) \ \frac{+A}{+(A \vee B)} \qquad (+\vee\text{I.}_2) \ \frac{+B}{+(A \vee B)}$$

$$(+\vee\text{E.}) \ \frac{+(A \vee B) \qquad \overset{[+A]}{\overset{+\vdots}{\varphi}} \qquad \overset{[+B]}{\overset{+\vdots}{\varphi}}}{\varphi} \ \text{if } (+\Diamond\text{E.}) \text{ and } (\oplus\Diamond\text{E.}) \text{ were not used to derive } \varphi.$$

**Epistemic Multilateral Logic.** Let *epistemic multilateral logic* be the natural deduction calculus over the three bilateralist rules for conjunction, the four coordination principles, the eight rules for $\Diamond$ and $\neg$, (MP), and (Ded*).

---

(II) **Epistemic multilateral logic (EML)** is a natural deduction calculus over the following rules of inference:

  – (Rejection), $(\text{SR}_1)$ and $(\text{SR}_2)$.
  – (Ded*) and (MP).
  – (Assertion) and (WI).
  – $(+\Diamond\text{E.})$, $(+\Diamond\text{I.})$, $(\oplus\Diamond\text{E.})$, $(\oplus\Diamond\text{I.})$.
  – $(\oplus\neg\text{E.})$, $(\oplus\neg\text{I.})$, $(-\neg\text{E.})$, $(-\neg\text{I.})$.
  – $(+\wedge\text{E.}_1)$, $(+\wedge\text{E.}_2)$, $(+\wedge\text{I.})$, $(+\vee\text{E.})$, $(+\vee\text{I.}_1)$, $(+\vee\text{I.}_2)$.

---

This logic preserves the bilateralist defence of classical negation.

**4.2.1.** PROPOSITION. *For all propositions A and B, the following are tautologies in epistemic multilateral logic:*

  • $+(A \rightarrow \neg\neg A)$.
  • $+(\neg\neg A \rightarrow \neg A)$.
  • $+((A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A))$.

**Proof:**
The proofs of the double-negation tautologies are instructive in that they demonstrate how the introduction/elimination rules together with the coordination principles explain the meaning of negation, so I present them on the next page.

$$\cfrac{[+A]^1 \quad \cfrac{\cfrac{[-\neg\neg A]^2}{\oplus\neg A}\,(-\neg\text{E.})}{-A}\,(\oplus\neg\text{E.})}{\cfrac{\cfrac{\bot}{+\neg\neg A}\,(\text{SR}_1)^2}{+(A \to \neg\neg A)}\,(\text{Ded*})^1}\,(\text{Rejection})$$

$$\cfrac{[+\neg\neg A]^1 \quad \cfrac{\cfrac{[-A]^2}{\oplus\neg A}\,(\oplus\neg\text{I.})}{\neg\neg A}\,(-\neg\text{I.})}{\cfrac{\cfrac{\bot}{+A}\,(\text{SR}_1)^2}{+(\neg\neg A \to A)}\,(\text{Ded*})^1}\,(\text{Rejection})$$

The proof of contraposition is more involved and can be found in Appendix A. □

It is easy to check that (Ded*) also entails standard axioms for the conditional (see Appendix A). This, together with the above result on negation means that EML validates standard axioms for the propositional calculus. Note in particular that this applies to all substitution-instances of classical tautologies; e.g. $+(\neg\neg A \to A)$ is a EML-tautology also for $A$'s involving epistemic $\Diamond$s. Therefore, the *logic of strongly asserted content* is classical.

---

(III)  On strongly asserted premisses EML contains all classical tautologies and maintains classical *modus ponens*. Thus, in fact, **EML validates classical inference**. It sanctions as valid all classically valid arguments, without restriction.

---

Similar to WBL, this sets apart multilateral logic from both paraconsistent and paracomplete logics. While some paraconsistent logics (e.g. Priest 1979) also include all classical tautologies, they do not reproduce classical inference because they reject certain instances of *modus ponens*. Paracomplete logics (Field, 2008), on the other hand, are similar to multilateral logic in their treatment of the Deduction principle. They, however, deny certain instances of classical tautologies.

Thus, EML defends classical inference against the criticism levelled against it by Michael Dummett (Dummett, 1991, 2002): the introduction and elimination rules governing the meaning of the connectives are evidently harmonious, pure, and simple.

## 4.3   Soundness

In this section, I prove that EML is sound for an embedding into the modal logic S5. This is a *technical* result that serves to demonstrate that epistemic multilateral logic is consistent. This result should not be taken to show anything about the *meaning* of the operator $\Diamond$. Its meaning is given by its rules for introduction and elimination which do *not* specify an S5-$\Diamond$.

The Soundness result is in particular useful to show independence results: a formula that is not semantically entailed by S5 cannot be a theorem of EML (such as, e.g., $+(\Diamond p \to p)$).

**4.3.1.** PROPOSITION (Facts about S5).

1. $\models^{S5} \Box \Diamond \varphi \leftrightarrow \Diamond \varphi$.

2. $\models^{S5} \Diamond \Diamond \varphi \leftrightarrow \Diamond \varphi$.

3. $\models^{S5} \Box \varphi \rightarrow \Diamond \varphi$.

4. *S5 is closed under substitution: If $\varphi \models \psi$ and $\sigma : \boldsymbol{At} \rightarrow \boldsymbol{Fml}$ is a substitution of propositional atoms for arbitrary formulae, then $\sigma(\varphi) \models \sigma(\psi)$, where $\sigma(\varphi)$ is the formula obtained by uniformly replacing every propositional atom $p$ occurring in $\varphi$ with $\sigma(p)$ and same for $\sigma(\psi)$.*

**4.3.2.** DEFINITION (Satisfaction). Let $V = (W^V, V^V, R^V, w^V)$ be an S5-model. Then define:

- $V \models +A$ iff $V, w^V \models^{S5} \Box A$.

- $V \models \oplus A$ iff $V, w^V \models^{S5} \Diamond A$.

- $V \models -A$ iff $V, w^V \models^{S5} \Diamond \neg A$.

- $V \not\models \bot$.

**4.3.3.** THEOREM (Soundness). *Epistemic multilateral logic is S5-sound.*

**Proof:**
The proof proceeds by induction on the length $n$ of derivations. Write $\Gamma \vdash^D \varphi$ if $D$ is a derivation of $\varphi$ from the premisses in $\Gamma$ and show that if $\Gamma \vdash^D \varphi$ for some $D$ with $|D| = n + 1$, then $\Gamma \models \varphi$. The base case is trivial and most induction steps are straightforward. I present here only the interesting case of (Ded*) to demonstrate how the restriction strategy avoids trivialising the $\Diamond$ operator. The remaining steps are in Appendix A.

So, assume that the last step in the derivation $D$ is an application of (Ded*) to show $+(A \rightarrow B)$. Then there is a finite subset $\Gamma' \subseteq \Gamma$ such that all formulae in $\Gamma'$ are prefixed with a $+$, and a derivation $D'$ that does not use $\Diamond$-elimination rules such that $\Gamma', +A \vdash^{D'} +B$. By induction, $\Gamma', +A \models +B$. That is, for any model $V$ of $\Gamma \cup \{+A\}$, $V, w^V \models^{S5} \Box B$.

The proof proceeds in two steps. Call a formula *propositional* if it does not contain any $\Diamond$s.

i. Show that (Ded*) is sound if $A$ and $B$ are propositional and $\Gamma'$ can be split in $\Gamma' = \Delta \cup \Theta$ such that: for all $+C \in \Delta$, $C$ is propositional; and for all $+D \in \Theta$, $D = \Diamond X \rightarrow X$ for some propositional $X$.

ii. Show that any application of (Ded*) can be reduced to (i.).

So, first assume that $A$ and $B$ are propositional (i.e. do not contain $\Diamond$s) and $\Gamma' = \Delta \cup \Theta$ as above. It is to show that for any model $V$ of $\Gamma'$, $V, w^V \models^{S5} \Box(A \rightarrow B)$. Suppose there is a countermodel, i.e. a $V$ with $V, w^V \models^{S5} \neg\Box(A \rightarrow B)$. Let

$v \in W^V$ be a witness for this, i.e. $V, v \models^{\text{S5}} A \wedge \neg B$. Note that for all $+C \in \Delta$ it holds that $V, v^V \models C$, since $V, w^V \models^{\text{S5}} \Box C$.

Now consider the model $V'$ such that: $W^{V'} = \{v\}$, $V^{V'}(v) = V^V(v)$, $R^{V'} = \{(v, v)\}$, $w^{V'} = v$. Note that all $C$ with $+C \in \Delta$ are assumed to be propositional. That is, the fact that $V, v^V \models C$ is dependent *only* on the valuation $V^V(v)$ and not on any other worlds in $W^V$. Thus it is also the case that $V', w^{V'} \models^{\text{S5}} C$ for all $C$ with $+C \in \Delta$.

Also note that $V', w^{V'} \models \Diamond X \rightarrow X$ for any $X$, since $V'$ has precisely one world, and so $V', w^{V'} \models \Diamond X$ iff $V', w^{V'} \models X$. Thus $V' \models \Theta$. Hence $V' \models \Gamma'$. By construction, $V', w^{V'} \models \Box A$ and $V', w^{V'} \models \neg \Box B$. This contradicts that $\Gamma' \cup \{+A\} \models +B$. Thus there is no such countermodel $V$. This shows (i.).

For (ii.), relax the assumption so that $A$, $B$ and the formulae in $\Gamma'$ may no longer be propositional. Note that the proof $D'$ does not use $(\oplus\Diamond E.)$ and $(+\Diamond E.)$ but may include $(\oplus\Diamond I.)$, $(+\Diamond I.)$.

Then, reduce (ii) to (i) as follows. For each formula $Z$ that occurs anywhere (with any sign) in $\Gamma'$ or $D'$ pick an unused atom $c_Z$ (this is easily possible, since $\Gamma'$ and $D'$ are finite). Then construct a finite set $\Sigma$ as follows. For all formulae $X$, $Y$, $Z$ occurring anywhere in $D'$ add the following to $\Sigma$:

a. $+(c_{\neg X} \rightarrow \neg c_X)$, $+(\neg c_X \rightarrow c_{\neg X})$, $+(c_{\neg\neg X} \rightarrow c_X)$ and $+(c_X \rightarrow c_{\neg\neg X})$.

b. $+(c_{X \wedge Y} \rightarrow (c_X \wedge c_Y))$ and $+((c_X \wedge c_Y) \rightarrow c_{X \wedge Y})$.

c. $+(c_X \rightarrow c_{X \vee Y})$, $+(c_Y \rightarrow c_{X \vee Y})$ and $+((c_{X \vee Y} \wedge c_{X \rightarrow Z} \wedge c_{Y \rightarrow Z}) \rightarrow c_Z)$.

d. $+(c_{X \rightarrow Y} \rightarrow (c_X \rightarrow c_Y))$ and $+((c_X \rightarrow c_Y) \rightarrow c_{X \rightarrow Y})$.

e. $+(c_X \rightarrow c_{\Diamond X})$.

f. $+(\Diamond c_{\Diamond X} \rightarrow c_{\Diamond X})$.

Now let $\Gamma'' = \{+c_Z \mid Z \in \Gamma'\}$. Now check that $D'$ can be transformed into a proof $D''$ such that $\Gamma'' \cup \Sigma \cup \{+c_A\} \vdash^{D''} +c_B$. Construct $D''$ as follows: For every application of a rule in $D'$ with consequent $+X$ $(-X, \oplus X)$ show that there is a derivation of $+c_X$ $(-c_X, \oplus c_X)$ from the substituted antecedent of that rule.

For instance, if $D'$ contains an application of (Assertion) to move from $+X$ to $\oplus X$, show that there is derivation of $\oplus c_X$ from $+c_X$. This is just an application of (Assertion) again. Indeed, applications of structural rules—(Rejection), $(\text{SR}_1)$, $(\text{SR}_2)$ and (WI)—immediately carry over from $D'$ to $D''$. The operational rules need some more work.

Consider (MP). This requires use of use one of the formulae we added to $\Sigma$. Suppose $D'$ uses (MP) to move from $+(X \rightarrow Y)$ and $+X$ to $+Y$. Then for $D''$ we need to construct an inference from $+c_{X \rightarrow Y}$ and $+c_X$ to $+c_Y$. Note that above in (d.) we have added $+(c_{X \rightarrow Y} \rightarrow (c_X \rightarrow c_Y))$ to $\Sigma$. Thus in $D''$ we can reason:

$$\cfrac{+c_X \qquad \cfrac{+c_{X \rightarrow Y} \qquad +(c_{X \rightarrow Y} \rightarrow (c_X \rightarrow c_Y))}{+(c_X \rightarrow c_Y)} \text{(MP)}}{+c_Y} \text{(MP)}$$

Similarly, (d.) allows translation of (Ded*). If in $D'$, (Ded*) was used to move from $+X \vdash +Y$ to $+(X \to Y)$ we have by an induction argument that $+c_X \vdash +c_Y$. Then:

$$
\cfrac{\cfrac{[+c_X]}{\vdots \atop +c_Y}}{\cfrac{+(c_X \to c_Y)}{} \text{ (Ded*)} \qquad +((c_X \to c_Y) \to c_{X \to Y})}{+c_{X \to Y}} \text{ (MP)}
$$

It is then easy to check that (a.), (b.) and (c.) can be used to translate applications of $(+\wedge\text{I.})$, $(+\wedge\text{E.})$, $(+\vee\text{E.})$, $(+\vee\text{I.})$, $(-\neg\text{E.})$, $(-\neg\text{I.})$, $(\oplus\neg\text{I.})$ and $(\oplus\neg\text{E.})$: the formulae in $\Sigma$ in combination with (MP) and (WI) precisely allow the appropriate inferences. To show the method, I demonstrate the case for $(-\neg\text{E.})$:

If $D'$ uses $(-\neg\text{E.})$ to move from $-\neg Z$ to $\oplus Z$, we need to derive $\oplus c_Z$ from $-c_{\neg Z}$ for $D''$. This goes as follows, with premisses added in (a.).

$$
\cfrac{\cfrac{[+c_{\neg\neg Z}]^2 \quad +(c_{\neg\neg Z} \to c_Z)}{+c_Z} \text{ (MP)} \qquad \cfrac{\cfrac{-c_{\neg Z}}{\oplus\neg c_{\neg Z}} \text{ ($\oplus\neg$I.)} \qquad \cfrac{[+\neg c_{\neg Z}]^1 \quad +(\neg c_{\neg Z} \to c_{\neg\neg Z})}{+c_{\neg\neg Z}} \text{ (MP)}}{\oplus c_{\neg\neg Z}} \text{ (WI)}^1}{\oplus c_Z} \text{ (WI)}^2
$$

It is left to treat applications of $(\oplus\Diamond\text{I.})$ and $(+\Diamond\text{I.})$ in $D'$.

Suppose in $D'$ there is an application of $(\oplus\Diamond\text{I.})$ to move from $\oplus Z$ to $\oplus\Diamond Z$. Note that we have added the premiss $+c_Z \to c_{\Diamond Z}$ to $\Sigma$ in (e.). Then, we can replace this application of $(\oplus\Diamond\text{I.})$ by an application of WI as follows:

$$
\cfrac{\cfrac{[+c_Z]^1 \quad +(c_Z \to c_{\Diamond Z})}{+c_{\Diamond Z}} \text{ (MP)} \qquad \oplus c_Z}{\oplus c_{\Diamond Z}} \text{ (WI)}^1
$$

Finally suppose in $D'$ there is an application of $(+\Diamond\text{I.})$ to move from $\oplus Z$ to $+\Diamond Z$. With $+(\Diamond c_{\Diamond Z} \to c_{\Diamond Z})$ (f.), we can replace this application of $(+\Diamond\text{I.})$ as follows.

$$
\cfrac{\cfrac{\cfrac{\oplus c_Z}{\oplus c_{\Diamond Z}} \text{ (as above)}}{+\Diamond c_{\Diamond Z}} \text{ ($+\Diamond$I.)} \qquad +(\Diamond c_{\Diamond Z} \to c_{\Diamond Z})}{+c_{\Diamond Z}} \text{ (MP)}
$$

Thus we have $\Gamma'' \cup \Sigma \cup \{+c_A\} \vdash +c_B$. Note that the elements of $\Gamma'' \cup \Sigma$ are as described in (i): $\Gamma'' \cup \Sigma = \Theta \cup \Delta$ with $\Theta$ being all formulae added in (f.) and $\Delta$ being all formulae added in (a.)–(e.). Thus we obtain $\Gamma'', \Sigma \models +(c_A \to c_B)$ by the argument of (i).[11]

---

[11]Technically, since the argument in (i) used the induction hypothesis of the whole Soundness proof, and since $D''$ is a longer proof than $D'$, this is not quite right. To spell this out properly, one would first prove Soundness for all rules except (Ded*) and then prove Soundness of the full calculus by an induction on the number of times (Ded*) is applied in a proof; since moving from $D'$ to $D''$ does not add new applications of (Ded*) this would then be correct. But this would needlessly complicate exposition here.

Now substitute for every $Z$, $c_Z \mapsto Z$ in $\Gamma''$, $\Sigma$, $+c_A$, $+c_B$. The initial $\Gamma'$ is the result of this substitution on $\Gamma''$. Let $\Sigma'$ be the result on $\Sigma$. Since S5 is closed under substitution (Fact 4), $\Gamma' \cup \Sigma' \models +(A \to B)$.

Note that $\Sigma'$ contains only S5-tautologies. Thus for *any* formula $\xi$, $\Gamma' \cup \Sigma \models^{S5} \xi$ iff $\Gamma' \models^{S5} \xi$. Hence in particular $\Gamma' \models +(A \to B)$. Since $\Gamma \supseteq \Gamma'$ it follows that $\Gamma \models +(A \to B)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark: One might wonder why the above argument does not work for the rules excluded from (Ded*): $(+\Diamond E.)$ and $(\oplus \Diamond E.)$. This is because translating these rules in the proof $D''$ would require us to add $+(c_{\Diamond Z} \to c_Z)$ to $\Sigma$. This, however, does not substitute to an S5-tautology, so the final step in the proof fails.

## 4.4   Epistemic Puzzles

The literature raises a number of puzzles that demonstrate some perplexing requirements for a semantics for *might*. I demonstrate how the EML-treatment of *might* accounts for these phenomena. Most of the prior literature is in search of a truth-functional semantics for *might*. To those authors, I cannot offer anything here. Since EML is an inferentialist theory, I will not say more about the meaning of *might* than is encoded in the inference rules governing its introduction and elimination. In fact, it seems to me that it is the very search for truth-conditions that makes the epistemic puzzles, well – puzzling.

**Yalcin's Puzzle.**   The Classicality result of EML is surprising, since there is a body of evidence that purports to show that the logic of epistemic *might* cannot be classical. The *locus classicus* is Seth Yalcin's 2007 paper *Epistemic Modals* (Yalcin, 2007). He observed that the epistemic modal *might* differs in its embedding behaviour from self-reported ignorance of the contrary. Compare (56) with (57).

(56)  #a.  It is raining and I don't know that it is raining.

      b.  Suppose it is raining and I don't know that it is raining.

      c.  If it is raining and I don't know that it is raining, . . .

(57)  #a.  It is raining and it might not be raining.

     #b.  Suppose it is raining and it might not be raining.

     #c.  If it is raining and it might not be raining, . . .

(56a) is a version of Moore's paradox. The typical explanation for the infelicity of (56a) is a pragmatic one: asserting that *it is raining* pragmatically entails knowledge that *it is raining*, which then contradicts *I don't know that it is raining*. Such pragmatic entailment is suspended in certain embeddings, e.g. under *suppose* or *if*. The typical explanation therefore predicts that (56b) and (56c) should be felicitous, as is indeed the case.

Comparing this with (57) shows that *might* is not reducible to *I don't know that not.* Moreover, it seems that the explanation of why (57a) is bad cannot be the usual pragmatic one of (56a). For the pragmatic explanation would predict that (57b) and (57c) should be felicitous, which they are not. However, if one were pursue a *semantic* explanation that renders sentences of the form *p and it might not p* false, classical logic rules would entail that *if it might be that p, then p*, trivialising *might*. Thus, we are caught in a dilemma: the phenomenon that *p and it might not be p* is bad seems to resist pragmatic explanations, but a semantic explanation would violate classical logic.

EML can provide a semantic explanation that preserves classicality. To a multilateralist, the sentence *p and it might not p* gives rise to *incompatible attitudes.* To have incompatible attitudes is absurd and can also not be supposed. For example, *suppose I assent to p and dissent from p* is absurd. The following derivation shows that *p and it might not p* inferentially reduces to simultaneously having assented and dissented from *p*.

$$\cfrac{\cfrac{+(p \wedge \Diamond \neg p)}{+p}\,(+\wedge\text{E.}) \qquad \cfrac{\cfrac{\cfrac{\cfrac{+(p \wedge \Diamond \neg p)}{+\Diamond \neg p}\,(+\wedge\text{E.})}{\oplus \neg p}\,(+\Diamond\text{E.})}{-p}\,(\oplus\neg\text{E.})}{\phantom{x}}\,(\text{Rejection})}{\bot}$$

Thus, *suppose p and it might not be p* is absurd for the same reasons that *suppose I assent to and dissent from p* is absurd: one cannot have two mutually exclusive attitudes towards the same proposition, not even hypothetically.

However, while $+A$ and $+\Diamond\neg A$ are incompatible, $+(\Diamond A \to A)$ is *not* an EML-tautology. This can be derived from the soundness proof above: there are S5-models where $\Box(\Diamond A \to A)$ fails, so there can be no EML proof of $+(\Diamond A \to A)$. However, $-(A \wedge \Diamond\neg A)$ *is* an EML tautology, as can be demonstrated by applying Smilean *reductio* to the above proof.[12]

Thus the multilateralist can explain Yalcin's Puzzle by way of the unassertibility of the paradoxical sentence, without simultaneously committing to that sentence being contradictory. This allows EML to explain the puzzle while preserving classical logic on its asserted fragment.

It is worth to compare this explanation to Moore's paradox. Note that *assent* does not seem to *entail* knowledge as, e.g., the following is felicitous.

(58)   Suppose that I assent to $p$ without knowing that $p$.

It seems that someone who assents to $p$ without knowing that $p$ is misbehaving. One *should not* assent to $p$ without knowing $p$. This is, roughly, the knowledge

---

[12]Accordingly, $\Diamond\neg(A \wedge \Diamond\neg A)$ is an S5-tautology. This may be difficult to see; it is akin to the Drinker paradox of first order logic. $\Diamond\neg(A \wedge \Diamond\neg A)$ is equivalent to $\Diamond(A \to \Box A)$. This is true because either $A$ is true in all accessible worlds or it fails in at least one. If the latter, any world where $A$ fails witnesses $A \to \Box A$. Otherwise the actual world witnesses $\Box A$, so also $A \to \Box A$.

norm of assertion (Williamson, 2000). This norm, however, is not explained as a conflict of attitudes. Accordingly, it is perfectly fine to *suppose* that one has violated the norm (as in (58)). This separates Moore's paradox (a violation of a pragmatic norm, or similar) from this explanation of Yalcin's Puzzle (a semantic conflict of attitudes).

**Assent and Certainty.**   One immediate consequence of this solution to Yalcin's puzzle is that EML sanctions as valid that $+p \vdash +\Box p$, if we take $\Box$ to abbreviate $\neg\Diamond\neg$. That is, if I assent to $p$ I also assent to *must p* (for an epistemic reading of *must*), i.e. that I am certain about what I assent to. However, it would be a mistake to conclude from this that I assent to the *conditional* $+(p \rightarrow \Box p)$—that I am certain about everything that is the case. The inference $p \vdash \Box p$ is a typical feature of dynamic logics of epistemic modality (Veltman, 1985; Groenendijk *et al.*, 1996; Willer, 2013). Now, EML validates $+p \vdash +\Box p$ (statically) as follows.

$$\dfrac{+p \qquad \dfrac{\dfrac{\dfrac{\dfrac{[-\neg\Diamond\neg p]^1}{\oplus\Diamond\neg p}\ (-\neg\mathrm{E.})}{\oplus\neg p}\ (\oplus\Diamond\mathrm{E.})}{-p}\ (\oplus\neg\mathrm{E.})}{\dfrac{\bot}{+\neg\Diamond\neg p}\ (\mathrm{SR})^1}}{}\ (\mathrm{Rejection})$$

Note that this proof, too, eliminates a $\Diamond$ and so one cannot here apply (Ded*) to conclude $\vdash +(p \rightarrow \Box p)$. Thus, again, the restrictions on the Deduction principle in EML separate a plausible inference from an implausible conditional. This is not just a feature of this particular proof. The Soundness proof demonstrates that EML has models where $+(p \rightarrow \Box p)$ fails, so indeed $+(p \rightarrow \Box p)$ is not a theorem.

   Furthermore, in a monolateral logic one may have a meta-inference of contraposition, i.e. $p \models q$ entails that $\neg q \models \neg p$. In such logics, $p \models \Box p$ would immediately deliver the unacceptable $\neg\Box p \models \neg p$, i.e. $\Diamond\neg p \models \neg p$, trivialising $\Diamond$. However, this meta-inference makes no sense in multilateral logic, since force-indicators do not embed: $+A \vdash +B \Rightarrow \neg{+}B \vdash \neg{+}A$ is ill-formed in the consequent (and $+A \vdash +B \Rightarrow +\neg B \vdash +\neg A$ is in general false).

   A related argument claims that some conditionals of the form *if A, then it must be B* are just *prima facie* true, but their contraposition is not (Kolodny & MacFarlane, 2010; MacFarlane, 2014). Consider, for instance, (59).[13]

(59)    If it is raining, the streets must be wet.

This seems correct, but the contrapositive *if the streets might not be wet, then it is not raining* is incorrect, since one may assent to *the streets might not be wet* whilst being ignorant about the weather. However, I claim that the *if … then* of

---

[13]Frank Veltman (1985) and Seth Yalcin (2012) present similar cases, but use probabilistic terms such as *likely*. It is not obvious what an equivalent case with *might/must* would look like.

(59) should not be analysed as the → of EML. In this case, EML does not validate its contraposition and thus is untroubled by this argument.

Linguistic evidence supports the claim that natural language *if ... then* is ambiguous between different interpretations with different assertibility conditions (Dancygier & Sweetser, 2005). Consider (60).

(60)  a.  If Alicia is not tenured, then she is close to be tenured.

      b.  If Alicia is not tenured, then she must be close to be tenured.

The conditional in (60a) expresses a fact about Alicia's job situation. So, if one has no specific knowledge of her situation, one should not assert (60a). But, in contrast, if one judges her to be competent and mature, then one may still assert (60b), which seems to express some heuristic reasoning based on what one knows about Alicia's prowess. Conditionals like (60b) have therefore been called *inferential conditionals* (Krzyżanowska *et al.*, 2013). The *must* in the consequent is said to indicate the epistemic nature of such conditionals, i.e. that they express inferences about what one believes and why (Dancygier & Sweetser, 2005).

Thus, *if ... then* is ambiguous between different formal operators; at least between expressing a fact as (60a) and an inference as (60b).

Now, my stance is the inferentialist one, so the meaning of a logical constant is given by its rules for introduction and elimination; in the case of → these are (Ded*) and (MP). Thus, to challenge EML by applying contraposition to (59), one needs to show that the conditional in (59) can be introduced by (Ded*). If this is not the case, then whatever the meaning of *if ... then* is in (59), it is not the meaning that is conferred onto →.

Evidently, (59) is justified by the fact that one cannot simultaneously have rain and dry streets. If one, accordingly, accepts *it must not be the case that it is raining and the streets are not wet* as a commonsense premiss, then EML derives *the streets are wet* from *it is raining* by Classicality and thus, analogously to the proof of $+p \vdash +\Box p$ above, also *the streets must be wet*. However, this latter inference crucially involves a step that eliminates a $\Diamond$. Thus, the application of (Ded*) is excluded. But this means that the conditional in *if it is raining, the streets must be wet* cannot be introduced by (Ded*).

Note that, indeed, $+\Box\neg(p \wedge \neg q)$ and $-(p \rightarrow \Box q)$ are jointly EML-consistent, as demonstrated by the Soundness proof. One may respond here that *if it is raining, then the streets must be wet* is a clear and obvious truth that need not be reduced to simpler facts. I would disagree. This conditional involves complex epistemic vocabulary in a non-trivial way and it is not at all obvious what kind of conditional it expresses, or what licenses its assertion (cf. Stojnić, 2017).

Thus, while EML validates the *inference* expressed in (59) (which, arguably, explains why (59) seems to be assertible), it does not validate an analysis of (59) using →. Thus, EML does not sanction the troublesome application of contraposition to (59). I leave open here whether there is an embeddable operator corresponding to the *if ... then* of (59) and, if so, what its semantics are.

**Disagreement about Epistemic Modals.**   Consider the following example, discussed (in slight variants) by Kai von Fintel & Anthony Gillies (2008), Malte Willer (2013) and Peter Hawke & Shane Steinert-Threlkeld (2016), among others.

(61)   *Context: Mark hasn't been able to find his house keys. While searching, he went out and checked Sue's car where he did not find the keys*

    a. Mark: I'm so annoyed. I must have accidentally left my keys on the bus.

    b. Sue:   They might be in my car.

    c. Mark: No, I already checked your car.      (Hawke & Steinert-Threlkeld, 2016)

This is a puzzle. As Hawke & Steinert-Threlkeld put it, 'it is hard to identify a plausible candidate proposition [JJS: that Sue asserts in (61b)] that both the speaker [JJS: Sue] is licensed in asserting and that a reasonable listener [JJS: Mark] could be licensed in disputing' (p2). Indeed, examples like (61) generally cause trouble for truth-functional theories of *might* (von Fintel & Gillies, 2008, 2011). It seems that Sue in (61b) was entirely justified in her assertion, maybe even *truthful* (Kratzer, 1991). And yet, Mark is equally justified in his disagreement.

One's immediate reaction may be to say that *might* expresses not something about the world, but about the speaker's epistemic state; in that case Sue in (61) would be roughly saying that it is compatible with her epistemic state that the keys are in the car. But consider the case where Mark has *not* previously checked the car and, prompted by (61), now believes that *the keys are in the car or on the bus*. As Hawke & Steinert-Threlkeld (2016) correctly point out, Mark has learned something about *the keys*, not about *Sue*.

After gathering additional evidence (more than I can present here), von Fintel & Gillies (2008) conclude that it must somehow be the case that *might* is asymmetrically ambiguous: the speaker can select a meaning that justifies assertion, but a hearer can select a meaning that justifies rejection. But they also suggest a 'loophole': 'one could tell a story that makes the utterance of *might* claims not subject to the usual norms of assertion, perhaps because uttering a *might* claim is not in fact an assertion but a different kind of speech act' (p95). Epistemic multilateral logic *almost* follows this suggestion: while EML takes a *might* claim to be a (strong) assertion, the meaning of this claim is explained through the speech act *weak assertion*. This goes as follows.

Consider (61). Let $p = $ *the keys are in Sue's car*. Mark in (61a) strongly asserts something that entails the strong rejection of $p$ (if the keys are on the bus, they cannot be in the car). Sue objects: Her utterance (61b) is an assertion of *might $p$* which is explained as the weak assertion of $p$. Thus, she expresses that, as far as she is concerned, $p$ should not yet be strongly rejected. She is, indeed, entirely justified in doing so, given her epistemic state. Then, in (61c) Mark's *no* rejects Sue's assertion, i.e. he makes the speech act $-\Diamond p$. By an earlier result, he is then taken to assent to *not $p$*. The second part of the utterance (61c) justifies this strong rejection of $p$ (see Chapter 7 for more on such justifications).

## 4.5   Liars and Deniers

If EML is to serve as a defence of classical logic, there is one clear objection. Since this logic derives the strong assertion of every substitution-instance of any classical tautology, it would also derive that $+(l \vee \neg l)$ for a Liar sentence $l$. Since EML also validates certain instances of the disjunctive syllogism, this is paradoxical. I demonstrate below how this trouble is eliminated—on entirely independent and principled grounds—by placing further restrictions on the Deduction principle. This is in apparent contradiction to a celebrated result by Andrew Bacon (2015), who gives a general recipe for generating revenge paradoxes in classical logics. For a bi/multi-lateralist solution, this revenge paradox is the *Denier*: the sentence that says of itself that it is correctly rejectable (Murzi & Carrara, 2015) or, equivalently, that says of itself that it is not correctly assertible (Bacon, 2015).

Note that, technically, the language of EML is only modal-propositional and phrasing Liars and Deniers requires first order logic and some arithmetic (for coding). For the purposes of the discussion, I assume that there is a suitable extension of EML to first order logic in which all the inference rules on modal-propositional formulae also hold for first order formulae; so, in this section $A$ and $B$ denote formulae in a modal first order language. Based on this assumption, I sketch a solution to the paradoxes.

**The Denier.**   The *Denier sentence* is the sentence $d =$ *this sentence is correctly rejectable* (Murzi & Carrara, 2015). That is, if a predicate $\mathcal{D}$ for (correct) rejectability is added to a sufficiently strong multilateral first order theory, standard fixed point arguments show that there is a formula $d$ such that $+(d \leftrightarrow \mathcal{D}(\text{"}d\text{"}))$ is a theorem (where $A \mapsto \text{"}A\text{"}$ is some coding procedure).

Murzi & Carrara (2015) give $\mathcal{D}$ a semantics that makes the assertion of $\mathcal{D}(\text{"}A\text{"})$ equivalent to the rejection of $A$. To an inferentialist, this corresponds to the following pair of introduction/elimination rules.

$$(+\mathcal{D}\text{I.}) \ \frac{-A}{+\mathcal{D}(\text{"}A\text{"})} \qquad (+\mathcal{D}\text{E.}) \ \frac{+\mathcal{D}(\text{"}A\text{"})}{-A}$$

These are equivalent natural deduction formulations of the sequent rules Murzi & Carrara state. The *Denier paradox* then unfolds as follows.

Murzi & Carrara (2015) argue that if it is fair to burden the truth-conditional logician with a truth predicate and the Liar paradox, then it is also fair to trouble the multilateralist with a rejectability predicate and the Denier paradox. Just as much as truth-conditionalism is concerned with truth and falsity, multilateralism is, arguably, concerned with assertibility and rejectability.

This seems fair, so the above derivation is worth a closer look. Informally, the proof has two steps. (i) $+d$ is absurd, so $-d$ by Smilean *reductio*; and (ii) $-d$ is equivalent to the assertion that $d$ is correctly rejectable—which is the assertion that $d$. Hence contradiction. Clearly, the reasoning in (i) is unobjectionable, but the application of $(+\mathcal{D}\mathrm{I.})$ in (ii) is concerning. As seen time and again during the construction of WBL and EML, moving from a rejection to an assertion is a troublesome matter and any such inference deserves scrutiny.

I have maintained that one can reject out of ignorance, i.e. one may refrain from assenting to a proposition when one does not know whether it is true. In such circumstances, one might want to reject a sentence, but *not* simultaneously assert that this rejection was correct; it would hence be a mistake to apply $(+\mathcal{D}\mathrm{I.})$ in this situation. This observation can be cashed out in the following inference.

$$\frac{\dfrac{\dfrac{+\Diamond\neg\mathcal{D}(\text{``}p\text{''})}{\oplus\neg\mathcal{D}(\text{``}p\text{''})}\;(+\Diamond\mathrm{E.})}{-\mathcal{D}(\text{``}p\text{''})}\;(\oplus\neg\mathrm{E.})\qquad \dfrac{[-p]^1}{+\mathcal{D}(\text{``}p\text{''})}\;(+\mathcal{D}\mathrm{I.})}{\dfrac{\bot}{+p}\;(\mathrm{SR}_2)^1}\;(\text{Rejection})$$

That is, under $(+\mathcal{D}\mathrm{I.})$ the rejection of a sentence is incompatible with the assertion that a sentence *might not be correctly rejectable.* By Smilean *reductio* this means that the assertion *p might not be correctly rejectable* is in fact taken to express assent to *p*. Clearly, this is an unacceptable trivialisation of *might.* Even someone not following me on rejections out of ignorance must recognise this as a defect.

However, this defect might just as well be considered a defect of EML. One may contend that $\mathcal{D}$ is more basic than $\Diamond$ and thus that the above derivation is evidence against the EML treatment of $\Diamond$ and not against $(+\mathcal{D}\mathrm{I.})$. However, the only assumption that the derivation makes about $\Diamond$ is rather minimal: that *it might not be the case that p* is grounds for weakly rejecting *p*.

To be clear, I agree that *correctly rejectable* is a property that sentences may have. But I reject that this property must be governed by $(+\mathcal{D}\mathrm{I.})$ and $(+\mathcal{D}\mathrm{E.})$. There seems to be no reason to assume that an attitude I can take towards a sentence is reducible to assenting to a predication on that sentence—here, that dissent is reducible to assent to correct rejectability. As I argued early in Chapter 2, it does not seem to be the case that rejection is reducible to a defect (or property) *of* the sentence in question; rather, it expresses an attitude that a speaker has *towards* a sentence. Thus, I want to suggest another way—at least equally plausible, or so I claim—of assigning meaning to the predicate $\mathcal{D}$.

**Disquotational Assertibility.** Consider the rule $(\oplus\mathcal{D}I.)$ as a replacement for $(+\mathcal{D}I.)$.

$$(\oplus\mathcal{D}I.) \ \frac{-A}{\oplus\mathcal{D}(\text{``}A\text{''})}$$

This expresses that if one has rejected $A$, one is implicitly taken to have *weakly* asserted $\mathcal{D}(\text{``}A\text{''})$. Now compare: $(+\mathcal{D}I.)$ says that if I have rejected $A$, I am taken to have *asserted* that $A$ is correctly rejectable; $(\oplus\mathcal{D}I.)$ says that if I have rejected $A$, then I *cannot assert* that $A$ is *not* correctly rejectable. The latter is plainly correct and, moreover, expresses something about how my rejections restrict what I can assert—a typical and plausible style of inference in a logic for rejection.

Now, is there any reason to *prefer* $(\oplus\mathcal{D}I.)$ over $(+\mathcal{D}I.)$? I claim that there is. While the notion of rejectability might not be particularly well understood, there is the dual notion of assertibility that *is* (comparatively) clear. So, now, consider a predicate $\mathcal{A}$ expressing *correct assertibility* as defined by the following rules.

$$(+\mathcal{A}I.) \ \frac{+A}{+\mathcal{A}(\text{``}A\text{''})} \qquad (+\mathcal{A}E.) \ \frac{+\mathcal{A}(\text{``}A\text{''})}{+A}$$

That is, if one asserts that $A$, one is also taken to assert that $A$ is correctly assertible; and if one asserts that $A$ is correctly assertible, then one is taken to have asserted that $A$. The latter is justified, since truth is taken to be a necessary condition for assertibility. Thus, asserting that $A$ is correctly assertible entails assent to $A$'s truth which entails assent to $A$. I take these rules to be *prima facie* correct.

Now, it is part and parcel of the bi/multi-lateralist model that what is not correctly assertible is correctly rejectable (Smilean *reductio*). Thus, the predicate $\mathcal{D}$ can be derivatively defined as $\mathcal{D} \equiv \neg\mathcal{A}$. Then, $(+\mathcal{A}I.)$ and $(+\mathcal{A}E.)$ are equivalent to $(\oplus\mathcal{D}I.)$ and $(\oplus\mathcal{D}E.)$:

$$(\oplus\mathcal{D}I.) \ \frac{-A}{\oplus\mathcal{D}(\text{``}A\text{''})} \qquad (\oplus\mathcal{D}E.) \ \frac{\oplus\mathcal{D}(\text{``}A\text{''})}{-A}$$

The proofs are in Appendix A. Thus, there is good reason to prefer $(\oplus\mathcal{D}I.)$ and $(\oplus\mathcal{D}E.)$ over $(+\mathcal{D}I.)$ and $(+\mathcal{D}E.)$ when it comes to giving meaning to the predicate for rejectability: they are equivalent to good rules for assertibility, which is a notion we understand much better. But these rules do not entail the Denier paradox: while one derives $\oplus d$ and $-d$ in analogy to the derivation of absurdity above, this is not a contradiction in EML. The result is simply that one can assert neither the Denier nor its negation (as is desired). However, there is now the Liar paradox to worry about.

**The Liar.**   As defined above, $\mathcal{A}$ is a disquotational predicate and thus gives rise to Liar paradoxes. That is, standard fixed-point arguments show that there is a sentence $l$ with $+(l \to \neg\mathcal{A}(\text{``}l\text{''}))$ and $+(\neg l \to \mathcal{A}(\text{``}l\text{''}))$. Note that $l$ is actually equivalent to the Denier $d$, if $\neg\mathcal{A} = \mathcal{D}$. Then, $l$ entails the classical Liar paradox as follows.

$$
\cfrac{+(l \vee \neg l) \quad \cfrac{\cfrac{[+l]^1 \quad (+\mathcal{A}\text{I.})}{+\mathcal{A}(\text{``}l\text{''})} \quad \cfrac{\cfrac{\cfrac{[+l]^1 \quad +(l \to \neg\mathcal{A}(\text{``}l\text{''}))}{+\neg\mathcal{A}(\text{``}l\text{''})} \text{(MP)}}{\cfrac{\oplus\neg\mathcal{A}(\text{``}l\text{''})}{-\mathcal{A}(\text{``}l\text{''})} (\oplus\neg\text{E.})} \text{(Assert.)}}{\bot}(\bot) \quad \cfrac{\cfrac{[+\neg l]^2 (\text{Assert.})}{\cfrac{\oplus\neg l}{-l}(\oplus\neg\text{E.})} \quad \cfrac{\cfrac{[+\neg l]^2 \quad +(\neg l \to \mathcal{A}(\text{``}l\text{''}))}{+\mathcal{A}(\text{``}l\text{''})}\text{(MP)}}{+l}(+\mathcal{A}\text{E.})}{\bot}(\bot)}{\bot}(+\vee\text{E.})^{12}
$$

It is the final application of $(+\vee\text{E.})$ that appears to be at fault here. The minor premises of $(+\vee\text{E.})$ in the above proof entail that the Liar and its negation should be rejected, which seems unobjectionable. Note that WBL and EML already place restrictions on which subderivations are suitable for disjunction elimination, due to specificity arguments. Thus, restricting $(+\vee\text{E.})$ further appears to be a promising point of attack.

**Disquotation and Specificity.**   The predicate $\mathcal{A}$ introduces unspecificity in broadly the same way as $\lozenge$ does. This can be seen as follows. A simple application of $(\text{Ded*})$ to $(+\mathcal{A}\text{I.})$ entails that $+(p \to \mathcal{A}(\text{``}p\text{''}))$, i.e. assent to *everything that is the case is assertible*. We know this not to be the case, so it is a mistake to assent to it. One may take this to evince that $(+\mathcal{A}\text{I.})$ is bad after all. But given earlier troubles, the Deduction principle is a more likely culprit.

Indeed, the *specificity argument against Deduction* developed in Chapter 2 can again be put to work here. The present restriction placed on the Deduction principle and related principles license the following version of classical *reductio* (proof in Appendix A).

$$
(\text{CNI*}) \ \cfrac{\begin{array}{c}[+A]\\ +\vdots\\ \bot\end{array}}{+\neg A} \text{ if } (+\lozenge\text{E.}) \text{ and } (\oplus\lozenge\text{E.}) \text{ were not used to derive } \bot.
$$

Then, $+\neg\mathcal{A}(\text{``}A\text{''})$ entails $+\neg A$, per the following derivation.

$$
\cfrac{\cfrac{\cfrac{+\neg\mathcal{A}(\text{``}A\text{''})}{\cfrac{\oplus\neg\mathcal{A}(\text{``}A\text{''})}{-\mathcal{A}(\text{``}A\text{''})}(\oplus\neg\text{E.})}(\text{Assertion}) \quad \cfrac{[+A]^1}{+\mathcal{A}(\text{``}A\text{''})}(+\mathcal{A}\text{I.})}{\bot}(\text{Rejection})}{+\neg A}(\text{CNI*})^1
$$

This means that classical *reductio* (even in a restricted form) entails that assent to a negation is a necessary condition for claiming a sentence unassertible. But then the notion of assertibility is unduly trivialised: there are other grounds for unassertibility than falsity. In particular, both $p$ and its negation may be judged unassertible, but only one of $p$ and $\neg p$ can be judged assertible.

Thus, by the usual arguments involving (CNI) and specificity, the information contained in $+\neg \mathcal{A}("p")$ is *unspecific* and thus the relevant instance of (CNI) must be excluded by amending the rules sensitive to specificity: (Ded*), (WI) and $(+\vee E.)$.

$$(\text{Ded**}) \quad \frac{\begin{array}{c} [+A] \\ +\vdots \\ +B \end{array}}{+(A \to B)} \text{ if no } \Diamond \text{ was eliminated, and no } \mathcal{A} \text{ was introduced to derive } +B.$$

$$(\text{WI*}) \quad \frac{\begin{array}{cc} [+A] & \\ +\vdots & \\ +B & \oplus A \end{array}}{\oplus B} \text{ if no } \Diamond \text{ was eliminated, and no } \mathcal{A} \text{ was introduced to derive } +B.$$

$$(+\vee E.*) \quad \frac{\begin{array}{ccc} & [+A] & [+B] \\ & +\vdots & +\vdots \\ +(A \vee B) & \varphi & \varphi \end{array}}{\varphi} \text{ if no } \Diamond \text{ was eliminated and no } \mathcal{A} \text{ was introduced to derive } \varphi.$$

This in particular invalidates the final step in the proof of the Liar paradox, for reasons entirely independent of the Liar paradox itself. Thus, it is provable that $+(l \vee \neg l)$, $-l$ and $--\neg l$ without entailing a contradiction. Thus, as suggested in Chapter 2 already, the Liar and its negation are rejected, but this does not preclude assent to their disjunction.

One may object that I have not solved the Liar at all, since one properly develops the Liar paradox on a disquotational *truth* predicate. However, as argued here and in Chapter 2, the multilateral inference rules are not motivated by preserving truth, but by preserving certain attitudes one is taken to have. Thus, direct talk of truth is somewhat foreign to EML. Indeed, my defence of classical logic is not defending it as the logic that preserves truth, but as the logic that preserves assent. Now note that, trivially, every disquotational predicate is inferentially equivalent to $\mathcal{A}$. Thus, if one were to include a disquotational predicate $T$ for 'truth' into EML, this $T$ would not be a truth predicate at all, since there truths that are not assertible and so a proper truth predicate is *not* inferentially equivalent to an assertibility predicate. If, on the other hand, one includes a truth predicate that is not disquotational, there is no liar paradox associated with that predicate.

# 4.6  Conclusion

This chapter concludes the project begun in Chapter 2: to defend classical logic from an inferentialist perspective. While the *bi*lateral version from Chapter 2 is arguably *ad hoc* and possibly fails to satisfy the harmony criterion, these concerns are resolved in the *multi*lateral setting. By considering the speech acts *strong assertion*, *weak assertion* and *weak rejection*, one can give perfectly symmetrical rules that define the meanings of negation and *might*. In this epistemic multilateral logic, classical logic is the logic of strongly asserted content. Moreover, it gives an interpretation of *might* that reconciles epistemic modality with classical logic, explaining Yalcin's Puzzle and other challenges.

The move from bilateralism to multilateralism is independently motivated through a linguistic analysis of the adverb *perhaps*. I have argued that *perhaps* naturally complements the analysis of *yes* and *no* from Chapters 2 and 3. Including a third speech act and another embeddable operator leads to a stronger version of the *specificity argument against Deduction*: that certain inferences underspecify the grounds for rejecting a sentence, while the Deduction principle only applies to specific inferences. A Soundness proof shows that the restrictions placed on the Deduction principle avoid this problem.

This result is programmatic. It seems natural to further extend the multilateral language with additional operators and/or speech acts. These would then lead to new sources of unspecificity, which require further restrictions of the Deduction principle. I have explored one such extension by adding a predicate for *assertibility* to epistemic multilateral logic. In this extension, solving the specificity problem can potentially resolve the Liar and Denier paradoxes. But to fully solve these paradoxes, multilateralism will need to be extended to a full first order language.

## Part Two

# Linguistic Aspects

# Chapter 5

# Taste Disagreement

[I]n many orders of beauty, particularly those of the finer arts, it is requisite to employ much reasoning, in order to feel the proper sentiment; and a false relish may frequently be corrected by argument and reflection.
—David Hume, *An Enquiry Concerning the Principles of Morals*

Fast food is extremely good. I'm tired of people criticizing it.
—Someone on the Internet

A thesis on disagreement in dialogue can hardly ignore the problem of *taste disagreement*. This phenomenon is the subject of a lively debate on supposedly subjective properties, their linguistics and their metaphysics. Allegedly, taste judgements give rise to so-called *faultless* disagreements: situations in which two disagreeing speakers are essentially right about what they are saying. This is taken to support an account of truth that is relativised to judges.

Against this, I defend the view that the *linguistic data* do not rule out *realism about taste*, i.e. that judgements like *this is tasty* are true or false *simpliciter*. I regiment this position in a semantics of *reframable predicates* that squares with the linguistic data on taste judgements. The conclusion to draw is that one can argue about taste *as if* it were an objective matter and that linguistic arguments about the metaphysics of taste need to be rethought.

This chapter is an expansion of *How to be a realist about taste.*

## Chapter Contents

## 5.1   Faultless Disagreement

A persuasive argument for relativism about truth, advanced by Max Kölbel (2004; 2008), rests on the notion of *faultless disagreement.* A faultless disagreement, roughly, is a situation in which one person judges that *p*, another judges that ¬*p*, and neither has made a mistake in doing so. It is easy to see how such situations support arguments that the truth of *p* ought to be relative to a judge. Examples of faultless disagreements are usually judgements of taste or other supposedly subjective qualities. A paradigmatic example is (62).

(62)   a. Sean:    This chili is not tasty.

b. Lauren: You're wrong. This chili is tasty.

**The Null Hypothesis.**   The naïve analysis says that *tasty* is just like any other predicate, i.e. there is nothing fundamentally special about predicates of personal taste *vis-à-vis* other predicates. That is, if one would analyse *this chili is red* as red(⟦*this*⟧), one would analyse *this chili is tasty* as tasty(⟦*this*⟧) (without relativising truth). This is realism about taste, as it entails that taste judgements have *simpliciter* truth conditions. I take this to be the *null hypothesis* of a semantic theory of predicates of personal taste.

The null hypothesis is not usually countenanced. The debate on disagreements about taste revolves around the tension between *relativist* and *contextualist* semantics for predicates of personal taste. Relativists (e.g. Kölbel (2004), Lasersohn (2009)) claim that *truth is relative to judges* and thus that *the chili is tasty* is false-for-Sean, but true-for-Lauren. Contextualist semantics (e.g. Glanzberg (2007), Moltmann (2010)) maintain that truth is absolute, but that *tasty* is a two-place predicate that takes a judge *parameter*.

Both relativists and contextualists can argue that they are, in some sense, more parsimonious than the other. A relativist is conservative over existing semantic theory, since it is not *logical form* that is revised, but *how* this logical form is interpreted truth-conditionally. A contextualist may reply that a revision of what truth is, or of what truth-conditions are, is *ipso facto* less parsimonious than any additional semantic machinery.[1]

It is my contention that the null hypothesis is more basic than either relativism or contextualism. Thus, it ought to be ruled out prior to a discussion of relativism *vs.* contextualism. My goal in this chapter is to show that realism about taste is not in fact ruled out.

**Outset.**   I analyse the phenomenon of *taste disagreement* as a *linguistic* phenomenon. Kölbel's (2004) original phrasing of the faultless disagreement argument

---

[1]Oddly, relativists argue that contextualists cannot account for (62) (MacFarlane, 2007) and contextualists argue that relativists cannot account for (62) (Moltmann, 2010). There, at least, the realist has an advantage: (62) is a disagreement about *facts* just like any other disagreement.

is about one person *believing* a proposition *p* and another believing that ¬*p*. Then it is intuitive to say that the two are *in disagreement*, even though they have never voiced their beliefs. Disagreement in dialogue can be analysed using various linguistic tests like I am doing here, but disagreement in thought is *prima facie* rather more opaque. Thus I follow a large part of the prior literature in focussing on the *linguistic* data (e.g. Moltmann (2010); Egan (2012); López de Sa (2015); Marques (2016)).

However, this means that an anti-realist about taste can retreat: it is entirely compatible with my analysis that speakers act linguistically *as if* matters of taste are objective, while these matters are not objective metaphysically. This is fair, but it is then up to the anti-realist to provide a non-linguistic reason to reject the null hypothesis. Before turning to the linguistic data, I address some arguments that look like such reasons. My goal is not to give a positive argument for realism, but merely to point out that it is not ruled out.

Then, in the next section, I present some data that elucidates the phenomenon of taste disagreement and discuss how the realist would understand these cases. In particular, I present cases showing that there are non-faultless disagreements about taste. Afterwards, I introduce the notion of *framing* and show how it can aid a realist explanation of some further data. To formalise a realist semantics, I review a dialogue model that helps to conceptualise the lessons learned from the data in Section 5.3. My discourse semantics of *reframable predicates* is developed in Section 5.4. I wrap up in Section 5.6.

**Argument 1.** A commonsense property of taste judgements is that it is idle to argue about them (Wright, 2012). Disputes about *facts* might eventually be resolved by determining one claim to be true. Such a resolution seems impossible for judgements involving *tasty*, hence there are no facts about taste.

**Rebuttal.** One should not confuse verifiability with objectivity. The claim that anything unverifiable cannot be objective truth is already an anti-realist position and thus begs the question. Our intuition that it is idle to argue about taste might in the end stem from the unverifiability of such propositions, not from their non-objectivity. Thus, realism is not ruled out.

In general, a metaphysical realist countenances that there are unknowable truths (Williamson, 2000). Thus, assuming that taste judgements are objective but unverifiable does not burden the realist with additional commitments.

**Argument 2.** We have a strong intuition that neither Sean nor Lauren have made a *mistake* in (62). Kölbel (2004) claims that it is a mistake to make a false judgement. The realist, however, is committed to the claim that either Sean or Lauren said something false—thus that either one is mistaken. Thus, Kölbel rules out realism.

**Rebuttal.** The realist does not need to account for the alleged *fact* that neither Sean nor Lauren made a mistake in (62), but only needs to account for the

*intuition* that neither speaker made a mistake.

As above, the realist can claim that taste judgements have an objective truth value, but are impossible (or extremely difficult) to verify. Thus, if realism is right and either of Sean or Lauren in (62) made a mistake, we would still not be able to *tell*. Our intuitive judgement that *neither made a mistake* might be a *withheld judgement* on who made a mistake. That is, or so the realist may argue, because we *cannot attribute* a mistake to either speaker, our intuition is blind to any actual mistake.

Moreover, there is an ambiguity in the interpretation of *mistake* potentially affecting our intuitions. Supposedly, both Sean and Lauren in (62) have correctly applied their *individual norm* for tastiness to reach their taste judgement. We can say that *neither made a mistake* in that application (Schafer, 2011). Yet, it is not ruled out that there is an *objective* norm of tastiness according to which either Sean or Lauren is mistaken about *objective* tastiness (Baker & Robson, 2015).

**Argument 3.**   It seems highly intuitive to say that there is no tastiness without a taster. Realism is committed to *simpliciter* tastiness and hence implausible.

**Rebuttal.** The realist position is not as eccentric as it might seem. Consider, for example, that qualities like *salty* or *spicy* are matters of taste, but correspond to some degree to impersonal facts like the amount of salt or capsaicin in a dish. Plausibly, *a dish without any salt is not tasty* could be an objective *fact* and even part of an objective norm for tastiness.

To be *without any salt* is an objective matter of chemistry; thus, it is quite plausible that *distilled water is not tasty because it has no salt* is a judgement true *simpliciter*. If so, then there are at least *some* objective taste judgements. As long as this is not ruled out, realism about taste should also not be ruled out.

**Argument 4.**   There is no adequate (formal) semantic analysis of predicates of personal taste from the realist position.

**Rebuttal.** In the remainder of this chapter, I defend the following claim. A realist semantics for taste is not ruled out by the data if we make these assumptions: (a) speakers are correct about their individual taste judgements, (b) they can yet fail to make objectively true taste judgements, and (c) they nonetheless operate under the *assumption* that their taste judgements are true.

## 5.2   Data and Analysis

**Non-faultless Disagreement.**   A necessary condition for faultless disagreement is the *parity* requirement: that speakers disagree while holding each other's judgements in equally good standing (Wright, 2012). The relativist argues that taste disagreements satisfy the parity requirement and that *equally good standing*

entails that both judgements are true. I now argue that there are at least some examples that do not satisfy parity.

It is entirely possible to attribute a *fault* to someone making a taste judgement. For (62), consider this continuation:

(63)  a. Sean:    This chili is not tasty.

  b. Lauren: You're wrong. This chilli is tasty.

  c. Lauren: You are just not accustomed to spicy food.

*Prima facie* Lauren *is* pointing out a mistake in Sean's judgement. Lauren is not, however, disputing that the chili is not tasty *to Sean*. Her objection disappears when Sean marks his judgement explicitly as subjective.

(64)  a. Sean:    This chili is not tasty.

  b. Lauren: You're wrong. This chilli is tasty.

  c. Lauren: You are just not accustomed to spicy food.

  d. Sean:    Well, it is not tasty *to me.* I want something else.

The realist can analyse (64) as follows: Sean's judgement is individually correct, since he correctly applied his *individual* norm for tastiness (including, possibly, that spiciness rules out tastiness), but Lauren argues that he has failed to make a *true* judgement (possibly according to an *absolute* norm where tastiness and spiciness are compatible).

Of course, it is also not ruled out by (64) that both Sean and Lauren are arguing from individual norms and that there is no absolute norm (i.e. realism fails). My point is that the phenomenon of taste disagreement is not decisive either way. It is easy to find further examples like (63).

(65)  a. Sean:    Coffee is not tasty.

  b. Lauren: Yes it is. It's an acquired taste.

(66)  a. Sean:    This wine is tasty.

  b. Lauren: No it's not. All wine tastes the same to you.

(67)  a. Sean:    This game is not fun.

  b. Lauren: Yes it is. You're not doing it right.

Claiming that these are *faultless* seems to put the cart before the horse—the theory before the data. Lauren is attributing a mistake of sorts to Sean in each of these examples. She is, however, not denying that Sean *individually* finds *coffee not tasty*, *this wine tasty* and *this game not fun* (i.e. as judged according to Sean's individual norms). Rather, she is saying that Sean is *incompetent*, i.e. that there are other norms that are privileged above Sean's—or even objectively correct.[2]

---

[2]Similar arguments *cannot* be made about perspectival predicates like *left*. One cannot reply to *this arrow is pointing left* with *no, you are not standing on the right side of it.*

In the latter case, Sean would have said something *false*. Of course, this might just be part of Lauren's conversational *strategy*: the examples might be analysed as Lauren *positioning* her judgement as privileged without it *objectively* being privileged. The relativist can defend that there is no objectivity here. My point is again that the data do not decide either way.

**Excluding These Cases.**    The relativist has a response. Not *every* taste disagreement needs to be faultless for the faultless disagreement argument to be performed. In fact, a *single* case where speakers disagree faultlessly (in the strong sense where one judges $p$ and another $\neg p$) is sufficient. Thus, we might attempt to restrict our attention to disagreements about taste that do not involve objections as in (63)–(67). Then, supposedly, we have a more narrow class of disagreements that are unequivocally faultless and realism is ruled out by this class. I claim that it is not possible to isolate such a class.

Say we restrict our attention to disagreeing taste judgements made by speakers who are unquestionably competent and experienced (e.g. established critics in the relevant area). Then, one might argue, the responses Lauren makes in (63)–(66) are not available. But this would not go far enough. Merely having refined their tastes does not preclude speakers from accusing each other of incompetence. The following discourse (68) could occur between two highly experienced evaluators.

(68)  a. Sean:    This game is no fun.

   b. Lauren: Yes it is. You're not doing it right.

   c. Sean:    I *am* doing it right.

The point of disagreement here is that Lauren and Sean differ on how the game ought to be played in order to judge whether it is fun. Even if Lauren and Sean are highly qualified at judging whether something is fun, this is a reasonable point of disagreement.

To avoid cases like (68), we need to make Sean and Lauren agree on all relevant circumstances and criteria of evaluation. But this means that we need to make Sean and Lauren agree on *a norm of judgement*—since every difference in their individual norms licenses an objection like (68b). Thus, I put the following dilemma to the notion of faultless disagreement:

   – The disagreeing speakers share a norm. Then they are expected to make the same judgements.
   – The disagreeing speakers do not share a norm. Then they can argue that one particular norm is privileged.

The dilemma shows that in any disagreement on taste, the speakers will have different norms. Then, at any point in the dialogue, they can violate parity, i.e. the speakers can make explicit that they do *not* hold each other's judgements in equally good standing. In that case, the disagreement did not satisfy parity to

begin with. Speakers may choose not to do so. But the fact that the option to violate parity *is generally available* means that there is no isolable class of taste disagreements satisfying parity. But as long as there we cannot isolate such a class, we cannot be sure that there is a taste disagreement that is a counterexample to realism.

**Framing.** Kate McFarland (2015) develops the useful distinction of objective *vs.* subjective *frames* of argument. She convincingly argues that one can argue about *any* matter as long as the argument is *framed* objectively (i.e. as a matter of objective fact), but one cannot argue in subjective frames. The arguments in (63) and (65)–(68) above are objectively framed. In (64d), Sean *reframes* the argument to be subjective, and the point of disagreement disappears.

Indeed, it is a common observation that one can only disagree with objectively framed taste judgements, but not with subjectively framed ones (see e.g. MacFarlane (2007)).

(69)    a. Sean:    This chili is tasty.

      b. Lauren: No, it's not.

(70)    a. Sean:    This chili is tasty to me.

    #b. Lauren: No, it's not.

In (69), Sean is presenting *this chili is tasty* objectively, and Lauren can coherently disagree with this. Not so in (70) where Sean phrases his judgement subjectively. It is important to be precise here, as the various responses in (71) show.

(71)    a.    Sean:    This chili is tasty to me.

      b.    Lauren: But it really is not tasty.

   ?b.′    Lauren: I disagree, it's not tasty to me.

  #b.″    Lauren: No, it's not—it's not tasty to me.

  #b.‴    Lauren: You're wrong—it's not tasty to me.

One might consider (71b) to be a disagreement with a subjectively framed taste judgement. However, Lauren is *not* denying that the chili is tasty to Sean, but herself reframes the argument to an objective frame.

Torfinn Huvenes (2012) argues hat Lauren in (71b′) *can* disagree with (71a), but he himself says that it sounds 'stilted' (p173). However, even in (71b′), Lauren is not denying the content of (71a) and doing so in (71b″,b‴) is infelicitous.[3] While (71b′) hence seems to reveal something curious about the semantics and pragmatics of *I disagree*, this is not immediately relevant here. It suffices to observe that Lauren cannot reasonably deny that *this chili is tasty to Sean.*

---

[3]Lauren might also say *It's disgusting—you cannot possibly like this!* This seems to be a figurative use of *cannot* and not a denial of Sean's subjective judgement.

There is one tacit assumption in all this: that when someone judges *this is tasty to me* and they are not somewhat impaired, then they are right about that. That is, it is *actually* tasty to them.

---

(I) If a sufficiently unimpaired person judges *this is tasty to me*, then it is actually tasty to them. I refer to this as the **correctness of individual judgement**.

---

This assumption underlies the faultless disagreement argument: that I cannot be mistaken about my own tastes, i.e. my own belief that *this is tasty to me* is indeed *true*. It may be possible to imagine circumstances where a sufficiently impaired speaker is mistaken about their own beliefs about taste (e.g. intoxication of some sort; marijuana seems to be a prime example). But then these are circumstances where the faultless disagreement argument does not apply, because it rests on the idea that taste judgements are made without mistake. Thus, defenders of the faultless disagreement argument should stipulate to (I), as I do too.

**Framing and Realism.**   Subjective frame is how the realist cashes out the commonsense intuition that leads to the parity requirement: that both speakers can acknowledge that, in the end, one taste judgement is just as valid as the other. That something is tasty *to someone*, but also not tasty *to someone else* is plausible, even to a realist about taste. In particular, speakers can coherently maintain that their judgements differ from someone else's.

(72)   a. Lauren: This chili is tasty to me.

     b. Sean:    Alright, but not to me.

One can also attribute a taste judgement to a third person in what has been called *exocentric* judgements (Lasersohn, 2005).

(73)   Lauren: Victor finds this tasty.

The attribution need not be explicit, but can be implicit in the context; the next example is adapted from an example by Lasersohn (2005).

(74)   [*Sean returns with his child from a visit to an amusement park.*]

     Lauren: How were the rides?

     Sean:    Well, the merry-go-round was fun, but the water slide was a little too scary.

Clearly, Sean is attributing the judgements *fun* and *scary* to *his child* and not himself. But he need not explicate this.

In summary, two speakers can accept that they both were correct about their individual judgement even if these differ, i.e. the chili is tasty to the one, but not to the other. However, or so the realist story goes, they can still be in disagreement about *the objective facts*: whether the chili is tasty not *to* anyone, but *simpliciter*.

**The Objectivity Assumption.** Realism can be reconciled with the framing data as follows. The subjectively framed utterance *this is tasty to me* is an entirely different judgement than the one expressed in *this is tasty*. The former indicates by way of adding *to me* that the relevant evaluative norm is an individual one and the judgement it expresses is correct *ipso facto* (as per I). The latter that other (privileged, objective) norms are relevant.

However, the data also show a link between objective and subjective frame. A speaker who agreed to an objectively framed taste judgement cannot subjectively disagree. In particular, the vaguely Moorean assertion (76) sounds bad.

(75)　　a. Lauren: This chili is tasty. $\rightsquigarrow$ it is tasty to Lauren.

(76)　#a. Sean: This chili is tasty, but I don't find it tasty.

This too makes sense to the realist. You cannot simultaneously judge that *p* and that this judgement is mistaken. Accordingly, you cannot accept that the evaluative norms that led you to a particular belief are objectively false. To the realist, this entails that speakers are bound to the assumption that their individual norms are objectively correct.[4]

---

(II) The **objectivity assumption** is the belief, held by speakers, that their individual taste judgements coincide with *simpliciter* facts about taste.

---

In this case, I can hardly ask of my opponents—relativists and contextualists alike—that they stipulate to (II), since to them 'tasty *simpliciter*' is nonsense. Rather, I put forward (II) as a principle that a realist would derive from the data.

Thus, my response to the faultless disagreement argument is this. A realist can stipulate to the fact that speakers are right about *their own* taste, and moreover accept that if a speaker makes a *simpliciter* judgement is *also* not mistaken regarding their own tastes. But the inference that a speaker is therefore not mistaken about *the objective facts* of *simpliciter* tastiness is not thereby validated. A speaker's objectivity assumption might simply be wrong.

Note moreover that the assertion *this is tasty* is felicitous in circumstances where one finds something tasty and does not have any additional belief about someone else's taste judgement.

(77)　a. Sean: This is tasty.

　　　　$\not\rightsquigarrow$ Sean thinks it is tasty to someone else.

This is also explained by (II), assuming that belief licenses assertion: if a speaker judges *this is tasty to me* they are led to believe—per the Objectivity Assumption–that *this is tasty*. This then licenses the sincere assertion (77a).

---

[4]There are certain specific exceptions to this. Someone whose sense of taste has been affected by illness might concede that something is tasty, but that they are currently not in a position to make that judgement themselves.

**Impersonal Predicates and Reframable Predicates.**   Speakers have a number of linguistic reframing devices at their disposal. For instance, *to me* to force subjective frame, and *really* or *tout court* to force objective frame (Stojanovic, 2007). But framing can also be implicit in the context.

(78)   a. Sean:   Did you like the chili?

      b. Lauren: Yes, it was tasty.

  ??c. Sean:   No, it wasn't.

The pair (78b,c) seems fine on its own, but in the context of (78a), Sean's disagreement in (78c) appears odd. The question (78a) puts the discussion in (78b,c) in a subjective frame. That this is an instance of framing can be seen as follows. When Lauren's utterance is changed to reframe the discussion, Sean's disagreement is licensed again ((79b) is adapted from Stojanovic (2007)).

(79)  a. Sean:   Did you like the chili?

    b. Lauren: It was tasty!

              And it's not just that I liked it—it was tasty *tout court.*

    c. Sean:   No, it wasn't. I didn't like it.

The possibility to frame and reframe an argument distinguishes predicates of personal taste from *impersonal* predicates. One cannot frame impersonal judgements subjectively.

(80)   a. Sean:   This table is wooden.

      b. Lauren: No, it's made of plastic.

  #c. Sean:   Well, it is wooden to me.

    c.′ Sean:   Well, it looks wooden to me.

In (80c), Sean appears—at best—as an incompetent language user. One might think that (80c″) is a subjective reframing, but it seems more principled to say that *looking wooden* is itself a (personal) predicate that can be reframed. Thus, I claim that *tasty* belongs to a class of predicates that I call *reframable*.

---

(III) A **reframable predicate** is a predicate that can be reframed; i.e. it can felicitously co-occur with reframing phrases such as *to me* and *tout court.*

---

Note that the classes of taste predicates and reframable predicates are not co-extensional. Aside from *looking wooden*, predicates that are also reframable (but are not predicates of personal taste) are *unusual* and *invisible.*

(81)  a. I find it unusual to eat insects—I'm European.

    b. This number is invisible to me—I'm colour blind.

    c. Black holes are invisible *tout court.*

# 5.3 Stalnakerian Disagreement

To account for the data on subjective framing, I need a semantic framework that explains why some assertions *cannot be disagreed with.* Stalnaker's (1978; 1988) model of assertion does the job. In Chapter 2, I rehearsed his perspective on disagreement, but there is an added subtlety that I need to elaborate here.[5]

**Stalnaker's Effects of Assertion.** According to Stalnaker's analysis, the *essential effect* of an assertion is to update the conversational *common ground*: the set of propositions the interlocutors assume to be mutually accepted. However, as argued in Chapter 2, asserting does not update the common ground immediately, but does so only in the absence of disagreement. Put differently, asserting that $p$ *proposes* to make $p$ common ground, pending a disagreement move.

However, there is an added subtlety that cannot be ignored. The essential effect described above is the *second* effect that an assertion has on the context. The *first* one is that after someone has asserted that $p$, the common ground now records that this particular speech event has happened.[6] Clearly, this update does not require unanimity. Indeed, Stalnaker continues:

> More exactly, rejection of an assertion blocks the *second* kind of effect that assertions have on the context. The first kind of effect cannot be blocked or withdrawn. (Stalnaker, 1978, p87, fn9)

Now it makes sense to say that an assertion can only be disagreed with if its contents are not yet in the common ground. For if they are, the assertion's essential effect is fulfilled and there is simply nothing to disagree *with.* I now show how that works out for subjectively framed taste judgements.

**Disagreement and Subjective Frame.** The relevant pair, examples (69) and (70), is here repeated.

(69)  a. Sean:   This chili is tasty.

   b. Lauren: No, it's not.

(70)  a. Sean:   This chili is tasty to me.

   #b. Lauren: No, it's not.

Lauren cannot disagree with (70a). The Stalnakerian framework offers an explanation: for subjectively framed assertions like (70a) the two effects of assertion coincide. This can be seen as follows.

---

[5]Egan (2012, 2014), López de Sa (2015) and Khoo & Knobe (2016) offer broadly similar analyses of Stalnaker's account.

[6]Like Stalnaker, I abstract away from the fact that assertions might be misunderstood (Clark, 1996; Ginzburg, 2012); see also the simplifying assumption of Chapter 3.

- The first effect of (70a) is to make it common ground that *Sean asserted 'this chili is tasty to Sean'.*

- We assume that Sean is sincere. Then, the fact that *Sean asserted 'this chili is tasty to Sean'* entails that *Sean believes that this chili is tasty to Sean.*

- We take for granted that Sean is *correct* about his own tastes (I).[7]

- Thus, *Sean believes that this chili is tasty to Sean* entails *this chili is tasty to Sean.*

- Hence *this chili is tasty to Sean* is in the common ground.

This means that the essential effect of Sean's assertion in (70) is already achieved. Lauren's disagreement move is incoherent because there is nothing under discussion that she could disagree with. This argument is not new in itself, but I show in Section 5.4 how a realist can formally conceptualise this.

The framework also explains what Lauren is doing in (71b,b′), repeated here.

(71)   a.  Sean:    This chili is tasty to me.

  b.  Lauren: But it really is not tasty.

 ?b.′ Lauren: I disagree, it's not tasty to me.

According to the model elaborated here, Lauren is simply not disagreeing with Sean. That is, she is not cancelling the essential effect of Sean's utterance (71a). Thus, she is not making a disagreement move in the Stalnakerian sense. Rather, (71b) seems to be a comment on the current topic of the conversation (*to whom is this chili tasty?*). In particular, as long as long as Lauren has no reason to assume that Sean is impaired in his judgement, she has to take it at face value that *the chili is tasty to Sean* (as shown by the derivation above).

This leaves the 'stilted' (71b′) as a linguistic curiosity; apparently one can utter *I disagree* without making a disagreement move (strictly speaking). However, from the Stalnakerian perspective, it seems reasonable to suppose that Lauren takes (71a) to pragmatically invite her to share Sean's judgement and it is this invitation that she rejects.

## 5.4   Realist Semantics for Taste

I now show how to be a realist about taste. First, I axiomatise the key pieces of the Stalnakerian account in a commitment framework (also see Chapter 3). Then, I axiomatise two principles discussed in Section 5.2. This allows me formalise the informal arguments I made there and in Section 5.3. Afterwards I specify the formal semantics of *tasty* as a *reframable predicate.*

---

[7]As said, it might be possible to construct circumstances in which we cannot easily take this for granted. However, talk of faultless disagreement presupposes that the speakers are correct about their individual taste. Thus, present purposes, this assumption is innocent.

**Assertion and Commitment.** I somewhat simplify the model of Chapter 3 by (i) ignoring the temporal dimension; (ii) assuming that speakers are cooperative; and (iii) assuming that speakers always understand each other. The temporal dimension, non-cooperativity and potential misunderstandings are not relevant to the analysis here. Moreover, I add belief modals ($\mathcal{B}_S$) to formalise a sincerity assumption. The minimal axioms required to perform my argument are the following.

**5.4.1.** DEFINITION (Commitment Principles). For every pair of speakers $A \neq B$ and all formulae $\varphi$ and $\psi$:

- $C_A(\varphi \rightarrow \psi) \rightarrow (C_A\varphi \rightarrow C_A\psi)$ (K).
- $C_A\varphi \rightarrow C_A C_A\varphi$ (4).
- $C_A\varphi \rightarrow C_B C_A\varphi$ (Understanding).
- $C_A\varphi \rightarrow \mathcal{B}_A\varphi$ (Sincerity).

The justification for these goes as follows. Axiom (K) makes commitments closed under logical inference, i.e one cannot coherently assert something that contradicts an earlier assertion.[8]

Axiom (4) together with (Understanding) expresses that the *first effect* of assertion obtains. If $A$ undertakes a commitment to $p$, $C_Ap$, then (4) entails $C_A C_A p$ and (Understanding) entails $C_B C_A p$. That is, $A$ and $B$ share a commitment to the proposition that $C_A p$. That means that the fact that $A$ has undertaken this commitment is common ground. Thus, these axioms express the (simplifying) assumption that the first effect of assertion obtains immediately.

Finally, (Sincerity) formalises that I am only concerned with *sincere* assertions, i.e. I assume that speakers believe what they say. More sophisticated theories of dialogue would weaken (Sincerity) to a defeasible inference to account for lying (Asher & Lascarides, 2003). I agree in principle, but insincere assertions are non-cooperative moves and not relevant here.

**The Objectivity Assumption.** As discussed in Section 5.2, speakers who assert a subjectively framed judgement cannot deny the same judgement in objective frame. I argued there that the realist therefore stipulates the following: speakers are bound to the assumption that their individual judgements are objectively correct (II). This is formally expressed in the following *objectivity axiom*.

Here and henceforth $P(x, S)$ means that $x$ is $P$ to $S$, and $P(x)$ means that $x$ is $P$ *simpliciter*.[9] The appropriate semantics for such predicates $P$ is developed below.

---

[8]A consequence of this is that speakers may not be fully aware of their own commitments, e.g. if they fail to draw some inference. This is not unexpected in natural dialogue as speakers generally have limited attention, even regarding their own past utterances (Walker, 1996b).

[9]My object language abstracts away from the application of evaluative norms and puts *x is tasty according to S's norm* simply as the two-place predication $\texttt{tasty}(x, S)$.

**5.4.2.** DEFINITION (Objectivity). For every speaker $S$ and every predicate of personal taste $P$, the following holds.

- $\mathcal{B}_S \forall x(P(x, S) \leftrightarrow P(x))$.

Let me get one potential confusion out of the way. I require that the representational language contain the predicate `tasty` as both one-place and two-place. That is, I consider *tasty* to be a multigrade predicate (Oliver & Smiley, 2004). In first order logic this means that, in fact, there are two predicate symbols $\texttt{tasty}_1$ and $\texttt{tasty}_2$ which have different arity. This is a mere artefact of the first-order formalism and should not be ascribed *linguistic* significance. In particular, I do not consider *tasty* to be lexically ambiguous; I do not stipulate two lexical entries $tasty_{subjective}$ and $tasty_{objective}$. This would only be needed if different meanings of *tasty* were required to be available in the same context. In the semantics I develop later in this section, the meaning of *tasty* in a context is unique.

Contrast this with the paradigmatically ambiguous *bank*. The sentence *Sean is at the bank* can, in the same context, be interpreted as Sean being at a bank-as-in-riverbank or as him being at a bank-as-in-financial-institution. This is not what I am proposing for *tasty*.

It seems that a large part of the prior literature presupposes that an appropriate semantics of predicates of personal taste ought to assign them predicate symbols of fixed arity. I contend that this needlessly reproduces the restrictions of the first order language; *prima facie* these restrictions are artificial.

**Correctness.**   The informal argument in Section 5.3 requires the assumption that speakers are individually correct about their own taste (I). Formally, this can be understood as follows: if someone believes that something is tasty to them, then it is true that it is tasty to them.

**5.4.3.** DEFINITION (Correctness). For every speaker $S$ and every predicate of personal taste $P$, the following holds.

- $\forall x : (\mathcal{B}_S P(x, S)) \rightarrow P(x, S)$.

This axiom secures that my realist semantics respects a crucial intuition on taste judgements. Taste judgements are made *individually correctly*. To wit: if one judges, one acquires a belief; if this belief is true, one has judged correctly. Thus, the axiom (Correctness) states that taste judgements are individually correct.

Under this assumption, the informal result of Section 5.3 follows formally.

**5.4.4.** PROPOSITION (Subjective Frame). *Assume $P$ is a predicate of personal taste and $k$ is a constant symbol denoting a referent appropriate for $P$. Then $C_A P(k, A) \models C_A P(k, A) \wedge C_B P(k, A)$.   That is, if $C_A P(k, A)$ then $P(k, A)$ is common ground.*

**Proof:**

$C_A P(k, A) \models C_B C_A P(k, A) \models C_B \mathcal{B}_A P(k, A) \models C_B P(k, A)$ by Understanding, Sincerity, and Correctness. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Thus, the essential effect of assertion obtains immediately for subjectively framed taste judgements and there is no opportunity for disagreement.

Together, the axioms (Correctness) and (Objectivity) go a long way in accounting for the data. They give due to the explanations from Section 5.2: two opposing taste judgements are both individually correct by (Correctness), but the speakers are yet in dispute about the facts by (Objectivity). They also yield formal explanations of the relevant data. As a case in point, the cases in (75), (76) and (77) are also easily validated formally.

(75)  a. Lauren: This chili tasty.

$\rightsquigarrow$ it is tasty to Lauren.

$\qquad C_S P(k) \models P(k, S)$

(76)  #a. Sean: This is tasty, but I don't find it tasty.

$\qquad C_S(P(k) \wedge \neg P(k, S)) \models \mathcal{B}_S \bot$

(77)  a. Sean: This is tasty.

*is felicitous if Sean has no knowledge of other people's tastes*

$\qquad \mathcal{B}_S P(k, S) \models \mathcal{B}_S P(k)$

**Disputing Taste Judgements.**  The axioms also predict that disagreements about taste arise as expected. The proofs of the following propositions are trivial.

**5.4.5.** PROPOSITION (Disagreement in Objective Frame). *Let S be a speaker and P be a predicate of personal taste. If $\mathcal{B}_S \neg P(k, S)$, then S cannot accept $P(k)$ (i.e. S's beliefs exclude $P(k)$).*

**5.4.6.** PROPOSITION (Parity in Subjective Frame). *Let A and B be speakers and P be a predicate of personal taste. In general, $\mathcal{B}_B P(k)$ does not entail that B cannot accept $\neg P(k, A)$.*

Proposition 5.4.5 shows that a speaker who has made a certain judgement is expected to disagree with the contrary objective judgement. However, nothing prevents them from having made their judgement and simultaneously accepting that someone else made a different one (Proposition 5.4.6).

The objectivity assumption also explains what *mistake* one speaker can now attribute to another (as in the examples of Section 5.2): the objectivity assumption can itself be put under discussion. This reflects the informal realist argument

that in a taste disagreement, one's evaluative norms can be challenged. That is, if Lauren tells Sean that he made a *false* taste judgement, then it is Sean's objectivity assumption that is challenged by Lauren.

**Framing Semantics.**   As said, I contend that it is reasonable to treat the predicate `tasty` as having varying arity. It would be possible to revise the first order language to include predicates of varying arity (Oliver & Smiley, 2004), but such revisionism would lead too far afield. It is possible to keep with the usual first order language by determining the framing of an utterance from a contextual parameter.

As I discussed in Section 5.2, most notably in (78), the framing of a taste judgement, and how it can be argued about, varies with the availability of a contextually salient judge. If there is one, the predicate is framed subjectively for that judge. Thus, I model framing by including an *optional* contextual parameter $j$ for a judge (equivalently, one can assume a parameter for a possibly empty set of judges). That is, a discourse is *framed subjectively* if there is a salient judge and *framed objectively* otherwise.

In Section 5.2, I contrasted predicates of personal taste with impersonal predicates. Impersonal predicates can only be objectively framed, but predicates of personal taste are *reframable* and allow for both frames. Thus, the latter's semantics depends on the frame, i.e. on whether a judge is available. Impersonal predicates, such is their nature, are invariant under framing.

---

(IV) **Framing Semantics**

$$\llbracket tasty \rrbracket^{c,w,j} = \lambda y_e \lambda x_e.\mathtt{tasty}^w(x, y).$$
$$\llbracket tasty \rrbracket^{c,w} \ \ = \lambda x_e.\mathtt{tasty}^w(x).$$

$$\llbracket wooden \rrbracket^{c,w,j} = \lambda x_e.\mathtt{wooden}^w(x).$$
$$\llbracket wooden \rrbracket^{c,w} \ \ = \lambda x_e.\mathtt{wooden}^w(x).$$

---

That is, in subjective frame, `tasty` takes two arguments, an object and a judge. If the subjective framing stems from the context, e.g. from the question under discussion (Roberts, 2012), the judge can be filled in by an elliptical pronoun $pro_j$ that picks up $j$, i.e. $\llbracket pro_j \rrbracket^{c,w,j} = j$ (Stephenson, 2007).[10] No such judge is required for objectively framed judgements.

In Section 5.2, I moreover argued that predicates can be *reframed*. The following semantic postulates show how this works with the phrases *to x* (fixing subjective frame) and *tout court* (fixing objective frame). As Lasersohn (2005) proposes for *to x*, I model reframing phrases as intensional adverbs. That is, when applied to a predicate, they fix the context in which that predicate is evaluated.

---

[10]Stephenson provides independent evidence for similar ellipses.

(V) **Reframing Semantics**

$$[\![\text{P } to \ s]\!]^{c,w,j} = [\![\text{P}]\!]^{c,w,s}(s).$$

$$[\![\text{P } to \ s]\!]^{c,w} \ = [\![\text{P}]\!]^{c,w,s}(s).$$

$$[\![\text{P } tout \ court]\!]^{c,w,j} = [\![\text{P}]\!]^{c,w}.$$

$$[\![\text{P } tout \ court]\!]^{c,w} \ = [\![\text{P}]\!]^{c,w}.$$

Thus, *to* takes an argument $s$ and puts a predicate in subjective frame with judge $s$ and *tout court* puts a predicate in objective frame, regardless of the context, by dropping the judge parameter. Similar semantics apply to other reframing phrases like *x finds…* or *…for x* (subjective) and *really* or *objectively* (objective).

Note that these phrases *fix* the frame and cannot undo each other. This can be appreciated by considering the derivations in (82) and (83). In both cases, it makes no difference to start with or without a contextual judge $j$.

(82)   $[\![\textit{this is tasty to s tout court}]\!]^{c,w,(j)}$

$=[\![\textit{tasty to s tout court}]\!]^{c,w,(j)}([\![\textit{this}]\!]^c) = [\![\textit{tasty to s}]\!]^{c,w}([\![\textit{this}]\!]^c)$

$=[\![\textit{tasty}]\!]^{c,w,s}(s)([\![\textit{this}]\!]^c) = \texttt{tasty}^w([\![\textit{this}]\!]^c, s).$

?a. This is tasty to Sean *tout court*.

(83)   $[\![\textit{this is tasty tout court to s}]\!]^{c,w,(j)}$

$=[\![\textit{tasty tout court to s}]\!]^{c,w,(j)}([\![\textit{this}]\!]^c)$

$=[\![\textit{tasty tout court}]\!]^{c,w,s}(s)([\![\textit{this}]\!]^c) = [\![\textit{tasty}]\!]^{c,w}(s)([\![\textit{this}]\!]^c)$

$=(\lambda x_e.\texttt{tasty}^w(x))(s)([\![\textit{this}]\!]^c) = (\texttt{tasty}^w(s))([\![\textit{this}]\!]^c).$

*a. This is tasty *tout court* to Sean.

In the complex expression *tasty to Sean tout court*, *tout court* makes no contribution and sounds odd in (82a). Similarly, *tasty tout court to* is evaluable in principle, but *to* makes no contribution, as *tasty tout court* is read as an impersonal predicate. Thus, the second noun phrase is stranded and the sentence in (83a) is ungrammatical.

Now, consider the following further interpretations (for simplicity I assume that $[\![\textit{this}]\!]$ has been resolved to an appropriate constant symbol $k$).

(84)   a. This table is wooden.          $\texttt{wooden}^w(k)$

?b. This table is wooden *tout court*. $\texttt{wooden}^w(k)$

*c. This table is wooden to me.

The sentence (84c) is again ungrammatical, since it tries to frame something subjectively that is necessarily objective. Formally this means that two arguments

are given for a one-place predicate like it is the case in (83a). The sentence (84b) sounds odd, again because *tout court* makes no semantic contribution.

The prediction that (84b) is odd, but not ungrammatical, is correct. Utterances like (84b) can be felicitous if the speaker wishes to stress the impersonal nature of a judgement (*this dress does not <u>look</u> blue and black, it is blue and black <u>tout court</u>*). Conceivably, there might be similar circumstances where utterances like (82a) are used to stress that someone cannot be mistaken about their own taste (*I don't <u>think</u> that this is tasty to me, it is tasty to me <u>tout court</u>.*). The predictions for reframable predicates are as expected

(85)   a. This chili is tasty to me.     $\texttt{tasty}^w(k, i)$

         b. This chili is tasty *tout court.* $\texttt{tasty}^w(k)$

Reframing is a novelty of this account. Objectively framed taste judgements are, on my account, not true or false given a judge, but true or false *simpliciter*, whereas subjectively framed judgements take a judge parameter. This means that this semantics is realist in objective frame, but contextualist (in the style of Stephenson (2007)) in subjective frame. However, this account does not consider an utterance like *this chili is tasty* to be ambiguous in a given context. Likewise, *tout court* and *to me* are not disambiguation phrases, but are semantically significant *reframing* phrases.

**The Null Hypothesis?**   This semantics maintains the null hypothesis: *this is tasty* is analysed as $\texttt{tasty}(\llbracket this \rrbracket)$ and truth is absolute. One might think that my semantics goes beyond the null hypothesis, since predicates of personal taste *are* special due to being reframable. But this is does not violate the null hypothesis; the semantics merely stipulates that predicates of personal taste belong to a class of predicates that I call *reframable*. But other predicates belong to that class as well. The predicate *tasty* (as well as other predicates) also belongs to other classes, e.g. it is a *gradable adjective* and a *vague predicate*. The null hypothesis that *tasty* is just like any other predicate is not in contradiction to the fact that *tasty* belongs to some additional classes—because any other predicate does too.

## 5.5   Contextualism *vs.* Reframing

The framing semantics given in the previous section is rather similar to some contextualist proposals. Thus it is useful to compare it to contextualism more closely. Here, I point out one advantage that the framing semantics has over a pure contextualist approach. A contextualist semantics for *tasty* goes as follows (Stojanovic, 2007; Moltmann, 2010).

$$\llbracket tasty \rrbracket = \lambda x_e \lambda j_e.\texttt{tasty}(x, j).$$

In a contextualist account, *tasty* is a two-place predicate: it takes a judge and an object as arguments. Sæbø (2009) shows how the judge parameter can be supplied in subjective frame through framing devices like *to me*. The challenge for contextualism is to satisfy the *disagreement requirement*: to predict that disagreements like (62) are actually disagreements. If Sean's utterance *this is not tasty* and Lauren's utterance *this is tasty* both are assigned semantics relative a judge, they are completely different propositions. One can only speak of a disagreement if the two taste judgements are assigned contradictory logical forms.

All proposals I am aware of solve this by remaining with the above semantics and filling in an appropriate *abstract* judge in objective frame. The problem with this approach can be appreciated by comparing the following two proposals.

- Moltmann (2010) inserts a first-person generic judge, i.e. *this chili is tasty* is like *one finds this chili tasty* with the same semantics as a sentence like *one can see a number in this picture.*

- Crespo & Fernández (2011) expand on Moltmann's proposal. They insert a generic judge that in particular instantiates as any of the dialogue participants.

Moltmann (2010) accounts for the disagreement requirement by assigning absolute truth-conditions to the first-person generic *one finds this tasty*. That is, Sean and Lauren in (62) are in disagreement about the truth of *one finds this chili tasty.*

However, the truth of *one finds this chili tasty* does not entail that *anyone specific* finds the chili tasty. This seems to miss the requirement that someone who agrees to *this chili is tasty* also agrees that the chili is tasty *to them* (see (86a)). In Moltmann's model, Sean may accept *one finds this chili tasty* without finding it tasty himself in much the same way as he can say *one can see me from over there* without accepting that he can see himself (this example is due to Moltmann herself). Crespo & Fernández (2011) raise an equivalent criticism against Moltmann: they claim that she cannot account for the infelicity of (76), repeated here in (86b).

(86)    a. Sean: This is tasty.

          ⤳ It is tasty to Sean.

    #b. Sean: This is tasty, but I don't find it tasty.

Crespo & Fernández (2011) fix this complaint. On their account, the generic *one finds this chili tasty* explicitly instantiates for both dialogue participants. This means that on their account the content of (87a) entails the content of (87b).

(87)  a. Lauren: This chili is tasty.

    b. Lauren: This chili is tasty to me and to you.

This means in particular that the truth-conditions of the proposition expressed in (87a) are not absolute (as they are on Moltmann's account), since they depend on

who is participating in the dialogue. But the speakers still *share* a truth-conditional interpretation of *this chili is tasty*. This suffices to satisfy the disagreement requirement.

This fix, however, overstates the claim made in an objectively framed judgement. Note that (87a) and (87b) have different felicity conditions. In (87a) it seems to be sufficient that Lauren finds the chili tasty herself. Not so in (87b) where one assumes that Lauren has some additional information regarding someone else's taste judgement. Thus, the proposal of Crespo & Fernández (2011) overestimates what Lauren is saying in (87a). In particular, the following seem reasonable in some contexts, but their account would predict infelicity.

(88)  a.  Sean: This is tasty, but I can see that you don't like it.

   a.$'$  Sean: This is tasty, but nobody else seems to think so.

There seems to be a third option. Take the middle ground between Moltmann and Crespo & Fernández. That is, make it so that *the speaker* must be an instance of the first person generic *one*, but leave open that *the addressee* is not. But this means that *one finds this tasty* does not have a truth-conditional interpretation shared by the speakers. This misses the disagreement requirement.

While this comparison is only between two closely related proposals, the analysis points towards a general problem. Having a mandatory judge parameter in the semantics of *tasty* means that something needs to fill this parameter in objectively framed judgements. One can include the dialogue participants among the implicitly inserted (abstract) judges, or not. Doing so overstates the semantics of objectively framed judgements. Not doing so either (i) misses the disagreement requirement, or (ii) means that speakers can coherently accept an objectively framed judgement without ratifying it as subjectively correct as well. This problem seems to be one of principle, going beyond the two accounts discussed here.

This particular problem does not occur in framing semantics. In *objective frame*, the semantics of *tasty* as a reframable predicate have it as a one-place predicate. This trivially yields the result that it is the single proposition $\texttt{tasty}(c)$ that is under discussion in a taste disagreement. In subjective frame, however, one speaker commits to $\texttt{tasty}(c, A)$ and the other to $\texttt{tasty}(c, B)$. Since these are different propositions, there is no disagreement.

## 5.6   Conclusion

This chapter addresses a point that has largely been overlooked in the debate on taste disagreement: that realism about taste is not ruled out by the linguistic data. I have discussed the data on taste disagreements and argued that they are compatible with realism. These arguments are formalised in a semantics for predicates of personal taste as *reframable predicates*; i.e. predicates that *can* take a judge parameter, but do not *need* to. This semantics explains the data in the

context of an appropriate discourse model enriched with two basic assumptions: the *Objectivity Assumption* that a subjective judgement entails belief in objective truth, and the *Correctness Assumption* that if someone judges something tasty, it is actually tasty to that someone.

The issues raised by subjective and objective framing, represented in examples (69)–(76) are explained by the semantics for reframing phrases like *tout court* or *to me*. These combine with the Stalnakerian framework which provides a notion of when utterances cannot be disagreed with. The framing semantics also captures implicitly subjective judgements through the contextual judge parameter and the framing semantics.

Thus, the realist can explain the data in a formal semantics. I conclude that the realist null hypothesis is not ruled out. This result is a linguistic one; a metaphysical anti-realist can maintain that speakers argue *as if* taste is objective or under the *pretence* that it is, but that *actually* there are no objective matters of taste. The anti-realist, however, is then in need of a new, non-linguistic argument.

# Chapter 6

# Intonation

Sometimes the line between agreement and disagreement is only drawn by intonation, e.g. in cases involving verbal irony (154), scalar items (139), or ambiguous uses of non-contrary predicates (@135). To account for this, I develop a novel, general theory of intonation from the ground up. My account analyses the semantics and pragmatics of intonation along three dimensions: the *placement of the focal stress*, the *overall prosodic tune*, and the *discourse context*. I first use data involving two distinct tunes to argue for modelling focus and tune jointly. Then, I combine some independently motivated models to give a theory of focus that is sensitive to tunes. I formally regiment this theory in in SDRT (Asher & Lascarides, 2003) and give formal analyses of some instructive examples. I then demonstrate the flexibility of my approach by extending it to ironic intonation.

For the examples marked with @, there is a correctly intonated audio version available. I thank Bob Ladd for recording these. See `http://jjsch.github.io/audio/tunes.pdf` for the full list of audio data.

This chapter is an expansion of *Understanding focus: Tune, placement and coherence* (co-authored with Alex Lascarides) with the results of *Interpreting English pitch contours in context* (co-authored with Alex Lascarides) and *Towards a formal semantics of verbal irony*.

## Chapter Contents

## 6.1   Tune, Placement, Coherence

The intonation of an utterance contributes to its meaning, but in highly complex ways. Some researchers aim for compositional semantics of discretised *pitches* (e.g. via the ToBI annotation scheme; Silverman *et al.*, 1992). But others provide theoretical and empirical evidence for a non-discrete *spectrum* of possible *tunes* with varying meaning, which cannot be treated compositionally (Ladd, 1980; Bolinger, 1982; Calhoun, 2007). In any case, there are strong intuitions regarding the felicity and meaning of some tunes in some contexts: (@89) demonstrates the basic intuition that in the context of a *wh*-question, focal placement in an answer follows the structure of the question; in (@90), the intonation leads to an *as-opposed-to* implicature (Pierrehumbert & Hirschberg, 1990).[1]

(@89)      a.  Harvey: Who likes Michael?

         b.  Jessica: $\underline{\text{Rachel}}_{\text{H}}$ likes Michael.$_{\text{LL\%}}$

     #b.′ Jessica: Rachel likes $\underline{\text{Michael}}_{\text{H}}$.$_{\text{LL\%}}$

(@90)      a. Louis:   Is Harvey going to fire me?

         b. Donna: $\underline{\text{Harvey}}_{\text{L+H}}$ is not going to fire you.$_{\text{LH\%}}$

              $\leadsto$ *but someone else is*

The judgements on these and similar cases appear to be robust enough to be amenable to systematisation and formal treatment, at least in principle.

**Outset.**  Following Mark Steedman (2014) and others, my goal is to associate each intonational form with a *single* logical form, with its distinct implicatures in distinct contexts being derivable via independently motivated principles of pragmatics. My account is distinct from prior work, however, in achieving a *combination* of two things: (i) making the logical form derived from form sensitive to *both* the tune and accent placement; and (ii) providing formally precise logical derivations from these logical forms to pragmatic implicatures, via existing axioms of pragmatic inference that have already been used to model other pragmatic phenomena (e.g. presupposition and anaphora). Existing work that formally derives implicatures ignores the effects of different tunes (e.g. Roberts, 2012), or applies to only one tune (e.g. Reese, 2007). On the other hand, existing work where logical forms are sensitive to both accent placement and tune offer no formal derivation from those logical forms to implicatures (e.g. Steedman, 2014).

    Primarily, this chapter aims to expand semantic research on *focus* by linking it to the semantics of *tunes*. I cannot capture the full *spectrum* of tunes here and

---

[1]In all examples, the underlined word is where the nuclear accent is placed with its pitch type in subscript. This discretised notation obscures intensity: the examples should be intonated with strong and prominent stress; for fall–rise, the most intense part can be on the low or high pitch.

at first only consider two exaggerated and clearly discernible tunes:

- a *falling* tune with a single clearly discernible high pitched nucleus, which I annotate as H LL%; and

- a *fall–rise* tune with a single clearly discernible low–high pitched nucleus, which I annotate as L+H LH%.[2]

Thus, the relevant data consist of a single clause with only one accent where exactly two attributes vary: (i) the placement of that accent; and (ii) whether the overall tune is falling or fall–rise. This means that I ignore several relevant phenomena, including boundary tones within clauses and multiple-focus constructions. A tune that in a discretising system like ToBI would be annotated with pre-nuclear pitches is a *different tune* than one with only one accent and thus is outside the present scope. To be perfectly clear: my tune semantics are *not* compositional semantics for the *pitches* H* or L*+H but holistic semantics for non-discretised *tunes.*

This proposal breaks with much of the extant semantic theory about focus and intonation. Over the course of our joint work, Alex Lascarides and I have come to the conclusion that the basics in this area need to be rethought. The ideas in this chapter offer a first step towards a new approach. But this also means that I cannot here encompass the whole extent of puzzles and problems that have been raised in a substantial prior literature.

This chapter proceeds as follows. The remainder of this section argues for a main tenet of this work—that focus and tune need to be modelled jointly. In Sections 6.2 and 6.3, I informally motivate and describe a semantics for the above tunes that draws on familiar pragmatic concepts, such as discourse coherence and presupposition. This semantics is formalised within within Segmented Discourse Representation Theory (SDRT, Asher & Lascarides 2003), by exploiting its existing model of interaction between discourse coherence, the interpretation of presuppositions (Asher & Lascarides, 1998) and perlocutionary effects (Asher & Lascarides, 2013). I rehearse the basics of SDRT in Section 6.4 and demonstrate how it can be amended with a *Hindsight Logic* (Schlöder & Lascarides, 2015). The formal tune semantics is in Section 6.5 In Section 6.5 I demonstrate the predictive power of the model by analysing some of the examples that motivated it. Afterwards, I show the flexibility of this account by extending it to an additional tune.

**Congruence, Focus and Tune.** I now argue that modelling focus without tunes underspecifies the meaning of the accent placement. That is, I argue for assigning meaning along two dimensions: the *accent placement* and the *meaning of the tune itself* (in the sense of Bob Ladd's intonational lexicon (1980, ch7)).

---

[2]Gunlogson (2003) and Steedman (2014) assign the final rise a distinct semantics from the rest of the tune. Following Calhoun (2007) I prefer a holistic approach, but see my *A formal semantics of the final rise* (Schlöder, 2015) for a proposal on how to execute a separate semantics for the final rise in SDRT.

Example (@89) motivates the principle of *question–answer congruence* (Halliday, 1967; Büring, 2007; Roberts, 2012): i.e. that focus indicates the *wh*-question an assertion answers. This *congruent* question is the one obtained by placing the *wh*-element on the focal constituent. So (@89b) succeeds in answering (@89a) but (@89b′) does not. However, contrary to congruence, there are felicitous answers to *wh*-questions where accent placement does *not* match the *wh*-element of the question: (@91ab) is infelicitous, but (@91ab′) and (@91ab″) are acceptable and indeed natural in a context where Jessica thinks anyone liking Michael is absurd.[3]

(@91)   a.   Harvey: Who likes Michael?

   #b.   Jessica: Nobody likes $\underline{\text{Michael}}_{\text{H·LL\%}}$

   b.′   Jessica: Nobody likes $\underline{\text{Michael}}_{\text{L+H·LH\%}}$

   b.″   Jessica: Nobody $\underline{\text{likes}}_{\text{L+H}}$ Michael.$_{\text{LH\%}}$

Craige Roberts (2012, p34) claims that placing stress on "likes" or on "Michael" is infelicitous in response to (@91). But she only marks focus—not tune—in her annotation. She would be right if (@91b′) and (@91b″) were uttered with a falling tune. Therein lies a methodological problem to which I want to raise attention. Considering focus without tune leads to confusion. This point is acknowledged (Beaver & Clark, 2009; Roberts, 2012), but I want to give it prominence.

One might claim that annotating focus without tune (as, e.g. underlining or capitalising the focal or stressed element) categorically denotes a falling tune. But this does not square with how focus is discussed *across* the literature. To show this, I now demonstrate that Angelika Kratzer's (1989) widely influential claims about (92) are correct only for the fall–rise tune.

(92)   a.   $\underline{\text{Paula}}$ does not live in Paris.

   b.   Paula does not live in $\underline{\text{Paris}}$.                                                 (Kratzer, 1989)

Kratzer claims that (92a) presupposes that *someone who is not Paula lives in Paris* (and this contrasts with the proffered content) whereas (92b) presupposes that *Paula lives somewhere that is not Paris* (which again contrasts the proffered content). But in fact, the felicity and presupposition vary when placing (92a) in different contexts with different tunes:

(@93)   a.   William: Does Paula live in Paris?

   b.   Edith:   $\underline{\text{Paula}}_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$

            ⤳ *someone (else) does live in Paris*

   #b.′   Edith:   $\underline{\text{Paula}}_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$

   b.″   Edith:   Paula does $\underline{\text{not}}_{\text{H}}$ live in Paris.$_{\text{LL\%}}$

            ⤳̸ *someone (else) does live in Paris*

---

[3]Roberts's (2012) QUD account is more sophisticated, allowing the accommodation of congruent questions; I discuss this further below.

(@94)  a. William: Who does not live in Paris?

b. Edith:  Paula_H does not live in Paris._{LL%}

↝̸ *someone (else) does live in Paris*

Focus placement and tune work jointly here: since (@93b″) is felicitous, it is not the falling tune in itself that is bad in (@93b′), but specifically the falling tune *with accent on 'Paula'.* Now, on Kratzer's reading, (92a) has a meaning equivalent to the *it*-cleft in (95a). It does in (@93b). But (@94b) shows that (92a) can be interpreted in a way similar to (95b) too.

(95)  a.  It is not Paula who lives in Paris.

b.  It is Paula who does not live in Paris.

Furthermore, (95a) is an anomalous response to the question (@94a) and an acceptable response to (@93a), but it is the other way around for (95b). This suggests that (95a) broadly corresponds to the fall–rise tune and (95b) to the falling tune (but Section 6.2 addresses how some cases defy this correspondence).

Since (@93a) is a *prima facie* more natural context than (@94a), it seems reasonable to assume that one tends to read (92a) in its null context with the fall–rise tune given in (@93b); this would explain Kratzer's intuitions. I conjecture that in general, the intuitions linked to examples like (92) where stress is annotated but no tune and no discourse context are given are actually the intuitions associated with a tune that is felicitous in the most natural discourse context.

Overall, discussion is clouded by the inherent ambiguity in considering focus without also considering tune. Kratzer's analysis of (92) has led Geurts & van der Sandt (2004) to claim that a negation outscoped by the focal element always leads to a reading where the negation is part of the presupposition, i.e. as in the cleft (95a). As the above analysis shows, this reading does not (always) arise and so their analysis ignores the possible reading (95b). Thus, there is research that is incompatible with the fall–rise tune (e.g. Roberts, 2012) and research that is incompatible with the falling tune (e.g. Geurts & van der Sandt, 2004). It is my goal here to resolve this methodological confusion and to take a step towards a model that considers focus and tune *jointly.*

**Focus as Interest.**  My model follows a substantial body of prior work that links intonation to *cognitive attitudes* (e.g. Pierrehumbert & Hirschberg, 1990; Hobbs, 1990; Steedman, 2014). Notably, Dwight Bolinger (1972, 1985) claims that accent placement indicates what is *interesting* to the speaker. This offers an alternative explanation to question–answer congruence for predicting (@89).

(@89)    a.  Harvey: Who likes Michael?

b.  Jessica: Rachel_H likes Michael._{LL%}

#b.′ Jessica: Rachel likes Michael_{H·LL%}

It is here quite intuitive to assume that 'Rachel' is significant (*interesting*) in the context of (@89a) while 'Michael' is *given* and hence not interesting. Thus, in (@89b) focus follows interest, but in (@89b′) it does not.

Bolinger's notion of interest can potentially explain (@91) too.

(@91)      a.   Harvey: Who likes Michael?

         #b.   Jessica: Nobody likes $\underline{\text{Michael}}_{\text{H}}$.$_{\text{LL\%}}$

          b.′  Jessica: Nobody likes $\underline{\text{Michael}}_{\text{L+H}}$·$_{\text{LH\%}}$

          b.″  Jessica: Nobody $\underline{\text{likes}}_{\text{L+H}}$ Michael.$_{\text{LH\%}}$

Assuming for the time being that the fall–rise tune here implicates that Jessica finds liking Michael absurd, the accent placement then indicates Jessica's interest: in (@91b′) she thinks it absurd that of all the people one could like, one would choose Michael; in (@91ab″), that of all the things one could feel towards Michael, it would be liking. In (@91b) this implicature is missing and 'Michael' is not interesting as in (@89b′).

Explaining focus in terms of *interest* may be compelling, but it is far too vague to systematically predict infelicity judgements or compute implicatures. It is not formally *definable* what it means for a speaker to find something interesting. In fact, part of Bolinger's (1972) point is that interest (and thus focus) is not predictable. However, one can identify conditions that are *necessary* for something to be marked as interesting. Bolinger (1972) mentions one such condition that *is* predictable: what is obvious cannot be interesting. My account formalises this necessary condition for interest and, conversely, takes the non-focal parts of an utterance to be *not marked as interesting* and function as a coherent presupposition (Jackendoff, 1972; Geurts & van der Sandt, 2004). These ideas can then be made predictive via independently motivated theories of presupposition (van der Sandt, 1992; Asher & Lascarides, 1998).

On this account, focus does not cleanly separate the new from the given content. Just because something is not *marked* as interesting does not mean that it *is not* interesting; in particular, it can be new. So, an interest-based account allows non-focal information to be discourse-new (contrary to what Büring (2006) describes as the "standard view" on focus projection). In such cases, the presupposition triggered by the non-focal parts needs to be accommodated. Thus, or so I claim, the semantics for intonation I develop here is sufficiently underspecified to model readings that might be treated via focus projection on other accounts (see Section 6.3 for details).

**Coherence.**    The next examples motivate interest interacting with *coherence*:

(@96)   a.  Harvey: Does Rachel like Michael?

         b.  Jessica: Rachel $\underline{\text{does}}_{\text{H}}$ like Michael.$_{\text{LL\%}}$

         b.′ Jessica: Rachel does $\underline{\text{not}}_{\text{H}}$ like Michael.$_{\text{LL\%}}$

(@97) a. Harvey: Does Rachel like Michael?

    b. Jessica: Rachel$_{L+H}$ does not like Michael.$_{LH\%}$

        ↝ *but someone else does.*

    b.′ Jessica: Rachel does not like$_{L+H}$ Michael.$_{LH\%}$

        ↝ *but she does something else regarding Michael.*

    b.″ Jessica: Rachel does not like Michael$_{L+H\cdot LH\%}$

        ↝ *but she likes someone else.*

(@98) *Context: Jessica and Katrina are both job hunting.*

    a. Harvey: Did you get a job?

    b. Jessica: Katrina$_{L+H}$ got a job.$_{LH\%}$

        ↝ *Jessica did not get one.*

        or *Jessica does not know yet whether she got offered one.*

    #b.′ Jessica: Katrina got a job$_{L+H\cdot LH\%}$

    #b.″ Jessica: Katrina got a job$_{H\cdot LL\%}$

The answers in (@96) are congruent; in Bolinger's terms, Jessica marks as interesting the polar issue raised by the question (@96a). But the answers in (@97) indicate interest in other elements, leading to certain implicatures. In (@97b), Jessica seems to indicate that whether Rachel likes Michael is less interesting than *who* likes him; in (@97b′) *what* Rachel thinks of Michael; and in (@97b″) *who Rachel* likes. Thus Jessica answers the question, but what the intonation marks as interesting is *coherently related* rather than *equivalent* to the polar issue raised by the question (@97a). That Jessica's marked interest must be *coherent* predicts why focus placement in her responses (@98) is more constrained. (@98b) is just like (@97b); i.e. Jessica places interest in her answer on the person corresponding to the agent in the prior yes/no question (roughly, on *who* got a job). But in (@98b′) and (@98b″), she places interest on *what Katrina got*, which bears no discernible relation to (@98a); it sounds *off-topic.*

## 6.2 Semantics for Focus and Tune

Here, I informally develop a model for intonation in terms of coherence, where the *background* and *foreground* derived from intonational form are dependent on *both* the accent placement and the tune.

**Background and Foreground.** Most accounts of focus separate a *foregrounded* (focal, rhematic) constituent from a *backgrounded* (given, thematic) one (e.g. Gussenhoven (1983); Krifka (1992); Beaver & Clark (2009) and others).[4] The accounts

---

[4]Vallduví & Engdahl (1996) also partition the background, making theirs a tripartite division.

differ on how these two parts interact with the context and with each other. Following Geurts & van der Sandt (2004), I assume that the background triggers a presupposition. But unlike them, I make it sensitive to the tune what precisely is being presupposed. Moreover, I add a clause to restrict *how* the foreground *relates* to the background presupposition. For the falling tune, this is as follows.

---

(I) **Focus Semantics** (falling tune)

Focal placement separates an utterance into a *foreground f* and a *background* $\varphi$, where a variable $x$ of the same type as $f$ occurs freely in $\varphi$. Updating a discourse with an utterance that has a falling tune with nuclear accent on $f$ proceeds as follows:

- Update with the *presupposition* $\varphi$; that is, its free variable $x$ must be resolved anaphorically (it is either bound or accommodated as $\exists x.\varphi$).

- Update the result with the *proffered* content $(\lambda x.\varphi)(f)$ (and all its presuppositions), such that the proffered content and $\varphi$ are coherently connected to form a common topic (this means that the proffered content must *elaborate* or *continue* on the presupposition (Asher & Lascarides, 2003)).

---

I treat presuppositions as anaphora (van der Sandt, 1992), i.e. a presupposition must be bound to an available unit in the discourse context or accommodated by coherently relating it to such a unit (Asher & Lascarides, 1998). That is, the free variable $x$ in the background $\varphi$ is treated as an anaphor, which needs to be bound or existentially closed. For example, recall (@89ab).

(@89)   a. Harvey: Who likes Michael?

b. Jessica: <u>Rachel</u><sub>H</sub> likes Michael.<sub>LL%</sub>

In (@89b), the Focus Semantics sets $f = $ *Rachel* and $\varphi = x$ *likes Michael* (the foreground triggers a presupposition via the proper name, but by Rule (I) this updates the context *after* $\varphi$). After $\varphi$ updates the context, the proffered content $\varphi(f)$ must attach to it with *Elaboration* (making their common topic $\varphi$) or *Continuation* (their common topic is a generalisation of their distinct but related contents). The question (@89a) presupposes *someone likes Michael*, and so by the dynamic semantic approach to anaphora, the background $x$ *likes Michael* binds to this, with $x$ bound to the existential quantifier. The proffered content then attaches to this background with *Elaboration*. That is, colloquially, the utterance is interpreted as *someone likes Michael—specifically, Rachel does.*

Other accounts also predict that the background is presupposed. To wit, $?\lambda x.\varphi(x)$, where $\varphi$ is the background defined in Rule (I), is the *wh*-question that is required by a Question Under Discussion (QUD) account (Roberts, 2012), and $\{x \mid \varphi(x)\}$ is the corresponding set of alternatives in Alternative Semantics (Rooth, 1992). Under the reasonable assumptions that *wh*-questions presuppose that there

is at least one true answer and that alternatives sets are non-empty, both accounts generate the presupposition that there is an $x$ such that $\varphi(x)$. In this sense, presupposing the background is a *minimal* assumption (Geurts & van der Sandt, 2004). One might therefore think that drawing on coherence relations is redundant. But that would be too hasty. As the remainder of this discussion should show, coherence is needed: it accurately constrains how presuppositions get resolved in context (Asher & Lascarides, 1998), and it also characterises the semantic dependencies between background and focus in ways that are sensitive to the tune.

**Binding and Accommodation.** Geurts & van der Sandt (2004) argue convincingly that standard models for how presuppositions get bound or accommodated make the right predictions for focus (though further below I argue against some of their claims). The above semi-formal analysis of (@89) is an example of binding; (@99) is an example of accommodation.

(@99)  a. Harvey: Does anybody like Michael?

b. Jessica: <u>Rachel</u><sub>H</sub> likes Michael.<sub>LL%</sub>

Unlike (@89a), the question (@99a) does not generate an existential presupposition, so *x likes Michael* is accommodated: in standard dynamic semantics, this is equivalent to adding an existential quantifier $\exists x$. This accommodation is similar to Roberts's (2012) idea that focus triggers the presupposition of a congruent *wh*-question and that this question can be *accommodated* as a QUD. Supposedly, accommodating this question would then also accommodate the existential presupposition that the question has an answer.

However, there must be constraints on what can be accommodated in what context. Otherwise, *any* accent placement is acceptable. In (@99) accommodation of $\exists x.x$ *likes Michael* is permitted because it *coherently relates* to the question (@99a). Roberts (2012, pp14–15) allows a question to be accommodated if any complete answer to it partially answers the previous QUD. In (@99), Jessica's utterance requires *Who likes Michael?* to be accommodated, and this succeeds because each complete answer to it also answers (@99a).

So, by and large, Rule (I) makes the same predictions as the QUD model. But the following examples motivate replacing congruence with something else.

(@100)  a. Harvey: Who likes Michael?

#b. Jessica: Rachel likes <u>Michael</u><sub>H·LL%</sub>

(@101)  a. Harvey: Does Rachel like Michael?

b. Jessica: <u>Rachel</u><sub>L+H</sub> does not like Michael.<sub>LH%</sub>

#b.′ Jessica: <u>Rachel</u><sub>H</sub> does not like Michael.<sub>LL%</sub>

Roberts (2012, p2) claims her model predicts the infelicity in (@100). To predict

that (@100b) is infelicitous, accommodating *Who does Rachel like?* must be ruled
out. But Roberts's constraint on accommodating questions is satisfied: according
to her formal definitions of answerhood,[5] each complete answer to *Who does
Rachel like?* entails either that Rachel likes Michael, or that she does not, and so
partially answers (@100a). Variations of these definitions may be able to avoid
this, but I will not pursue a fix here. There is a more significant difficulty: *Both*
(@101b) and (@101b′) would involve accommodating *Who does not like Michael?*;
so merely computing how this question relates to question (@101a) cannot tease
apart their differences in felicity. So it is not clear how Roberts's model makes
the right predictions in (@100) and (@101).

My model accounts for these cases by assigning a different semantics to fall–rise
tunes. Then, a weakened notion of question–answer congruence (see Section 6.3)
makes different predictions, depending on the tune.

**Negation and Contrast.**    Earlier I argued that (@93b) has the truth-conditions
of (95a) and (@94b′) of (95b).

(@93)    a.  William: Does Paula live in Paris?

       b.  Edith:    $\underline{\text{Paula}}_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$

              ↝ *someone (else) does live in Paris*

  #b.′ Edith:    $\underline{\text{Paula}}_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$

(@94)    a.  William: Who does not live in Paris?

       b.  Edith:    $\underline{\text{Paula}}_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$

              ↝̸ *someone (else) does live in Paris*

(95)    a.  It is not Paula who lives in Paris.

        ↝ *someone (else) does live in Paris*

       b.  It is Paula who does not live in Paris.

        ↝̸ *someone (else) does live in Paris*

However, observe now that in (@102b) the fall–rise tune is acceptable and its
meaning is *not* that of the *it*-cleft (95a).

(@102)    a.  William: Who does not live in Paris?

       b.  Edith:    $\underline{\text{Paula}}_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$

              ↝ *But is this what you wanted to know?*

Thus the fall–rise tune does not *mandatorily* result in a presupposition corre-
sponding to that of (95a). The model must account for this. Moreover, the

---

[5]To Roberts, a *complete* answer to $\lambda x.\varphi(x)$ decides all propositions that can be obtained by
substituting an entity $e$ for $x$ in $\varphi$. A partial answer decides at least one.

as-opposed-to reading of the fall–rise tune in (@93b) can arise in the absence of any overt negation:

(@103)  a. William: Does Paula live in Paris?

      b. Edith:     <u>Paula</u>$_{\text{L+H}}$ lives in Paris.$_{\text{LH\%}}$

           $\rightsquigarrow$ *but someone (else) does not*

There is no negation in (@103b), so it would be odd to attribute its implicature to determining the relative scope of a presupposition to a linguistically-introduced negation. Rather, this as-opposed-to reading derives from *adding* a negation to the background content, and determining its relative scope. I regiment this as follows: if the tune is fall–rise, the *polarity* of the background is left underspecified. Moreover, to obtain the intuitive readings in (@94b) and (@103b), the semantics specify that the foreground is in *contrast* to the background. Thus the way the polarity in the presupposition gets resolved depends on how it coherently relates to its discourse context, and how it can support a contrast with the proffered content. This semantics is expressed as follows.

---

(II) **Focus Semantics** (fall–rise tune, first attempt)
The focus placement separates an utterance into a *foreground f* and a *background $\varphi$* with free variable $x$. Updating a discourse with an utterance that has a fall–rise tune with nuclear accent on $f$ proceeds as follows:

– Update with the *presupposition* $y(\varphi)$ where $y$ is an underspecified variable of type *polarity*; i.e. $y \in \{\top, \neg\}$.

– Update with the *proffered* content $(\lambda x.\varphi)(f)$ (and all its presuppositions) such that the proffered content *contrasts with* the presupposition.

---

Rule (II) is a significant departure from prior accounts. Recall example (@93ab).

(@93)  a. William: Does Paula live in Paris?

      b. Edith:     <u>Paula</u>$_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$

           $\rightsquigarrow$ *someone (else) does live in Paris*

On the face of it, (@93b) corresponds to the alternatives set *people who do not live in Paris* or the QUD *Who does not live in Paris?*, respectively. Both accounts can, in principle, predict the implicature (as Rule (II) does) by accommodating instead the alternatives set *people who live in Paris* or the QUD *Who lives in Paris?* respectively. It is not clear *how* these accounts would compute this question / set. But whatever the grounds for doing so, it cannot be congruence: (@93b) is not congruent to *Who does not live in Paris?* and is not an alternative from *people who live in Paris.* Rule (II), which makes no mention of congruence, predicts the implicature of (@93b) as follows.

**Maximising Contrast.** The utterance (@93b) presupposes $y(x$ *does not live in Paris*) according to (II). The only available referent to bind $x$ to is *Paula*, but this is blocked (regardless of how $y$ is resolved) by the requirement that it contrast the proffered content: *Paula does live in Paris, but Paula does not live in Paris* ($y = \neg$) is a contradiction and *Paula does not live in Paris, but Paula does not live in Paris* ($y = \top$) is not a contrast because one cannot contrast a statement with itself.

Thus, the presupposition must be accommodated as $\exists x.y(x$ *does not live Paris*). Both $y = \top$ and $y = \neg$ would be permissible here, but $y = \neg$ establishes a stronger contrast to the proffered content *Paula does not live in Paris.* Therefore, $y = \neg$ is preferred on the independently motivated principle that people interpret discourse in a way that *maximises coherence* (Asher & Lascarides, 2003). By double negation elimination, this results in the presupposition that *someone else lives in Paris.* To see the maximisation of contrast, compare these approximate paraphrases of the two alternatives for resolving the value of $y$:

(104)   *There is someone who does not live in Paris*

*but that someone is not Paula and Paula does not live in Paris.*

(105)   *There is someone who lives in Paris, but Paula does not.*

While (104) *can* contrast 'someone' with 'Paula', the contrast in (105) is better. Thus, the pragmatically preferred interpretation of (@93b) can be paraphrased as (105) rather than (104). I elaborate how to formalise such reasoning in Section 6.4

Similar reasoning about maximising contrast captures the implicature in (@103) as well. However, it takes additional machinery from Section 6.3 to explain (@93b′). Now consider (@102).

(@102)   a. William: Who does not live in Paris?

b. Edith:   $\underline{\text{Paula}}_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$

$\rightsquigarrow$ *But is this what you wanted to know?*

The *wh*-question (@102a) presupposes *some e does not live in Paris.* By Rule (II) (@102b) presupposes $y(x$ *does not live in Paris*). This presupposition can be bound to the presupposition of (@102a) by making $y = \top$ and $x = e$. This is the preferred resolution, since binding a presupposition is better than accommodation.

However, the proffered content must *contrast* with this resolution, and so binding $e = x =$*Paula* is blocked: this would result in the presupposition *Paula does not live in Paris* and the identical, therefore non-contrasting, proffered content *Paula does not live in Paris.* Thus, $e \neq$*Paula*, resulting in the reading (104). In other words, Edith implicates that her answer *Paula* does not resolve William's question: *Paula lives in Paris, but this is not who I take your question to be about*—the desired implicature of (@102b).

Typically, $y = \neg$ results in a better contrast with the foreground and is thus often a part of the pragmatically preferred interpretation. In particular, $y = \neg$ is

preferred in the null context, yielding the as-opposed-to reading of fall–rise. Only in highly particular contexts, such as (@102a), does binding $x$ to an available antecedent and resolving $y$ to $\top$ yield a more coherent discourse. Thus, the Focus Semantics for L+H LH% crucially rest on the principle that a discourse is interpreted in a way that *maximises coherence.* This principle explains why and how the underspecified background content gets resolved to capture the intuitive implicatures of the fall–rise tune in its various contexts.

Finally, note that the negation that $y$ resolves to can be *metalinguistic.*

(@106)  a. William: We bought po-tah-toes.

b. Edith:    We bought <u>po-tay-toes</u><sub>L+H·LH%</sub>   $\rightsquigarrow$ *not "po-tah-toes"*

Clearly, in (@106b) Edith is not denying the propositional content of (@106a) (Carston, 1996). Rule (II) can account for these cases by allowing the $\neg$ to be metalinguistic and the $x$ to resolve to prior *use* or *mention.* This option also accounts for fall–rise signalling a speaker taking issue with the *presentation* of a proposition: While Edith logically gives a *positive* answer to (@107a), her answer implicates that "in the US", while true, mischaracterises the circumstances.

(@107)  a. William: Do you live in the US?

b. Edith:    I live in <u>New York City</u><sub>L+H·LH%</sub>

$\not\rightsquigarrow$ *not in the US.*

$\rightsquigarrow$ *not "the US".*

**Uncertainty.**    Rule (II), as it stands, fails to model *uncertainty readings.*

(@108)  a. William: Did Paula eat all the cookies?

b. Edith:    Paula ate <u>some</u><sub>L+H</sub> of the cookies.<sub>LH%</sub>

$\rightsquigarrow$ *but not all of them*;

or *but Edith is not sure whether it was all the cookies.*

(@109)  a. William: Is Michael coming to the party?

b. Edith:    He is <u>invited</u><sub>L+H·LH%</sub>

$\rightsquigarrow$ *but he is not coming*;

or *but Edith does not know whether Michael is coming.*

Edith's utterances in (@108b) and (@109b) are ambiguous: they can be interpreted as indirect *negative answers* or as indicating that Edith *is uncertain about* the answer to William's question; in the latter case, Edith is giving information that she has marked (by way of intonation) as perhaps relevant but insufficient to resolve the question.

Due to cases like these, the fall–rise tune, or the pitches into which some try to decompose it, have been associated with *uncertainty*: Ward & Hirschberg (1985)

associate it with the whole tune; Pierrehumbert & Hirschberg (1990) with the L+H accent; and Šafářová (2005) with the final rise. Moreover, Pierrehumbert & Hirschberg (1990) associate this uncertainty exclusively with scalar items, but a scalar relationship is not always necessary:

(@110)  *Context: Jessica and Katrina applied for the same job.*

      a. Harvey: Did they offer the job to Katrina?

      b. Jessica: Well, they didn't offer it to <u>me</u>$_{\text{L+H·LH\%}}$

Jessica's utterance is ambiguous, like (@109). It can express that Katrina got the job offer, or that Katrina may have gotten it, but Jessica is not sure about this. But there is nothing scalar in (@110).

Which of the two readings is preferred seems to vary with the intensity of the intonation and the steepness of the rises (Ward & Hirschberg, 1988), as well as contextual knowledge of the speaker's knowledge or intentions. The annotation for fall–rise tunes underspecifies intensity and steepness, so the semantics needs to be amended to make both readings *available*. The first version of the Focus Semantics for fall–rise (Rule II) only predicts the indirect answer reading; i.e. that Edith is saying that *Paula did not eat all the cookies (only some)* in (@108) and that *Michael is not coming (despite being invited)* in (@109).

To account for such uncertainty readings, the Focus Semantics (II) needs to be amended to allow a *third* option for the underspecified polarity $y$: the option to resolve to an uncertainty modal $\Diamond$ (as introduced in the semantics of Marie Šafářová (2005)).[6] Since my account includes underspecified auxiliaries anyway (see below), this incurs no additional formal overhead.

---

(II′)  **Focus Semantics** (fall–rise tune, final version)
    As before, the focus placement separates an utterance into a *foreground f* and a *background $\varphi(x)$*. Updating a discourse with an utterance that has a fall–rise tune with nuclear accent on $f$ proceeds as follows:

    – Update with the *presupposition $y(\varphi)$* where $y$ is an underspecified variable of type *modality*; i.e. $y \in \{\top, \Diamond, \neg\}$.

    – Update with the *proffered* content $(\lambda x.\varphi)(f)$ (and all its presuppositions) such that the proffered content *contrasts with* the presupposition.

---

Note that, as before, *maximising contrast* typically favours the $\neg$ reading: "something is not P but f is P" ($y = \bot$) is typically a better contrast than both "something is P but f is not that something despite being P" ($y = \top$) and "something might be P but f is (definitely) P" ($y = \Diamond$). So the predictions of the first version of the fall–rise semantics, outlined earlier, are replicated by the final version.

---

[6]This is not quite the $\Diamond$ from Chapter 4, but a dynamic operator. That is, if something is $P$, it is *definitely $P$* $(P(x) \models \Box P(x))$, but $\Diamond \neg P(x) \wedge P(x) \not\models \bot$ (Groenendijk *et al.*, 1996).

**Ambiguities.** Thus, all else being equal, the better contrast is established by $y = \neg$. This means that the semantics in (II′) *favours* interpreting (@108b), (@109b) and (@110b) as indirect negative answers. But the reading where $y = \Diamond$ is available to interpret (@109b) (for instance) as *maybe Michael is coming, but he (definitely) is invited*, and this reading arises if the indirect answer reading is pragmatically *blocked*, for instance by the common ground knowledge that Edith cannot know for sure that Michael is coming.

More generally, real world knowledge can substantially affect how $y$ resolves in context. Example (@111) is a case where the uncertainty reading is preferred:

(@111)  a.  Amy: Does Paula like opera?

      b.  Bob: She likes <u>Wagner</u><sub>L+H·LH%</sub>

             ↝̸ *Paula does not like opera.*

             ↝ *maybe Paula likes opera.*

      b.′  Bob: She likes <u>Wagner</u><sub>H·LL%</sub>

             ↝ *Paula likes opera.*

Axioms of rationality and cooperativity predict that responses to polar questions provide evidence for a positive answer or for a negative answer (Asher & Lascarides, 2003, pp403–405); when the evidence proffered is conclusive, a particular answer is implied. Combining this expectation with the real world knowledge that *liking Wagner* is strong evidence for *liking opera* predicts that Bob has offered evidence for a *positive* answer; so the reading *Paula does not like Opera, but she likes Wagner* (making $y = \neg$) is dispreferred. Further, $y = \Diamond$ is preferred to $y = \top$, because $y = \Diamond$ produces a better contrast: *maybe Paula likes opera but she (definitely) likes Wagner* vs. *Paula likes opera* ?*but she (also) likes Wagner*. Thus in this context, the fall–rise intonation conveys that Bob does not quite commit to a positive answer.

This contrasts with (@111b′) uttered with falling intonation, where the (same) evidence for a positive answer, provided by real world knowledge about Wagner and opera, commits Bob to a positive answer (*Paula likes opera, specifically, she likes Wagner*). These differences are predicted by the semantics: the fall–rise tune demands a contrast between given and proffered content, while a falling tune demands elaboration or continuation. However, such readings are a matter of degree: (@112b) is arguably ambiguous as to which answer it implicates because one cannot decide with sufficient confidence whether Bob's assumptions about the commonsense relations between musicals and opera lead him to believe that *liking musicals* is positive evidence for *liking opera*, or negative evidence.

(@112)  a. Amy: Does Paula like opera?

      b. Bob: She likes <u>musicals</u><sub>L+H·LH%</sub>

**Focussed Quantifiers.**   Explaining focus as triggering a presupposition of the background has been criticised (Dryer, 1996; Rooth, 1999). Even Geurts & van der Sandt (2004, pp28–30) see an issue with it when a quantifier is focussed: a naïve reading of presupposition semantics for focus can yield the faulty readings in (@113). Here, $x$'s type appears as a subscript.

(@113)   a. Polyphemus: <u>Nobody</u><sub>H</sub> likes Michael.<sub>LL%</sub>

                 ??background: $x_{\mathbf{entity}}$ *likes Michael.*

      b. Polyphemus: <u>Somebody</u><sub>H</sub> likes Michael.<sub>LL%</sub>

                 ??background: $x_{\mathbf{entity}}$ *likes Michael.*

The background in (@113a) contradicts the proffered content: $x_{\mathbf{entity}}$ denotes an individual in the model and so cannot be nobody. The background in (@113b) makes the proffered content not new (or interesting). In both cases presupposing the background is absurd, and so Geurts & van der Sandt (2004, pp28–30) argue instead for a *polarity* focus, thereby yielding for both (@113a) and (@113b) the tautological presupposition *either nobody likes Michael or somebody did.*

However, the Focus Semantics specify to replace the foreground constituent with a free variable of the *same type.* Now, a semanticist (or cyclops) who assumes that the focal constituent 'nobody' in (@113a) is of type *entity* might in the end lose an eye to Odysseus. Wishing to avoid that fate, one may want to take a closer look at the arguments of Geurts & van der Sandt.

> The non-logical part of the semantic content of words like 'somebody' and 'nobody' is so general that it is unlikely to attract the focus of a statement; 'somebody' cannot be used to mean 'some person, as opposed to some vehicle' (say). What remains to be focused is the negative part of 'nobody' and the corresponding positive component of 'somebody'               (Geurts & van der Sandt, 2004, p29)

However, this is wrong: there are cases where (i) *somebody* gets an 'as opposed to' reading; and cases where (ii) focus on an existential quantifier is not polar (i.e. not contrasting with 'nobody' or 'nothing'). To see (i) consider (@114) from Walker (1996a) (where I added an appropriate tune) and a variant (@115).

(@114)   a. William: There is a man in the garage.

      b. Edith:    There is <u>something</u><sub>L+H</sub> in the garage.<sub>LH%</sub>         (Walker, 1996a)

(@115)   a. William: There is something in the garage.

      b. Edith:    There is some<u>body</u><sub>L+H</sub> in the garage.<sub>LH%</sub>

Intuitively, (@114b) means *it (merely) might be a man* and so corrects (@114a), contrary to what Geurts & van der Sandt claim. However, this reading may still be attributed to a Quantity implicature: since *something* is less informative than *a man*, one may infer the desired reading via a Gricean reasoning. But one cannot analyse (@115) this way—*somebody* is not less specific than *something*—and yet

its meaning is exactly what Geurts & van der Sandt (2004) deny. Dialogue (@116) from Schlöder & Fernández (2015b) is an example of case (ii): Danny's denial move with focus on 'some' cannot be about polarity: the issue is not *some* vs. *none*, but *some* vs. *all* (a potential Quantity implicature notwithstanding).

(@116)  a. James: (...) we're all mad, aren't we?

   b. Danny: Well, some$_{L+H}$ of us.$_{LH\%}$

   (BNC, HUV, 1468–1469; constructed tune (original audio unavailable))

I propose that in (@114–@116), the second speaker takes issue with the first speaker's choice of *specific quantifier*. So, I allow quantifiers as foregrounds, and hence as free variables in the presupposition triggered by the Focus Semantics:[7]

(@117)  a. Polyphemus: Nobody$_H$ likes Michael.$_{LL\%}$

   background: $x_{\textbf{quantifier}}z.(z \textit{ likes Michael})$.

   b. Polyphemus: Somebody$_H$ likes Michael.$_{LL\%}$

   background: $x_{\textbf{quantifier}}z.(z \textit{ likes Michael})$.

The presupposition in (@117) can be accommodated to form a tautology, since *there is a quantifier $x$ such that $x(p)$ is true of any proposition $p$*. This replicates the tautology reading of Geurts & van der Sandt. However, this semantics is also compatible with the more specific readings of the backgrounds in (@114–@116). Note moreover that a similar phenomenon applies to modal operators.

(@118)  a. William: Do I have to attend class?

   b. Edith:  You may$_H$ attend class;$_{LL\%}$ we do not take attendance.

   background: *William $x_{\textbf{aux}}$ attend class*.

If modal operators can be foregrounds $f$, then there is little harm in also having foregrounded quantifiers. This also means that to explain (@118), the model needs to have variables ranging over modal auxiliaries. So the underspecification added in the Focus Semantics (II') does not incur formal overhead by underspecifying a modal. This treatment of modals is not particular to my account either. To the best of my understanding, an Alternative Semantics would raise a set of alternative modalities and a QUD-account would take (@118b) to answer *What are William's obligations regarding attending class?*.

**Embedded Focus.**   Geurts & van der Sandt (2004) avoid further potential counterexamples to a presupposition account by making the presupposition trigger a *default*. That is, the presupposition of the background gets what they call *suspended* in certain cases.

   However, making the Focus Semantics sensitive to the tune explains the cases of apparent suspension while keeping the Focus Semantics *universal*: i.e. the

---

[7]I assume standard semantics for (generalised) quantifiers. To wit $[\![x_{\textbf{GQ}}z.\varphi]\!] = 1$ iff $\{z \mid \varphi\} \in [\![x_{\textbf{GQ}}]\!]$. E.g. $[\![\forall]\!]$ is the singleton universal set and $[\![\exists]\!]$ is the set of all non-empty sets.

background is *always* presupposed. Geurts & van der Sandt (2004, p12) motivate suspension by claiming that (119a) presupposes *somebody stole the tarts*, while this presupposition is inconsistent when (119a) is embedded as in (119b).

(119)  a.  <u>Fred's wife</u> didn't steal the tarts.

   b.  I'm still not convinced that the tarts were stolen, but surely

   <u>Fred's wife</u> didn't steal them.                    (Geurts & van der Sandt, 2004)

But they do not consider tune. (@120) shows that, indeed, (119a) with its most natural intonation triggers *someone stole the tarts*, whereas (119b) does not. But there is no need to suspend a presupposition to make this prediction.

(@120)  ?a.′  <u>Fred's wife</u>$_H$ didn't steal the tarts.$_{LL\%}$

   a.″  <u>Fred's wife</u>$_{L+H}$ didn't steal the tarts.$_{LH\%}$

   b.′  I'm still not convinced that the tarts were stolen, but surely

   <u>Fred's wife</u>$_H$ didn't steal them.$_{LL\%}$

   b.″  I'm still not convinced that the tarts were stolen, but surely

   <u>Fred's wife</u>$_{L+H}$ didn't steal them.$_{LH\%}$

In most natural contexts, it sounds odd to utter (@120a′): one is inclined to read (119a) as belonging to a context in which Fred's wife is suspected of tart stealing, and a fall–rise tune as in (@120a″) is the intuitive tune for denying this suspicion (and implicate that someone else stole them). (@120a′) would be felicitous only in the odd context where the issue *who did not steal the tarts* is salient; for instance when many people stole tarts and one wonders whether anyone did not.

   According to the Focus Semantics (Rule II′), (@120a″) generates the presupposition $y(x$ *did not steal the tarts*$)$ and its preferred coherent interpretation yields $y = \neg$: *there is someone who did steal the tarts* forms the best available contrast with the proffered content (*someone stole the tarts but Fred's wife did not*).

   Now, (@120b′) generates the presupposition *surely x didn't steal the tarts*; this is consistent and coherent with the antecedent clause (that the speaker is not convinced the tarts were stolen). Similarly, (@120b″) generates the presupposition $y(surely\ x\ didn't\ steal\ the\ tarts)$. Setting $y = \neg$ and globally accommodating $x$ to $\exists x$ yields the consistent and coherent interpretation (reading *not surely not* as *might*): *Someone might have stolen the tarts. I'm still not convinced that the tarts were stolen, but surely Fred's wife didn't steal them.*

## 6.3   Pragmatics of Focus and Tune

In this section, I take up Bolinger's suggestion that focus marks *interest*: I give a formal condition on which parts of an utterance can be considered interesting and discuss aspects of interest that go beyond this condition. Then, I discuss the *perlocutionary effects* of the two tunes under consideration.

**Given and Interest.**  Any adequate semantics of focus must explain the infelicity of (@89b′).

(@89)   a.  Harvey: Who likes Michael?

   #b.′ Jessica: Rachel likes <u>Michael</u>$_{\text{H·LL\%}}$

I earlier outlined a potential explanation: intuitively, one cannot place focus on 'Michael' because in the context of (@89a), 'Michael' is *given* and hence not interesting. This can be formalised as follows. Say that to be interesting it is *necessary* (but not sufficient) to be *not given*. To make such a constraint predictive, a precise definition of 'given' is needed. Because focus marks the *speaker's* interest, I now take the backgrounded content of the current utterance to contribute to what is considered given. Moreover, I restrict the given information to the (salient) part of the discourse context which is being addressed by the current utterance.

---

(III) **Relevant Segment**
The *relevant segment* for a foreground–background pair $\langle \varphi, f \rangle$ is the segment of the prior discourse that the proffered content most coherently relates to.

(IV) **Givenness**
The given information is the content that results from (coherently) updating the relevant segment of the prior discourse with the (presupposed) background of the current utterance.

---

In all examples here, the relevant segment is the last utterance in the prior discourse. But in general, choosing the most coherent antecedent discourse unit to attach the content of the current discourse unit to is based on commonsense inference (Hobbs *et al.*, 1993; Asher & Lascarides, 2003).

   Now, Rules (III) and (IV) entail that if the most coherent way to update the discourse with the background content results in a meaning that entails the proffered content, then what is proffered is *given*. Rule (V) makes this anomalous:

---

(V) **Necessary Condition for Interest**
A foreground–background pair $\langle \varphi, f \rangle$ is *not interesting* if the proffered content $\varphi(f)$ is given. This means an utterance is not interesting if:

 − for the falling tune, which presupposes $\varphi$ with free variable $x$, it is given that $x = f$.

 − for the fall–rise tune, which presupposes $y(\varphi)$ with free variable $x$ and modality variable $y$, it is given that $y = \top$ and $x = f$.

---

At first glance, this seems to repeat the old wisdom that focal information cannot be given (Büring, 2006; Beaver & Clark, 2009), but, note that my definition of

Givenness (IV) *substantially* differs from prior accounts. It does not say that what can be background is *constrained* by what is given, but instead *adds* the background to *compute* what is given (compare in particular with Schwarzschild (1999)). Thus, the background may contain *new* information (i.e. that is not in the prior context), which is *accommodated* as given. This is contrary to the received view on givenness, where it is stipulated that everything that is not focal (i.e. given) must be not discourse-new.

Moreover, the prior definitions of givenness take the entire prior discourse to be given (Büring, 2006). The definition (IV) only takes one part of the prior discourse—the relevant segment—to be given. This is because attention in discourse is limited (Walker, 1996b) and a fall–rise tune can be used to repeat prior information.

(@121)  a. Harvey: Who likes Michael?

      b. Jessica: <u>Rachel</u>$_H$ likes Michael.$_{LL\%}$

$$\vdots$$

      y. Harvey: Who likes Michael again?

      z. Jessica: <u>Rachel</u>$_{L+H}$ likes Michael.$_{LH\%}$

          $\rightsquigarrow$*you should know that*

Later in this section I show how the fall–rise tune leads to the *you should know* implicature.

The possibility to accommodate new information as given mirrors the capacity for presuppositions to contain *new* information that gets accommodated. Thus, what Rule (V) essentially says is that you cannot *mark as interesting* content that you (also) *present as given*. As said, this is not a *definition* of interest. Instead, it uses notions already available—presupposition, coherence, *etc*—to *constrain* what *can* be (marked as) interesting.

**How this works.** Rules (III), (IV) and (V) together are a weakening of question–answer congruence. They predict that (@89b′) sounds odd.

(@89)    a. Harvey: Who likes Michael?

      #b.′ Jessica: Rachel likes <u>Michael</u>$_H$.$_{LL\%}$

In the context of (@89a), the most coherent way to instantiate $x$ in the presupposition *Rachel likes $x$* that is triggered by (@89b′) is to bind it to *Michael*. But then the proffered content is entailed, which is ruled out by the necessary condition for interest. Similarly, Rule (V) predicts that (@122b) sounds odd.

(@122)    a. Harvey: Who does Rachel like?

      #b. Jessica: <u>Rachel</u>$_{L+H}$ likes Michael.$_{LH\%}$

The most coherent way to instantiate $y$ and $x$ in *$y(x$ likes Michael)* is to resolve $x$ to *Rachel* and $y$ to $\top$, because *Rachel likes Michael* is an answer to (@122a), but

*Rachel does not like Michael* and *maybe Rachel likes Michael* are partial answers and thus dispreferred. Note that in computing what is given, the proffered content is not considered; so it does not matter that the resolution *Rachel likes Michael* cannot contrast the proffered content in (@122b). Thus the proffered content with both tunes is given, violating Rule (V).

Computing given information via coherence plays a further useful role in predicting felicity judgements: if the background does not attach to the relevant segment—i.e. one *cannot compute* what is given—then the dialogue is incoherent.

(@98)   *Context: Jessica and Katrina are both job hunting.*

      a.   Harvey: Did you get a job?

  #b.$'$   Jessica: Katrina got a job$_{\text{L+H·LH\%}}$

  #b.$''$   Jessica: Katrina got a job$_{\text{H·LL\%}}$

No coherent interpretation can be computed of the presuppositions triggered by (@98b$'$,b$''$) in the context (@98a), as there is no $x$ that would validate *Katrina got an $x$* as a coherent response to (@98a). Thus the independently motivated principle that accommodating a presupposition must coherently connect it to its context (Asher & Lascarides, 1998) explains (@98b$'$,b$''$).

Moreover, just like presuppositions generally, the coherent interpretation of given information need neither be *unique* nor *final*—but there must be *at least one* way of establishing its coherence. Example (@93) demonstrates this. It is a case where a highly salient coherent interpretation of given information gets overridden by proffered content:

(@93)   a. William: Does Paula live in Paris?

      b. Edith:    Paula$_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$

             $\rightsquigarrow$ *but someone (else) does.*

The presupposed content of (@93b) is $y(x$ *does not live in Paris*). The most coherent update of (@93a) with the (underspecified) presupposition *on its own* yields $x = $ *Paula* but does not resolve $y$ uniquely: $y = \top$ and $y = \neg$ both supply (full) answers to the question, while $y = \Diamond$ provides a partial answer (and so arguably is less coherent and hence dispreferred). Either way, the presupposition is coherent (though ambiguous), with $x = $ *Paula.*

However, when updating this with the proffered information, the Focus Semantics (II$'$) for fall–rise demands a *contrast*, and so the defeasible inference that $x = $ *Paula* is overridden: $x$ gets existentially bound and $y$ resolves to $\neg$ (resulting in *someone else lives in Paris*). This (still) coherently attaches to the question (as a commentary rather than an answer) but it also contrasts the proffered content (which in turn attaches to (@93a) as an answer), as demanded by the Focus Semantics.

Since the tunes trigger different presuppositions, the focal placement alone does not determine whether the foreground can be interesting. Rather, interest is

governed by the focal placement and tune in combination. As opposed to (@93b), (@93b′) is infelicitous.

(@93)    a. William: Does Paula live in Paris?

        b. Edith:    $\underline{\text{Paula}}_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$

    #b.′ Edith:    $\underline{\text{Paula}}_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$

The presupposition of (@93b′) is *x does not live in Paris*; in the context of (@93a), the most coherent resolution is *x = Paula*. So by Rule (V), (@93b′) is not interesting and thus incoherent. In contrast, (@93b) generates the presupposition *y(x does not live in Paris)*, but does not resolve *y*. Thus (@93b) can be interesting, according to Rule (V). *Mutatis mutandis* this also explains why (@93b″) is coherent (where the focus is on the polarity).

(@93)  a. William: Does Paula live in Paris?

        b.″ Edith:    Paula does $\underline{\text{not}}_{\text{H}}$ live in Paris.$_{\text{LL\%}}$

Here, what is backgrounded is *Paula $x_{\textbf{polarity}}$ lives in Paris*. Since resolving $x = \top$ and $x = \neg$ results in equally coherent answers to (@93), $x = \neg$ is not given.

**Interest Beyond the Necessary Condition.**   Due to the underspecified variable *y* in the semantics for the fall–rise tune, whenever the polarity of a fall–rise utterance is not given, then the negated utterance's polarity is not given either. This is in particular the case when a polar question is the relevant segment.

(@123)  *Context: Jessica and Katrina applied for the same job.*

        a. Harvey: Did they offer the job to Katrina?

        b. Jessica: Well, they didn't offer it to $\underline{\text{me}}_{\text{L+H·LH\%}}$

        b.′ Jessica: They offered it to $\underline{\text{me}}_{\text{L+H·LH\%}}$

But there are some clear counterexamples to this.[8]

(@124)    a. Julian: Is Nicholas coming to my talk?

        b. Alex:    $\underline{\text{Ernie}}_{\text{L+H}}$ is coming to your talk.$_{\text{LH\%}}$

    #b.′ Alex:    $\underline{\text{Ernie}}_{\text{L+H}}$ isn't coming to your talk.$_{\text{LH\%}}$

The necessary condition for interest predicts the felicity of (@124b), but *not* the infelicity of (@124b′). (@124b) passes the test since the background polarity is not given—but then the background polarity of (@124b′) is not given either.

However, (@124b′) is infelicitous because there *is no relevant segment.* In (@124b) the proffered content attaches to the question (@124a) as a *Plan-Correction* (Asher & Lascarides, 2003, p320), i.e. Alex is not following Julian's conversational strategy, but cooperatively amends it according to what she infers about Julian's *goals.* Validating this relation requires knowledge of these goals.

---

[8]I owe this point to the audience of the 2017 Rutgers Semantics & Pragmatics Colloquium, in particular to Sam Carter who pressed me on this.

It is accommodable in (@124) that Julian's goal is to know how challenging the question session will be; by answering *Ernie is coming* Alex is giving relevant, maybe even *better*, information towards Julian's goal. But it is much harder to accommodate that *Ernie is not coming* attaches as a Plan-Correction to (@124a)—or indeed with any other relation. In comparison, the context in (@123) is such that the proffered content in both responses attaches as an *indirect answer* (@123a).

Cases like (@124) show that computing the relevant segment (and thus what is given) can rely on commonsense knowledge. In a highly specific context (e.g. where Nicholas goes where Ernie goes), (@124b′) may also be an indirect answer to (@124a) and would be felicitous. The following example is another case in point.

(@125)   a. Edith:    Is my daughter in hospital?

  #b. William: Ernie$_{L+H}$ is in hospital.$_{LH\%}$

  #b.′ William: Ernie$_{L+H}$ is not in hospital.$_{LH\%}$

In the context of (@125a), commonsense knowledge entails that Edith is not interested in the health of anyone else. Thus both (@125b) and (@125b′) are infelicitous despite satisfying the necessary condition for interest. Their infelicity is explained like (@124b′): there is no coherence relation that would make *Ernie is in the hospital* relevant to Edith's salient goals.

This shows that the necessary condition is at best a small component of what it means to be *interesting*. At its heart, interest is paralinguistic and subject to individual variation. Example (@125) already points to this, but Bolinger (1985) offers a stronger case. He suggests the Boston Strangler might utter (@126a) while a sane individual might describe the situation as (@126b):

(@126)  a. I'm looking for a girl$_H$ to strangle.$_{LL\%}$

  b. He's looking for a girl to strangle$_H$.$_{LL\%}$

For the Boston Strangler, it is a matter of course that he will strangle someone, and *who* exactly he will strangle is his matter of interest. To another person, it is more interesting that he will *strangle* someone. My model of focus and tune as it stands does not account for such variations.

**Beyond Interest.**   Bolinger (1972) presents some peculiar cases similar to (@127) to challenge syntactic approaches to focus. The necessary condition for interest can explain some of them.

(@127)   a. William: What did John do?

  #b. Edith:    John killed someone$_H$.$_{LL\%}$

  b.′ Edith:    John killed$_H$ someone.$_{LL\%}$

  b.″ Edith:    John killed a policeman$_H$.$_{LL\%}$

  #b.‴ Edith:    John killed$_H$ a policeman.$_{LL\%}$

Even though *killed someone* and *killed a policeman* are both discourse-new, the former requires focus on *killed* and the latter on *policeman*. This is puzzling for any account that attempts to explain focus placement via what is discourse-new.

However, the necessary condition for interest can make partial sense of this. Observe that in (@127b) the presupposed background is *John killed $x_{\text{quantifier}}$*. Since *John killed nobody* is a dispreferred response to (@127a), it is given that $x$ resolves to a non-empty quantifier. Thus *John killed someone* is given and so by Rule (V), (@127b) is anomalous and (@127b′) is the correct way to mark interest. However, *John killed a policeman* is not given, so (@127b″) is fine. However, explaining why (@127b‴) is infelicitous requires principles that go beyond interest.

There are constraints on what *can* be marked as interesting. The necessary condition (V) is one of them. But there are other linguistic constraints on focus that Bolinger was not concerned with. Notably, if there are multiple parts of a constituent that are not given (and so potentially interesting) there is an observed tendency to place focus on the right-most part.

(@128)　a. Stephen: What did you see?

　　　　b. Lucy:　　I saw a white <u>cat</u>$_{\text{H}}$·LL%

　#b.′ Lucy:　　I saw a <u>white</u>$_{\text{H}}$ cat.LL%

Neither 'white' nor 'cat' are given in (@128), so at least as far as the necessary condition for interest is concerned, it would be legitimate to put focus on either of them. However, there is a clear preference for (@128b) over (@128b′). This preference to the right is preserved even in the romance languages, where adjectives follow the noun they modify. (@129) is an example from Italian.

(@129)　a. Stefano: Che cosa hai visto?　　　　　*What did you see?*

　　　　b. Lucilla: Ho visto un gatto <u>bianco</u>$_{\text{H}}$·LL%　*I saw a <u>white</u> cat.*

　#b.′ Lucilla: Ho visto un <u>gatto</u>$_{\text{H}}$ bianco.LL%　*I saw a white <u>cat</u>.*

Of course, this does not mean that English speakers find nouns interesting while Italian speakers are fascinated by adjectives. Rather, the two languages' grammars affect the rules for where to place focus *among* the potentially interesting information. Note that both languages forbid the focus marking of given information, in accordance with Rule (V):

(@130)　a. Stephen: What kind of cat did you see?

　#b. Lucy:　　I saw a white <u>cat</u>$_{\text{H}}$·LL%

　　b.′ Lucy:　　I saw a <u>white</u>$_{\text{H}}$ cat.LL%

(@131)　a. Stefano: Hai visto una cosa bianca?　　*Did you see something white?*

　#b. Lucilla: Ho visto un gatto <u>bianco</u>$_{\text{H}}$·LL%　*I saw a <u>white</u> cat.*

　　b.′ Lucilla: Ho visto un <u>gatto</u>$_{\text{H}}$ bianco.LL%　*I saw a white <u>cat</u>.*

Comparing (@128) and (@129) can be used to support focus projection: by projecting focus to the head of a noun phrase, both (@128b) and (@129b) focus *white cat.* This is, in fact, necessary if one subscribes to the view that all discourse-new information must be focussed.

However, I do not think that such projection is necessary. The logical forms that my Focus Semantics assigns to these examples can be paraphrased as follows:

(132)   (@128b): *L saw something white—it was a cat.*

   (@129b): *L saw a cat—it was white.*

These logical forms are, to all intents and purposes, equivalent; both semantically and pragmatically (e.g. where accessibility to anaphora are concerned). Thus I contend that the Focus Semantics are sufficiently flexible to avoid the need for focus projection. However, this departs from Bolinger's account: I do not, and indeed cannot, say that 'cat' in (@128b) or 'bianco' in (@129b) is *the* or *the most* interesting constituent. Rather, it is *one* of the constituents that *can* be marked as interesting, and other principles (that are possibly syntactic and may not relate to interest at all) govern the placement among these.

**Non-Focal Tune Semantics.**   There is another way in which intonation affects a discourse. The tunes do not only affect the Focus Semantics, but also make a contribution aside from this. Here is a case in point. Responses (@133b) and (@133b′), from Ward & Hirschberg (1985), have the same illocutionary force: they are both positive, indirect answers to (@133a).

(@133)   a. Amy: Did you read the first chapter?

   b. Bob: I read the entire dissertation$_{\text{H} \cdot \text{LL\%}}$

   b.′ Bob: I read the entire dissertation$_{\text{L+H} \cdot \text{LH\%}}$

      $\rightsquigarrow$ *I thought you knew this already.*

However, intuitively, they carry distinct *perlocutionary (meta-communicative) effects*: (@133b′) carries the implicature paraphrased above, while (@133b) does not. Similarly, (@134b) and (@134b′) both deny (@134a), but with similar distinct perlocutionary effects:

(@134)   a. Amy: You can't afford that.

   b. Bob: I'm a millionaire$_{\text{H} \cdot \text{LL\%}}$ $\not\rightsquigarrow$ *I thought you knew this already*

   b.′ Bob: I'm a millionaire$_{\text{L+H} \cdot \text{LH\%}}$ $\rightsquigarrow$ *I thought you knew this already*

In (@133) and (@134), the distinct perlocutionary consequences of the two tunes do not yield different inferences about the speaker's *illocutionary* act. This is because semantic relationships between the linguistic contents—entailment between *reading the entire dissertation* and *reading its first chapter*, and divergence between *being unable to afford something* and *being a millionaire*—suffice to infer the same specific coherence relations between Amy's move and Bob's response. In (@135ab) and

(@135ab′), however, where world knowledge supplies no (prior) logical relationship between *being a liar* and *being a fool*, both the illocutionary and the perlocutionary effects are different, as discussed by Ladd (1980):

(@135)  a. Amy: Harry is the biggest liar in town.

      b. Bob:  The biggest <u>fool</u>$_H$, maybe.$_{LL\%}$

      b.′ Bob:  The biggest <u>fool</u>$_{L+H}$, maybe.$_{LH\%}$                    (Ladd, 1980)

This is a case where intonation makes the difference between agreement and disagreement. In (@135ab), Bob is not denying that *Harry is the biggest liar*; he *adds* that *Harry may be the biggest fool (as well).* But in (@135ab′), Bob does not agree with (@135a): he offers the proposition that Harry is the biggest fool *instead*, and moreover implicates something like *before you said what you did, I thought you knew this.*

In general, tunes give voice to certain *cognitive attitudes* (Liberman & Sag (1974); Ladd (1980) and many others). I attribute the following perlocutionary consequences to the falling and fall–rise tunes (largely in line with my earlier work (Schlöder & Lascarides, 2015) and that of others):

---

(VI) **Tune Semantics**

The falling tune marks a proposition as *informative.*

The fall–rise tune makes a contribution that can be glossed as: what you just said (i.e. the content of the relevant segment) leads me to believe that you do not know what I'm saying now, but I thought you did know what I'm saying now.[9]

---

I show in Section 6.5 that the denial move in (@135ab′) is derivable from the formalisation of the semantics in (VI). The Tune Semantics *voice* what the interpretation as a denial implicates: that Bob thinks that believing Harry is a liar normally means you do not believe he's a fool. *Ceteris paribus*, the analysis of (@136),[10] is the same; with the only difference being that there may already be a commonsense relationship between *being good at badminton* and *not being a klutz* (whereas *fool* and *liar* are largely independent).

(@136)  a. Amy: Alan's such a klutz.

      b. Bob:  He's good at <u>badminton</u>$_{L+H \cdot LH\%}$

In earlier work, I conjectured that focus and pitch can relate to certain scalar implicatures (Schlöder & Fernández, 2015b). The corpus data gathered in this

---

[9]If a fall–rise is interpreted discourse-initially, then these semantics can either attach to non-linguistic antecedents (such as actions indicating particular beliefs) or simply fail to apply.

[10](@136b) is usually given in the form "He's a good <u>badminton</u>$_{L+H}$ <u>player</u>$_{L+H \cdot LH\%}$" To fit my simplifying assumptions (one accent per tune), I give here an equivalent single–focus construction.

paper showed that scalar items can appear in utterances with rejection force, but that inferring the correct force is difficult. Consider, for instance, examples (137) and (138), taken from this paper:

(137)  a. Sandra: That's brilliant.

   b. Danny: Well I thought that was quite good.   (BNC, KDV, 5306–5307)

     $^?\rightsquigarrow$ *good, not (necessarily) brilliant*

(138)  a. Bev:   It was good, weren't it? [*sic*]

   b. Wendy: It's brilliant.        (BNC, KE6, 1642–1643)

     $^?\rightsquigarrow$ *not merely good*

According to the analysis in *Pragmatic rejection* (Schlöder & Fernández, 2015b), (137b) is a rejection *by* implicature (*quite good* implicates *not brilliant*), and (138b) is a rejection *of* the same implicature. However, only about 50% of untrained annotators judged such (and similar) examples as actually having rejection force. This points towards an ambiguity, and it seems that intonation could be a disambiguating device. As a case in point, consider the following variant of (137).

(139)  a. William: *The Matrix* is a brilliant movie.

   b. Edith:   It's good$_{\text{H·LL\%}}$

      $\rightsquigarrow$ *I agree, it is good.*

   b.' Edith:   It's good$_{\text{L+H·LH\%}}$

      $\rightsquigarrow$ *it is not brilliant, but (just) good*$^{\textbf{exh}}$.

Apparently, the unintonated surface form 'it's good' is ambiguous between (139b) and (139b'), but adding a tune resolves the ambiguity. This is explained by the Tune Semantics in the same way as they explain (@135).

**Summary: Intonated Discourse Update.**  In sum, interpreting an utterance with a particular intonation updates a discourse as follows.

 i. The grammar produces a foreground–background pair $\langle f, \varphi \rangle$, where the (unique) accent is on $f$ (recall that I only consider tunes with only one accent) and $\varphi$ features a free variable $x$ of the same semantic type as $f$.

 ii. From $f$ and $\varphi$, the *proffered* and (underspecified) *presupposed* contents are computed, according to the Focus Semantics (I, II').

 iii. To interpret the utterance, first compute the *relevant segment* via the general and independently motivated principles for computing an update of the (prior) discourse context with the (underspecified) proffered content (III).

 iv. Then, compute whether and how the *presupposed* content attaches to the relevant segment, again via independently motivated principles for presupposition resolution. The result is the *given* information, defined in (IV).

v. Check whether the given information and proffered content satisfy the necessary condition for interest (V).

vi. If all is well, update the discourse with both the presupposed and proffered content (again via reasoning about discourse coherence), ensuring that the result is consistent with the coherence relations entailed by the focus semantics (I and II′) and the meaning postulates for the tunes (VI).

Steps (i–ii) serve to define the logical form of the utterance given its intonational form (though I forego *deriving* these within the grammar): $\varphi$ and $f$ are simply computed by $\lambda$-abstracting the constituent with the nuclear accent. Steps (iii–v) jointly are a *check* on the felicity conditions of those (underspecified) logical forms. This check makes use of notions related to coherent discourse update, but does not amount to *actually updating* the current context; this is executed in Step (vi) only if the felicity condition in (v) is satisfied.

## 6.4   SDRT and Hindsight Logic

I now give a brief overview over the fundamental principles of Segmented Discourse Representation Theory (SDRT, Asher & Lascarides 2003) and the Hindsight Logic I developed with Alex Lascarides (Schlöder & Lascarides, 2015). I use SDRT here because it has already been used extensively to model the interaction between discourse coherence, presuppositions and cognitive states. Moreover, the informal analysis in the preceding sections rests on a principle that selects the *maximally coherent* specific interpretation of an underspecified logical form. SDRT offers a highly sophisticated implementation of this principle.

**Discourse Structure.**   SDRT models discourse structure by connecting the contents of utterances with *coherence relations*; e.g. *Narration*, *Elaboration* and *Correction*. Logical forms in SDRT consist of a set of *labels* $\pi_1$, $\pi_2$, . . . that each represent a unit of discourse, and an *assignment function* $\mathcal{F}$ that associates each label $\pi$ with a logical form $\varphi$, representing the unit's interpretation; written $\mathcal{F}(\pi) = \varphi$ or $\pi : \varphi$. The content $\varphi$ assigned to a label can contain coherence relations among labels, and so the relations induce an *order* on the labels: say that $\pi_1$ *immediately outscopes* $\pi_2$ if $R(\pi_2, \pi_3)$ or $R(\pi_3, \pi_2)$ is a part of the formula $\varphi$ where $\mathcal{F}(\pi_1) = \varphi$. The *outscoping* relation on labels is the reflexive transitive closure of immediate outscoping. A *coherent* logical form—known as a Segmented Discourse Representation Structure or SDRS—is an assignment function where the outscoping order is a partial order with a unique maximal element ('root'). Put differently, a coherent discourse consists of a single segment of non-circularly connected subsegments.

   Cue phrases (e.g. *then, therefore, so, but*) can entail such coherence relations, but they frequently are left underspecified: in natural language, such cues might be ambiguous among different alternative coherence relations, their scope (i.e. which

segments are linked) may be ambiguous, or cue phrases may be entirely absent. In these cases, the precise discourse structure must be inferred via commonsense reasoning about both linguistic and non-linguistic information. Consider, for instance, (140). Simplifying somewhat (ignoring presuppositions, tense and so on), (140′) disambiguates (140) to an interpretation that is equivalent to having *so* connecting the clauses; (140″) is equivalent to *because*. Without a connective, the interpretation of (140) is ambiguous.

(140) The meeting was cancelled. {So | Because} Nicholas stayed at home.

$$\pi_1 : (\iota x)(meeting(x) \wedge cancel(e_1, x))$$
$$\pi_2 : (\iota y)(stay(e_2, n) \wedge home(y) \wedge at(e_2, y))$$

(140′) $\pi_0 : Result(\pi_1, \pi_2)$

(140″) $\pi_0 : Explanation(\pi_1, \pi_2)$

SDRSs are assigned a dynamic semantics, where one starts to unpack its content from its (unique) root label. The dynamic semantics of a coherence relation $R(\pi_1, \pi_2)$ is defined in terms of its arguments' contents (i.e. $\mathcal{F}(\pi_1)$ and $\mathcal{F}(\pi_2)$).

For instance, the general rubric for *veridical* relations like *Explanation* and *Result* is given in (141), where $\mathcal{C}$ and $\mathcal{C}'$ are the contexts of interpretation (typically, sets of world-assignment pairs), $\wedge$ corresponds to *dynamic conjunction* (and so $[\![\varphi \wedge \psi]\!] = [\![\varphi]\!] \circ [\![\psi]\!]$), and $\varphi_{R(\pi_1, \pi_2)}$ is content that is specific to the coherence relation $R$ and is specified in terms of $\mathcal{F}(\pi_1)$ and $\mathcal{F}(\pi_2)$:

(141) $\mathcal{C}[\![R(\pi_1, \pi_2)]\!]\mathcal{C}'$ iff $\mathcal{C}[\![\mathcal{F}(\pi_1) \wedge \mathcal{F}(\pi_2) \wedge \varphi_{R(\pi_1, \pi_2)}]\!]\mathcal{C}'$.

**Glue Logic, Logical Form and MDC.** The dynamic semantics capture how to evaluate a *representation* of the discourse against the real world or a model. How this representation is *constructed* is another problem. This task is carried out in a separate but related *glue logic* that uses *defeasible* principles to construct logical form (Asher & Lascarides (2003) provide detailed motivation for this separation).

The glue logic consists of *default axioms* that model how commonsense reasoning with both linguistic and non-linguistic information yield (defeasible) inferences about which available unit(s) in the context the current unit connects to, which coherence relations connect them, and how other semantic elements that are left underspecified by linguistic form get resolved to specific values (e.g. anaphora or relative semantic scope of presuppositions). The glue logic thus reasons over *underspecified logical forms* (ULFs), which in turn express *partial descriptions* of fully specific logical forms (SDRSs). The default axioms thus support defeasible inferences from ULFs to more specific and pragmatically preferred forms (that may still be underspecified).

For instance, pronouns introduce a condition $x = ?$, which means that $x$ must be co-referent with an available antecedent, but it is left underspecified *exactly*

*which* antecedent. More generally, a ULF uses a variable ? of an appropriate sort whenever a value needs to be resolved to form a (fully specific) SDRS. This is in particular the case for coherence relations in examples like (140). Thus, for instance, the ULF expression $\lambda :?(\alpha, \beta)$ means that $\beta$ is connected to $\alpha$, forming part of the segment labelled by $\lambda$, but the value of their coherence relation is not known. I will somewhat abuse notation on ? and only sometimes indicate the type of the variable that is underspecified in subscript.

The glue logic's *default axioms* are expressed with a defeasible conditional $>$ ($\varphi > \psi$ means *If $\varphi$ then normally $\psi$*; also see Chapter 3). For example, IQAP is a glue-logic default axiom which stipulates that normally, a response to a question is an answer.

$$(\text{IQAP}) \ (\lambda :?(\alpha, \beta) \wedge \text{interrogative}(\alpha) \wedge \text{spk}(\alpha) \neq \text{spk}(\beta)) > \lambda : IQAP(\alpha, \beta)$$

In words, if $\beta$ is connected to $\alpha$ but it is not (yet) resolved by what coherence relation, $\alpha$ is an interrogative and $\alpha$ and $\beta$ are said by different people, then normally, the relation between them is *Indirect Question Answer Pair* (IQAP).[11] Another general default axiom stipulates that when the contents associated with $\alpha$ and $\beta$ satisfy the necessary semantic consequences of $R(\alpha, \beta)$, then normally they connect with $R$ (Asher & Lascarides, 2003, p403).

The conditional $>$ defines a nonmonotonic proof theory $\mathrel{\vert\!\sim}_G$ (Asher & Lascarides, 2003, ch5), which validates a number of intuitively compelling patterns of defeasible inference, such as defeasible *modus ponens* ($\varphi, \varphi > \psi \mathrel{\vert\!\sim}_G \psi$, but not $\varphi, \varphi > \psi, \neg\psi \mathrel{\vert\!\sim}_G \psi$) and Specificity (If $\varphi \vdash_G \varphi'$ then $\varphi, \varphi > \neg\psi, \varphi' > \psi \mathrel{\vert\!\sim}_G \neg\psi$). The glue logic proof theory $\mathrel{\vert\!\sim}_G$ is then used to define Discourse Update:

**6.4.1.** DEFINITION (Update). Let $\Gamma$ be a ULF for the discourse context and $\pi : \mathcal{K}$ a ULF representing new information. Then $update(\Gamma, \pi : \mathcal{K})$ is the set of all (and only) those SDRSs that satisfy the glue logic consequences of attaching $\mathcal{K}$ to some available segment $\alpha$ in $\Gamma$. More formally: $K \in update(\Gamma, \pi : \mathcal{K})$ iff $K$ is an SDRS and there is an available segment $\alpha$ in $\Gamma$ where for all formulae $\varphi$

$$\text{If } \Gamma, \pi : \mathcal{K}, \lambda :?(\alpha, \pi) \mathrel{\vert\!\sim}_g \varphi, \text{ then } K \models_g \varphi.$$

That is, SDRT computes a *set* of potential interpretations of the current (underspecified) information about the discourse. This set contains all fully specified SDRSs that model the overt information (i.e. literal semantic content as well as cue phrases like *because* or *so*) and all its defeasible consequences.

However, some SDRSs in *update* are better than others: they are *more coherent* and thus *preferred* interpretations of the discourse (Asher & Lascarides, 2003, p233). Thus, SDRT's central principle for discourse interpretation is this: one (always) interprets the discourse in a way that *Maximises Discourse Coherence*

---

[11]SDRT has it that a direct answer is in particular (also) an indirect answer.

(MDC). For instance, Asher & Lascarides (1998) use MDC to predict when a presupposition gets locally accommodated: in some contexts, MDC overrides the default that presuppositions are accommodated globally, when it is more coherent to accommodate locally.

As seen multiple times in the prior discussion, discourse coherence is a matter of *degree* (or preference), so potential interpretations are *ranked* according to their coherence. Here is a semi-formal unpacking of the principle MDC:

---

(VII) **Maximise Discourse Coherence (MDC).**

An SDRS $K$ is at least as coherent as an SDRS $K'$, $K' \leq^c K$, if and only if all of the following hold:

1. *Prefer consistency:* If $K'$ is consistent, then so is $K$.

2. *Prefer rich structure:* $K$ has at least as many coherence relations as $K'$.

3. *Prefer better relations:* For every rhetorical relation $R(\pi_1, \pi_2)$ that $K'$ and $K$ share: $R(\pi_1, \pi_2)$ is at least as coherent in $K$ as it is in $K'$.

4. *Prefer resolution:* $K$ resolves at least as many underspecifications and anaphora as $K'$ does.

5. *Prefer flat structure:* $K$ has at most as many labels as $K'$ unless $K'$ has a *semantic clash* and $K$ does not.[12]

---

While MDC constrains interpretations to be maximally coherent, it does not entail that there is a *unique* maximally coherent interpretation, even in context.

Regarding point 3, note that I have talked of the *Contrast* connection as varying in quality. Plainly, some contrasts are better than others. This goes as follows: a *Contrast*$(\pi_1, \pi_2)$ means that the contents of $\pi_1$ and $\pi_2$ are *structurally similar* but *semantically dissimilar*. That is, it is required for *Contrast*$(\pi_1, \pi_2)$ that one can compute a partial, structural isomorphism between $K_{\pi_1}$ and $K_{\pi_2}$. To judge the quality of that *Contrast*, one then judges the semantic dissimilarity of the contents; the more dissimilar, the better the contrast. Semantic dissimilarity can be cashed out in two principal ways (Asher, 1999; Asher & Lascarides, 2003): (i) if one content defies an expectation of the other (e.g. if it defeats some commonsense consequence

---

[12]The SDRS (142c) is an example of a semantic clash: the content of $\pi_2$ is entailed by *Parallel* but hypothesized by *If*(-then); (142b) is thus more coherent (despite having more labels).

(142)  a. $\pi_1$: If a shepherd goes to the mountains,

   $\pi_2$: he normally brings his dog.

   $\pi_3$: He brings a good walking stick too.

  b. $\pi_0$: *If*$(\pi_1, \pi)$

   $\pi$ : *Parallel*$(\pi_2, \pi_3)$

  c. $\pi_0$: *If*$(\pi_1, \pi_2) \wedge$ *Parallel*$(\pi_2, \pi_3)$

as in (143a)); or (ii) if the isomorphic parts are syntactically outscoped by modifiers with different polarity (e.g. *does* vs. *does not* in (143b) or *loves* vs. *hates* in (143c)).

(143)   a.   I invited Michael, but he did not come.

      b.   Paula does live in Paris, but Jessica does not.

      c.   John loves Mary, but Paula hates her.

Monotonic entailment between two contents counts as semantic similarity and is thus incompatible with *Contrast*; at the same time, *Contrast* is veridical—so outright *contradictory* contents are not permitted either. That is, neither *p, but p* nor *not p, but p* are contrasting. Thus, the best possible contrasts would be (i) commonsense defeaters, or (ii) *is/is not* pairs.

The static glue logic from Asher & Lascarides (2003) does not axiomatise MDC, and so the maximally coherent SDRSs in $update(\Gamma, \mathcal{K})$ are selected by an extra-logical ordering procedure that implements MDC. In later work, Asher & Lascarides (2011) have developed a *dynamic* glue logic that is expressive enough to phrase MDC as an axiom. For simplicity, I remain here with the classical, static version where MDC is an extra-logical ordering procedure.

**Presuppositions.**   Following van der Sandt (1992), SDRT assumes that the grammar derives an ULF in which proffered content is separated from presupposed content and their relative semantic scope is underspecified. For instance, the presupposition trigger *regret* yields the logical form for (144a) given in (144b):

(144)   a.   A man didn't regret smoking.

      b.   proffered:      $\pi_1 : \exists x(man(x) \wedge \neg regret(e, x, {}^{\wedge}smoke(e', x))$

          presupposed: $\pi_2 : smoke(e', x)$

Asher & Lascarides (1998) stipulate glue logic axioms that put constraints on how the presupposed and proffered contents coherently relate to their context (and each other). These axioms ensure the following: (i) an utterance's presuppositions update the discourse before its proffered content; and (ii) these axioms make presuppositions (defeasibly) *bind* to a prior unit over coherently accommodating it. Similar defaults do not apply to proffered content.

For example, simplifying somewhat, the following ULFs are generated for the two sentences in (145a) (where presupposed content is marked with $\partial$).

(145)   a.   A man had a health scare. But he didn't regret smoking.

      b.   $\pi_1 : \exists x \exists y(man(x) \wedge \textit{health-scare}(y) \wedge have(e_1, y, x))$

      c.   $\pi_2 : Contrast(\pi, \pi_3) \wedge \pi =?$

        $\pi_3 : \neg regret(e_2, z, {}^{\wedge}smoke(e_3, z))$

        $\pi_4 : \partial smoke(e_3, z) \wedge z =?$

      d.   $\pi_2 : Background(\pi_1, \pi_4) \wedge Contrast(\pi_1, \pi_3)$

The label $\pi_1$ in (145a) records the (fully specified) content of *A man had a health scare*. Then, *but he didn't regret smoking* resolves to the three underspecified logical forms in (145b). Sentence-initial *but* introduces the *Contrast* relation in $\pi_2$ whose first argument is not yet resolved ($\pi =?$). The pronoun *he* introduces the variable $z$ in $\pi_3$ and $\pi_4$.

Then, given number and gender constraints (which are omitted here), the only candidate for resolving $z =?$ is $z = x$. Thus, since by MDC resolving an underspecification is more coherent than not, a maximally coherent interpretation must attach $\pi_3$ and $\pi_4$ to $\pi_1$ in a way that makes $x$ available.

First, the presupposition $\pi_4$ must be resolved and attached to $\pi_1$. Both $\pi_1$ and $\pi_4$ describe *states*, and so the glue logic axioms validate a (defeasible) inference that they connect as $Background(\pi_1, \pi_4)$.

Then, the underspecification $\pi$ from the *Contrast* segment is resolved to $\pi_1$. This makes it possible to resolve $z = x$ in $\pi_3$ and forms a better quality *Contrast* (and a flatter structure) than any alternative. For instance, SDRT update also countenances that there is a distinct segment $\pi_5 : Background(\pi_1, \pi_4)$ and that $\pi_2$ resolves to $\pi_2 : Contrast(\pi_5, \pi_3)$. However, this structure has one more label (i.e. it is less flat). Moreover, the *Contrast* in (145d) is rather good, since after having a health scare, one is expected to regret not taking care of one's health (at least according to some commonsense knowledge) and $\pi_3$ defies that expectation.

So the final SDRS is as shown in (145d): this correctly entails that smoking occurred, even though *smoking* was syntactically outscoped by *not*.

**Cognitive Modelling Logic.** SDRT's glue logic inferences are additionally aided and constrained by information about the participants' *cognitive states*. Following Hamblin (1971), SDRT makes a speaker $S$ who conveys a message $\varphi$ publicly committed to $\varphi$ (Asher & Lascarides, 2008; Lascarides & Asher, 2009); see also Chapter 3. Now, the *cognitive modelling logic* (CML)[13] features some familiar modal operators (KD45 operators for belief, $\mathcal{B}_S$ for a speaker $S$; and KD45 operators for public commitment, $C_S$) as well as higher-order operators for intentions ($I_S$). For reasons of exposition I remain here with these operators instead of embedding the SDRT CML in my revised commitment framework from Chapter 3; however, see below for some discussion on how the two models can be combined. In the CML, one can formalise some basic pragmatic principles:[14]

**Sincerity.** (a) $C_S\varphi > \mathcal{B}_S\varphi$.
      (b) $\mathcal{B}_S\neg\varphi > \neg I_S C_S\varphi$.

    Normally, you believe what you commit to; and you do not intend to commit to what you do not believe.

---

[13]The name is ambiguous. It is not a logic for the modelling of cognitive processes and it makes no claim to cognitive plausibility. Rather, it is a logic for modelling *abstractly* that interlocutors keep track of each other's cognitive states.

[14]The precise axioms vary across the SDRT literature. The ones presented here are a synthesis of the ones used by Asher & Lascarides (2013) and Schlöder & Lascarides (2015).

**Intention Transfer.** $C_S\varphi > C_S I_S C_H \varphi$.

>   Normally, you intend to make your commitments shared.

**Cooperativity.** $C_S I_S \varphi > I_H \varphi$.

>   Normally, intentions are kept aligned.

**Sincere Questions.** (a) $\mathsf{interrogative}(\varphi) \to (C_S\varphi > \neg\mathcal{B}_S\mathsf{resolved}(\varphi))$.
(b) $\mathsf{interrogative}(\varphi) \to (\mathcal{B}_S\mathsf{resolved}(\varphi) > \neg I_S C_S \varphi)$.

>   Normally, questions sincerely ask for unknown information.

**Hindsight Logic.**   In earlier, unpublished work, Alex Lascarides attempted to model intonational meaning as communications about how an *earlier* utterance has led to a *revision* of the cognitive model *in hindsight*. In our joint work (Schlöder & Lascarides, 2015), we give the formal details of such a model. The semantics for prosody developed in this chapter largely supersedes this earlier paper. However, the *hindsight logic* remains useful to model *some* of the meaning of a tune.

To allow for hindsight inferences in the cognitive modelling logic, we add complex operators $[s_S(\pi)]$ ('after $S$ uttered the discourse unit $\pi$') and $[s_S(\pi)]^{-1}$ ('before $S$ uttered $\pi$'). These do not make the logic dynamic. These operators operators are not *updates*, but function as *accesses* to different states (at different times) in the cognitive model. See Chapter 3 for further discussion on why the commitment model remains static.

For instance, these complex operators can formally express *surprise*: the term $[s_H(\pi)]^{-1}\mathcal{B}_S\neg I_H C_H \mathcal{K}_\pi$ expresses that *before* H's speech act $\pi$, S thought that H didn't intend to commit to $\mathcal{K}_\pi$. That is, H's action defied S's prior expectations. The following axioms regiment how these operators facilitate hindsight inferences (Schlöder & Lascarides, 2015).

**Commitment.** Let $\pi_1 \ldots \pi_n$ be elementary discourse units spoken by $S_1 \ldots S_n$, and $\Gamma_n$ be the context after $\pi_n$ (i.e. their ULFs plus salient facts and axioms). Let $\vdash_G$, $\vdash\!\!\!\sim_G$ be the monotonic and nonmonotonic proof theories of the glue logic. Let $\vdash_C$ and $\vdash\!\!\!\sim_C$ be the ones for the cognitive modelling logic.
>   – If $\Gamma_n \vdash_G \varphi$, then $\Gamma_n \vdash_C [s_{S_1}(\pi_1)] \ldots [s_{S_n}(\pi_n)] C_{S_n} \varphi$.
>   – If $\Gamma_n \vdash\!\!\!\sim_G \varphi$, then $\Gamma_n \vdash\!\!\!\sim_C [s_{S_1}(\pi_1)] \ldots [s_{S_n}(\pi_n)] C_{S_n} \varphi$.

**Persistence.** If $\Gamma \vdash\!\!\!\sim_C C_A \varphi$ and $A \neq S$, then $\Gamma \vdash\!\!\!\sim_C [s_S(\pi)] C_A \varphi$.

>   A person's public commitments are unaffected by other speakers.

**Conservativity.** $([s_S(\pi)]\mathcal{B}_{S'}\varphi) \to (\mathcal{B}_{S'}\varphi \lor \mathcal{B}_{S'}((C_S K_\pi) > \varphi))$.

>   Beliefs after an utterance are either carried over from before, or are inferred from that utterance.

**Hindsight.** If $\Gamma_n \vdash\!\!\!\sim_C [s_{S_1}(\pi_1)] \ldots [s_{S_n}(\pi_n)] \mathcal{B}_S [s_{S_i}(\pi_i)]^{-1}\varphi$,
then $\Gamma_n \vdash\!\!\!\sim_C [s_{S_1}(\pi_1)] \ldots [s_{S_{i-1}}(\pi_{i-1})] \mathcal{B}_S \varphi$
Beliefs about 'before'-operators cancel up to a corresponding 'after'-operator.

**Foresight.** $(\mathcal{B}_{S'}[s_S(\pi)]\varphi) > ([s_S(\pi)]\mathcal{B}_{S'}\varphi)$.

     If a speaker believes that after the act $\pi$, the proposition $\varphi$ holds, they normally have that belief in foresight.

These axioms ensure that inferences about previous dialogue states work correctly. Note that the context $\Gamma_n$ in (Hindsight) does not change. The logic models hindsight inferences that interlocutors can make *about* previous cognitive states *from* their current knowledge $\Gamma_n$, which *extends* their prior knowledge $\Gamma_{i-1}$. That is, the axiom (Hindsight) may apply in $\Gamma_n$, despite $\Gamma_{i-1} \not\mid\sim_C [s_{S_1}(\pi)]\ldots[s_{S_{i-1}}(\pi_{i-1})]\mathcal{B}_S\varphi$. Thus a speaker A can derive that their interlocutor Bob must have held a particular belief earlier on *without* being aware of that belief at the time. For instance, Bob's contribution in (146), if it expresses surprise, triggers the following inference:

(146)  $\alpha$: Amy: I want you to come to the ball with me.

      $\pi$: Bob:  Really?!

     $[s_A(\alpha)][s_B(\pi)]C_B[s_A(\alpha)]^{-1}\mathcal{B}_B\neg I_A C_A \mathcal{K}_\alpha$.

$\mid\sim$  $[s_A(\alpha)][s_B(\pi)]\mathcal{B}_B[s_A(\alpha)]^{-1}\mathcal{B}_B\neg I_A C_A \mathcal{K}_\alpha$ (Sincerity a).

$\mid\sim$  $\mathcal{B}_B\mathcal{B}_B\neg I_A C_A \mathcal{K}_\alpha$ (Hindsight).

$\mid\sim$  $\mathcal{B}_B\neg I_A C_A \mathcal{K}_\alpha$ (Belief modal).

$\rightsquigarrow$  "At the beginning of the dialogue, Bob thought Amy wouldn't say that!"

Thus, if Amy performs this inference, she learns something about what Bob thought *before* she uttered (146$\alpha$). Expanding the CML to hindsight inferences is straightforward and does not change how it is used: any conclusion drawn in the basic CML can also be drawn in hindsight logic when prefixed with $[s_{S_1}(\pi_1)]...[s_{S_n}(\pi_n)]$ where $\pi_n$ is the latest move in the dialogue.

    To all intents and purposes, it suffices to describe the proof theory of the CML. However, Asher & Lascarides (2003) also assign a Kripkean model theory to the CML. To also include hindsight inferences in this model theory, it must be changed from a *single* commitment structure to a *tree* that represents the prior moves (and their resultant commitment structures) and the anticipated future moves in the dialogue. Thus, one can take the *Tree Models* from Chapter 3 and *label* each transition in a tree with the discourse segment that effected the transition. Then, the complex operators $[\alpha]$ and $[\alpha]^{-1}$ are *accesses* to past nodes. That is, $M, x, w \models [\alpha]^{-1}\varphi$ if $M, y, v \models \varphi$ where $(y, v)$ is the node before the transition with $\alpha$ and similar for $[\alpha]$.

    As described in Chapter 3, a tree model is a *static representation* of one particular state in the dialogue, including past states and future projections. So, now, every time SDRT-update is performed, one computes a new commitment state—but also *a new tree model* to maintain the hindsight logic axioms. It would not suffice to of just compute the state corresponding to the current discourse interpretation, since *all accessible* information might need to be *enriched* with the

hindsight inferences stemming from the current information; not doing so might invalidate Conservativity and Hindsight (on temporally earlier nodes) or Foresight (on temporally later nodes).

The upshot is this: in Chapter 3 I did not give a procedure to update Tree Models, but SDRT-update as applied to the CML and hindsight logic is such a procedure. *Prima facie*, this seem infeasible since Tree Models can be vast (infinite, even). However, given that $\Gamma_n$ is finite, it suffices to compute a *minimal* tree that contains the information in $\Gamma_n$ as closed under the CML axioms. This tree will have finitely many branches of finite length.

## 6.5   Formalised Account and Worked Examples

**Formalised account.**   Definition 6.5.1 formalises the Focus Semantics (I, II′).

**6.5.1.** DEFINITION (Focus Semantics).   Let $\langle \varphi, f \rangle$ be the foreground–background pair of the current utterance. The discourse update associated with $\langle \varphi, f \rangle$ is an update with $\pi^b : \partial \mathcal{K}_{\pi^b}$, $\pi^f : \mathcal{K}_{\pi^f}$ and $\pi : R^{tune}(\pi^b, \pi^f)$, where:

- $\mathcal{K}_{\pi^f}$ is the ULF corresponding to $\varphi(f)$.

- If the tune is falling, then $R^{tune} = Continuation \vee Elaboration$ and $\mathcal{K}_{\pi^b}$ is the ULF corresponding to $\varphi(x)$, where $x$ is free in $\varphi$ and of the same type as $f$, and the semantic index (an eventuality term) of $\varphi$ is syntactically distinct in $\mathcal{K}_{\pi^b}$ and $\mathcal{K}_{\pi^f}$ (although they can denote the same eventuality).

- If the tune is fall–rise, then $R^{tune} = Contrast$ and $\mathcal{K}_{\pi^b}$ is the ULF corresponding to $?_{mod}(\varphi(x))$, where $x$ is free in $\varphi$ and of the same type as $f$, and $?_{mod}$ underspecifies *modality*—that is, it can resolve to $\top$, $\Diamond$ or $\neg$. Again, the semantic index of $\varphi$ is syntactically (but not necessarily semantically) distinct in $\mathcal{K}_{\pi^b}$ and $\mathcal{K}_{\pi^f}$.

- Via SDRT-update and MDC compute the segment $\alpha^r$ of the prior discourse where $\mathcal{K}_{\pi^f}$ most coherently attaches (the *relevant segment*). If there are multiple such segments let $\alpha^r$ be the minimal segment (w.r.t. the order induced by outscoping) that outscopes all most coherent attachment points.

Givenness (Rule IV) corresponds to updating the relevant segment with the (presupposed) background content, which in SDRT is the maximally coherent interpretation of their combination. Accordingly, the necessary condition for interest (Rule V) is simply that the foreground is not entailed by this.

**6.5.2.** DEFINITION (Givenness and Interest).   The *given information* is what is inferable from updating the relevant segment $\alpha^r$ with the (presupposed) ULF $\pi^b : \partial \mathcal{K}_{\pi^b}$ representing the background content (in a maximally coherent way).

Write $\Gamma \restriction \alpha^r$ for the subset of $\Gamma$ that contains $\alpha^r$ and all segments outscoped by $\alpha^r$ (i.e. the sub-SDRS of $\Gamma$ whose root is $\alpha^r$). Then, $\varphi$ is *given* iff for every

maximally coherent SDRS $K \in update(\Gamma \upharpoonright \alpha^r, \pi^b : \mathcal{K}_{\pi^b})$, $K \models_g \varphi$. And the proffered content $\mathcal{K}_{\pi^f}$ (i.e. $\varphi(f)$) is *interesting* only if it is not given. That is, only if there is a maximally coherent $K \in update(\Gamma \upharpoonright \alpha^r, \pi^b : \mathcal{K}_{\pi^b})$ such that $K \not\models_g \mathcal{K}_{\pi^f}$.

The cognitive effects of the tunes (Rule VI) is expressed in SDRT's cognitive logic. The constraint that the foreground of a falling tune is *informative* is formalised as the speaker's belief that this content is not yet mutually accepted. The cognitive contribution of fall–rise as follows: the attachment point $\alpha^r$ is taken to be the utterance that the current speaker S is marking as triggering a change in S's beliefs about H. The hindsight formula expresses that before $\alpha^r$, S did not believe that H wouldn't know $\mathcal{K}_{\pi^f}$, but that $\alpha^r$ changed that belief.[15]

**6.5.3.** DEFINITION (Tune Semantics). The different tunes impose the following restrictions in the cognitive modelling logic, where cg stands for *it is common ground that* (Asher & Lascarides, 2008) and $S$ and $H$ are, respectively, speaker and hearer of the current utterance (labelled $\pi$ with foreground content labelled $\pi^f$).

**fall:** $C^{tune} = \mathcal{B}_S \neg \mathrm{cg}(K_{\pi^f})$.

**fall–rise:** $C^{tune} = C_S\Big(([s_H(\alpha^r)]^{-1} \neg \mathcal{B}_S \neg \mathcal{B}_H \mathcal{K}_{\pi^f}) \wedge (\mathcal{B}_S \neg \mathcal{B}_H \mathcal{K}_{\pi^f})\Big)$.

Definition 6.5.4 combines these as specified in Section 6.3:

**6.5.4.** DEFINITION (Intonated Discourse Update). Let $\Gamma$ be the prior context and $\langle \varphi, f \rangle$ be the background–foreground pair of the current utterance.

   i. Compute the ULFs $\mathcal{K}_{\pi^f}$ and $\mathcal{K}_{\pi^b}$ as in Definition 6.5.1.

   ii. If the glue logic supports no coherence relation between $\mathcal{K}_{\pi^f}$ and an available unit $\alpha^r$ in $\Gamma$, break.

   iii. If the glue logic supports no coherence relation between $\mathcal{K}_{\pi^b}$ and $\alpha^r$, break.

   iv. If $\mathcal{K}_{\pi^f}$ is not *interesting*, (i.e. it is given according to Definition 6.5.2), break.

   v. Add $C^{tune}$ (Definition 6.5.3) to the cognitive modelling logic.

   vi. Do discourse update on $\Gamma$ with $\pi^b : \partial K_{\pi^b}$, $\pi^f : K_{\pi^f}$ and $R^{tune}(\pi^b, \pi^f)$.

The constraints in (ii–iv) can all be expressed as monotonic axioms in the glue logic: they are then effectively a part of the *lexical semantics of tunes*. One can then understand these semantics as affecting the standard SDRT-update (Definition 6.4.1) and 'intonated discourse update' as just the usual discourse update. Here, I remain with the procedure outlined in Definition 6.5.4 for the sake of clarity and simplicity.

---

[15]This logical form for fall–rise differs from the one of Schlöder & Lascarides (2015) in the wider scoping of the negations. This weakened form is more appropriate for interrogative antecedents $\alpha$; the earlier paper was only concerned with responses to assertions.

**Worked Examples**    First, I demonstrate how this formal semantics for intonation makes the right predictions for (@135ab) vs. (@135ab′), i.e. that the former is an agreement move, but the latter a disagreement move.

(@135)  a. Amy: Harry is the biggest liar in town.

b. Bob: The biggest <u>fool</u>$_H$, maybe.$_{LL\%}$

b.′ Bob: The biggest <u>fool</u>$_{L+H}$, maybe.$_{LH\%}$

Simplifying somewhat (treating "maybe" as equivalent to $\Diamond$, omitting tense information, treating subject pro-drop like a pronoun and considerably over-simplifying the semantics of "biggest X in town") the ULF of (@135a) is as in (147a) and the Focus Semantics (Definition 6.5.1) yields the ULFs (147b) and (147b′) for (@135b) and (@135b′):

(147) a.  $\alpha^r : liar(e_{\alpha^r}, h) \wedge biggest(e_{\alpha^r})$                          *Harry is the biggest liar.*

b.  $\pi \; : ?_{fall}(\pi^b, \pi^f)$

$\pi^b : \partial \Diamond (P(e'_\beta, x) \wedge biggest(e'_\beta)) \wedge P =?$          *he is maybe the biggest P.*

$\pi^f : \Diamond (fool(e_\beta, x) \wedge biggest(e_\beta))$                *he is maybe the biggest fool.*

b′. $\pi' : Contrast(\pi'^b, \pi'^f)$

$\pi'^b : \partial ?_{mod} \Diamond (P(e'_\beta, x) \wedge biggest(e'_\beta)) \wedge P =?$   *y(he is maybe the biggest P).*

$\pi'^f : \Diamond (fool(e_\beta, x) \wedge biggest(e_\beta))$                *he is maybe the biggest fool.*

In (147ab), one must first ensure that $\pi^b$ can coherently update the context $\alpha^r$ to yield given information that does not entail $\pi^f$. MDC prefers binding $P$ to the available antecedent *liar* (and $x$ to the available antecedent Harry) rather than resolving $P$ via existential quantification. Thus the background can be coherently attached, but does not violate the interest check, since *fool* is not part of the prior context (and thus cannot be given).

Then, resolving $P$ to *liar* validates the relation *Accept* between $\alpha^r$ and $\pi^b$. Then, the underspecified relation $?_{fall}$ that is introduced by the tune semantics resolves to *Continuation*: the glue logic axioms do not validate inferring *Elaboration* ('liar' and 'fool' are conventionally not in an entailment relation) but do validate *Continuation* (because $\pi^b$ and $\pi^f$ share a common topic that can be paraphrased as *what Harry is*). Since flat structures are preferred by MDC, the final discourse structure is:

$$\pi_0 : Accept(\alpha^r, \pi^b) \wedge Continuation(\pi^b, \pi^f).$$

The dynamic semantics of *Accept* and *Continuation* are veridical. Thus, they entail that Bob is committed to Harry being the biggest liar and (also) maybe the biggest fool. In addition, given the tune, the cognitive modelling logic derives $\mathcal{B}_B \neg cg(\mathcal{K}_{\pi^f})$: Bob believes that *Harry is maybe the biggest fool* is not (yet) common ground.

Now consider (@135ab′), where the ULFs are $\alpha^r$ and (147b′). Again start by computing how (just) the presupposed background content $\pi'^b$ would update $\alpha^r$.

As with (147ab), binding $P$ to the available antecedent *liar* and binding $x$ to Harry is preferred (via MDC). The underspecified modality $?_{mod}$ can resolve to $\top$, $\Diamond$ or $\neg$. Any of these resolutions produce a coherent update, and render the proffered content $\pi'^f$ not given. So the example passes the interest test. However, when updating the context with both $\pi'^b$ and $\pi'^f$ a clearly preferred resolution of the modality emerges.

For simplicity, assume that $\Diamond\Diamond\varphi$ is equivalent to $\Diamond\varphi$ and $\neg\Diamond$ is $\Box\neg$; so resolving $?_{mod}$ to $\top$ or to $\Diamond$ are equivalent, and resolving it to $\neg$ leads to a 'definitely not' reading. Then $?_{mod} = \neg$ maximises contrast, since saying what Harry is *vs.* what he is *not* yields a better contrast than contrasting two things that he is. So by MDC, $?_{mod}$ resolves to $\neg$. As a consequence, the contents of $\alpha^r$ and $\pi'^b$ satisfy the necessary consequences of *Correction* (for the latter entails the negation of the former). So SDRT's glue logic yields the logical form (148).

(148)   $\pi_0 : Correction(\alpha^r, \pi'^b) \wedge Contrast(\pi'^b, \pi'^f)$.

$\quad \alpha^r : liar(e_{\alpha^r}, h) \wedge biggest(e_{\alpha^r})$             *Harry is the biggest liar.*

$\quad \pi'^b : \neg\Diamond(liar(e_{\alpha^r}, h) \wedge biggest(e_{\alpha^r}))$    *Harry is (def.) not the biggest liar.*

$\quad \pi'^f : \Diamond(fool(e_\beta, h) \wedge biggest(e_\beta))$         *Harry is maybe the biggest fool.*

There is no commonsense knowledge that liars are not fools. However, the Tune Semantics entails that Bob considers these terms to be somewhat at odds. To wit, he is communicating that *if you say he is a liar, you do not think he's a fool.* This derivation in the cognitive modelling logic is as follows. First, the tune semantics in Definition 6.5.3 triggers a (hindsight) derivation from Bob's utterance to Bob's beliefs about Amy before Amy said (@135a):

$\vdash \quad [s_A(\alpha^r)][s_B(\pi)]C_B[s_A(\alpha^r)]^{-1}\neg\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi$. ($C^{fallrise}$).

$\vdash\!\!\!\sim \quad [s_A(\alpha^r)][s_B(\pi)]\mathcal{B}_B[s_A(\alpha^r)]^{-1}\neg\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi$. (Sincerity a).

$\vdash\!\!\!\sim \quad \mathcal{B}_B\neg\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi$. (Hindsight).

$\vdash\!\!\!\sim \quad \neg\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi$. ($\mathcal{B}_B$ is KD45).

Second, there is a derivation about how that belief changes, given Amy's move:

$\vdash \quad [s_A(\alpha^r)][s_B(\pi)](\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi)$. ($C^{fallrise}$) .

$\vdash\!\!\!\sim \quad [s_A(\alpha^r)]\big((\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi \vee \mathcal{B}_B(C_B\mathcal{K}_\pi > \neg\mathcal{B}_A\mathcal{K}_\pi)\big)$. (Conservativity).

$\vdash\!\!\!\sim \quad [s_A(\alpha^r)](\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi)$. ($\vee$-Elimination).
     The second disjunct is (normally) false, as $C_B\mathcal{K}_\pi \vdash\!\!\!\sim I_A C_A \mathcal{K}_\pi$ and so $C_B\mathcal{K}_\pi > \neg\mathcal{B}_A\mathcal{K}_\pi$ clashes with Sincerity b.

$\vdash\!\!\!\sim \quad (\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi \vee \mathcal{B}_B(C_A\mathcal{K}_{\alpha^r} > \neg\mathcal{B}_A\mathcal{K}_\pi)\big)$. (Conservativity).

$\vdash\!\!\!\sim \quad \mathcal{B}_B(C_A\mathcal{K}_{\alpha^r} > \neg\mathcal{B}_A\mathcal{K}_\pi)$ ($\vee$-Elimination, given the above derivation).

$\approx$   "That you told me $\alpha^r$ tells me that you do not think that $\pi$."

Thus, by way of intonation, Bob conveys a strong contrast between *liar* and *fool*: the above derivations show how the tune implicates that Bob believes a commitment to *Harry being a liar* normally entails you do not believe he is a fool. This additionally supports the *Contrast* relation in (148). This contrast is *absent* in (@135ab)—neither does the focus semantics entail *Contrast*, nor does the tune carry the 'surprise'-type perlocutionary effects that are associated with fall–rise intonation. Now, in (@134) the tune changes the perlocutionary effects but not the illocutionary ones.

(@134)  a.  Amy: You can't afford that.

b.  Bob: I'm a <u>millionaire</u>$_{\text{H·LL\%}}$

$\not\leadsto$ *I thought you knew this already*

b.′  Bob: I'm a <u>millionaire</u>$_{\text{L+H·LH\%}}$

$\leadsto$ *I thought you knew this already*

The glue logic (defeasibly) yields *Correction* in both (@134ab) and (@134ab′), on the grounds that the contents of the sentences satisfy *Correction*'s necessary consequences. The perlocutionary derivation is slightly different from the above, however. The hindsight derivations from the fall–rise tune is the same as above.

$$\mathrel{|\!\sim}\ \ (\neg\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi) \wedge (\mathcal{B}_B(C_A\mathcal{K}_{\alpha^r} > \neg\mathcal{B}_A\mathcal{K}_\pi)).$$

However in contrast to (@135), in (@134) the second conjunct is already given, since *cannot afford* and *millionaire* are, conventionally, contrary (i.e. $\Box(\mathcal{K}_\pi > \neg\mathcal{K}_{\alpha^r})$). Thus a commitment to *cannot afford* normally means that one does not believe *millionaire*: i.e. $C_A\mathcal{K}_{\alpha^r} > \neg\mathcal{B}_A\mathcal{K}_\pi$ is derivable independently of the tune. However, one can reason further with the first conjunct that is derivable from the tune, by including the fact that millionaires can afford things.

$\mathrel{|\!\sim}\ \neg\mathcal{B}_B\neg\mathcal{B}_A\mathcal{K}_\pi$ ($\wedge$-elimination from the above formula).

$\mathrel{|\!\sim}\ \neg\mathcal{B}_B\neg\mathcal{B}_A\neg\mathcal{K}_{\alpha^r}$. (by $\Box(\mathcal{K}_\pi > \neg\mathcal{K}_{\alpha^r})$)

$\mathrel{|\!\sim}\ \neg\mathcal{B}_B\neg\neg I_A C_A\mathcal{K}_{\alpha^r}$ (Sincerity b).

$\mathrel{|\!\sim}\ \neg\mathcal{B}_B I_A C_A\mathcal{K}_{\alpha^r}\ \ \ \approx$ "I didn't believe you would commit to $\alpha^r$."

Thus, the Tune Semantics entail the desired perlocutionary effect of the fall–rise intonation in (@134ab′). *Mutatis mutandis* this is also the prediction of made for (@133ab′), using *Sincere Questions (b)* instead of *Sincerity (b)*.

(@133)  a.  Amy: Did you read the first chapter?

b.′  Bob: I read the entire <u>dissertation</u>$_{\text{L+H·LH\%}}$

$\leadsto$ *I thought you knew this already.*

Finally, here is how my account predicts felicity judgements and implicatures in the context of *wh*-questions. First consider the example that seems to contradict question-answer congruence.

(@91)  a.  Harvey: Who likes Michael?

   b.′ Jessica: Nobody likes Michael<sub>L+H·LH%</sub>

Definition 6.5.1 yields the ULF of (@91b′) given by $\pi$, $\pi^b$ and $\pi^f$ below, and these must update Harvey's move (where the presupposition triggered by the *wh*-question is already resolved). Note in particular that this is a case where $\alpha^f$ and $\alpha^b$ are equally coherent attachment points for the proffered content $\pi^f$. The segment $\pi^f$ may attach equally well to $\alpha^f$ as an answer and to $\alpha^b$ as a Correction (or both simultaneously). Thus the *relevant segment* here is the supersegment $\alpha^r$:

(149)    $\alpha^r : Background(\alpha^f, \alpha^b).$

   $\alpha^f :?\lambda y.like(e_\alpha, y, m)$                    *Who likes Michael?*

   $\alpha^b : \exists y.like(e_\alpha, y, m)$                     *someone likes Michael*

   $\pi\ : Contrast(\pi^b, \pi^f).$

   $\pi^b : \partial?_{mod}(\neg\exists z.like(e_\beta, z, x)) \wedge x =?$          *y(nobody likes x)*

   $\pi^f : \neg\exists z.like(e_{\beta'}, z, m)$                 *nobody likes Michael*

First check that proffered content is not given (Definition 6.5.2). The glue logic predicts that $\pi^b$ *binds* to $\alpha^b$ (on the grounds that binding presuppositions is maximally coherent), thereby setting $?_{mod} = \neg$, and $x = m$. This resolution is entailed by $\alpha^b$ and is thus a coherent binding. This also satisfies the necessary condition for interest: $\mathcal{K}_{\pi^f}$ is not entailed by this result. But $x = m$ cannot be a part of the *final* update with $\pi^f$, whatever the resolution of $?_{mod}$: (150c) is inconsistent with $\pi^f$, and the resolutions in (150a) and (150b) are entailed by $\pi^f$ and therefore fail to establish a contrast.

(150)  a. $?_{mod} = \top \rightsquigarrow \pi^b$: *nobody likes Michael.*

   b. $?_{mod} = \Diamond \rightsquigarrow \pi^b$: *maybe nobody likes Michael.*

   c. $?_{mod} = \neg \rightsquigarrow \pi^b$: *somebody likes Michael.*

So when executing the update of the context with both the background and proffered contents, the final logical form must *accommodate x* via an existential quantifier (rather than binding it to $m$). In the usual ways MDC predicts that resolving $?_{mod}$ to $\neg$ maximises the contrast with $\pi^f$ (*someone likes someone (other than Michael), but nobody likes Michael*), to yield (151):

(151)    $\pi_0 : Background(\alpha^f, \pi^b) \wedge Contrast(\pi^b, \pi^f) \wedge$
          $Correction(\alpha^b, \pi^f) \wedge QAP(\alpha^f, \pi^f)$

   $\alpha^f :?\lambda y.like(e_\alpha, y, m)$                         *Who likes Michael?*

   $\alpha^b : \exists y.like(e_\alpha, y, m)$                          *someone likes Michael*

   $\pi^b : \exists x.\exists z.like(e_\alpha, z, x)$         *there is someone that somebody likes.*

   $\pi^f : \neg\exists z.like(e_\alpha, z, m)$                           *nobody likes Michael*

The discourse *structure* is one where Jessica is answering Harvey's question but correcting its presupposition. Note that, due to the resolution of $\pi^b$, Jessica tacitly acknowledges that there was a liking, but she denies that anyone likes Michael. While this sounds slightly odd for *liking*, the relevance of this acknowledgement becomes more apparent when considering a telic verb like *vote*. Replacing 'like' by 'vote' in (@91) yields the interpretation that Jessica acknowledges that *there was a vote* but that *nobody voted for Michael*.[16]

In addition, the cognitive contribution for fall–rise again derives something that can be paraphrased as *I did not think I would need to tell you this*, explaining why, intuitively, Jessica seems aghast at the idea of someone liking Michael.

Moreover, note that if (@91b$'$) were uttered with a falling tune, then the background ULF is $\partial\neg\exists y.like(y,x)$, and the most coherent way to interpret this sets $x = m$, so that it corrects $\alpha^b$. But this entails the proffered content, and so by the necessary condition for interest (Definition 6.5.2) it is anomalous.

In (@89b$'$), however, both tunes are infelicitous. I start with the falling tune.

(@89)     a.  Harvey: Who likes Michael?

　　　　　#b.$'$ Jessica: Rachel likes $\underline{\text{Michael}}_{H\cdot LL\%}$

As before, the question (@89a) yields the SDRS rooted at $\alpha^r$ in (152); (@89b$'$) yields the following ULFs labelled by $\pi$, $\pi^b$ and $\pi^f$:

(152)   $\alpha^r$ :  $Background(\alpha^b, \alpha^f)$

　　　　$\alpha^b$ :  $\exists y.like(e_\alpha, y, m)$

　　　　$\alpha^f$ :  $?\lambda y.like(e_\alpha, y, m)$

　　　　$\pi$ :  $?_{fall}(\pi^b, \pi^f)$

　　　　$\pi^b$ :  $\partial like(e'_\beta, r, x) \wedge x =?$

　　　　$\pi^f$ :  $like(e_\beta, r, m)$

As before, the interest test amounts to examining the most coherent way of interpreting $\pi^b$ (alone) in the context of $\alpha^r$, and checking that this interpretation is coherent and does not entail $\pi^f$. Here, the most coherent update resolves $z$ to $m$, as this can attach as an answer to $\alpha^f$. So the interest test fails: the most coherent interpretation of the background entails the proffered content. Thus, the intonation is anomalous.

If Jessica's response in (@89) has a fall–rise tune, then the ULFs are as above except that $\pi^b$ features the underspecified modality $?_{mod}$ and *Contrast* replaces $?_{fall}$. The most coherent update of $\alpha^r$ with $\pi^b$ still resolves $z$ to $m$ and $?_{mod}$ to $\top$ because this results in an answer to (@89a), whereas $?_{mod} = \neg$ or $?_{mod} = \Diamond$ provide only *partial* answers (MDC makes complete answers more coherent). Thus, by failing the interest test, it is also predicted to be anomalous.

---

[16]I choose 'like' for the examples in the discussion, because it does not carry any uniqueness presupposition; in contrast, 'X voted for Y' usually entails that X did not vote for anyone else.

in flection is so-          -o   inte re   st   i   ng
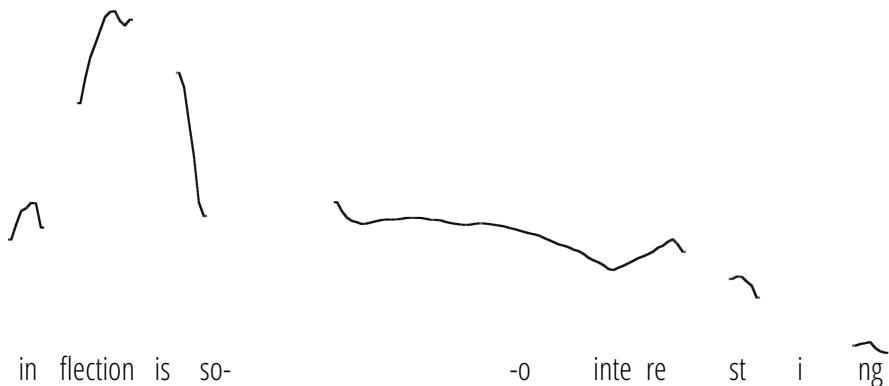
Figure 6.1: Tune of example (@153).

## 6.6 Verbal Irony

In this section I demonstrate that the framework developed in this chapter is flexible enough to be extended to further tunes and prosodic features. I present semantics for a tune signalling *verbal irony*.

**An Ironic Tune.**  Empirical data suggests that verbal irony is linked to *prosodic contrasts* (Bryant, 2010). Here, I consider one tune realising such a contrast: a *sustained low tune* with a steep fall and a final fall, which robustly leads to ironic readings. I will annotate this tune as ↓ LL%, but see Figure 6.1 for a precise contour. Example (@153) with the tune in Figure 6.1 is an example of an ironic utterance with this tune (note in particular the stretched vowel in 'so').[17]

(@153)  a. Abed: Inflection is <u>so-o</u>$_\downarrow$ interesting.$_{LL\%}$

$\rightsquigarrow$ *inflection is very uninteresting.*

The sustained low tune can make the difference between acceptance and rejection. (154) is one such case.

(154)  a. Amy: Are you going to Mike's show tonight?

b. Bob:  I'll <u>definitely</u>$_H$ go to that.$_{LL\%}$ $\rightsquigarrow$ *will go*

b.′ Bob:  I'll <u>de-efinitely</u>$_\downarrow$ go to that.$_{LL\%}$ $\rightsquigarrow$ *won't go*

Tune is the only variable that distinguishes (154b) from (154b′). So the semantics for the sustained low tune should predict such ironic readings. Note that I do not claim that the tune in (@153) and (154b′) is the *only* tune that marks irony—or indeed that irony requires *any* particular intonation (cf. Bryant & Fox Tree, 2005). I only claim that this tune is *one* particular way of cueing irony.

---

[17]This example is from the TV show *Community*.

**Implicit Negation.**    There are many competing explanations of verbal irony, but they—by and large—revolve around a common intuition: that the speaker of an ironic utterance means *the opposite* or *the inverse* of the literal content of their utterance. On a Gricean account, verbal irony is the flouting of the Maxim of Quality, i.e. the speaker asserts *something recognisably false* and therefore means the opposite; on the *echoic* account an ironic utterance *mentions* a sentence and indicates *dissent* from it (Sperber & Wilson, 1981); on the *joint pretence* account, an ironic utterance invites one's interlocutors to consider a situation in which the utterance would be true and notice how absurd this situation is (Kumon-Nakamura *et al.*, 1995); my own earlier work (Schlöder & Lascarides, 2015) followed Robert Martin (1992) and Rachel Giora (1995) in considering irony to be a form of *implicit negation*.

However, it is not sufficient to just add a negation to the ironic utterance's propositional content. The negation of the literal content of (154b′) is *it is not the case that Bob will definitely go to the show* which resolves to *Bob might not go to the show*. The correct reading is however that *Bob will (definitely) not go*. Giora (1995) can potentially account for this: she makes the (implicit) negations stemming from irony be subject to a preference for *contrary* negation over *contradictory* negation (Horn, 1989, chs4–5). This is in particular realised as a preference for narrow over wide scoped negation: *definitely not* is contrary to *definitely*, but *not definitely* is merely contradictory to *definitely*.

This preference, however, is ambiguous. There are multiple different possibilities to form a contrary negation. In the case of (154b′), at least two such readings are available.

(155)  a.  Bob will definitely not go to Mike's show.

   b.  Bob will definitely go somewhere that is not Mike's show.

Both (155a) and (155b) are contrary to the literal content of (154b′). However, (155b) overstates what Bob expresses in (154b′) because it entails that Bob will go somewhere—but this does not seem to be part of the meaning of (154b′). It is unclear what privileges (155a) over (155b) (or other options) if all we have is a general preference for contrariness.

While there might be ways to spell this out, there is another option. The *position of the fall* in the sustained low tune seems to indicate the placement of the negation. For instance, (155b) would be the appropriate interpretation for a third possible answer to (154a).

(156)  a.  Amy: Are you going to Mike's show tonight?

   b.″ Bob: Yeah, I'll definitely go to <u>tha-at</u>$_\downarrow$.LL%

(156b″) is felicitous in a context where there is a salient alternative activity that Bob could attend and the context moreover suggests that Bob would prefer this one (e.g. because there is a much better show overlapping with Mike's).

**Implicit Contrary Negation.** I model this as follows. Just as the fall–rise tune adds a negation in the background, the sustained low tune adds a *contrary negation in the foreground.*

**6.6.1.** DEFINITION (Contrary Negation). Let $\sim$ be a meta-operator (i.e. an operator on logical forms; its application is computed when logical form is constructed) that specifies *contrary negation*:

– if $f$ is a modal or quantifier, $\sim f$ is $f\neg$.

– if $f$ is on a scale, $\sim f$ is an item from the opposite end of the scale;

– if $f$ is a bivalent predicate, then $\sim f$ is $\neg f$;

– if $f$ is an entity, then $\sim f$ is a meta-variable such that for any predicate $P$, $P(\sim f) = \sim P(f)$.

Then, the Focus Semantics for the ↓ LL% tune can be put as follows.

---

(VIII) **Focus Semantics** (sustained low tune with steep fall)

The focus placement (i.e. position of the steep fall) separates an utterance into a *foreground f* and a *background $\varphi$* with free variable $x$ of the same type as $f$. Updating a discourse with an utterance that has a sustained low tune with fall on $f$ proceeds as follows:

– Update with the *presupposition $\varphi$*.

– Update with the *proffered* content $(\lambda x.\varphi)(\sim f)$ (and all its presuppositions) such that the proffered content *contrasts* with the presupposition.

---

This rule can be regimented in SDRT in the same way as the Focus Semantics for the falling and fall–rise tunes; the operator $\sim$ is an operator in the glue logic that constructs logical form.

By specifying the relative scope of the contrary negation $\sim$ this semantics avoids ambiguities. I show this here for (154b′) and (156b″): (In the logical forms I again simplify or ignore a number of ancillary details, including tense, possessive case, and the presuppositions associated with proper names.)

(154)  a.  Amy: Are you going to Mike's show tonight?

presupp: $\exists s.\mathtt{of}(m, s) \wedge \mathtt{show}(s)$

proffers: $?\mathtt{go}(b, s)$

b.′  Bob: Yeah, I'll <u>de-efinitely</u>↓ go to that.<sub>LL%</sub>

presupp: $\exists x_{aux}.x(\mathtt{go}(b, s))$

proffers: $\sim \Box \mathtt{go}(b, s) \equiv \Box \neg \mathtt{go}(b, s)$

(156)   b.″ Bob: Yeah, I'll definitely go to <u>tha-at</u>$_\downarrow$.$_{\text{LL}\%}$

        presupp: $\exists x_{entity}.\square\mathsf{go}(b, x)$

        proffers: $\square\mathsf{go}(b, \sim s) \equiv \square\neg\mathsf{go}(b, s)$

Thus, the proffered contents of (154b′) and (156b″) are the same: both are—by way of irony—negative answers to Amy's question in (154a). But while in (154b′) the presupposition effected by the Focus Semantics (VIII) is a tautology (because for any $p$ there is a modality $\nabla$ such that $\nabla p$), the presupposition of (156b″) entails that Bob is going somewhere—just not to Mike's show. This is precisely the difference between the two competing contrary negations in (155).

    Finally, also note that the necessary condition for interest comes to bear here as well. It rules out the anomalous (157b).

(157)     a. Jeff:   Inflection is not interesting at all.

     #b. Abed: Inflection is <u>so-o</u>$_\downarrow$ interesting.$_{\text{LL}\%}$

The Focus Semantics for $\downarrow$ LL% backgrounds a modifier on *interesting*; the most coherent interpretation binds this to the antecedent modifier *not at all* which is contrary to the literal interpretation of *so*—i.e. equivalent to $\sim so$. Thus, the foreground is given and (157b) is ruled out by the necessary condition for interest.

## 6.7   Conclusion

To model cases where the difference between agreement and disagreement is (merely) the tune with which an utterance is intonated, I have here developed a novel theory of intonational meaning. This theory is not a narrow explanation of the few cases that led me to study this phenomenon, but a general and programmatic approach to the study of intonation and focus.

    The prime novelty of this account is that it models focus and tune jointly: both the *placement* of the stress and the *tune* that sets the stress influence the Focus Semantics. I have argued that doing so is necessary, as theorising about focus without tune leads to confusion. In particular, if one considers at least the falling and the fall–rise tune, a question–answer congruence account is insufficient, as seen in examples (@91), (@93), (@100) and (@101). To resolve this shortcoming, I have proposed a new notion of *givenness* to then weaken congruence to a *necessary condition for interest*: that what is presented as interesting cannot also be presented as given. This condition is buttressed on independently motivated principles of discourse coherence that are also used to model anaphora and presupposition. In stating the necessary condition, I do not propose a model theory of interest—not even a definition of it. Rather, the necessary condition is one small fragment of that notion that seems to be formalisable and thus predictable.

    While there surely is ambiguity and confusion in mapping a raw acoustic signal to a specific tune (Calhoun, 2007), there need not be ambiguity in mapping

a specific tune to a logical form. In particular the fall–rise tune often conveys content indirectly, but what exactly is implicated varies radically from one context to another. This variation is not modelled as an ambiguity (as different fall–rises with different semantics), but captured in that the Focus Semantics underspecify certain semantic elements. These are then resolved to specific values via reasoning about discourse coherence and its interaction with compositional and lexical semantics, real world knowledge and cognitive states—in particular, the distinct Tune Semantics that communicate something about the speaker's cognitive state.

I have demonstrated the flexibility of this approach by sketching how it applies to a tune signalling verbal irony. However, the theory I offer here is merely a small step towards a new, general and *comprehensive* account of focus and intonation. In particular, I leave open how one would model tunes that in a discretising annotation system like ToBI would be annotated with multiple / pre-nuclear accents. Nonetheless, I contend that combining focus with tune is a challenge that the field of formal semantics *needs* to take up, even if not in the way I have proposed here. However, I do hope that the work presented here may serve a first step—towards a complete account of intonation that models tune jointly with focus by exploiting principles that link coherence, cognitive modelling and presupposition.

# Chapter 7

# Why-Questions

Sometimes, speakers neither outright agree nor outright disagree with a proposal put forward by their interlocutors. Instead, they may prompt their interlocutor to back up their proposal with reasons by asking a *Why*-question. As mentioned in Chapter 3, someone who undertakes a commitment thereby displays that they can vindicate their commitment. Such vindication may be prompted with *Why*-questions. Thus, a conceptually sound analysis of *Why*-questions is required to properly understand the notion of commitment. The analysis of *Why* questions, however, is difficult, due to their open-ended nature. As a matter of fact, speakers can disagree on whether something is a proper answer to a *Why*-question.

In this chapter, I present an analysis of *Why*-questions that models them as eliciting a premiss of an *enthymeme* (an incomplete argument) that is correct according to some non-explicit *topos* (a general reasoning scheme). This acknowledges and explains the open-endedness of *Why*: the answer space of a *Why*-question is constrained only by the contextually available topoi, and speakers may have different opinions on which topoi are contextually available. I then adapt the model from Chapter 6 to explain the phenomenon of focus in *Why*-questions.

This chapter is based on, and draws from, *Why?* (co-authored with Ellen Breitholtz and Raquel Fernández) and further unpublished collaborative work with Ellen Breitholtz and Raquel Fernández.

## Chapter Contents

# 7.1   Reasons, Enthymemes, Topoi

For example: Aunt Minnie is in the hospital. Why? Because she slipped on the ice and broke her hip. That satisfies people—but it wouldn't satisfy someone who came from another planet and knew nothing about it. First you need to understand why when you break your hip you go to the hospital. Well, because her husband has seen that she had her hip broken and called the hospital up, and they sent someone to get her. All this is understood by people.

So when you are explaining a why, you have to be in some framework such that you allow something to be true. Otherwise you are perpetually asking why. Why did the husband call up the hospital? Because a husband is interested in his wife's welfare. But not always: some husbands aren't interested in their wife's welfare (when they're drunk and they're angry). Why did she slip on the ice? Well ice is slippery, everybody knows that, no problem. But then you ask, why is ice slippery?
—Richard Feynman on *Fun to Imagine (BBC)*

Nardole: Why is she here?
The Doctor: Because she isn't anywhere else.
—*Doctor Who* Series 10, Episode 2

In this section, I briefly introduce the *micro-rhetorical* framework that has been developed by Ellen Breitholtz in her dissertation and subsequent papers. Based on this framework, I give a definition of what it means to give a *reason* for something. The analysis of *Why*-questions in the following sections rests on this definition.

**Giving Reasons.**   Participants in a dialogue reason about certain propositions and circumstances. On one hand, interlocutors are generally expected to back up the assertions they make with arguments when prompted. On the other hand, a relevant utterance is made *for* a reason, e.g. by relating to information pertinent to the purpose of the dialogue. Such reasons are not always explicit, but can be elicited with *Why*-questions (Jackson & Jacobs, 1980; Breitholtz, 2010; Schlöder & Fernández, 2015a). The following dialogues exemplify the basic phenomenon.

(158) a. Brenda: He's in hospital.

     b. Carla:   Why?

     c. Brenda: Because he's not very well.                    (BNC, KBF, 3394–3396)

(159) a. Gillian: Do you want mum to come to Argos with me tomorrow morning?

          (three lines omitted)

     b. Robert: Why are you asking me?

     c. Gillian: Cos you said you'd come to Argos with me. (BNC, KC8, 191–196)

In (158), Brenda makes an assertion and Carla asks for a reason that backs the truth of the *proposition* expressed in (158a); note that this may (but need not)

entail that Carla is doubting the content of Brenda's assertion. Contrast this with (159). There, Gillian asks a question and Robert inquires about Gillian's reason for doing so. In both cases, the initial speaker then supplies a *reason* that is marked with the particle 'because'. On a commitment account of assertion, committing to a proposition (by way of asserting it) entails a commitment to vindicating this proposition;[1] then, the reply (158c) *vindicates* the commitment from (158a). If one also models questions as the undertaking of commitments (as, e.g. Asher & Lascarides, 2003), the same can be said about (159c).

Saying that *Why*-questions elicit reasons, however, is hardly explanatory unless there is a tangible definition of what counts as a reason—and, in particular, of what counts as a *good* reason. Certainly, the answers given by Brenda and Gillian in the initial examples (158) and (159) are not arbitrary. Not *any* utterance would be an acceptable answer to the *Why*-questions in these examples and, moreover, speakers may disagree about whether some particular response is acceptable. Similarly, not every utterance that expresses *q because p* is immediately acceptable to its addressee. To provide a suitable formal analysis of reasons, I make use of the micro-rhetorical approach of Ellen Breitholtz (2014).

**Enthymemes and Topoi.** Observe that the arguments expressed by the first and third utterances in (158) and (159) are logically incomplete: they indicate that the third utterance *is* a reason for the first, but not what *warrants* this connection. In an Aristotelian fashion, one may call such arguments *enthymemes.*

An enthymeme is an argument of the form *p hence q* (write $p \therefore q$) which requires the listener to supply one or more underpinning premises. They are recognised as correct due to a known, salient *topos* that expresses a general pattern of accepted reasoning that in particular licenses $p \therefore q$. That is, enthymemes consist of two parts, a premiss and a conclusion. This distinguishes them from syllogisms (or formal arguments in general) which have *three* parts: a specific statement (minor premiss), a general law (major premiss) and a conclusion. In an enthymeme, there is no general (major) premiss.

It has been observed that enthymematic reasoning is widespread in natural dialogue, and it has been linked to clarification and cognitive load management (Jackson & Jacobs, 1980; Breitholtz & Villing, 2008; Breitholtz, 2014). These prior observations are not directly about *Why* questions, but they provide a compelling and useful framework. Consider the example in (160).

(160)  Amy: Let's walk along Walnut Street. It's shorter.  (Walker, 1996b)

This utterance may be made in the context of two colleagues on their way to work, where several routes are possible. Amy suggests to take one of them and provides

---

[1]Ellen Breitholtz points out to me that such vindications might be developed *post hoc*: committing to being able to vindicate something is possible without precisely knowing how one would go about this. Rather, one commits to being willing to engage in an *activity* of vindication that may include longer discussion and *ad hoc* arguments.

a reason supporting this. Breitholtz (2014) proposes the following analysis. The two propositions convey the enthymeme in (161).

(161)      It (Walnut Street) is shorter
           ————————————————————————
           ∴ Let's walk along Walnut Street

In this case, Amy counts on her interlocutor being able to supply something that underpins (160). These kinds of underpinnings (i.e. general principles of accepted reasoning) are often referred to as *topoi* in the literature on rhetoric and argumentation. On this view, speakers have access to a vast set of topoi that derives from their commonsense knowledge, everyday experiences, truisms, folk wisdom *etc* (Ducrot, 1980, 1988). Some topoi may be applied to various subjects, while others are specific to a particular subject (Anscombre, 1995).[2]

However, the topoi accessible to one individual do not constitute a logical calculus. A set of topoi may contain outright contradictions or principles of inference which jointly entail contradictions (Anscombre, 1995). Crucially, therefore, topoi *tolerate exceptions.* As in the epigraph, *husbands care about their wives* is a commonsense topos, but does not always hold for all husband-wife pairs. Moreover, clearly, speakers might draw upon different topoi to license an enthymeme and might differ on which topoi they judge applicable in the situation at hand; thus they can disagree about whether some enthymeme is correct—or about *why* it is.

**What Reasons Are.**    A topos that could be drawn upon to license the argument in (160) could be something like *if a route is shorter (than other options), it is preferred.* This topos would underpin the enthymeme (161). However, it is not clear why to interpret (160) one would try to license an enthymeme in the first place. In terms of discourse relations, the intuitive reading of (160) is *Explanation*:[3] *it's shorter* explains why the speaker suggests to *walk along Walnut Street.* But to *infer* the discourse relation of *Explanation*, one must already know that *it's shorter* can explain (in some sense) the suggestion to *walk along Walnut street* (Asher & Lascarides, 2003, p159). Otherwise, we may also interpret *it's shorter* as merely a commentary on or continuation of *let's walk along Walnut street.*

Thus, the topos *if a route is shorter (than other options), it is preferred* does double duty: it is licensing the enthymeme in (161) and it is *also* the reason why (160) is interpreted as conveying that enthymeme. Thus, I define *reason* as follows.

---

(I) A **reason** for $q$ is some proposition $p$ such that making $p$ and $q$ salient makes salient a topos that licenses the enthymeme $p \therefore q$.

---

[2]The publications of Ducrot and Anscombre are not readily available in English; I thank Ellen Breitholtz for communicating to me some relevant arguments and results from these works.

[3]SDRT distinguishes *Explanation* and *Explanation\**; the former explains propositional contents, the latter explains the making of certain speech acts. So, technically, *Explanation\** is correct here. For further discussion on this, see below.

Then, (160) is explained as follows: the utterance makes salient both the suggestion to *walk along Walnut street* and that *Walnut street is shorter (than other salient alternatives)*; the joint saliency of these makes salient topoi related to route-selection and short routes—among them in particular *shorter routes are preferable.* Thus, the conditions for inferring *Explanation* are satisfied, which leads to the enthymeme (161).[4] This step would not be required if it were linguistically cued (e.g. with *because*) that the correct discourse relation is *Explanation.*

The double duties of topoi can diverge. That is, an interlocutor can infer an *Explanation* relation between two discourse segments due to some salient topos, but deny that this topos is correct (or that it is appropriate in the particular current circumstances). The following example is a case in point. The second speaker explicitly mentions the principle that he takes to back the conditional statement in (162a), but denies that principle.[5]

(162)   a. Donald: I'm self-funding my campaign, I tell the truth.

       b. John:    'I'm rich, therefore I tell the truth' has [. . . ] no cause and effect between the two.      (*Last Week Tonight*, Feb. 29th, 2016)

The explication of the topos in (162b) suggests that John has interpreted Donald's utterance as the enthymeme in (163) and immediately denies that this is correct.

(163)     Donald self-funds his campaign
       ∴ Donald tells the truth
   Topos: *rich people tell the truth*

An important consequence of the definition in (I) is that the acceptability of a given reason (e.g. in response to a *Why*-question) does not depend on an inferential relationship being *valid* (in some objective sense, e.g. logical validity), but merely on it being subjectively *acceptable* to the interlocutors (according to some made-salient topos). Moreover, example (162) demonstrates that made salient topoi are not unique. An alternative reading would be the enthymeme (164).

(164)     Donald self-funds his campaign
       ∴ Donald tells the truth
   Topos: *people who are not in anyone's pocket tell the truth*

Thus, what topos is made salient depends on the individual and their rhetorical

---

[4]Formally, this is executed as follows. The necessary condition for *Explanation*$(\alpha, \beta)$ is that $\beta$ explains $\alpha$. In the present discussion this means that the content of $\beta$ is a reason for the content of $\alpha$. If an appropriate topos for $\beta \therefore \alpha$ is salient, this is the case. Now, a central principle of discourse interpretation in SDRT is that, all else being equal, satisfaction of the necessary conditions for a discourse relation is *sufficient* to infer that relation. Thus, *Explanation*$(\alpha, \beta)$, is inferred whenever $\beta$ is a reason for $\alpha$ and nothing else blocks this inference.

[5]This example is from the TV show *Last Week Tonight*. (162a) is an excerpt from a speech by Donald Trump; (162b) is John Oliver's commentary.

resources. As my analysis will show, topoi are dynamic and negotiable: they can
be accommodated, elicited and themselves be discussed.

Moreover, if there is no explicit cueing, it may remain underspecified which
part of an utterance states a reason and which part states what one gives a reason
for. Consider (165), due to Moore & Pollack (1992).

(165)  a.  Bush supports big business—he's sure to veto House Bill 1711.

        b.  Bush supports big business—[thus] he's sure to veto House Bill 1711.

        c.  Bush supports big business—[after all] he's sure to veto House Bill 1711.

As Moore & Pollack point out, (165a) is ambiguous. It may indicate that *Bush
supports big business* explains why (the speaker holds the belief that) *Bush is sure
to veto House Bill 1711* (as cued in (165b)) or that *Bush is sure to veto House Bill
1711* explains (or evinces) why (the speaker holds the belief that) *Bush supports
big business* (as cued in (165c)). The ambiguity can be resolved when considering
what is common ground: if we already accepted as fact that *Bush supports big
business*, I can leverage this fact to explain (my belief that) *Bush is sure to veto
House Bill 1711*, and *vice versa*.

However, neither proposition needs to be accepted fact for (165a) to be
felicitously uttered: I can present both *Bush supports big business* and *he is
sure to veto House Bill 1711* as new information and moreover present one as
evidence for the other (ambiguously, without further context). Now note that in
*any* situation where (165) is uttered, even a politically ignorant speaker can infer
that House Bill 1711 is somehow opposed by big business.

The definition in (I) accounts for this as follows: making salient that *Bush
supports big business* and that *Bush sure to veto House Bill 1711* makes salient
the topos *presidents veto (exactly) the bills that go against their interests*. This
topos supports both readings of (165a) and therefore underspecifies what explains
what. But both possible applications of the topos entail that the specific House
Bill 1711 is something that goes against Bush's interests, with these interests
specified as *supporting big business*. Thus, (165a) implicates that *House Bill 1711
undermines big business*.

**Formalisation.**   For the purposes of applying this account in SDRT and my
extensions of its cognitive modelling logic (Chapters 3 and 6), I formalise this
account in a default logic.

---

(II) The **general scheme for topoi** is this:

$$\varphi > (\mathcal{A} \to \mathcal{B}),$$

where $\mathcal{A}$ and $\mathcal{B}$ are underspecified logical forms, $\varphi$ is a description of the
circumstances in which the topos applies, and $>$ is a default conditional.

---

Then, whenever we are in a $\varphi$-situation, an inference from *any* full specification of $\mathcal{A}$ to *any* full specification of $\mathcal{B}$ is a correct enthymeme according to the topos $\varphi > (\mathcal{A} \to \mathcal{B})$.

Such topoi can be added to the world knowledge which is used to construct logical form in the glue logic (see Chapter 6, Section 6.4). This permits that some topoi contradict each other, e.g. when there are active topoi $\varphi > (\mathcal{A} \to \mathcal{B})$ and $\varphi > (\mathcal{A}' \to \mathcal{B}')$ that apply in the same circumstances $\varphi$, and there are full specifications of $\mathcal{A}$ and $\mathcal{A}'$ that are mutually consistent, but full specifications of $\mathcal{B}$, $\mathcal{B}'$ that are not. In such cases, *both* inferences are considered defeated (Asher & Lascarides, 2003, p190). Thus, this situation would not trivialise the logic.

As any world knowledge, it only affects the construction of logical form if it is salient (or accessible) in the discourse (Asher & Lascarides, 2003, ch3). However, I do not formalise saliency here. In any case, it should stand to reason that in a situation matching $\varphi$, if $A$ and $B$ are full specifications of $\mathcal{A}$ and $\mathcal{B}$, respectively, then $\varphi > (\mathcal{A} \to \mathcal{B})$ is salient. Any formal account of saliency and world knowledge, i.e. an account that formally explains how salient facts are retrieved from world knowledge, should do that job.

Note that when the micro-rhetoric framework is thusly integrated into SDRT, a separate definition of *enthymeme* is not required. The enthymeme $A \therefore B$ corresponds to the discourse structure (166).

(166)   $\alpha : A$

      $\beta : B$

      $\lambda : Explanation(\beta, \alpha).$

Also, conversely, if $\varphi > (\mathcal{A} \to \mathcal{B})$ is an active premiss in the glue logic, $\varphi$ is the case and $A$ and $B$ instantiate $\mathcal{A}$ and $\mathcal{B}$, respectively, the glue logic can infer *Explanation*. Thus, all the pieces of the framework can be expressed in SDRT. For reasons of exposition, however, I remain with the semi-formal terminology of enthymemes and topoi to analyse the data.

Moreover, note that since $>$ is closed on the right (Asher & Lascarides, 2003, p189), topoi can be composed to license complex inferences. If $\varphi > (\mathcal{A} \to \mathcal{B})$ and $\varphi > (\mathcal{B} \to \mathcal{C})$ are active topoi, then $\varphi > (\mathcal{A} \to \mathcal{C})$ is too. This allows simple topoi to compose to license complex enthymemes.

Breitholtz (2014) models this differently. She takes both enthymemes and topoi to be functions from *situations* to *situation types*. An enthymeme $p \therefore q$ expresses that in a situation satisfying $p$, $q$ holds, i.e. for all situations $s$, if $s$ is a $p$-situation, then $s$ is also a $q$-situation. Topoi have the same form and license enthymemes as follows: if from the function representing the topos one can compute the function representing the enthymeme, then the enthymeme is licensed by the topos. In the typical case of a topos being a general principle, this computation would be to *restrict* the topos to the specific situation in which the enthymeme is used. Since functions can be composed, topoi can be composed on this account as well.

# 7.2   Responding to *Why*

Based on the definition in (I), I now give a definition of what makes a good answer to a *Why*-question. I discuss the initial consequences of this definition by analysing some instructive examples.

**Reasons and *Why*.**   Given the definitions from the previous section, it is now straightforward to define how *Why*-questions are interpreted.

---

(III)   **Discourse Semantics for *Why*-questions** (first version)

An **answer to** *Why q?* is an utterance proffering $p$ such that $p$ is a reason for $q$.

---

That is, a *Why*-question *elicits an enthymeme.* Hence, what counts as a satisfactory answer to a *Why*-question crucially depends on what topoi both interlocutors are willing to accept in the circumstances of the discourse. That is, as also pointed out in the epigraph to Section 7.1, the process of resolving *Why*-questions rests on what kinds of explanations the interlocutors are willing to take for granted.

Thus, for instance, (158) is analysed as follows: the propositions *he's in hospital* and *he's not very well* jointly make salient that *people who are unwell go to hospital.* This then licenses that *he's not very well* ∴ *he's in hospital.* Note in particular that not all—probably not even many—people who are unwell go to hospital. Hence, this topos is not generally and universally available. In the constructed case (167), B seems rather rash in her inference.

(167)    a. Amy: John is not very well.

?b. Bob:  Oh, he is in hospital?

This means in particular that *he's not very well* is alone not sufficient to make the topos *unwell people go to hospital* salient. As specified in (I), it is both *he's in hospital* and *he's not very well* that *jointly* make that topos salient.

A similar explanation can be given for (159), if we represent *the fact that Gillian asked the question* as a proposition that can be queried by *Why*-question as defined in (III).

(159)   a. Gillian: Do you want mum to come to Argos with me tomorrow morning?

(three lines omitted)

b. Robert: Why are you asking me?

c. Gillian: Cos you said you'd come to Argos with me.

An appropriate topos seems to be *people want to have a say in activities they partake in.*

Note that my definition of what makes an answer to a *Why*-question differs from accounts that analyse *Why*-questions as they occur in scientific (or otherwise

rigorous) inquiry. On such accounts, an answer to a *Why*-question is an explanans for some explanandum; on the covering theorem account (Hempel, 1965; Hintikka & Halonen, 1995) an explanans is an interpolant in a derivation of the explanandum from the established theory (containing laws and empirical facts). This means that the explanans logically entails the explanandum (from the background theory).

However, my goal here is different: to analyse the *linguistic* phenomena associated with *Why*-questions. Thus, it is not directly relevant here what is a good explanans (or how one would formalise that notion), but only what interlocutors would generally accept as good answers. In particular, the collection of topoi that interlocutors draw from may not be a coherent theory (some topoi might outright contradict each other), but the scientific background theory in an account of explanans/explanandum is assumed to be coherent.

Nonetheless, one may attempt to model systematic inquiry as a special case of general dialogue. Note that it varies with context which topoi are acceptable or available. Then, it seems reasonable to assume that in contexts of rigorous inquiry the only acceptable topoi are those that express modes of inference satisfying a (discipline-dependent) notion of explanation. This echoes Bas van Fraassen's insight that even in rigorous inquiry, one must consider the context to interpret a *Why*-question; e.g. physicists and engineers, when presented the same explanandum, may differ entirely on what counts as an explanans (van Fraassen, 1980, ch5). On my account, this contextual influence is reflected in which topoi are acceptable.

**Topoi.**   In fact, which topoi are available (and what is made salient) varies with the speakers. Consider (168).

(168)  a. Jessie: I roasted it and we couldn't eat it on the Sunday and

     b. Anon: Could not? Why could you not eat it?

     c. Jessie: That was bull beef.

     d. Anon: Oh right.        (BNC, K65, 284–299; backchannel utterances omitted)

In (168c), Jessie gives an answer to a *Why*-question, i.e. Jessie gives what she construes to be a reason for *being unable to eat the roast*. Speaker A indicates that he accepts this as an answer, but it may be rather obscure to outsiders what makes (168c) an answer to (168a). The analysis is aided by Harry's elaboration:

(168)  a. Jessie: I roasted it and we couldn't eat it on the Sunday and

     b. Anon: Could not? Why could you not eat it?

     c. Jessie: That was bull beef.

     d. Anon: Oh right.

     e. Harry: our second class beef, you see.

     f. Jessie: Then I, I put it in a saucepan and I stewed it the next day

                          (BNC, K65, 284–299; backchannel utterances omitted)

The additional information in (168e,f) suggests the following enthymeme:

(169)    $\underline{x \text{ is bull beef}}$

   $\therefore$ Jessie could not eat roasted $x$

  Topos: *tough meat needs to be stewed*

The dialogue in (168) suggests that what *makes* (168c) an answer to (168b) is the more general statement indicated in (168e,f). In particular, to someone who is unaware of the topos of (169), answering (168c) to (168b) would seem like a *non sequitur*.

Note that in a discourse of the form *Why q? p!* it is already cued that *q* is supposed to be interpreted as a reason for *p*. So the problem of inferring *that* one needs to interpret the enthymeme *p* $\therefore$ *q* (as in (160) above) does not occur here. However, this also means that if an addressee cannot find or accommodate an appropriate topos to validate the cued discourse relation of *Explanation*, they may need to ask for clarification. An example is (170).

(170)  a. Anon:  Why didn't you record anything?

    (two lines omitted)

  b. Cassie: Er this is my third tape and you've recorded a whole side!

  c. Anon Yeah, so?

    (one line omitted)

  d. Cassie  Well you're supposed to record as much as you can.

               (BNC, KP4, 1738–1743)

Cassie is giving an answer to the *Why*-question in (170a) that Anon cannot validate or accommodate as an answer, as evinced by (170c). However, her answer in (170d) reveals Cassie's reasoning: she did not record anything, because she filled three and a half tapes already (the BNC was recorded on cassette tape decks) and she needs to be parsimonious now to record more later. Arguably, a very general topos like *to maximise use of a limited resource, one must ration it* applies here.

**Conditionals.** An interesting special case arises when a *Why*-question queries a reason for a *conditional* statement. Consider (171).

(171)  a. Dorothy:  If you feel cold you'd be dead.

  b. Christopher: Why?

  c. Dorothy:  You just are. Part of you being alive is that you're warm.

               (BNC, KBW, 11065–11068)

The utterance (171a) states an inferential relationship without giving grounds for the relationship. The *Why?* in (171b) asks for a reason for this relationship. If we understand (171a) to also express an enthymeme, this is as follows.

(172) $\dfrac{-}{\therefore (x \text{ is cold} \therefore x \text{ is dead})}$

Put differently, (171b) asks for a reason why *x is cold* is a reason for *x is dead*. Then, *Why?* elicits a backing for an already stated premiss–conclusion pair (i.e. it elicits a reason for the enthymeme *itself*). Simply put, asking for the grounds of a conditional statement elicits an underpinning premiss, but the application of this premiss is *itself* enthymematic. Thus, (171c) may be interpreted as follows.

(173) $\therefore \dfrac{\dfrac{\text{Living things are warm.}}{x \text{ is cold}}}{\therefore x \text{ is dead}}$

The topos licensing this enthymeme seems to be some loose instance of contrapositive reasoning.

This in particular serves to illustrate the fact that enthymemes can be nested: In principle, this situation would now license the elicitation of a backing to support the enthymeme in (173) again (and so on). It is a truism that one can always ask another *Why*-question; here, any further *Why*-question will add another level of nesting in the enthymeme currently under discussion.

**Achilles and the Tortoise.** Lewis Carroll (1895) presents a particularly confounding argument for how hard it is to justify even a simple inference—because one can always ask one more time what licenses an inference. In Carroll's tale, Achilles tries to convince the sceptical Tortoise of the inference (174).

(174) $\dfrac{A \wedge A \to B}{\therefore B}$

The Tortoise accepts all the premisses and asks *why it must accept the conclusion*. Achilles answers that *if A is true and (if A then B) is true, then B is true.*

(175) $\therefore \dfrac{\dfrac{(A \wedge (A \to B)) \to B}{A \wedge A \to B}}{\therefore B}$

The Tortoise again accepts the premisses, but asks why the conclusion must follow. Achilles nests the enthymeme once more.

(176) $\therefore \dfrac{\dfrac{((A \wedge (A \to B)) \to B) \to \big((A \wedge (A \to B)) \to B\big)}{(A \wedge (A \to B)) \to B}}{\therefore \dfrac{A \wedge A \to B}{\therefore B}}$

Clearly, nothing stops the Tortoise from asking *Why?* again (and again) and thus the enthymeme can be nested further *ad infinitum*.

By common consent, something went wrong here. But what? The en-thymematic account offers one potential answer: Achilles is continuing to provide further *facts* to the Tortoise. But an inference requires a *topos*—a general principle—to be licensed. The regress would stop once a topos is made salient that would license the nested enthymeme. Achilles provides further facts that make salient a topos corresponding to *modus ponens*. However, the Tortoise clearly does not recognise *modus ponens*—and did not do so to begin with. Thus, providing further facts that form enthymemes with this topos will not help Achilles to defend his original point.

**Resistance.** As mentioned before, sometimes *Why*-questions can be a form of *resistance* (Bledin & Rawlins, 2016). That is, if one speaker asserts that *p*, another can ask *Why p?* without thereby agreeing to *p*. Rather, the purpose of the *Why*-question in these cases is to elicit further information which the speaker requires *before* agreeing (or disagreeing) (Schlöder & Fernández, 2015a).

This can be appreciated as follows. Sometimes, the rejection of a premiss in an enthymeme entails that the *conclusion* of that enthymeme does not count as mutually accepted. A case in point where a *Why* question seems to defeat an earlier assertion is (177).

(177)   a. Carla:    Got the junior tap and the senior tap.

b. Brenda: Yeah but you'll get that next year again.

c. Carla:    Why?

d. Brenda: Because you got honours didn't you? In grade three. ⟨pause⟩

e. Carla:    No cos junior tap was for grade three.

f. Brenda: Have you done grade four tap?          (BNC, KBF, 12258–12264)

Brenda makes an assertion in (177b) that is not immediately acceptable to Carla, so she *resists* by asking for a reason in (177c). Brenda supplies a reason in (177d), completing the enthymeme in (178).

(178)   $\underline{\text{C got grade three honours}}$
∴ C will get junior and senior tap

Then, in (177e), Carla denies that this is a correct inference. Apparently Brenda concedes this: instead of arguing the point of (178) she is looking for a *different* premiss that would allow her to infer the conclusion of (178). This evinces that the proposition asserted in (177b) is still not accepted by Carla.

This use of *Why*-questions can be explained in a commitment account of assertion (see Chapter 3): Say that a commitment to *p necessarily requires* the willingness and ability to vindicate *p*. If one cannot vindicate one's commitments (because, e.g. the reason one had in mind is defeated), then the original commitment is defective and must be retracted. This explains the discourse in (177).

## 7.3 Attachment in Discourse

In some of the examples above, *Why q?* is not asked discourse initially, but *q* is part of the prior context. That is, in a theory of coherence relations, these *Why*-questions *relate* to this prior context. I now explain how this works.

**Propositional and Meta-Discursive *Why*-Questions.** To explain how *Why*-questions relate to a discourse, SDRT has a discourse relation *Explanation$_q$*. This relation connects an antecedent to a question asking for a reason for this antecedent. That is, *Explanation$_q$*$(\alpha, \beta)$ and *Question-Answer-Pair*$(\beta, \gamma)$ if and only if *Explanation*$(\alpha, \gamma)$ (Asher & Lascarides, 2003, p351).

This is, however, not quite right yet. As mentioned already, *Why*-questions can inquire about propositions as well as about speech acts.

(179) a. Richard: They'll check every single doctor.

b. Anon    Why is that?    (BNC, KSV, 2979–2980)

(159) a. Gillian: Do you want mum to come to Argos with me tomorrow morning?

b. Robert: Why are you asking me?

In (159b), R is asking for G's reason for *making this particular speech act*; I call this the *meta-discursive* use of *Why*-questions. In contrast, the *Why*-question in (179b) inquires about the *content* of the speech act in (179a). This latter reading apparently only arises for propositional content, so I will refer to this as the *propositional* use.

This is straightforward to model. SDRT offers a *Explanation\** relation that is a pendant to *Explanation*, but is used in cases like (180) where one explains why one makes a particular speech act.

(180) a. Amy: Are you alright?

b. Amy: [I'm asking because] you seem a bit off today.

Thus to model cases like (159), another discourse relation *Explanation$_q^*$* is required that corresponds to *Explanation\**. In cases like (179b) and (159b) where it is explicit whether the *Why*-question is propositional or meta-discursive, it is straightforward to decide whether the *Why*-question attaches as *Explanation$_q$* or *Explanation$_q^*$*. However, the matter is more complicated with bare *Why?*.

**Bare *Why?*** It seems that assertive antecedents generally result in a propositional reading, but question antecedents a meta-discursive one.

(158) a. Brenda: He's in hospital.

b. Carla:    Why?

[*Why is he in hospital?*    ~~*Why are you telling me?*~~]

c. Brenda: Because he's not very well

(181)   a. Anon: Do you love me ⟨unclear⟩?

     b. Bnon: Why?

       [*Why are you asking?*]

     c. Anon: ⟨unclear⟩ I love you so much.                     (BNC, KCM, 1057–1060)

The meta-discursive reading also arises with imperative mood antecedents (that express a command), as in (182).

(182)   a. Richard: Put it on the door handle.

     b. Anon:    Why?

       [*Why are you telling me to do this?*]

     c. Richard: Cos I'm gonna use it tomorrow.                 (BNC, KSV, 1930–1932)

However, there are cases where bare *Why?* is ambiguous. Consider (183).[6]

(183)   a. Amy: I'll have you know that I'm upset.

     b. Bob:  Why?

       [*Why are you upset?* OR *Why are you saying that?*]

     c. Amy: I had a terrible day at work.

     c.′ Amy: So you be careful around me today.

The bare *Why?* in (183b) seems to be ambiguous between propositional and meta-discursive readings. This is evinced by the fact that both (183c) and (183c′) are *prima facie* permissible answers; (183c) takes up the propositional reading, whereas (183c′) takes up the meta-discursive one.

**Resolving the Ambiguity.**   (183) suggests that overt speech act marking is significant here. Compare (183) with (184).

(184)    a. Amy: I'm upset!

     b. Bob:  Why?

       [*Why are you upset?* ~~OR~~ ~~*Why are you saying that?*~~]

     c. Amy: I had a terrible day at work.

   #c.′ Amy: So you be careful around me today.

The answer in (184c′) does not seem to be acceptable here. Thus, the meta-discursive reading is not available. Hence, evidently, the *I'll have you know* in (183a) makes the crucial difference. It seems that it makes the speech act being performed available as an antecedent for *Why?*.

    Now, I account for this by specifying *preferences* about how bare *Why?* attaches: all else being equal, propositional antecedents yield propositional readings

---

[6]I thank Robin Cooper for helpful discussion on this example.

(IVa); and non-propositional antecedents yield meta-discursive readings (IVb). The glue axioms in (IV) express these preferences such that they overlap in (183).

---

(IV) **Glue Axioms for *Why?***

a. $(\lambda : R(\alpha, \pi) \wedge (R = Explanation_q \vee R = Explanation_q^*) \wedge (\mathcal{K}_\alpha \rightarrow prop(\mathcal{K}_\alpha))$
$> R = Explanation_q$.

b. $(\lambda : R(\alpha, \pi) \wedge (R = Explanation_q \vee R = Explanation_q^*) \wedge \neg(\alpha : prop(\mathcal{K}_\alpha))$
$> R = Explanation_q^*$.

---

These are read as follows. The term *Why?* itself is an ambiguous cue for the discourse relations *Explanation$_q$* and *Explanation$_q^*$*. Thus, if the content of $\pi$ is a *Why*-question that attaches to $\alpha$, the first two conjuncts in (IVa) and (IVb) are true. Then, the third conjuncts express, respectively:

a. $(\mathcal{K}_\alpha \rightarrow prop(\mathcal{K}_\alpha))$: the content of $\alpha$ entails $\alpha$'s core proposition (in particular, the content is itself propositional);

b. $\neg(\alpha : prop(\mathcal{K}_\alpha))$: the content of $\alpha$ is *not* $\alpha$'s core proposition (but the content may still be propositional).

Thus, in (158) and (184) the propositional reading is preferred by (IVa) and in (181) and (182) the meta-discursive one by (IVb). The content of (183a) however is $\alpha : \mathtt{inform}(\llbracket \text{A is upset} \rrbracket)$.[7] This satisfies *both* the antecedents of (IVa) and (IVb), because then $\mathcal{K}_\alpha$ entails $\llbracket \text{A is upset} \rrbracket$ (the core proposition), but it is not the case that $\alpha : \llbracket \text{A is upset} \rrbracket$. Thus, the precise attachment is underspecified in the logical form of (183b).

These rules are default (non-monotonic) *preferences* and can be defeated. To appreciate this point, consider the following example that has been pointed out to me by Jonathan Ginzburg.

---

[7]This is not the place to fully articulate a theory of performatives in SDRT. For the present purposes, the following ideas are sufficient. A performative is a monotonic linguistic cue to the glue logic about what speech act is being performed; that is, it restricts the possible discourse relations with which the utterance can attach. Here, an assertive performative (*I'll have you know* or, more properly, *I am hereby asserting to you*) introduces the predicate $\mathtt{inform}$ on the content of $\pi$, i.e. $\pi : \mathtt{inform}(\mathcal{K}_\pi)$. The contribution of $\mathtt{inform}$ is governed by the following axiom scheme:

$\pi : \mathtt{inform}(\mathcal{K}_\pi) \rightarrow (\mathcal{K}_\pi \rightarrow prop(\mathcal{K}_\pi) \wedge ((R(?, \pi) \wedge right\text{-}veridical(R)) \vee (R(\pi, ?) \wedge left\text{-}verdicial(R)))).$

That is, to attach as an assertion, the segment $\pi$ must attach veridically, i.e. with a discourse relation that entails the content of $\pi$ (Asher & Reese, 2007) Moreover, it must entail its propositional content so that uttering *I hereby assert that p* entails a commitment to *p*.

This means that *I hereby assert* is a monotonic cue for assertion, whereas indicative mood is merely a defeasible cue (e.g. an indicative mood utterance with final rise may attach non-veridically as a question, and committing to the corresponding segment may also not entail a commitment to the core proposition. The predicate $\mathtt{inform}$ is only relevant in the glue logic; it has no bearing on the truth-conditional evaluation of an utterance.

(185)   a. Amy: You're upset.

      b. Bob:  Why?

          [~~*Why am I upset?*~~ OR *Why are you saying that?*]

Here, a bare *Why?* in response to an assertion gets a meta-discursive reading (in apparent contradiction to (IVa)). This is because the speaker A has no epistemic access to speaker B's internal state. As a matter of fact, A cannot know for sure *that* B is upset, let alone *why*. Thus the propositional reading is absurd and the preference expressed in (IVa) is defeated.

**Veridicality and Presuppposition.**   Some prior literature asserts that *Why q?* presupposes *q* (e.g. Bromberger (1992), Hintikka & Halonen (1995)). This is not *per se* wrong, since these authors only speak of *Why*-questions with core propositions expressed in indicative mood. For such questions, this is correct.

(186)   a.  Why is John coming to the party?

          presupposes: *John is coming to the party*

However, this is not true for subjunctive *Why*-questions like (187).

(187)   a.  Why would John be coming to the party?

          ~~presupposes: *John is coming to the party*~~

Less mundanely, however, bare *Why?* also does not always trigger presuppositions. Recall (177).

(177)   a. Carla:   Got the junior tap and the senior tap.

      b. Brenda: Yeah but you'll get that next year again.

      c. Carla:   Why?

          ~~presupposes: *Carla will get junior tap and senior tap next year again*~~

I argued above that Carla after (177c) has not agreed with (177b), but rather is *resisting* (177b). Thus, bare propositional *Why?* does not generally presuppose the proposition it is inquiring about.

    This shows two things. First, that the discourse relation $Explanation_q$ is *not* left-veridical, i.e. $Explanation_q(\alpha, \beta)$ does not entail the truth of the content of $\alpha$. Second, that bare *Why?* is not elliptical on its antecedent. Analysing (177c) as *Why will I get that next year again?* where everything but the bare *Why?* is elided/unpronounced would predict a presupposition that (177c) does not trigger. This may be surprising since some, e.g. Jason Merchant (2005), seem to claim that *any* bare *wh*-phrase is elliptical. Merchant does not mention *Why?* explicitly, but Fernández *et al.* (2004), for instance, explicitly include it in their study on sluicing.

## 7.4   Focus, Presuppositions and Enthymemes

In this section, I discuss which presuppositions are triggered by *Why*-questions and defend that the presuppositions triggered by *Why*-questions affect the space of acceptable answers. This, when combined with the presuppositional theory of focussing from Chapter 6, explains the role of focus in *Why*-questions.

***Why* and Presupposition.**   The examples (188a) and (189a) are a minimal pair where one question triggers a presupposition that the other does not. Famously, the novel *Fight Club* (by Chuck Palahniuk) has a twist ending. With this knowledge, one may find (188b) felicitous, but (189b) sounds rather odd.

(188)    a. Amy: Why is John reading *Fight Club* again?

             presupposes: *John has read* Fight Club *before*

     b. Bob:  Because the ending changes the entire story.

(189)    a. Amy: Why is John reading *Fight Club*?

    #b. Bob:  Because the ending changes the entire story.

One may countenance the topos that *knowing a twist ending makes a story worth re-reading*, e.g. because it is fun to catch allusions to the twist. This topos licenses (188b) as a reason for John reading *Fight Club*, because the presupposition of (188) entails that John knows the ending of *Fight Club*. Plainly, this topos is active (and, arguably, salient) in (188) and licenses the relevant enthymeme. But this is not the case in (189b). The information that *John knows the ending* is not available and thus the topos cannot apply (and is also not salient).

Thus, the presuppositions triggered by a *Why*-question must be available for the enthymematic interpretation of an answer. However, there is one caveat to this claim. *Why q?* in indicative mood also triggers the presupposition that *q*. But one cannot simply reply *because q* to *Why q?*.

(190)    a. Amy: Why is John reading *Fight Club*?

    ?b. Bob:  Because he is.

While the answer is not outright infelicitous, Bob is certainly rather unhelpful in (190). Surely, *John is reading Fight Club* is not a reason for *John is reading Fight Club*. Thus only the presuppositions of *Why q?* that are not triggered by the *Why*-element itself should be made available. The discourse semantics for *Why*-questions must be amended to capture these facts.

---

(III′) **Discourse Semantics for *Why*-questions** (final version)
An **answer to** *Why q?* is an utterance that proffers the content *p* such that $psp \land p$ is a reason for *q*, where *psp* are the presuppositions triggered by *Why q?* that are not *q* itself.

---

Note that in certain contexts the dispreferred answer *q* to *Why q?* can be felicitous.

(191)  a. Amy: Why are you going to John's party?

      b. Bob:  Because I am.

Bob's response in (191) sounds better than the one in (190). However, this is not because *Bob is going to John's party* is a reason for *Bob going to John's party*, but because (191b) seems to implicate that Bob does not owe Amy an answer.

**The Focus Puzzle.**   A question raised by Sylvain Bromberger (1992) is why *Why*-questions seem to interact with focus in a way that other *wh*-questions do not. He cites the examples in (192) without marking tune, but it seems that the H LL% tune is natural.[8]

(192)  a.   Why did $\underline{\text{Adam}}_\text{H}$ eat the apple?$_\text{LL\%}$

      a.′  Why did Adam eat the $\underline{\text{apple}}_\text{H}$?$_\text{LL\%}$

He makes two claims about these: (i) that (192a) implicates that someone else may have eaten the apple (but didn't) and (192a′) implicates that Adam could have eaten something else (but ate the apple); (ii) that they take different answers (p161). For instance, (192a) tolerates (193b), but (192a′) tolerates (194b′).

(193)     a. God:  Why did $\underline{\text{Adam}}_\text{H}$ eat the apple?$_\text{LL\%}$

           $\rightsquigarrow$ *and not someone else?*

      b. Snake: Because Eve does not like apples.

   #b.′ Snake: Because there was nothing else to eat.

(194)     a. God:  Why did Adam eat the $\underline{\text{apple}}_\text{H}$?$_\text{LL\%}$

           $\rightsquigarrow$ *and not something else?*

   #b. Snake: Because Eve does not like apples.

      b.′ Snake: Because there was nothing else to eat.

This is in contrast to other *wh*-questions. For instance, any felicitous answer to (195a) is a felicitous answer to (195a′) and *vice versa.*

(195)  a.   When did $\underline{\text{Adam}}_\text{H}$ eat the apple?$_\text{LL\%}$

      a.′  When did Adam eat the $\underline{\text{apple}}_\text{H}$?$_\text{LL\%}$

**Solving the Puzzle.**   The theory of focus and intonation of Chapter 6 only applies to assertions. So, to explain the data here, I extend it to indicative *wh*-questions by letting the Focus Semantics for H LL% apply to the existential closure of these questions. For a *wh*-question $?\lambda e.\varphi$, this is $\exists e.\varphi$. For *Why p?*,

---

[8] *Wh*-questions in English do not typically feature a final rise (Hedberg *et al.*, 2004).

where nothing needs to be closed, it is just $p$. Note that this existential closure is just the proposition that is typically presupposed by a *wh*-question.

(196) a. Why did <u>Adam</u>$_H$ eat the apple?$_{LL\%}$

   presupposes (*wh*): *Adam ate the apple.*

   presupposes (focus): *x ate the apple.*

(197) a. When did <u>Adam</u>$_H$ eat the apple?$_{LL\%}$

   presupposes (*wh*): $\exists t.Adam\ ate\ the\ apple\ at\ t.$

   presupposes (focus): $\exists t.x\ ate\ the\ apple\ at\ t.$

In (196) and (197) the presupposition effected by the focus semantics is benign and might as well be left out: it is entailed by the presupposition triggered by the *wh*-element and so does not change the interpretation of the utterance. This explains why the focal placement in (195) makes no difference. However, I have argued that presuppositions that are *not* the presupposition of the *wh*-element matter for the discourse semantics of *Why*-questions. Thus, the presupposition triggered by focus in a *Why*-question *is* relevant, because it interacts with the revised semantics (III′).

The examples to explain are (193) and (194), here repeated with the presupposition effected by the focus semantics added. When computing the presuppositions in a null context, the anaphorical $x$ are existentially closed.

(193)   a. God:   Why did <u>Adam</u>$_H$ eat the apple?$_{LL\%}$

      presupposes: *someone ate the apple.*

   b. Snake: Because Eve does not like apples.

   #b.′ Snake: Because there was nothing else to eat.

(194)   a. God:   Why did Adam eat the <u>apple</u>$_H$?$_{LL\%}$

      presupposes: *Adam ate something.*

   #b. Snake: Because Eve does not like apples.

   b.′ Snake: Because there was nothing else to eat.

These presuppositions may explain the intuitions responsible for Bromberger's observation (i). Now, regarding (ii), note that according to the semantics (III′), presuppositions factor into the interpretation of responses to *Why*-questions. Thus, to make (193b) felicitous there must be a salient topos licensing the following enthymeme.

(198)   $\underline{\text{Someone ate the apple} \wedge \text{Eve does not like apples.}}$

   $\therefore$ Adam ate the apple.

An appropriate topos seems to be *people don't eat food they don't like*. Then, if *Adam* and *Eve* are the only available referents, one can infer *Adam ate the apple* by

some version of disjunctive syllogism. The additional restriction on the available referents is required here, since (199b) does not seem to be a good answer.

(199)      a. God:   I thought Eve or Sue would eat the apple.

Why did <u>Adam</u><sub>H</sub> eat the apple?<sub>LL%</sub>

??b. Snake: Because Eve does not like apples.

The Snake's answer does not seem to fully answer God's question, because (198) cannot be licensed as above.

Now, to make (193b′) felicitous there must be a salient topos licensing the following enthymeme.

(200)      <u>Someone ate the apple ∧ there was nothing else to eat</u>

∴ Adam ate the apple.

This does not seem to be correct; Adam is not even mentioned in the antecedent of (200), so there is no way to infer that he ate anything at all—let alone the apple. Hence, the semantics in (III′) predicts infelicity.

The dual analysis explains (194). To judge (194b) felicitous one would require a topos for (201) and to judge (194b′) felicitous requires a topos for (202).

(201)      <u>Adam ate something ∧ Eve does not like apples.</u>

∴ Adam ate the apple.

(202)      <u>Adam ate something ∧ there was nothing else (but the apple) to eat</u>

∴ Adam ate the apple.

The enthymeme in (202) is trivially correct, but there seems to be no topos that would license (201). Again, the premises cannot possibly entail that the apple was eaten at all. This explains the felicity of (194b′) and the infelicity of (194b).

## 7.5   Conclusion

I have presented an analysis of the meaning of *Why*-questions in discourse. These questions are of particular importance to this thesis, since they can be used to *resist* a previous speech act (without outright accepting or rejecting it). My analysis explains this as a necessary condition on commitment structures: a commitment necessarily requires that one is willing and able to vindicate the commitment. This in particular includes the willingness and ability to answer *Why*-questions.

The analysis is grounded in a *micro-rhetorical* framework for dialogue interpretation. In this framework, one can define a *reason* as something that forms an enthymeme while simultaneously making salient what licenses the enthymeme. Then, a reason for *q* is a permissible answer to *Why q?*. I showed how this conception can straightforwardly be embedded in the SDRT framework.

Based on this groundwork, I have then investigated the problems presented by bare *Why?*, presuppositions in *Why*-questions and focus in *Why*-questions. The linguistic analysis shows that bare *Why?* is ambiguous between propositional and meta-discursive readings, but that this ambiguity can be captured and regimented in a systematic manner. Regarding presuppositions, I have argued that the presuppositions triggered by a *Why*-question—except the presupposition triggered by the *wh*-element itself—factor in the enthymematic interpretation of potential responses. When combined with an appropriate adaption of the focus semantics of Chapter 6, this also explains the data on focus.

# Chapter 8

# Conclusion

Philosophers and linguists have only recently begun to appreciate the problems posed by disagreement in dialogue. Certainly, the relevant data are far from understood and this dissertation is but a small step forward. However, there seems to be a budding consensus that countenancing disagreement data must entail revisions of classical logic or classical semantic theory. This thesis stands against this consensus.

The *weak bilateral logic* developed in Chapter 2 demonstrates that it is both useful (*contra* Frege) and possible (*contra* Dickie) to include rejected content in one's logic without thereby giving up on the classicality of the logic of assertion. However, this argument applies only to the impoverished language of propositional logic. The linguistic data shows that presupposition failures are grounds for rejection, but weak bilateral logic does not capture this fact, since presuppositions cannot be expressed in its language. The first step to make progress on this problem would be an extension of weak bilateral logic to a first order language. Then, or so one would hope, the bilateralist can speak of what exists and thus model certain existential presuppositions (possibly, this would also require an operator $\iota$ for the definite article *the*). However, extending the language leads to new versions of the *specificity argument against Deduction* which need to be addressed through appropriate restriction strategies.

Such restriction strategies are developed in Chapter 4, where weak bilateral logic is extended to *epistemic multilateral logic.* I demonstrate the restriction strategies for monomodal language and a disquotational predicate. This resolves the epistemic modal challenge against classical logic (Yalcin and others) and the Denier paradox (Bacon and Murzi & Carrara). Moreover, epistemic multilateral logic satisfies a desideratum that weak bilateral logic (arguably) does not: its inference rules are pure, simple and harmonious. But the development of epistemic multilateral logic directly leads to another open question: if it is fair game to extend weak bilateral logic with a speech act of *weak assertion*, are there other speech acts one can (or should) include?

Chapter 3 extends the results of Chapter 2 in another direction. The arguments related to weak rejection generalise when they are understood in a framework that analyses speech acts as the undertakings of certain commitments. In particular, this framework can model non-cooperative settings, formalise rejections of arbitrary speech acts, and derive strong rejection as the linguistically preferred interpretation of weak rejection. These results are derived from first principles that partially axiomatise what it means to commit. However, I only analyse assertion and rejection from these first principles, and it is left open whether other speech acts require additional principles. Thus, my axiomatisation remains partial. Another facet of commitment—that one must vindicate one's commitments—is addressed in Chapter 7.

The commitment framework moreover allows me to tackle the problem of *faultless disagreement* in Chapter 5. I demonstrate that a suitable commitment framework allows one to articulate formal semantics for taste predicates that square with the linguistic data, but do not require any revision of orthodox semantic theory (*contra* Kölbel, MacFarlane and others). It is left open, however, what conclusions a philosopher should draw from these semantics: do taste predicates have absolute truth values, or do speakers behave *as if* they do?

In Chapter 6 I analyse the problem of intonation in dialogue. Evidently, there are some utterances that admit agreement readings with one particular intonation, but disagreement readings with a different one. This is a linguistic problem that deserves attention, but also philosophers who use disagreement data should take notice: it may not always be obvious what the correct interpretation of an utterance is. In my development of an appropriately articulated theory of prosody, I found that one should not study focus placement without studying tune. In fact, or so I argue, much prior discussion on focus has been misled by a failure to do so. My analysis of the relevant data also casts doubt on theories of focus that use QUD (Roberts and others) or posit the strict identification of *new* with *focal* (Büring and others). However, my alternative proposal can only be considered a small step towards a comprehensive theory of focus and tune.

The results of this thesis are, by and large, foundational and beg for further research. Nevertheless: I set out to defend that it is possible and useful to countenance a speech act of *rejection* that is not negative assertion, to defend classical logic, to defend non-relative truth, to elucidate the relationship of intonation and rejection, and to explain how rejection interfaces with commitment. I hope the reader will be inclined to agree that I have gone some way towards these goals.

"Are you looking forward to the day
when you emerge from these cloistered walls
into what some call the world?"

"No sir."

John Williams, *Stoner*

# Appendix

# Appendix A

# Proofs

## A.1 Chapter 2

**Smilean inference is equivalent to $(-\neg E.)^*$:**

$\Leftarrow$ By $(+ \to I.)^*$, the antecedent of (SI) can be written as $+(\neg A \to \neg B), -A$. Then in WBL $- $ (SI) $+ (-\neg E.)$:

$+(\neg A \to \neg B), +\neg A \vdash +\neg B$ by $(+ \to E.)$.
$+(\neg A \to \neg B), +\neg A \vdash -B$ by $(+\neg E.)$.
$+(\neg A \to \neg B), +B, +\neg A \vdash \bot$ by (Rejection).
$+(\neg A \to \neg B), +B \vdash -\neg A$ by $(SR_1)$.
$+(\neg A \to \neg B), +B \vdash +A$ by $(-\neg E.)^*$.
$+(\neg A \to \neg B), -A, +B \vdash \bot$ by (Rejection).
$+(\neg A \to \neg B), -A \vdash -B$ by $(SR_1)$.

$\Rightarrow$ WBL entails that negation is classical. Thus:

If $\vdash^+ -\neg A$ then $\vdash +\neg\neg A$ by $(+\neg I.)^*$.
Thus $\vdash^+ +A$ by double negation elimination.

$\square$

**Rules for Rejected Premisses:**

$(-\vee I.)$ If $\vdash^+ -A$ and $\vdash^+ -B$, then $+A \vdash^+ \bot$ and $+B \vdash^+ \bot$. Thus $+(A \vee B) \vdash \bot$ by $(+\vee E.)$. Hence $-(A \vee B)$ by Smilean reductio.

$(-\vee E.)$ $-(A \vee B), +A \vdash +(A \vee B)$ by $(+\vee I.)$. Hence $-(A \vee B), +A \vdash \bot$ by (Rejection). And thus $-(A \vee B) \vdash -A$ by Smilean reductio.

$(-\vee E._2)$ is analogous.

$(-\wedge I._1)$ $-A, +(A \wedge B) \vdash +A$ by $(+\wedge E.)$, so $-A, +(A \wedge B) \vdash \bot$ by (Rejection). Thus $-A \vdash -(A \wedge B)$ by $(SR_1)$.

$(-\wedge I._2)$ is analogous.

$(-\wedge E.)$ Suppose $\varphi = +C$. Assume $-A \vdash +C$ and $-B \vdash +C$.
Then $-(A \wedge B), -A, -C \vdash \bot$ and $-(A \wedge B), -B, -C \vdash \bot$ by (Rejection).
Hence $-(A \wedge B), -C \vdash +A$ and $-(A \wedge B), -C \vdash +B$ by $(SR_2)$.
So, $-(A \wedge B), -C \vdash +(A \wedge B)$ by $(+\wedge I.)$, i.e. $-(A \wedge B) \vdash +C$ by $(SR_2)$.
The argument for $\varphi = -C$ is analogous by using $(SR_1)$ instead of $(SR_2)$ in the proof above.

$(-\to I.)$ $+A, -B, +(A \to B) \vdash +B$ by $(+ \to E.)$, so $+A, -B, +(A \to B) \vdash \bot$ by (Rejection). Thus $+A, -B \vdash -(A \to B)$ by $(SR_1)$.

$(-\to E._1)$ $-(A \to B), +B \vdash +(A \to B)$ by $(+ \to I.)^*$, so $-(A \to B), +B \vdash \bot$ by (Rejection). Thus $-(A \to B) \vdash -B$ by $(SR_1)$.

$(-\to E._2)$ This follows non-elementarily from the Classicality theorem:
If $\vdash {}^+ -(A \to B)$, then $+\neg(A \to B)$ by $(+\neg I.)^*$. Then $+A$ because $\to$ is a material conditional.

$\square$

## Proof of Lemma 2.3.2 (Contraposition):

- To show: $+(\neg A \to \neg B), +B \vdash +A$.

  $+(\neg A \to \neg B), -A \vdash -B$ by (SI).
  $+(\neg A \to \neg B), +B, -A \vdash \bot$ by (Rejection).
  $+(\neg A \to \neg B), +B \vdash +A$ by $(SR_2)$.

$\square$

## Proof of Proposition 2.3.3 (Classical negation):

- Double Negation Introduction (DNI): $\vdash +(A \to \neg\neg A)$.

  $+A, \vdash -\neg A$ by $(-\neg I.)$.
  $+A, \vdash +\neg\neg A$ by $(+\neg I.)^*$.
  $\vdash +(A \to \neg\neg A)$ by $(+ \to I.)^*$.

- Double Negation Elimination (DNE): $\vdash +(\neg\neg A \to A)$.

  $\vdash +(\neg A \to \neg\neg\neg A)$ (DNI).
  $\vdash +\neg\neg A \vdash +A$ by Contraposition.
  $\vdash +(\neg\neg A \to A)$ by $(+ \to I.)^*$.

- Law of Non-Contradiction (LNC): $\vdash +\neg(A \wedge \neg A)$.
  Show that $(A \wedge \neg A)$ is absurd to assert, and since this inference requires no further premisses, $(+\neg\text{I.})^*$ applies.

  $+(A \wedge \neg A) \vdash +A$ by $(+\wedge\text{E.})$.
  $+(A \wedge \neg A) \vdash +\neg A$ by $(+\wedge\text{E.})$.
  $+(A \wedge \neg A) \vdash -A$ by $(+\neg\text{E.})$.
  $+(A \wedge \neg A) \vdash \bot$ by (Rejection).
  $\vdash -(A \wedge \neg A)$ by $(\text{SR}_1)$.
  $\vdash +\neg(A \wedge \neg A)$ by $(+\neg\text{I.})^*$.

- Classical Reductio: $\vdash +((A \rightarrow (B \wedge \neg B)) \rightarrow \neg A)$.

  - By LNC and $(+ \rightarrow\text{E.})$, $+A, +(A \rightarrow (B \wedge \neg B)) \vdash \bot$.
  - By Smilean Reductio, $+(A \rightarrow (B \wedge \neg B)) \vdash -A$.
  - By $(+\neg\text{I.})^*$, $+(A \rightarrow (B \wedge \neg B)) \vdash +\neg A$.
  - By $(+ \rightarrow\text{I.})^*$, $\vdash +((A \rightarrow (B \wedge \neg B)) \rightarrow \neg A)$.

  $\square$

**Proof of Proposition 2.3.4 (Classicality):**

$\Leftarrow$ Suppose $+A \vdash^D +B$. We want to transform the proof $D$ into a proof in classical propositional logic. First note that by $(+\neg\text{I.})^*$ any step in the proof with a rejected premiss $-C$ can be replaced by $+\neg C$. Further note that all rules of WBL express classical valid forms of reasoning when mapping $+C \mapsto C$ and $-C \mapsto \neg C$. Thus, applying this mapping to $D$ yields a classically valid proof. No

$\Rightarrow$ Since the rules for asserted (plussed) $\vee$, $\wedge$ and $\rightarrow$ are classical and Proposition 3.4 shows that negation is classical, any proof in classical logic can be reproduced on the plussed fragment of weak bilateral logic.

$\square$

**Proof of Theorem 2.3.6 (Soundness):**
The proof proceeds by induction on the length $n$ of derivations. Write $\Gamma \vdash^D \varphi$ if $D$ is a derivation of $\varphi$ from the premisses in $\Gamma$. The base clause, $n = 1$ is trivial as $\Gamma \vdash^D \varphi$ with $|D| = 1$ only if $\varphi \in \Gamma$.

Now, for each $k \leq n$ and any set of premisses $\Gamma$, assume that if $\Gamma \vdash^D \psi$, $|D| = k$, then $\Gamma \models \psi$. Then assume that $\Gamma \vdash^D \varphi$ for some $D$ with $|D| = n + 1$ and show that $\Gamma \models \varphi$.

- $(+ \rightarrow \text{I.})^*$. Assume $\Gamma \vdash^D \varphi$ by an application of $(+ \rightarrow \text{I.})^*$, i.e. $\varphi = +(A \rightarrow B)$ for some $A$, $B$ and $\Gamma \cup \{+A\} \vdash^{D'} +B$ where $D'$ uses only asserted premisses from $\Gamma$. Let $\Gamma'$ be the asserted formulae in $\Gamma$. Then $\Gamma' \cup \{+A\} \vdash^{D'} +B$.

Assume that $\Gamma' \not\models +(A \to B)$. Then there is a model $V$ of $\Gamma'$ and a point $y \in \omega$ such that $V \models_y A \wedge \neg B$. Construct an $\omega$-pointed model $V'$ where *every* point is $y$, i.e. for any $x$ and atom $p$, $V' \models_x p$ iff $V \models_y p$. Because $\Gamma'$ contains only asserted formulae, $V' \models \Gamma'$. Also, because $V \models_y A$, $V' \models +A$. Hence $V' \models \Gamma' \cup \{+A\}$. By assumption, $V' \models +B$, but by construction $V' \models +\neg B$. Contradiction.

- $(+\neg I.)^*$. Proof in Chapter 2.

- Every point in an $\omega$-pointed model is a model of classical propositional logic. Hence the rules $(+\wedge I.)$, $(+\wedge E._1)$, $(+\wedge E._2)$, $(+\vee I._1)$, $(+\vee I._2)$, $(+\to E.)$ are trivially sound. For instance, if $(+\wedge I.)$ is the last rule applied in $D$, then $\varphi = +(A \wedge B)$ and there are derivations $D'$ and $D''$ of $+A$ and $+B$ respectively from $\Gamma$. By induction hypothesis, $\Gamma \models +A$ and $\Gamma \models +B$. That is, every point in every model of $\Gamma$ satisfies $A$ and $B$, hence every point in every model of $\Gamma$ satisfies $A \wedge B$. Thus $\Gamma \models +(A \wedge B)$.

- $(+\vee E.)^*$ Suppose the last step in $D$ is an application of $(+\vee E.)$ to derive $\varphi$. Then, there are shorter derivations $D', D'', D'''$ such that $\Gamma \vdash^{D'} +(A \vee B)$ and $\Gamma, +A \vdash^{D''} \varphi$ and $\Gamma, +B \vdash^{D'''} \varphi$. Do a case distinction on the sign of $\varphi$.

  If $\varphi = +C$ for some $C$, then by the Soundness of $(+ \to I.)^*$, $\Gamma \models +(A \vee B)$ and $\Gamma \models +(A \to C)$ and $\Gamma \models +(B \to C)$. Because every point in an $\omega$-pointed model is a model of classical logic, $\Gamma \models +C$.

  If $\varphi = \bot$ then the subderivations can be transformed by Smilean *reductio* to derivations of $+\neg(A \to A)$. Then, by the Soundness of $(+ \to I.)^*$, $\Gamma \models +(A \vee B)$ and $\Gamma \models +A(\to \neg(A \to A))$ and $\Gamma \models +(B \to \neg(A \to A))$. Because every point in an $\omega$-pointed model is a model of classical logic, $\Gamma \models +\neg(A \to A)$. But $\neg(A \to A)$ is classically unsatisfiable, so $\Gamma$ is unsatisfiable. Thus $\Gamma \models \bot$.

  If $\varphi = -C$ then by $(+\neg I.)$, $\Gamma, +A \vdash +\neg C$ and $\Gamma, +B \vdash +\neg C$. Then, as in the case where $\varphi = \neg C$, obtain $\Gamma \models +\neg C$. Thus, by the Soundness of $(+\neg E.)$, $\Gamma \models -C$.

- $(-\neg I.)$. If there is a derivation of $+A$ from $\Gamma$ then by induction, all points in all models of $\Gamma$ satisfy $A$. Since there is at least one point in every model, $\Gamma \models -\neg A$.

- $(+\neg E.)$. If there is a derivation of $+\neg A$ from $\Gamma$ then by induction, all points in all models of $\Gamma$ satisfy $\neg A$. Since there is at least one point in every model, $\Gamma \models -A$.

- (Rejection). By definition, no $\omega$-pointed model can satisfy $+A$ and $-A$. Hence if there are subderivations of $+A$ and $-A$ from $\Gamma$, then by induction, $\Gamma$ has no model, i.e. $\Gamma \models \bot$.

- $(SR_1)$. Assume that the last step in the derivation $D$ is an application of $(SR_1)$ to show $-A$. Then there is a derivation $D'$ of shorter length with $\Gamma, +A \vdash^{D'} \bot$. By induction, $\Gamma \cup \{+A\} \models \bot$, i.e. $\Gamma \cup \{+A\}$ is unsatisfiable. If already $\Gamma$ is unsatisfiable, then trivially $\Gamma \models -A$. If $\Gamma$ is satisfiable then let $V$ be any model of $\Gamma$. By assumption, $V \not\models +A$, i.e. there is a $x \in \omega$ such that $V \models_x \neg A$, hence $V \models -A$.

- $(SR_2)$. Assume that the last step in the derivation $D$ is an application of $(SR_1)$ to show $+A$. Then there is a derivation $D'$ of shorter length with $\Gamma, -A \vdash^{D'} \bot$. By induction, $\Gamma \cup \{-A\} \models \bot$, i.e. $\Gamma \cup \{-A\}$ is unsatisfiable. If already $\Gamma$ is unsatisfiable, then trivially $\Gamma \models +A$. If $\Gamma$ is satisfiable then let $V$ be any model of $\Gamma$. By assumption, $V \not\models -A$, i.e. for all $x \in \omega$, $V \models_x A$, hence $V \models +A$.

- (SI). By induction, we know that $\Gamma \models +(\neg A \to \neg B)$ and $\Gamma \models -A$. Let $x \in V$ such that $V \not\models_x A$. Then, $V \models_x \neg A$. Since $V \models_x (\neg A \to \neg B)$ then also $V \models_x \neg B$. Thus $V \models -B$.

- The rules on rejected premisses are derivative, so there is nothing to show.

This concludes the induction. $\qquad\square$

**Proof of Lemma 2.3.7:**

**Auxiliary lemma.** If $\Gamma$ is a WBL-consistent set of only asserted formulae, then $\Gamma' = \{A \mid +A \in \Gamma\}$ is classically satisfiable.

**Proof:** Assume $\Gamma'$ is not classically consistent. Then there is a proof of $\bot$ from the premisses in $\Gamma'$. By Classicality this proof can be carried out in weak bilateral logic with the premisses from $\Gamma$. Therefore, by contraposition, if $\Gamma$ is WBL-consistent, $\Gamma'$ is classically consistent. By the satisfiability theorem in classical logic, if $\Gamma$ is WBL-consistent, $\Gamma'$ is classically satisfiable.

By the auxiliary, it suffices to show that $\Gamma \cup \{+\neg A\}$ is consistent. Assume it is not. Then $\Gamma, +\neg A \vdash \bot$, i.e. $\Gamma \vdash -\neg A$ by $(SR_1)$. Since $\Gamma$ only contains asserted formulae, $\Gamma \vdash +\neg\neg A$ by $(+\neg I.)$. Hence $\Gamma \vdash +A$ by (DNE), which contradicts the assumption that $\Gamma \cup \{-A\}$ is consistent. $\qquad\square$

**Proof of Theorem 2.3.8:**
Show a Model Existence result first. That is, for every consistent $\Gamma$ there is an $\omega$-pointed model of $\Gamma$.

**Proof of Model Existence.**
Use Lemma 2.3.7 to construct an $\omega$-pointed model by separating the asserted

from the rejected formulae in $\Gamma$. Then let all the asserted formulae hold at every point and let every rejected formula fail at a single point.

So let $\Gamma^+$ be the set of asserted formulae in $\Gamma$ and let $\{-A_i \mid i \in \omega\}$ be an enumeration of the rejected formulae in $\Gamma$. For all $i \in \omega$, define $\Gamma^i = \Gamma^+ \cup \{-A_i\}$. As $\Gamma^i \subseteq \Gamma$ for all $i$, the $\Gamma^i$ are consistent. By Lemma 2.3.7 there are interpretations $I_i$ of classical propositional logic that satisfy $\{\neg A_i\} \cup \{A \mid +A \in \Gamma^+\}$. Then define an $\omega$-pointed model $V$ by setting $V(x) = \{p \mid I_x(p) = \text{True}\}$. With this construction, clearly, $V \models \Gamma$.

Then show completeness by case distinction on the force of $\varphi$:

- Assume $\Gamma \models +A$ and $\Gamma \not\vdash +A$. Then $\Gamma \cup \{-A\}$ is consistent and hence there is a $V$ with $V \models \Gamma \cup \{-A\}$. Then $\Gamma \not\models +A$.

- Assume $\Gamma \models -A$ and $\Gamma \not\vdash -A$. Then $\Gamma \cup \{+A\}$ is consistent and hence there is a $V$ with $V \models \Gamma \cup \{+A\}$. Then $\Gamma \not\models -A$.

$\square$

# A.2 Chapter 3

**Proof of Lemma 3.4.5:**
KD45 modal logic is sound and complete for Kripke frames that are serial, transitive and Euclidean; these properties correspond to axioms (D), (4) and (5), respectively. That is, a model is a triple $(W, R, V)$ where $W$ is a set of worlds, $V$ is a mapping from worlds to sets of atoms and $R \subseteq W^2$ is a relation satisfying:

(S) $\forall x \exists y (xRy)$  (T) $\forall x, y, z (xRy \wedge yRz \rightarrow xRz)$  (E) $\forall x, y, z (xRy \wedge xRz \rightarrow yRz)$.

Now show the lemma by contraposition. Suppose $\Gamma \not\vdash^{\text{KD45}} \square(A \rightarrow B)$. Then there is a model $M = (W, R, V)$ and a world $w \in M$ with $M \models \Gamma$ and $M, w \not\models \square(A \rightarrow B)$, i.e. $M, w \models \Diamond(A \wedge \neg B)$. Let $v \in W$ be a witness for this, i.e. $wRv$ and $M, v \models A \wedge \neg B$.

Consider the following model $M' = (\{w, v\}, \{(v, v), (w, v)\}, V \restriction \{w, v\})$. Trivially, M' is serial and satisfies (T) and (E). Now verify that it is also a model of $\Gamma$. Let $\varphi \in \Gamma$. Then there is a propositional $C$ such that $\varphi = \square C$. Since $M \models \Gamma$, $M, w \models \square C$, hence $M, v \models C$. Thus $M', v \models C$ and therefore $M', w \models \square C$ and $M', v \models \square C$. Hence $M'$ is a KD45-model of $\Gamma$. Now observe that $M', w \models \square A$ but $M', w \not\models \square B$. Hence $\Gamma, \square A \not\vdash^{\text{KD45}} \square B$. $\square$

**Proof of Theorem 3.4.3:**
The previous lemma establishes the soundness of modalised $(+ \rightarrow \text{I.})^*$. The only interesting cases here are (Rejection), $\neg$-elimination and -introduction, the Smilean reductios and the Smilean inference rules. The soundness of the other rules can

be established as in the proof of soundness on $\omega$-pointed models (Theorem 3.7). Show soundness by assuming that $M = (W, R, V)$ is an arbitrary model of KD45 and showing that it satisfies the rules of weak bilateral logic.

- (Rejection). If $M, w \models \Box A$ and $M, w \models \Box\neg\Box A$, then, by seriality, there is a $v \in W$, $wRv$, with $M, v \models \neg\Box A$. Then there is $v' \in W$, $vRv'$, and $M, v' \models \neg A$. By Transitivity, also $M, v' \models A$. Hence there is no such $M$.

- $(+\neg I.)^*$. Assume $\Box\neg\Box A$ holds in all models where the set of only necessitated formulae $\Gamma$ is true. Let $M = (W, R, V)$ be such a model and assume that there is some $w$ with $M, w \not\models \Box\neg A$, i.e. $M, w \models \Diamond A$. Let $v \in W$ be such that $wRv$ and $M, v \models A$. Let $M' = (W', R', V')$ be a model with $W' = \{w, v\}$, $R' = \{(w, v), (v, v)\}$ and $V'(v) = V(v)$. $R'$ clearly is serial, transitive and Euclidean. Since all formulae in $\Gamma$ are of the form $\Box B$ for propositional $B$ and $M, w \models \Gamma$, $M, v \models B$ for all $B$ with $\Box B \in \Gamma$. Hence also $M', w \models \Gamma$, as the valuation on $v$ is the same. But by construction also $M', v \models A$ and $vRv$, so $M, w \models \Box\Box A$. This contradicts the initial assumption that $\Box\neg\Box A$ is true in all models of $\Gamma$.

- $(+\neg E.)$. If $M, w \models \Box\neg A$, then $M, w \models \Diamond\neg A$ by (D). Then, by (5), $M, w \models \Box\neg\Box A$.

- $(-\neg I.)$. Analogous to $(+\neg E.)$.

- (SI). Assume $M, w \models \Box(\neg A \to \neg B)$ and $M, w \models \Box\neg\Box A$. The latter formula can be put as $M, w \models \Box\Diamond\neg A$. By Transitivity, it also holds that $M, w \models \Box\Box(\neg A \to \neg B)$. Hence $M, w \models \Box\Diamond\neg B$, which is equivalent to $M, w \models \Box\neg\Box B$.

Show the soundness of the Smilean reductios by deriving them from the KD45-axioms. Modal logic satisfies a version of (CNI) not expressible in bilateralist logics, namely $\Gamma, \varphi \vdash \bot \Rightarrow \Gamma \vdash \neg\varphi$. This cannot be put in the bilateralist language, because the force markers cannot embed under a negation sign.

- (SR$_1$). Suppose $\Gamma, \Box A \vdash \bot$. Then $\Gamma \vdash \neg\Box A$, hence by (5), $\Gamma \vdash \Box\neg\Box A$.

- (SR$_2$). Suppose $\Gamma, \Box\neg\Box A \vdash \bot$. Then $\Gamma \vdash \neg\Box\neg\Box A$. By the contrapositive of (5), $\Gamma \vdash \neg\neg\Box A$, i.e. $\Gamma \vdash \Box A$.

$\Box$

**Proof of Theorem 3.4.4 (i):**
The Classicality theorem of WBL result shows that the propositional logic underlying the modal formulation of weak bilateral logic is classical (in the same way that the points in an $\omega$-pointed model are classical propositional interpretations). Hence, only the frame conditions need checking. Letting the modal operators

embed transforms the Smilean reductios to axiom schemes on modal frames as
follows:

$$(\text{SR}_1)\ (\neg\Box A) \to (\Box\neg\Box A)  \quad (\text{SR}_2)\ (\neg\Box\neg\Box A) \to (\Box A).$$

So let $M = (W, V, R)$ be a model of the two schemes above. Show that $M$ satisfies
the axioms (D), (4) and (5).

- (D). Let $W \in W$ with $M, w \models \Box A$. We need to show that $M, w \models \Diamond A$.
  Assume this is false. Then $M, w \models \neg\Diamond A$, i.e. $M, w \models \Box\neg A$. By (+¬E.)
  then $M, w \models \Box\neg\Box A$. Then by (Rejection), $M, w \models \bot$. Contradiction.

- (4). Let $w \in W$ with $M, w \models \Box A$. We need to show that $M, w \models \Box\Box A$. By
  (Rejection), $\Box A, \Box\neg\Box A \vdash \bot$, hence $M, w \models \neg\Box\neg\Box A$. We can bracket this
  to read as $M, w \models \neg\Box(\neg\Box A)$ and apply (SR$_1$) outside the brackets. That
  is, $M, w \models \Box\neg\Box(\neg\Box A)$. Bracket this as $M, w \models \Box(\neg\Box\neg\Box A)$. Then apply
  the axiom scheme (SR$_2$) in the brackets to arrive at $M, w \models \Box\Box A$.

- (5). We need to show $M \models \neg\Box A \to \Box\neg\Box A$. This is exactly (SR$_1$).

$\square$

**Proof of Theorem 3.4.4 (ii):**
To prove this, one can use the completeness result on $\omega$-pointed models to show
that $\omega$-pointed models (where the force markers $+$ and $-$ are considered non-
embeddable) embed into KD45-models. That is, we can transform $\omega$-pointed
models into models of KD45 modal logic to show the non-embedded completeness
theorem above.

**Auxiliary Lemma.** Let $V$ be a $\omega$-pointed model. Then there is a KD45-model
$M^V$ such that $V \models +A$ iff $M^V \models \Box A$ and $V \models -A$ iff $M^V \models \Box\neg\Box\varphi$.

**Proof.** Define $M^V$ as follows: $M^V = (\omega, \omega^2, V)$. Trivially, the relation is serial,
transitive and Euclidean. Now observe: $V \models +A$ iff $\forall n : V \models_n A$ iff
$M^V \models \Box A$. $V \models -A$ iff $\exists n : V \models_n \neg A$ iff $M^V \models \Diamond\neg A$ iff $M^V \models \neg\Box A$ iff
$M^V \models \Box\neg\Box A$.

**Proof of 3.15.** Let $\Gamma$ be a set of rejectivist modal formulae. Show that if $\Gamma \models^{\text{KD45}}$
$\varphi$ then $\Gamma^b \vdash \varphi^b$ in weak bilateral logic. Assume not, i.e. $\Gamma \models^{\text{KD45}} \varphi$ and
$\Gamma^b \not\vdash \varphi^b$.

- Case 1: $\varphi = \Box A$. Then $\varphi^b = +A$. By assumption, $\Gamma^b \cup \{-A\}$ is
  consistent. By $\omega$-pointed Model Existence, there is an $\omega$-pointed model
  $V$ of the set $\Gamma^b \cup \{-A\}$. By the previous lemma, there is a KD45-model
  $M^V$ such that $M^V \models \Gamma$ and $M^V \models \Box\neg\Box A$. This contradicts the
  assumption that $\Gamma \models^{\text{KD45}} \varphi$, because then also $\Gamma \models^{\text{KD45}} \Box\Box A$.

- Case 2: $\varphi = \Box\neg\Box A$. Then $\varphi^b = -A$. By assumption, $\Gamma^b \cup \{+A\}$ is consistent. As above, there is an $\omega$-pointed model $V$ of $\Gamma^b \cup \{+A\}$. By the previous lemma, there is a KD45-model $M^V$ such that $M^V \models \Gamma$ and $M^V \models \Box A$. Then also $M^V \models \Box\Box A$. This contradicts the assumption that $\Gamma \models^{\mathrm{KD45}} \varphi$.

$\square$

**A.2.1.** DEFINITION (Admissibility Conditions for the Basic Principles). The Basic Principles correspond to the following structural axioms on tree models.

(a) $C_A\varphi > C_A\Box C_A\varphi$.

$\forall X \subseteq S \times W, \ (x,w) \in S \times W :$
$*(x, w, \{(y,v) \in S \times W \mid \forall v' \in W : vR_A^y v' \to (y, v') \in X\})$
$\subseteq \{(y,v) \in S \times W \mid \forall v' \in W :$
$\qquad vR_A^y v' \to \forall y' \in T(y,v')\big(\forall v'' \in W(v'R_A^{y'}v'' \to (y', v'') \in X)\big)\}$.

(b) $C_A\Diamond C_A\Diamond\varphi > C_A\Diamond\varphi$.

$\forall X \subseteq S \times W, \ (x,w) \in S \times W :$
$*(w, \{(y,v) \in S \times W \mid \forall v'\big(vR_A^y v' \to \exists y' \in T(y,v')\forall v'' \in W :$
$\qquad\qquad (v'R_A^{y'}v'' \to \exists y'' \in T(y',v'') : (y'',v'') \in X)\big)\})$
$\subseteq \{(y,v) \in S \times W \mid \forall v' \in W\big(vR_A^y v' \to \exists y' \in T(y,v') : (y',v') \in X\big)\}$.

(c) $C_A\Diamond\Box(C_A\varphi > C_B\varphi \wedge C_B\varphi > C_A\varphi)$:

$\forall x \in S\forall w, w' \in W : wR_A^x,w' \to \exists z \in T(x,w')\forall z' \in T(z,w') \ \forall X \subseteq S \times W :$

$\quad *(z', w', \{(y,v) \in S \times W \mid \forall v' \in W : vR_A^y v' \to (y, v') \in X\})$
$\quad \subseteq \{(y,v) \in S \times W \mid \forall v \in W : vR_B^y v' \to (y, v') \in X\}$

$\wedge *(z', w', \{(y,v) \in S \times W \mid \forall v' \in W : vR_B^y v' \to (y, v') \in X\})$
$\quad \subseteq \{(y,v) \in S \times W \mid \forall v' \in W : vR_A^y v' \to (y, v') \in X\}$

(d) $C_A\varphi > C_B C_A\varphi$:

$\forall X \subseteq S \times W, \forall(x,w) \in S \times W :$
$\quad *(x, w, \{(y,v) \in S \times W \mid \forall v' \in W : vR_A^y v' \to (y, v') \in X\})$
$\quad \subseteq \{(y,v) \in S \times W \mid \forall v' \in W : (\exists v'' : vR_B^y v'' \wedge v''R_A^y v') \to (y, v') \in X\}$

Analogously for the axioms for speaker $B$.

These look complicated, but are in fact just structural decodings of the truth-conditions of the defeasible conditionals involved. I demonstrate here the decoding for (a); the other proofs are analogous.

**A.2.2.** THEOREM (Soundness). *On tree models $M$ where the axioms for (a) from Definition A.2.1 holds: $\forall x \in S \ \forall w \in W : M, x, w \models C_A\varphi > C_A\Box C_A\varphi$.*

**Proof:**

Fix $x$ and $w$. It is to show that $*(x, w, [\![C_A\varphi]\!]) \subseteq [\![C_A \Box C_A \varphi]\!]$. Instantiate the axiom for (a) for $X = [\![\varphi]\!]$ to obtain the sets:

$$
* (x, w, \{(y, v) \in S \times W \mid \forall v' \in W : vR_A^y v' \to (y, v') \in [\![\varphi]\!]\})
$$
$$
= * (x, w, \{(y, v) \in S \times W \mid \forall v' \in W : vR_A^y v' \to M, y, v' \models \varphi\})
$$
$$
= * (x, w, \{(y, v) \in S \times W \mid M, y, v \models C_A\varphi\})
$$
$$
= * (x, w, \{[\![C_A\varphi]\!]\})
$$
$$
\subseteq \{(y, v) \in S \times W \mid \forall v' \in W
$$
$$
\left( vR_A^y v' \to \forall y' \in T(y, v')\Big(\forall v'' \in W(v'R_A^{y'} v'' \to (y', v'') \in [\![\varphi]\!])\Big)\right)\}
$$
$$
= \{(y, v) \in S \times W \mid \forall v' \in W
$$
$$
\left( vR_A^y v' \to \forall y' \in T(y, v')\Big(\forall v'' \in W(v'R_A^{y'} v'' \to M, y', v'' \models \varphi)\Big)\right)\}
$$
$$
= \{(y, v) \in S \times W \mid \forall v' \in W\left(vR_A^y v' \to \forall y' \in T(y, v')(M, y', v' \models C_A\varphi)\right)\}
$$
$$
= \{(y, v) \in S \times W \mid \forall v' \in W(vR_A^y v' \to \forall M, y, v' \models \Box C_A\varphi)\}
$$
$$
= \{(y, v) \in S \times W \mid M, y, v \models C_A \Box C_A\varphi\}
$$
$$
= [\![C_A \Box C_A\varphi]\!]
$$

This is precisely the truth-condition for $C_A\varphi > C_A \Box C_A\varphi$. $\qquad\qquad\square$

## A.3   Chapter 4

**Proof** of $(-\Diamond\text{I.}/+\neg\text{E.})$:

$$
\cfrac{+\neg A \qquad \cfrac{\cfrac{\cfrac{[+\Diamond A]^1}{\oplus A}\,(+\Diamond\text{E.})}{-\neg A}\,(-\neg\text{I.})}{}\,(\text{Rejection})}{\cfrac{\bot}{-\Diamond A}\,(\text{SR})^1}
$$

**Proof** of $(-\Diamond\text{E.}/+\neg\text{I.})$:

$$
\cfrac{-\Diamond A \qquad \cfrac{\cfrac{\cfrac{[-\neg A]^1}{\oplus A}\,(-\neg\text{E.})}{+\Diamond A}\,(+\Diamond\text{I.})}{}\,(\text{Rejection})}{\cfrac{\bot}{+\neg A}\,(\text{SR})^1}
$$

**Proof** of classical inference: *Frege's Propositional Calculus* is the following set of six axioms plus unrestricted Modus Ponens

1. $A \to (B \to A)$. Proof:

$$\cfrac{\cfrac{[+A]^1 \qquad [+B]^2}{+A} \text{ (Repetition)}}{\cfrac{+(B \to A)}{+(A \to (B \to A))} \text{ (Ded*)}^2} \text{ (Ded*)}^1$$

2. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$. Proof:

$$\cfrac{\cfrac{\cfrac{[+(A \to (B \to C))]^1 \qquad [+A]^3}{+(B \to C)} \text{ (MP)} \qquad \cfrac{[+(A \to B)]^2 \qquad [+A]^3}{+B} \text{ (MP)}}{+C} \text{ (MP)}}{\cfrac{+(A \to C)}{\cfrac{+((A \to B) \to (A \to C))}{+((A \to (B \to C)) \to ((A \to B) \to (A \to C)))} \text{ (Ded*)}^1} \text{ (Ded*)}^2} \text{ (Ded*)}^3$$

3. $(A \to (B \to C)) \to (B \to (A \to C))$. Proof:

$$\cfrac{\cfrac{\cfrac{[+(A \to (B \to C))]^1 \qquad [+A]^3}{+(B \to C)} \text{ (MP)} \qquad [+B]^2}{+C} \text{ (MP)}}{\cfrac{+(A \to C)}{\cfrac{+(B \to (A \to C))}{+((A \to (B \to C)) \to (B \to (A \to C)))} \text{ (Ded*)}^1} \text{ (Ded*)}^2} \text{ (Ded*)}^3$$

4. $(A \to B) \to (\neg B \to \neg A)$. Proof:

$$\cfrac{[+\neg B]^1 \qquad \cfrac{\cfrac{\cfrac{[+(A \to B)]^2 \qquad [+A]^3}{+B} \text{ (MP)} \qquad \cfrac{[-\neg A]^4}{\oplus A} \text{ (}-\neg\text{E.)}}{\oplus B} \text{ (Weak Inference)}^3}{-\neg B} \text{ (}-\neg\text{I.)}}{\cfrac{\cfrac{\bot}{+\neg A} \text{ (SR)}^4}{\cfrac{+(\neg B \to \neg A)}{+((A \to B) \to (\neg B \to \neg A))} \text{ (Ded*)}^2} \text{ (Ded*)}^1} \text{ (Rejection)}$$

Note that while there is a dischargeable assumption that is not strongly asserted, this does not preclude application of Weak Inferences or the Deduction principle: The assumption $-\neg A$ is outside the subproof relevant for Weak Inference and is discharged before (Ded*) is applied.

5. $\neg\neg A \to A$. Proof in Chapter 4.

6. $A \rightarrow \neg\neg A$. Proof in Chapter 4.

Thus, EML validates all of Frege's axioms plus unrestricted *modus ponens*. Thus, it validates classical inference. Note moreover that the usual propositional logic proofs then demonstrate that $\wedge$ and $\vee$, as defined by their introduction/elimination rules, obey classical inference as well.

**A.3.1.** THEOREM (Soundness). *Epistemic multilateral logic is S5-sound.*

**Proof:**
The proof proceeds by induction on the length $n$ of derivations. We write $\Gamma \vdash^D \varphi$ if $D$ is a derivation of $\varphi$ from the premises in $\Gamma$. The base clause, $n = 1$ is trivial as $\Gamma \vdash^D \varphi$ with $|D| = 1$ only if $\varphi \in \Gamma$.

Now, for each $k \leq n$ and any set of premises $\Gamma$, assume that if $\Gamma \vdash^D \psi$, $|D| = k$, then $\Gamma \models \psi$. Then assume that $\Gamma \vdash^D \varphi$ for some $D$ with $|D| = n + 1$ and show that $\Gamma \models \varphi$.

- (Rejection). No S5-model can satisfy $\Box A$ and $\Diamond \neg A$. Hence $+A$ and $-A$ cannot be jointly satisfied. Hence if there are subderivations of $+A$ and $-A$ from $\Gamma$, then by induction, $\Gamma$ has no model, i.e. $\Gamma \models \bot$.

- (SR$_1$). Assume that the last step in the derivation $D$ is an application of (SR$_1$) to show $-A$. Then there is a derivation $D'$ of shorter length with $\Gamma, +A \vdash^{D'} \bot$. By induction, $\Gamma \cup \{+A\} \models \bot$, i.e. $\Gamma \cup \{+A\}$ is unsatisfiable. If already $\Gamma$ is unsatisfiable, then trivially $\Gamma \models -A$. If $\Gamma$ is satisfiable then let $V$ be any model of $\Gamma$. By assumption, $V, w^V \not\models^{S5} \Box A$, i.e. $V, w^V \models \neg \Box A$ which is equivalent to $V, w^V \models \Diamond \neg A$. That is $V \models -A$.

- (SR$_2$). Analogously to (SR$_1$).

- (Assertion). Assume that the last step in $D$ is an application of (Assertion) to show $\oplus A$. Then there is a derivation $D'$ of shorter length of $+A$. By induction, $\Gamma \models +A$. That is, for every model $V$ of $\Gamma$, $V, w^V \models^{S5} \Box A$. By Fact 3, $V, w^V \models^{S5} \Diamond A$, i.e. $V \models \oplus A$.

- (Weak Inference). Assume that the last step in $D$ is an application of (Weak Inference) to show $\oplus B$. Then there is a subderivation of $+B$ from $+A$ satisfying the restrictions of the Deduction principle. Thus, by the soundness proof of (Ded*) (see Chapter 4), $\Gamma \models +(A \rightarrow B)$.

  Moreover, there is a shorter derivation of $\oplus A$. By induction, $\Gamma \models \oplus A$. Thus, for every model $V$ of $\Gamma$: $V, w^V \models^{S5} \Box(A \rightarrow B) \wedge \Diamond A$. By standard modal reasoning (Axiom K), this entails $V, w^V \models^{S5} \Diamond B$. That is, $V \models \oplus B$.

- $(+\Diamond I.)$ and $(+\Diamond E.)$. For any model $V$: $V \models \oplus A$ iff $V, w^V \models^{S5} \Diamond A$ iff (Fact 1) $V, w^V \models^{S5} \Box\Diamond A$ iff $V \models +\Diamond A$.

- $(\oplus\Diamond I.)$ and $(\oplus\Diamond E.)$. For any model $V$: $V \models \oplus A$ iff $V, w^V \models^{S5} \Diamond A$ iff (Fact 2) $V, w^V \models^{S5} \Diamond\Diamond A$ iff $V \models \oplus\Diamond A$.

- $(-\neg I.)$ and $(-\neg E.)$. For any model $V$: $V \models \oplus A$ iff $V, w^V \models^{S5} \Diamond A$ iff $V, w^V \models^{S5} \Diamond\neg\neg A$ iff $V \models -\neg A$.

- $(\oplus\neg I.)$ and $(\oplus\neg E.)$. For any model $V$: $V \models -A$ iff $V, w^V \models^{S5} \Diamond\neg A$ iff $V \models \oplus\neg A$.

- $(+\wedge I.)$ and $(+\wedge E.)$. For any model $V$: $V \models +A$ and $V \models +B$ $V \models \Box A$ and $V \models \Box B$ iff $V \models \Box(A \wedge B)$ iff $V \models +(A \wedge B)$.

- $(+\vee I.)$. For any model $V$: If $V \models +A$ then $V \models \Box A$, then $V \models \Box(A \vee B)$. Thus, $V \models +(A \vee B)$.

- $(+\vee E.)$ Suppose the last step in $D$ is an application of $(+\vee E.)$ to derive $\varphi$. Then, there are shorter derivations $D', D'', D'''$ such that $\Gamma \vdash^{D'} +(A \vee B)$ and $\Gamma, +A \vdash^{D''} \varphi$ and $\Gamma, +B \vdash^{D'''} \varphi$. Do a case distinction on the sign of $\varphi$. If $\varphi = +C$ for some $C$ let $\varphi' = C$; if $\varphi = \oplus C$ let $\varphi' = \Diamond C$; if $\varphi = -C$ let $\varphi' = \Diamond\neg C$; if $\varphi = \bot$, let $\varphi' = \neg(A \to A)$.

  By induction and the Soundness of (Ded*), $\Gamma \models +(A \vee B)$, $\Gamma \models +(A \to \varphi')$ and $\Gamma \models +(B \to \varphi')$. That is, $\Gamma \models \Box(A \vee B)$, $\Gamma \models \Box(A \to \varphi')$ and $\Gamma \models \Box(B \to \varphi')$. Hence $\Gamma \models \Box\varphi'$. Then, by checking the four cases, it can be seen that $\Gamma \models \varphi$.

- (MP). Assume that the last step in $D$ is an application of (MP) to show $\oplus B$. Then there are shorter derivations of $+A$ and $+(A \to B)$. By induction, $\Gamma \models +A$ and $\Gamma \models +(A \to B)$. That is, for any model $V$ of $\Gamma$: $V, w^V \models^{S5} \Box(A \to B)$ and $V, w^V \models^{S5} \Box A$. By standard modal reasoning (Axiom K), then $V, w^V \models^{S5} \Box B$, i.e. $V \models +B$.

$\square$

**Proof** that $(+\mathcal{A}I.)$ and $(+\mathcal{A}E.)$ are equivalent to $(\oplus\mathcal{D}I.)$ and $(\oplus\mathcal{D}E.)$. (If $\mathcal{A} \equiv \neg\mathcal{D}$).

- $(+\mathcal{A}I.)$ and $(+\mathcal{A}E.)$ entail $(\oplus\mathcal{D}I.)$.

$$\cfrac{-p \qquad \cfrac{\cfrac{[+\mathcal{A}(``p")]^1}{+p}\ (+\mathcal{A}\text{E.})}{} }{\cfrac{\cfrac{\cfrac{\bot}{-\mathcal{A}(``p")}\ (\text{SR}_1)^1}{\oplus\neg\mathcal{A}(``p")}\ (\oplus\neg\text{I.})}{\oplus\mathcal{D}(``p")}\ (\text{Identity})}\ (\text{Rejection})$$

- $(+\mathcal{A}\mathrm{I}.)$ and $(+\mathcal{A}\mathrm{E}.)$ entail $(\oplus\mathcal{D}\mathrm{E}.)$.

$$
\cfrac{\cfrac{\cfrac{\oplus\mathcal{D}(\text{``}p\text{''})}{\oplus\neg\mathcal{A}(\text{``}p\text{''})}\;(\text{Identity})}{-\mathcal{A}(\text{``}p\text{''})}\;(\oplus\neg\mathrm{E}.) \qquad \cfrac{[+p]^1}{+\mathcal{A}(\text{``}p\text{''})}\;(+\mathcal{A}\mathrm{I}.)}{\cfrac{\bot}{-p}\;(\mathrm{SR}_1)^1}\;(\text{Rejection})
$$

- $(\oplus\mathcal{D}\mathrm{I}.)$ and $(+\mathcal{D}\mathrm{E}.)$ entail $(+\mathcal{A}\mathrm{I}.)$.

$$
\cfrac{+p \qquad \cfrac{\cfrac{\cfrac{[-\mathcal{A}(\text{``}p\text{''})]^1}{-\neg\mathcal{D}(\text{``}p\text{''})}\;(\text{Identity})}{\oplus\mathcal{D}(\text{``}p\text{''})}\;(-\neg\mathrm{E}.)}{-p}\;(\oplus\mathcal{D}\mathrm{E}.)}{\cfrac{\bot}{+\mathcal{A}(\text{``}p\text{''})}\;(\mathrm{SR}_2)^1}\;(\text{Rejection})
$$

- $(\oplus\mathcal{D}\mathrm{I}.)$ and $(\oplus\mathcal{D}\mathrm{E}.)$ entail $(+\mathcal{A}\mathrm{E}.)$.

$$
\cfrac{\cfrac{+\mathcal{A}(\text{``}p\text{''})}{+\neg\mathcal{D}(\text{``}p\text{''})}\;(\text{Identity}) \qquad \cfrac{\cfrac{[-p]^1}{\oplus\mathcal{D}(\text{``}p\text{''})}\;(\oplus\mathcal{D}\mathrm{I}.)}{-\neg\mathcal{D}(\text{``}p\text{''})}\;(-\neg\mathrm{I}.)}{\cfrac{\bot}{+p}\;(\mathrm{SR}_2)^1}\;(\text{Rejection})
$$

**Proof** of (CNI*):

$$
\cfrac{\cfrac{\cfrac{\cfrac{[+A]^1}{\vdots}}{\bot} \quad \cfrac{[-\neg(A\to A)]^2}{\cfrac{\bot}{+\neg(A\to A)}\;(\mathrm{SR}_2)^2}\;(\text{Repetition})}{\cfrac{\oplus\neg(A\to A)}{-(A\to A)}\;(\oplus\neg\mathrm{E}.)} \qquad \cfrac{\cfrac{[-\neg A]^3}{\oplus A}\;(-\neg\mathrm{E}.)}{}\;(\mathrm{WI})^1 \qquad \cfrac{\cfrac{[+A]^4}{+A}\;(\text{Repetition})}{+(A\to A)}\;(\mathrm{Ded*})^4}{\cfrac{\bot}{+\neg A}\;(\mathrm{SR}_2)^3}\;(\text{Rejection})
$$

# Bibliography

Anscombre, JC (1995). La théorie des topoi: Sémantique ou rhétorique? *Hermés, La Revue*, 1(15):185–198.

Asher, N (1999). Discourse and the focus/background distinction. In: Bosch, P & van der Sandt, R (eds.), *Focus: Linguistic, Cognitive, and Computational Perspectives*, 247–267. Cambridge University Press.

Asher, N (2008). Troubles on the right frontier. In: Benz, A & Kühnlein, P (eds.), *Constraints in Discourse*, 29–52. John Benjamins Publishing Company.

Asher, N & Lascarides, A (1998). The semantics and pragmatics of presupposition. *Journal of Semantics*, 15(3):239–300.

Asher, N & Lascarides, A (2003). *Logics of Conversation.* Cambridge University Press.

Asher, N & Lascarides, A (2008). Commitments, beliefs and intentions in dialogue. In: Ginzburg, J, Healy, P & Sato, Y (eds.), *Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue*, 29–36.

Asher, N & Lascarides, A (2011). Reasoning dynamically about what one says. *Synthese*, 183(1):5–31.

Asher, N & Lascarides, A (2013). Strategic conversation. *Semantics and Pragmatics*, 6(2):1–62.

Asher, N & Morreau, M (1990). Commonsense entailment: A modal theory of nonmonotonic reasoning. In: *European Workshop on Logics in Artificial Intelligence*, 1–30.

Asher, N & Paul, S (2013). Infinite games with uncertain moves. In: *1st International Workshop on Strategic Reasoning*, 25–32.

Asher, N & Paul, S (2016). Evaluating conversational success: Weighted message exchange games. In: Hunter, J, Simons, M & Stone, M (eds.), *20th Workshop on the Semantics and Pragmatics of Dialogue*, 55–64.

Asher, N & Reese, B (2007). Intonation and discourse: Biased questions. In: Ishihara, S, Jannedy, S & Schwarz, A (eds.), *Interdisciplinary Studies on Information Structure*, vol. 8, 1–38.

Asher, N, Paul, S & Venant, A (2017). Message exchange games in strategic contexts. *Journal of Philosophical Logic*, 46(4):355–404.

Austin, JL (1962). *How to do Things with Words*. Clarendon Press.

Bacon, A (2015). Can the classical logician avoid the revenge paradoxes? *Philosophical Review*, 124(3):299–352.

Baker, C & Robson, J (2015). An absolutist theory of faultless disagreement in aesthetics. *Pacific Philosophical Quarterly*, 98(3):429–448.

Beaver, DI & Clark, BZ (2009). *Sense and Sensitivity: How Focus Determines Meaning*. Wiley–Blackwell.

Bellert, I (1977). On semantic and distributional properties of sentential adverbs. *Linguistic Inquiry*, 8(2):337–351.

Bendall, K (1979). Negation as a sign of negative judgment. *Notre Dame Journal of Formal Logic*, 20(1):68–76.

Bledin, J & Rawlins, K (2016). Epistemic resistance moves. In: Moroney, M, Little, CR, Collard, J & Burgdorf, D (eds.), *Proceedings of Semantics and Linguistic Theory 26*, vol. 26, 620–640.

Bolinger, D (1972). Accent is predictable (if you're a mind-reader). *Language*, 48(3):633–644.

Bolinger, D (1982). Intonation and its parts. *Language*, 58(3):505–533.

Bolinger, D (1985). Two views of accent. *Journal of Linguistics*, 21(1):79–123.

Brandom, R (1983). Asserting. *Noûs*, 17(4):637–650.

Breitholtz, E (2010). Clarification requests as enthyme elicitors. In: Łupkowski, P & Purver, M (eds.), *Proceedings of the 14th Workshop on the Semantics and Pragmatics of Dialogue*, 135–138.

Breitholtz, E (2014). *Enthymemes in Dialogue: A Micro-Rhetorical Approach*. Ph.D. thesis, University of Gothenburg.

Breitholtz, E & Villing, J (2008). Can Aristotelian enthymemes decrease the cognitive load of a dialogue system user? In: Ginzburg, J, Healy, P & Sato, Y (eds.), *Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue*, 94–100.

Bromberger, S (1992). *On What We Know We Don't Know.* University of Chicago Press.

Bryant, GA (2010). Prosodic contrasts in ironic speech. *Discourse Processes*, 47(7):545–566.

Bryant, GA & Fox Tree, JE (2005). Is there an ironic tone of voice? *Language and speech*, 48(3):257–277.

Büring, D (2006). Focus projection and default prominence. In: Molnár, V & Winkler, S (eds.), *The Architecture of Focus*, 321–346. de Gruyter.

Büring, D (2007). Semantics, intonation and information structure. *The Oxford Handbook of Linguistic Interfaces.*

Burnard, L (2000). *Reference Guide for the British National Corpus (World Edition).* Oxford University Computing Services.

Calhoun, S (2007). *Information Structure and the Prosodic Structure of English: A Probabilistic Relationship.* Ph.D. thesis, University of Edinburgh.

Carletta, J (2007). Unleashing the killer corpus: experiences in creating the multi-everything AMI Meeting Corpus. *Language Resources and Evaluation*, 41(2):181–190.

Carroll, L (1895). What the tortoise said to achilles. *Mind*, 4(14):278–280.

Carston, R (1996). Metalinguistic negation and echoic use. *Journal of Pragmatics*, 25:309–330.

Clark, HH (1996). *Using language.* Cambridge University Press.

Cohen, PR & Levesque, HJ (1990). Intention is choice with commitment. *Artificial intelligence*, 42(2-3):213–261.

Crabill, JD (2013). Suppose yalcin is wrong about epistemic modals. *Philosophical Studies*, 162(3):625–635.

Crespo, I & Fernández, R (2011). Expressing taste in dialogue. In: Artstein, R, Core, M, DeVault, D, Georgila, K, Kaiser, E & Stent, A (eds.), *Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue*, 84–93.

Dancygier, B & Sweetser, E (2005). *Mental spaces in grammar: Conditional constructions*. Cambridge University Press.

Dickie, I (2010). Negation, anti-realism, and the denial defence. *Philosophical Studies*, 150(2):161–185.

Dryer, MS (1996). Focus, pragmatic presupposition, and activated propositions. *Journal of Pragmatics*, 26(4):475–523.

Ducrot, O (1980). *Les échelles argumentatives*. Les Editions de Minuit.

Ducrot, O (1988). Topoï et formes topique. *Bulletin d'études de la linguistique française*, 22:1–14.

Dummett, M (1991). *The Logical Basis of Metaphysics*. Harvard University Press.

Dummett, M (2002). 'Yes', 'No' and 'Can't Say'. *Mind*, 111(442):289–295.

Dutilh Novaes, C (2015). A dialogical, multi-agent account of the normativity of logic. *dialectica*, 69(4):587–609.

Edgington, D (1995). On conditionals. *Mind*, 104(414):235–329.

Egan, A (2012). Relativist dispositional theories of value. *The Southern Journal of Philosophy*, 50(4):557–582.

Egan, A (2014). There's something funny about comedy: A case study in faultless disagreement. *Erkenntnis*, 79(1):73–100.

Ernst, T (2009). Speaker-oriented adverbs. *Natural Language & Linguistic Theory*, 27(3):497–544.

Farkas, DF & Bruce, KB (2010). On reacting to assertions and polar questions. *Journal of semantics*, 27(1):81–118.

Fernández, R, Ginzburg, J & Lappin, S (2004). Classifying ellipsis in dialogue: A machine learning approach. In: Lemnitzer, L, Meurers, D & Hinrichs, E (eds.), *Proceedings of the 20th International Conference on Computational Linguistics*, 240–246.

Field, HH (2008). *Saving Truth from Paradox*. Oxford University Press Oxford.

von Fintel, K & Gillies, A (2007). An opinionated guide to epistemic modality. In: *Oxford Studies in Epistemology*, vol. 2, 32–62. Oxford University Press.

von Fintel, K & Gillies, AS (2008). Cia leaks. *Philosophical review*, 117(1):77–98.

von Fintel, K & Gillies, AS (2011). "might" made right. In: Egan, A & Weatherson, B (eds.), *Epistemic Modality*, 108–130. Oxford University Press.

van Fraassen, BC (1980). *The Scientific Image.* Oxford University Press.

Frege, G (1919). Die Verneinung: Eine logische Untersuchung. *Beiträge zur Philosophie des deutschen Idealismus*, 1:143–157.

Frege, G (1923). Gedankengefüge. *Beiträge zur Philosophie des deutschen Idealismus*, 3(1):36–51.

Geach, PT (1965). Assertion. *The Philosophical Review*, 74(4):449–465.

Gentzen, G (1935). Untersuchungen über das logische Schließen I. *Mathematische Zeitschrift*, 39(1):176–210.

Geurts, B & van der Sandt, R (2004). Interpreting focus. *Theoretical Linguistics*, 30:1–44.

Ginzburg, J (2012). *The Interactive Stance.* Oxford University Press.

Giora, R (1995). On irony and negation. *Discourse Processes*, 19(2):239–264.

Glanzberg, M (2007). Context, content, and relativism. *Philosophical Studies*, 136(1):1–29.

Goodhue, D & Wagner, M (2015). It's not just what you say, it's how you say it: Intonation, *yes* and *no.* In: Ozyildiz, D & Bui, T (eds.), *Proceedings of the 45th Meeting of the North-East Linguistics Society.*

Goranko, V & Galton, A (2015). Temporal logic. In: Zalta, EN (ed.), *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, winter 2015 ed.

Grice, HP (1975). Logic and conversation. In: Cole, P & Morgan, JL (eds.), *Syntax and semantics*, vol. 3. New York: Academic Press.

Grice, HP (1991). *Studies in the Way of Words.* Harvard University Press.

Groenendijk, J, Stokhof, M & Veltman, F (1996). Coreference and modality. In: Lappin, S (ed.), *Handbook of Contemporary Semantic Theory*, 179–216. Blackwell.

Groenendijk, JAG & Stokhof, MJB (1984). *On the semantics of questions and the pragmatics of answers.* Ph.D. thesis, University of Amsterdam.

Gunlogson, C (2003). *True to Form: Rising and Falling Declaratives as Questions in English.* Routledge.

Gussenhoven, C (1983). Focus, mode and the nucleus. *Journal of Linguistics*, 19(02):377–417.

Halliday, MA (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, 3(02):199–244.

Hamblin, C (1970). *Fallacies*. Metheun.

Hamblin, CL (1971). Mathematical models of dialogue. *Theoria*, 37(2):130–155.

Harman, G (1986). *Change in View*. MIT Press.

Hawke, P & Steinert-Threlkeld, S (2016). Informational dynamics of epistemic possibility modals. *Synthese*.

Hedberg, N, Sosa, JM & Fadden, L (2004). Meanings and configurations of questions in english. In: *Speech Prosody*, 309–312.

Hempel, CG (1965). *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. Free Press.

Hintikka, J & Halonen, I (1995). Semantics and pragmatics for why-questions. *The Journal of Philosophy*, 92(12):636–657.

Hobbs, J (1990). The Pierrehumbert–Hirschberg theory of intonational meaning made simple: Comments on Pierrehumbert and Hirschberg. In: P R Cohen, JM & Pollack, ME (eds.), *Intentions in Communication*, 313–323. MIT Press.

Hobbs, JR, Stickel, ME, Appelt, DE & Martin, P (1993). Interpretation as abduction. *Artificial Intelligence*, 63(1-2):69–142.

Horn, L (1989). *A Natural History of Negation*. University of Chicago Press.

Hulstijn, J (2000). *Dialogue models for inquiry and transaction*. Ph.D. thesis, University of Twente.

Huvenes, T (forthcoming). On disagreement. In: Ichikawa, J (ed.), *Routledge Handbook of Epistemic Contextualism*. Routledge.

Huvenes, TT (2012). Varieties of disagreement and predicates of taste. *Australasian Journal of Philosophy*, 90(1):167–181.

Incurvati, L & Schlöder, JJ (2017). Weak rejection. *Australasian Journal of Philosophy*, 95(4):741–760.

Incurvati, L & Schlöder, JJ (under review). Weak assertion.

Incurvati, L & Smith, P (2009). Rejection and valuations. *Analysis*, 3–10.

Jackendoff, RS (1972). *Semantic Interpretation in Generative Grammar*. MIT Press.

Jackson, S & Jacobs, S (1980). Structure of conversational argument: Pragmatic bases for the enthymeme. *Quarterly Journal of Speech*, 66(3):251–265.

Johansson, I (1937). Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus. *Compositio mathematica*, 4:119–136.

Kamp, H (1973). Free choice permission. In: *Proceedings of the Aristotelian Society*, vol. 74, 57–74. JSTOR.

Kamp, H & Reyle, U (1993). *From Discourse to Logic*. Springer.

Khoo, J (2015). Modal disagreements. *Inquiry*, 58(5):511–534.

Khoo, J & Knobe, J (2016). Moral disagreement and moral semantics. *Noûs*, early access.

Kölbel, M (2004). Faultless disagreement. In: *Proceedings of the Aristotelian Society*, vol. 104, 53–73.

Kölbel, M (2008). Motivations for relativism. In: García-Carpintero, M & Köbel, M (eds.), *Relative Truth*, 1–38. Oxford University Press.

Kolodny, N & MacFarlane, J (2010). Ifs and oughts. *The Journal of Philosophy*, 107(3):115–143.

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

Kratzer, A (1991). Conditionals. In: von Stechow, A & Wunderlich, D (eds.), *Semantics (Handbücher zur Sprach- und Kommunkationswissenschaft 6)*, 651–656. Walter de Gruyter.

Krifka, M (1992). A compositional semantics for multiple focus constructions. In: Jacobs, J (ed.), *Informationsstruktur und Grammatik*, 17–53. Westdeutscher Verlag.

Krifka, M (2013). Response particles as propositional anaphors. In: Snider, T (ed.), *Proceedings of Semantics and Linguistic Theory 23*, vol. 23, 1–18.

Krifka, M (2015). Bias in commitment space semantics: Declarative questions, negated quetions, and question tags. In: D'Antonio, S, Moroney, M & Little, CR (eds.), *Proceedings of Semantics and Linguistic Theory 25*, vol. 25, 328–345.

Krzyżanowska, K, Wenmackers, S & Douven, I (2013). Inferential conditionals and evidentiality. *Journal of Logic, Language and Information*, 22(3):315–334.

Kumon-Nakamura, S, Glucksberg, S & Brown, M (1995). How about another piece of pie: The allusional pretense theory of discourse irony. *Journal of Experimental Psychology: General*, 124(1):3.

Ladd, DR (1980). *The Structure of Intonational Meaning: Evidence from English*. Indiana University Press.

Lascarides, A & Asher, N (2009). Agreement, disputes and commitments in dialogue. *Journal of Semantics*, 26(2):109–158.

Lasersohn, P (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, 28(6):643–686.

Lasersohn, P (2009). Relative truth, speaker commitment, and control of implicit arguments. *Synthese*, 166(2):359–374.

Liberman, M & Sag, I (1974). Prosodic form and discourse function. In: *10th Regional Meeting of the Chicago Linguistics Society*, 416–427.

López de Sa, D (2015). Expressing disagreement: a presuppositional indexical contextualist relativist account. *Erkenntnis*, 80(1):153–165.

MacFarlane, J (2007). Relativism and disagreement. *Philosophical Studies*, 132(1):17–31.

MacFarlane, J (2014). *Assessment Sensitivity: Relative Truth and its Applications*. Oxford University Press.

Marques, T (2016). Aesthetic predicates: a hybrid dispositional account. *Inquiry*, 59(6):723–751.

Martin, R (1992). Irony and universe of belief. *Lingua*, 87(1-2):77–90.

McFarland, K (2015). *Feigning Objectivity: An Overlooked Conversational Strategy in Everyday Disputes*. Ph.D. thesis, The Ohio State University.

McGee, V (1990). *Truth, Vagueness and Paradox*. Hackett Publishing Company.

Merchant, J (2005). Fragments and ellipsis. *Linguistics and Philosophy*, 27(6):661–738.

Mittwoch, A (1977). How to refer to one's own words: speech-act modifying adverbials and the performative analysis. *Journal of Linguistics*, 13(02):177–189.

Moltmann, F (2010). Relative truth and the first person. *Philosophical Studies*, 150(2):187–220.

Moore, JD & Pollack, ME (1992). A problem for RST: The need for multi-level discourse analysis. *Computational linguistics*, 18(4):537–544.

Moss, S (2015). On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics*, 8(5):1–81.

Murzi, J & Carrara, M (2015). Denial and disagreement. *Topoi*, 34(1):109–119.

Oliver, A & Smiley, T (2004). Multigrade predicates. *Mind*, 113(452):609–681.

Parsons, T (1984). Assertion, denial, and the liar paradox. *Journal of Philosophical Logic*, 13(2):137–152.

Pierrehumbert, J & Hirschberg, J (1990). The meaning of intonational contours in the interpretation of discourse. In: Cohen, PR, J, M & Pollack, ME (eds.), *Intentions in communication*, 271–311. MIT Press.

Pinker, S, Nowak, M & Lee, J (2008). The logic of indirect speech. *Proceedings of the National Academy of Science*, 105(3).

Pnueli, A (1977). The temporal logic of programs. In: *18th Annual Symposium on Foundations of Computer Science,*, 46–57.

Prawitz, D (1965). Natural deduction: a proof-theoretical study. *Almqvist & Wicksell*.

Priest, G (1979). The logic of paradox. *Journal of Philosophical Logic*, 8(1):219–241.

Priest, G (1997). Sylvan's box: A short story and ten morals. *Notre Dame Journal of Formal Logic*, 38(4):573–582.

Prior, AN (1960). The runabout inference ticket. *Analysis*, 21(2):38–39.

Rabern, B (2012). Against the identification of assertoric content with compositional value. *Synthese*, 189(1):75–96.

Reese, B (2007). *Bias in Questions*. Ph.D. thesis, University of Texas at Austin.

Restall, G (2005). Multiple conclusions. In: Hájek, P, Valdés-Villanueva, L & Westerståhl, D (eds.), *Logic, methodology and philosophy of science: Proceedings of the twelfth international congress*, 189–205. Kings College Publications.

Ripley, D (2011). Negation, denial, and rejection. *Philosophy Compass*, 6(1):622–629.

Roberts, C (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69.

Rooth, M (1992). A theory of focus interpretation. *Natural Language Semantics*, 1(1):75–116.

Rooth, M (1999). Association with focus or association with presupposition? *Focus: Linguistic, cognitive, and computational perspectives*, 232–244.

Rumfitt, I (1997). The categoricity problem and truth-value gaps. *Analysis*, 57(4):223–236.

Rumfitt, I (2000). "Yes" and "No". *Mind*, 109(436):781–823.

Rumfitt, I (2014). Truth and meaning. *Aristotelian Society Supplementary Volume*, 88(1):21–55.

Sæbø, KJ (2009). Judgment ascriptions. *Linguistics and Philosophy*, 32(4):327–352.

Šafářová, M (2005). The semantics of rising intonation and interrogatives. In: Maier, E, Bary, C & Huitink, J (eds.), *Proceedings of Sinn und Bedeutung 9*, 355–369.

van der Sandt, RA (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(4):333–377.

Schafer, K (2011). Faultless disagreement and aesthetic realism. *Philosophy and Phenomenological Research*, 82(2):265–286.

Schlöder, JJ (2014). *Uptake, Clarification and Argumentation*. Master's thesis, University of Amsterdam.

Schlöder, JJ (2015). A formal semantics of the final rise. In: Kaeshammer, M & Schulz, P (eds.), *Proceedings of the Student Session of the European Summer School in Logic, Language and Information 2015*, 117–127.

Schlöder, JJ (2017). Towards a formal semantics of verbal irony. In: Howes, C & Rieser, H (eds.), *Proceedings of the Workshop on Formal Approaches to the Dynamics of Linguistic Interaction*, 55–59.

Schlöder, JJ (under review). How to be a realist about taste.

Schlöder, JJ & Fernández, R (2014). Clarification requests on the level of uptake. In: Muller, P & Rieser, V (eds.), *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue*, 237–239.

Schlöder, JJ & Fernández, R (2015a). Clarifying intentions in dialogue: A corpus study. In: Purver, M, Sadrzadeh, M & Stone, M (eds.), *Proceedings of the 11th International Conference on Computational Semantics*, 46–51.

Schlöder, JJ & Fernández, R (2015b). Pragmatic rejection. In: Purver, M, Sadrzadeh, M & Stone, M (eds.), *Proceedings of the 11th International Conference on Computational Semantics*, 250–260.

Schlöder, JJ & Lascarides, A (2015). Interpreting English pitch contours in context. In: Howes, C & Larsson, S (eds.), *Proceedings of the 19th Workshop on the Semantics and Pragmatics of Dialogue*, 131–139.

Schlöder, JJ & Lascarides, A (under review). Understanding focus: Tune, placement and coherence.

Schlöder, JJ, Breitholtz, E & Fernández, R (2016). Why? In: Hunter, J, Simons, M & Stone, M (eds.), *Proceedings of the 20th Workshop on the Semantics and Pragmatics of Dialogue*, 5–14.

Schlöder, JJ, Venant, A & Asher, N (2017). Aligning intentions: Acceptance and rejection in dialogue. In: Truswell, R (ed.), *Proceedings of Sinn und Bedeutung 21*, to appear.

Schnieder, B (2010). Expressivism concerning epistemic modals. *The Philosophical Quarterly*, 60(240):601–615.

Schroeder, M (2008). What is the Frege–Geach problem? *Philosophy Compass*, 3(4):703–720.

Schwarzschild, R (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, 7(2):141–177.

Searle, JR (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.

Silverman, K, Beckman, ME, Pitrelli, JF, Ostendorf, M, Wightman, CW, Price, P, Pierrehumbert, J & Hirschberg, J (1992). ToBI: a standard for labeling English prosody. In: *2nd International Conference on Spoken Language Processing*, 867–870.

Skyrms, B (2010). *Signals: Evolution, Learning, and Information*. Oxford University Press.

Smiley, T (1996). Rejection. *Analysis*, 56(1):1–9.

Sperber, D & Wilson, D (1981). Irony and the use-mention distinction. *Philosophy*, 3:143–184.

Stalnaker, R (1978). Assertion. In: Cole, P (ed.), *Pragmatics (Syntax and Semantics 9)*. Academic Press.

Stalnaker, R (1988). Belief attribution and context. In: Grimm, RH & Merrill, DD (eds.), *Contents of Thought*, 140–156. University of Arizona Press.

Steedman, M (2014). The surface-compositional semantics of English intonation. *Language*, 90(1):2–57.

Steinberger, F (2009). *Harmony and Logical Inferentialism*. Ph.D. thesis, University of Cambridge.

Stephenson, T (2007). Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy*, 30(4):487–525.

Stojanovic, I (2007). Talking about taste: disagreement, implicit arguments, and relative truth. *Linguistics and Philosophy*, 30(6):691–706.

Stojnić, U (2017). One's modus ponens: Modality, coherence and logic. *Philosophy and Phenomenological Research*, 95(1):167–214.

Swanson, E (2010). On scope relations between quantifiers and epistemic modals. *Journal of Semantics*, 27(4):529–540.

Tennant, N (1997). *The Taming of the True*. Oxford University Press.

Tennant, N (1999). Negation, absurdity and contrariety. In: Gabbay, D & Wansing, H (eds.), *What is Negation?*, 199–222. Kluwer.

Vallduví, E & Engdahl, E (1996). The linguistic realisation of information packaging. *Linguistics*, 34(3):459–519.

Veltman, F (1985). *Logics for Conditionals*. Ph.D. thesis, University of Amsterdam.

Venant, A & Asher, N (2015). Dynamics of public commitments in dialogue. In: Purver, M, Sadrzadeh, M & Stone, M (eds.), *Proceedings of the 11th International Conference on Computational Semantics*, 272–282.

Venant, A, Asher, N & Degremont, C (2014). Credibility and its attacks. In: Muller, P & Rieser, V (eds.), *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue*, 154–162.

Walker, MA (1996a). Inferring acceptance and rejection in dialogue by default rules of inference. *Language and Speech*, 39(2-3):265–304.

Walker, MA (1996b). Limited attention and discourse structure. *Computational Linguistics*, 22(2):255–264.

Ward, G & Hirschberg, J (1985). Implicating uncertainty: The pragmatics of fall-rise intonation. *Language*, 61(4):747–776.

Ward, G & Hirschberg, J (1988). Intonation and propositional attitude: the pragmatics of L*+ H LH%. In: *Proceedings of the Fifth Eastern States Conference on Linguistics*, 512–522. Ohio State University Press Columbus, OH.

Willer, M (2013). Dynamics of epistemic modality. *Philosophical Review*, 122(1):45–92.

Williamson, T (2000). *Knowledge and its Limits*. Oxford University Press.

Wright, C (2012). Replies Part III: Truth, objectivity, realism, and relativism. In: Coliva, A (ed.), *Mind, Meaning, and Knowledge: Themes from the Philosophy of Crispin Wright*, 418–450. Oxford Scholarship Online.

Yalcin, S (2007). Epistemic modals. *Mind*, 116(464):983–1026.

Yalcin, S (2012). A counterexample to modus tollens. *Journal of Philosophical Logic*, 41(6):1001–1024.

Zimmermann, M (2004). Zum *Wohl*: Diskurspartikeln als Satztypmodifikatoren. *Linguistische Berichte*, 253–286.

Zimmermann, TE (2000). Free choice disjunction and epistemic possibility. *Natural language semantics*, 8(4):255–290.

# Samenvatting

Dit proefschrift gaat over het samenspel tussen de taalhandelingen van *bewering* (Engels: *assertion*) en *afwijzing* (Engels: *rejection*). Ten eerste verdedig ik dat het zinvol is deze twee taalhandelingen te onderscheiden. Dit gaat in tegen de oude consensus—teruggaand naar Frege—dat een afwijzing slechts een negatieve bewering is. Hiertoe introduceer ik de taalhandeling van *zwakke afwijzing* en verdedig deze tegenover haar critici. Op basis van deze taalhandeling ontwikkel ik een logica voor bewering and afwijzing—*zwakke bilaterale logica*—en toon het nut hiervan aan in de analyse van natuurlijke gevolgtrekkingen.

Om dit resultaat toegankelijk te maken voor een algemene theorie van talige interactie, ontwikkel ik een alternatieve karakterisering van zwakke afwijzing in een theorie van *openbare verbintenis*. Ik ontwikkel een algemene theorie van verbintenis gebaseerd op zichzelf staande basisprincipes die zelfs in de context van niet-coöperatieve dialogen gelden. Deze principes verklaren enkele dynamisch-interactieve eigenschappen van bewering en afwijzing en valideren bovendien zwakke bilaterale logica als *de logica die coöperatieve verbintenis behoudt*. Deze theorie wordt verder aangevuld door een theorie van *Waarom*-vragen die verklaart wat het betekent om een verbintenis te kunnen *rechtvaardigen*.

Vervolgens werk ik deze theorie verder uit en combineer ik het zwakke afwijzen met zijn omgekeerde taalhandeling: het zwakke beweren. Dit generaliseert zwakke bilaterale logica naar *epistemische multilaterale logica*. Deze logica heeft twee belangrijke voordelen: het verdedigt klassieke logica als *de logica van beweerde inhoud* tegen argumenten met betrekking tot harmonie, en het verklaart epistemische modale werkwoorden. Dit is reden om een groeiende verzameling van data opnieuw te evalueren die beweren dat klassieke logica onverenigbaar is met een bevredigende semantiek voor epistemische modale werkwoorden.

Een andere uitdaging voor de klassieke semantische theorie is het argument van *schuldloze onenigheid*. Het is welbekend dat oordelen over smaak (en vergelijkbare eigenschappen) taal- en interactie-eigenschappen hebben die ze lijken te onderscheiden van andere oordelen. Sommigen hebben betoogd dat dit een brede herziening

van de semantische theorie vereist. Ik laat zien dat dit niet noodzakelijk het geval is. Met behulp van de theorie van openbare verbintenis, ontwikkel ik een theorie van *herframebare predikaten* en specificeer de dialoogdynamiek van dergelijke predicaten. Als men dan accepteert dat smaakpredicaten herframebaar zijn—zoals de beschikbare data suggereren—kan men de relevante gegevens verklaren zonder de semantische theorie te herzien.

Ten slotte behandel ik een methodologisch probleem. Soms kunnen geschreven gegevens over bepaalde uitingen dubbelzinnig zijn wat betreft instemming en ongenoegen: *hoe* ze werden gezegd geeft de doorslag. Intonatie is zelden geannoteerd en als dat al zo is wordt vaak alleen de plaatsing van het nucleaire accent genoteerd, maar niet de *toon en melodie* die de focus plaatsen. Ik laat zien dat het weglaten van melodie tot verwarring kan leiden: de interpretatie van een uiting met dezelfde plaatsing van het accent kan variëren afhankelijk van de melodie. Daarom is er een verenigde theorie van focus en melodie nodig. Ik zet een eerste stap in het ontwikkelen van een formele pragmatische theorie van focus en melodie voor de Engelse taal. Mijn model legt een aantal interessante melodieën uit, waaronder een die verbale ironie uitdrukt. Deze theorie beschrijft in het bijzonder de relevante data over overeenstemming en onenigheid.

# Abstract

This thesis discusses the interplay between the speech acts of *assertion* and *rejection*. First, it defends that it is even sensible to distinguish these two speech acts, against an old consensus—going back to Frege—that rejection is merely negative assertion. To this end, I introduce the speech act of *weak rejection* and defend it against its critics. Based on this notion, I develop a logic of asserted and rejected content—*weak bilateral logic*—and demonstrate its usefulness in the analysis of natural language inferences.

To make this result accessible to a general theory of linguistic interaction, I give an alternative characterisation of weak rejection in a theory of *public commitment*. I develop a broad theory of commitment from independent, basic principles that hold even in non-cooperative dialogue contexts. These principles predict compelling dynamic-interactive properties of the speech acts of assertion and rejection, and moreover validate weak bilateral logic as *the logic that preserves cooperative commitment*. This theory is further rounded out by a theory of *Why*-questions that explains what it means to be able to *vindicate* one's commitments.

Then, I develop these investigations further by pairing weak rejection with its dual speech act: *weak assertion*. This generalises weak bilateral logic to *epistemic multilateral logic*. This logic has two major upshots: it defends classical logic as *the logic of asserted content* against arguments from harmony, while simultaneously explaining epistemic modals. This is grounds to re-evaluate a growing collection of data that alleges the incompatibility of classical logic and a satisfactory semantics for epistemic modals.

Another challenge to classical semantic theory is the *faultless disagreement* argument. Famously, judgements about taste (and similar properties) have linguistic and interactional properties that appear to distinguish them from other judgements. Some have argued that this calls for wide revisions of semantic theory. I show that this is not necessarily the case. Using a public commitment framework, I develop a theory of *reframable predicates* and specify the dialogue dynamics of such predicates. Then, if one accepts that taste predicates are reframable—as

suggested by the available data—one can explain the relevant phenomena without revising semantic theory.

Finally, I address a methodological issue. Sometimes, written records of certain utterances can be ambiguous between agreement and disagreement: *how* they were said makes the difference. Intonation is rarely annotated and, if so, it is frequently only the placement of the focal stress that is recorded, but not the *utterance tune* that sets the focus. I demonstrate that foregoing tune can lead to confusion: the interpretation of an utterance with the same focal stress can vary depending on the tune. Thus, one needs a unified theory of focus and tune. I develop a first attempt towards a formal discourse theory of focus and tune that already explains a number of interesting tunes, including one that signals verbal irony. This theory in particular explains relevant data on agreement and disagreement.

*Titles in the ILLC Dissertation Series:*

ILLC DS-2009-01: **Jakub Szymanik**
*Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language*

ILLC DS-2009-02: **Hartmut Fitz**
*Neural Syntax*

ILLC DS-2009-03: **Brian Thomas Semmes**
*A Game for the Borel Functions*

ILLC DS-2009-04: **Sara L. Uckelman**
*Modalities in Medieval Logic*

ILLC DS-2009-05: **Andreas Witzel**
*Knowledge and Games: Theory and Implementation*

ILLC DS-2009-06: **Chantal Bax**
*Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.*

ILLC DS-2009-07: **Kata Balogh**
*Theme with Variations. A Context-based Analysis of Focus*

ILLC DS-2009-08: **Tomohiro Hoshi**
*Epistemic Dynamics and Protocol Information*

ILLC DS-2009-09: **Olivia Ladinig**
*Temporal expectations and their violations*

ILLC DS-2009-10: **Tikitu de Jager**
*"Now that you mention it, I wonder...": Awareness, Attention, Assumption*

ILLC DS-2009-11: **Michael Franke**
*Signal to Act: Game Theory in Pragmatics*

ILLC DS-2009-12: **Joel Uckelman**
*More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains*

ILLC DS-2009-13: **Stefan Bold**
*Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.*

ILLC DS-2010-01: **Reut Tsarfaty**
*Relational-Realizational Parsing*

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION