Neg-Raising: Polarity and Presupposition

by

Jon Robert Gajewski

B.A. in Linguistics

University of Michigan 2000

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Linguistics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2005

Author . . . .

Department of Linguistics and Philosophy

August 10, 2005

Certified by . . . .

Irene Heim

Professor, Department of Linguistics and Philosophy

Thesis Supervisor

Accepted by . . . .

Irene Heim

Chair, Department of Linguistics and Philosophy

Neg-Raising: Polarity and Presupposition

by

Jon Robert Gajewski

B.A. in Linguistics

University of Michigan 2000

Submitted to the Department of Linguistics and Philosophy
on August 10, 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Linguistics

Abstract

In this thesis, I advance a semantic theory of Neg-Raising rooted in the work of Bartsch (1973) and Heim (2000) and defend it against syntactic and pragmatic alternatives. The primary source of support for my position on Neg-Raising comes from the natural way in which the approach explains a variety of facts about NPI-licensing in environments containing Neg-Raising predicates. In Chapter 2, a principled account is offered of a previously ill-understood contrast in NPI-licensing under stacked Neg-Raising predicates, first pointed out in Horn (1972). Also addressed are facts advanced in favor of the syntactic theory of Neg-Raising by Kiparsky and Kiparsky (1970) and Prince (1976).

Horn's (1989) attractive account of Neg-Raising is reviewed in detail in Chapter 3 and found to have deficiencies, particularly in the domain of NPI-licensing. The most compelling aspect of Horn's analysis is his derivation of Neg-Raising from general principles. The purposes of Chapters 4 and 5 is to develop an alternative analysis of Neg-Raising that attains a comparable depth of explanation. First, I compare the behavior of negated Neg-Raising predicates to that of negated definite plurals. Next, I show that there is a significant correlation across constructions between obeying the Excluded Middle and having the properties of definite plurals. Finally, I offer a tentative explanation of why definite plurals obey the Excluded Middle.

Thesis Supervisor: Irene Heim
Title: Professor, Department of Linguistics and Philosophy

Acknowledgments

First and foremost, I would like to thank my thesis advisor Irene Heim. Irene's work and professional conduct have always served as the highest model for my own. If not for Irene, I would not be a linguist

Thanks also to Kai von Fintel. The substance of my work owes the greatest debt to Kai's pioneering research. His clarity of thought, insight and pedagogical skill have been a constant source of inspiration.

Thanks to Danny Fox for his constant encouragement and frequent, stimulating conversations. Danny's help has been integral to the development of my ideas.

I also benefited greatly from interacting with the rest of the faculty at MIT, especially Sabine Iatridou, Shigeru Miyagawa, David Pesetsky, and Norvin Richards.

Thanks to my classmates Marketa Ceplova, Aniko Csirmaz, Marcelo Ferreira, Heejeong Ko, Andrew Nevins, Becky Norris, Joey Sabbagh, and Michael Wagner. Marcelo and Michael deserve special mention for how much they generously taught me. I also learned a lot from Valentine Hacquard, Youngjoo Lee, Luis Alonso-Ovalle and Lance Nathan.

In a thanks category of her own is Elena Guerzoni, who filled many roles in my life: Semantics TA, Roommate, Officemate, Advisor (unofficial) and Best Friend (official). Elena knows more about semantics (and in particular NPIs) than is seemly for a human; this thesis bears the marks of her influence. Equally important in my life has been Elena's generous friendship. In the last five years, Elena has become family.

Elena also introduced me to her husband, Martin Hackl, whom I thank for his shrewd counsel on semantics, the job market and Qi Gong.

Many thanks to all the residents of 22 Magazine St.#6, my home for the past five years. I benefited especially from the companionship of Valentine Hacquard, Michael Wagner, Rickard Sandberg, James John, Sonny Webb, Aniko Csirmaz, Marcelo Ferreira, Elena Guerzoni, Conny Krause and Meltem Kelepir. Thanks also to Joey Sabbagh and Diana Cresti for their support and friendship.

Thanks to the hardest rocking band on Earth, BTK.

Final and deepest thanks to my parents, Robert and Doreen, and my sister, Erin, for their love and encouragement. They believed in me when I didn't.

Contents

Chapter 1

Introduction

1.1 Neg-Raising

There is a class of sentence-embedding predicates in English and many other languages that intuitively validate instance of the following inference schema (at least on one reading).

(1) Not [Pred [S]] $\Rightarrow$ Pred [Not [S]]

These are known as Neg-Raising (NR) predicates. We talk about Neg-Raising *predicates* since this inference about negation is linked to specific predicates. That is, some sentence-embedding predicates validate (1) and others do not. Typical examples of those that do are *think*, *want* and *seem*.

(2) a. Bill doesn't think Mary is here

b. Bill thinks Mary isn't here

(3) a. Bill doesn't want Mary to leave

b. Bill wants Mary not to leave

(4) a. Mary doesn't seem to be here

b. Mary seems not to be here

Intuitively, (2a), (3a) and (4a) imply (2b), (3b) and (4b), respectively. This contrasts with the inferences associated with non-NR predicates such as *know*, *say* and *certain*.

(5) a. Bill doesn't know that Mary is here.

b. Bill knows that Mary isn't here.

(6) a. Bill didn't say that Mary was there.

b. Bill said that Mary wasn't there.

(7) a. It's not certain that Mary will leave.

b. It's certain that Mary will not leave.

None of the (a)-sentences in (5)-(7) implies its corresponding (b)-sentence. Horn 1978 offers the following list of NR predicates, suggesting that they generally fall within the labeled semantic classes:

(8) The classes of Neg-Raisers (Horn 1978)

a. [OPINION] think, believe, expect, suppose, imagine, reckon

b. [PERCEPTION] seem, appear, look like, sound like, feel like

c. [PROBABILITY] be probable, be likely, figure to

d. [INTENTION/VOLITION] want, intend, choose, plan

e. [JUDGMENT/OBLIGATION] be supposed, ought, should, be desirable, advise

To my knowledge, no one has offered a full explanation for why NR predictes fall within these classes and not others. Clearly this ought to be explained. We will, following Horn, address some though not all of the restrictions on what may be a NR predicate. I reproduce the list for sake of reference.

Accounting for the inference in (1) is the primary desideratum of a theory of NR. The literature has identified several other properties of NRPs that any theory of NR must account for and produced several theories to account for them. Chief among the properties of Neg-Raising predicates to be explained is their facilitation of long distance NPI-licensing. A detailed analysis of NPI-licensing plays a prominent role in the argument of this thesis.

1.2 NPI-licensing across NR predicates

It is well known that negation above a NR predicate is able to license in the complement of the NR predicate a certain kind of Negative Polarity Item (NPI) that has a more restricted distribution than the familiar *any/ever* type of NPI., cf. Lakoff 1969, Horn 1978.

(9) Punctual *until* (cf. Karttunen 1974, DeClerck 1995, de Swart 1996)

a. *Mary left until yesterday

b. Mary didn't leave until yesterday

(10) *In* (+indefinite time expression) (cf. Hoeksema 1996)

i. *Bill has left the country in (at least two) years.

ii. Bill hasn't left the country in (at least two) years.

A negation above a non-NR predicate (e.g., *claim*, *regret*, *know*) can license *any/ever* but not *until/in years*.

(11) a. Bill didn't claim/regret/know that Mary had ever left the country.

b. Mary didn't claim/regret/know that Bill had seen anything unusual.

(12) i. *Bill didn't claim/regret/know that Mary would arrive until tomorrow.

ii. *Mary didn't claim/regret/know that Bill had left the country in years.

This contrasts with negated Neg-Raising predicates which license *until/in years* as well as *any/ever*.

(13) a. Bill didn't think that Mary had ever left the country.

b. Mary didn't believe that Bill had seen anything unusual.

(14) i. Bill doesn't think Mary will leave until tomorrow.

ii. Mary doesn't believe Bill has left the country in years.

A successful theory of Neg-Raising will account for this asymmetry in the long-distance licensing of NPIs.

1.3 Thesis

In this thesis I argue for an analysis of Neg-Raising in terms of a lexical presupposition associated with Neg-Raising predicates. This analysis has antecedents in the groundbreaking article of Bartsch (1973) and the more recent suggestions of Heim (2000). The logic of the analysis is the following: We do not need to postulate an ambiguity for sentences containing a negated Neg-Raising predicate, such as (15), to account for the two putative readings (15a) and (15b). Instead a sentence containing the negation of a Neg-Raising predicate P embedding a clause S has the meaning ¬P(S). The fact that such a sentence is interpreted as if it meant P(¬S) can be made to follow from ¬P(S) and an auxiliary assumption. The auxiliary assumption is that the Excluded Middle holds for P, i.e., P(S) ∨ P(¬S). P(¬S) follows straightforwardly from ¬P(S) and [P(S) ∨ P(¬S)] by *modus ponens tollendo.*

(15) Bill doesn't think Sue is here

  i. It's not the case that Bill thinks Sue is here.

  ii. Bill thinks Sue is not here.

(16) `[ not [ NRP [ S ] ] ]`

  i. Assertion: ¬NRP(S)

  ii. Presupposition: NRP(S) ∨ ¬NRP(S)

This auxiliary assumption of the Excluded Middle arises as a presupposition associated with particular lexical items, namely the Neg-Raising predicates. The other "reading" of a sentence containing a negated Neg-Raising predicate, namely ¬P(S), cf. (15a), comes about when pragmatic factors override, or suspend, the presupposition of the Excluded Middle. So, for example, in a context where it has been explicitly denied that Bill has an opinion about whether or not Sue is here, we expect the "reading" in (15a) to emerge for (15).

This view of Neg-Raising has not been the consensus one, and not without reason. The syntactic theory of Neg-Raising, first proposed by Fillmore (1963), has enjoyed a good deal of attention and is supported by a number of compelling arguments.

Another influential theory of Neg-Raising is presented in Horn and Bayer (1984) and Horn (1989). While this theory rejects the syntactic account of Neg-Raising, it does not posit a presupposition in the manner of Bartsch (1973), but rather derives the strengthened reading of a sentence conatining a negated Neg-Raising predicate from general pragmatic principles (Horn's R-Principle).

In the next section we will take a careful look at these three theories in turn: §1.4.1 outlines Bartsch (1973) and Heim (2000), §1.4.2 the syntactic theory, and §1.4.3 Larry Horn's approach.

1.4 Three Theories of Neg-Raising

This sections sets the stage for the next by laying out the terrain of possible analyses for Neg-Raising. In the next section, I will argue that the advantages argued for by syntacticians and Horn can be claimed for the presuppositional theory. Where possible I will show that the presuppositional theory not only claims these advantages but improves on the empirical coverage of these alternative theories.

1.4.1 Presupposition

In the introduction to this chapter, we outlined the logic of the presuppositional approach to Neg-Raising. In this section we will look at two specific proposals for implementing this idea. We will pay special attention to the issue on which the two accounts differ: the grammatical status of the presupposition.

Bartsch 1973

Bartsch (1973) was the first to propose an analysis of Neg-Raising using the logic sketched in (16). According to her, the presupposition of the Excluded Middle arises as a general application condition on the use of certain clause-embedding expressions. In the normal case, she says, the conditions are fulfilled and Neg-Raising occurs. However, when the conditions are not fulfilled the requirement simply disappears and the expressions may be used anyway. This is a property that she attributes to pragmatic presuppositions - in contrast to semantic presuppositions, which *must* be fulfilled in order for an expression to be used at all.

She offers the following story as an example in which a Neg-Raising predicate can be used without having a Neg-Raised reading, i.e., negation is interpreted in it surface position.

(17) Peter has heard of Caesar and knows that he is a Roman general; Peter has also heard of Brutus as a Roman politician. He doesn't know, however, whether or not the two Romans lived at the same time. It is clear then that **(a) Peter does not think that Brutus murdered Caesar.** Peter also cannot contradict anyone who asserts that Brutus didn't murder Caesar. But from this one cannot naturally conclude that Peter agrees with him in his judgment about what passed between Caesar and Brutus.

In this paragraph it is made clear at the outset that Peter is not in a position to have an opinion as to whether Caesar murdered Brutus or not. So when the Neg-Raising predicate *think* is used the Excluded Middle presupposition is not in effect and Neg-Raising does not occur, i.e., (17a) is not taken to mean that Peter thinks Brutus did not kill Caesar.

I note in passing about Bartsch's example that the most natural way to pronounce (17a) is with stress on the negation: It is clear then that Peter does NOT think that Brutus murdered Caesar.

Heim 2000

Horn (1978) points out, in criticizing Bartsch (1973), that there is a conflict between Bartsch's assumption that the presupposition of the Excluded Middle is a general pragmatic application condition for the use of sentence-embedding predicates and the fact that Neg-Raising is a lexically-conditioned phenomenon. In other words if you want to argue that Neg-Raising follows from a presupposition of the Excluded Middle, then you have to bite the bullet and say that it is a lexical presupposition. Heim (2000) does just this in attempting to explain the interaction of Neg-Raising predicates and degree operators.

1.4.2 Syntactic Neg-Raising

The classic analysis of Neg-Raising, due to Fillmore 1963, is that there is a meaning-preserving transformation that can raise a negation from the position in which it is interpreted to a higher clause where it is pronounced. A number of arguments have been given in favor of this analysis, arguments demonstrating an analogy between Neg-Raising and more well-known movement operations. Arguments 1-3 are addressed in chapter 2 and shown not to argue against the presuppositional theory. I leave argument 4 as a topic for further research.

Argument 1: Cyclicity

Fillmore 1963 argues that Neg-Raising should be a movement rule since it displays a characteristic property of syntactic operations: it is cyclic. According to Fillmore, a negation can move indefinitely far away from its base position so long as it is separated from it only by NR predicates.

Fillmore's example in which the *not* was originally associated with *he did it*:

(18) I don't believe that he wants me to think that he did it.

This argument assumes there is no other way to obtain "cyclicity" than syntactic operations. We will challenge this below (§2.1.6).

Horn 1972 challenges this generalization suggesting that "cyclic" Neg-Raising is possible in (19i) but not (19ii). That is, the order of NR predicates determines whether long distance Neg-Raising is possible.[1]

(19) a. John doesn't think Mary wants him to go.

b. Mary doesn't want Bill to think she left.

[1]Note, however, that Fillmore's example involves *think* embedded under *want*. It is significant however that Fillmore does not use NPIs to diagnose Neg-Raising.

Argument 2: Factives

Kiparsky & Kiparsky 1970 observe that there are no factive Neg-Raisers. For example, (20a) has no reading equivalent to (20b):

(20) a. It doesn't bother me that he left early.

b. It bothers me that he didn't leave early.

Kiparsky & Kiparsky propose that this can be explained syntactically under their proposal that all factive complements are embedded under a (possibly silent) nominal *fact*.

They relate the gap in the class of Neg-Raisers to the lack of raising and ECM factives:

(21) a. *He regrets Bacon to be the real author.

b. *This makes sense to be Hoyle's best book.

It should be noted that Kiparsky & Kiparsky's claims only hold for emotive factives and not cognitive factives. While emotive factives are compatible with fact nominals, cognitive factives are not:

(22) a. Bill regrets/resents the fact that Bacon is the real author.

b. *Bill knows/realizes the fact that Bacon is the real author

Also, it seems that there is at least one ECM factive.

(23) Fred knows Bacon to be the real author.

The generalization about Neg-Raisers, however, holds for cognitive factives as well.

Argument 3: High NPIs

One particularly compelling argument in favor of the syntactic theory of Neg-Raising comes from certain facts about the licensing of NPIs. The syntactic theory of Neg-Raising predicts that in a sentence in which Neg-Raising has taken place, the negation, which is pronounced in a higher clause, is interpreted in the lower clause at LF.

Consequently, any phenomenon that is dependent on the scope of negation at LF should indicate to us that negation is present in the lower clause and, crucially, not present in the higher clause. The licensing of NPIs is one obvious phenomenon that depends on the scope of negation at LF (e.g. Ladusaw, 1979; Uribe-Etxebarria, 1994).

(24) An NPI α is licensed iff α occurs in the scope of a Downward-Entailing operator at LF

Given this, the syntactic theory of Neg-Raising predicts that an NPI should be licensed if it occurs in the lower clause, but not if it occurs in the higher clause (assuming there is no other DE operator in the higher clause).

The relevant data have been discussed by (at least) two authors. Lakoff (1969) presents (25) with judgment as marked.[2] Prince (1976) notes (26) as an argument in favor of the syntactic approach.

(25) * I didn't ever think that Bill would leave until tomorrow.

(26) a. I don't at all think that John will leave.

b. I don't think John will leave until next week.

c. * I don't at all think John will leave until next week.

The challenge these data present for a semantic/pragmatic theory of Neg-Raising is obvious. The hallmark of a semantic/pragmatic theory is that it assumes that the surface position is not deceiving: negation is interpreted in the position that it

[2] The conclusions Lakoff draws from this example are actually the opposite of Prince's. Lakoff seems to be taking it for granted that the presence of a negation in both the lower clause and the higher clause, at some point in the derivation, suffices to license NPIs in both clauses. This does not jibe well with more recent developments in the study of NPIs.

Further complicating Lakoff's conclusions are her judgments of the following sentences:

(i) It wasn't thought by anyone that John would leave until tomorrow

(ii) *I never thought that John would leave until tomorrow

My informants judgments paint a simpler picture, finding (i) ungrammatical and (ii) grammatical. Henceforth I will assume these latter judgments.

appears. Consequently, a semantic theory seems unavoidably to predict that an NPI in the higher clause ought to be licensed by the 'Neg-Raised' negation.

Argument 4: Slifting

Also compelling is the argument presented in Ross (1973). Ross's paper is on the phenomenon he dubs S(entence-)lifting, wherein, Ross proposes, a complement clause is fronted leaving behind the clause in which it is embedded. Typical examples are given in (27).

(27)

SLIFTED CLAUSE	SLIFTER CLAUSE
(a) Max is a Martian,	Fred thinks.
(b) Frogs have souls,	Osbert feels.
(c) Extraterrestrials exist,	it seems.

(28) [it seems [extraterrestrials exist]] = SLIFTING ⇒

[[extraterrestrials exist]$_1$ [it seems [t$_1$]]]

The prominent alternative analysis of this data, propose by Bresnan (1968) and Jackendoff (1971), is that these parenthetical clauses (such as *Osbert feels*) are adverbials adjoined to true main clauses (such as *Frogs have souls*). One simple argument in favor of the slifting account, advanced by Ross (1973), is that it straightforwardly accounts for the fact that a predicate allows slifting only if it syntactically selects for *that*-clauses.

(29) i. Bill wants Mary to be here

ii. *Bill wants that Mary is here

iii. Bill hopes that Mary is here

(30) i. *Mary is here, Bill wants

ii. Mary is here, Bill hopes

The adverbial account requires a more indirect explanation of these facts.

Now let's turn to the argument for the syntactic theory of Neg-Raising. Ross notes that in general negative slifters are ungrammatical.

(31) *Mary is here, Bill doesn't hope.

(32) *Mary is here, Bill doubts.

(33) *Mary is here, I deny.

However, Ross points out, there are two circumstances in which a negative slifter clause is grammatical: (I) when the negative slifter clause is negated, i.e., contains double negation as in (34),

(34) a. Mary is here, Bill doesn't doubt.

b. Mary is here, I don't deny.

and (II) when (i) the slift*ed* clause is itself negated and (ii) the predicate of the slifter clause is a Neg-Raising Predicate. Consider the following examples.

(35)

Slifted Clause not negated, Slifter predicate Neg-Raising
*Bill is standing in quicksand, I don't think

(36)

Slifted Clause negated, Slifter predicate not Neg-Raising
*Bill is not standing in quicksand, I don't claim

(37)

Slifted Clause negated, Slifter predicate Neg-Raising
Bill is not standing in quicksand, I don't think

(38) Neg-Raisers

i. Bill isn't here, I don't suppose

ii. Mary isn't here, it doesn't seem

iii. Bill won't come, it doesn't look like

(39) Non Neg-Raisers

i. *Bill isn't here, I'm not certain

ii. *Mary isn't here, it's not clear

iii. *Bill won't come, it's not obvious

Ross argues that the syntactic account of Neg-Raising can account for this pattern of judgments naturally under the following assumptions: (i) Neg-Raising is split into two operations, Neg-Copying and Neg-Deletion, which erases lower copies, (ii) the optional rule of Slifting is ordered between Neg-Copying and Neg-Deletion, and can thus bleed Neg-Deletion.

(40) Bill doesn't think Sue is here

i. INPUT

[Bill thinks [Sue is NOT here]]

ii. NEG-COPYING

[Bill NOT thinks [Sue is NOT here]]

iii. NEG-DELETION

[Bill NOT thinks [Sue is ~~NOT~~ here]]

(41) iii'. SLIFTING

[[Sue is NOT here]$_i$ [Bill NOT thinks t$_i$]]

iv'. NEG-DELETION cannot apply, output:

Sue isn't here, Bill doesn't think

Ross observes (p. 158) that it is a problem for this analysis that

> *Not*-DELETION would have to be prevented from applying unless this verb-complement structure had not been disturbed by SLIFTING. Further *not*-DELETION would have to be globally constrained so that it would not affect a lower *not* which had not been copied [...] It seems that this problem is not so much a feature of an analysis which makes use of SLIFTING as a feature of our present understanding of the ways in which rules can interact globally.

This analysis of Neg-Raising and Ross's discussion of movement rules is unsurprisingly prescient. It is common by now to analyze syntactic movement operations as the result of a copying operation and an interface operation that regulates the pronunciation of copies. Strong evidence has been presented for this analysis in the form of cases in which multiple copies of a single expression are pronounced (Nunes, 2004; Landau, to appear, cf.).

1.4.3 R-implicature

Horn 1989 represents the state-of-the-art in analyses of Neg-Raising. I will briefly summarize the analysis here, giving it a much fuller examination in chaper 3.

For Horn, Neg-Raising is an essentially pragmatic phenomenon. According to Horn, Neg-Raising predicates introduce something like an Excluded Middle implicature as a result of his **R**-principle. The **R**-principle, based on Grice's Maxim of Relevance, enjoins a hearer to read as much as possible into a statement; in this case to read a contrary negation into a contradictory negation.

(42) a. John doesn't believe Mary is here.

b. R-implicature of (42a): John believes Mary is here or John believes Mary is not here.

The calculation of this implicature is subject to several more-highly ranked pragmatic constraints. The first is to not create "pernicious ambiguities." If there is a high functional difference between a high negation and a low negation, then the implicature does not arise. The meaning that would result is too different from the literal meaning. For example, the position of negation in (28) carries too much information to permit an Excluded Middle implicature to arise:

(43) a. It's not certain that Bill is here.

b. It's certain Bill is not here.

The second pragmatic constraint has to do with the strength of negation. Bolinger observed that an NR sentence has a weaker negative force than a sentence in which the negation appears lower. Horn adds that this is more generally true.

(44) I think she's sad.
I think she's unhappy.
I think she's not happy.
I think she isn't happy.
I do not think she's happy.

I don't think she's happy.

According to Horn, "negative force weakens with the distance of the negative element from the constituent with which it is logically associated." Thus, says Horn, the further negation is from its associated clause the more uncertainty it conveys about the negation of that clause. Uncertainty, apparently, on the part of the speaker (this will be crucial below). Horn refers to this as the Uncertainty Principle (UP). When the meaning of a construction clashes with the UP, the implicature does not arise and there is no NR. This happens, according to Horn, in the case of factives.

The adherence to these higher-ranked pragmatic constraint, then, explains certain restrictions on the class of Neg-Raisers, e.g., the lack of factive Neg-Raisers. Horn also shows how the pragmatic theory might account for the NPI- licensing facts noted above. We will discuss analysis in some depth in chapter 3.

1.5 Summary

In sum, there are three plausible analyses of Neg-Raising that have been advanced: syntactic, semantic and pragmatic. Each analysis has its strengths and offers an explanation of the licensing of NPIs under negated Neg-Raising predicates. In the next chapter, we devise an analysis of the NPI facts within the semantic (presupposition) theory and show that it offers the best account of the facts. Arguments 2 and 3 from §1.4.2 are also addressed.

Chapter 2

Neg-Raising and Polarity

In this chapter I advance the cause of the presuppositional theory of Neg-Raising by showing that it offers an elegant account of a number of complex NPI-licesning facts. The standard for comparison in this chapter is the quite successful syntactic theory of Neg-Raising. The centerpiece of this chapter is an independently motivated account of Horn's argument against the syntactic theory based on failures of cyclicity. The presuppositional account of Neg-Raising explains both the apparent cyclicity of Neg-Raising and the existence of Horn's exceptions.

Proposal:
Adopt Bartsch/Heim approach to Neg-Raising in which Neg-Raising follows from a semantic presupposition of the Excluded Middle. Informally,

(45) $[\![$ **NRP** $]\!]$(p) is defined iff NRP(p) or NRP(¬p)
when defined,
$[\![$ **NRP** $]\!]$(p) =1 iff NRP(p)

This approach to Neg-Raising combined with an approach to NPI-licensing along the lines of Zwarts (1996b) and Zwarts (1996a) and standard assumptions about presupposition projection will suffice to explain the major characteristics of NPI licensing across Neg-Raising predicates.

2.1 NPI-licensing

One of the most important tests for Neg-Raising, and indeed for theories of Neg-Raising, has been the licensing of NPIs in the scope of a Neg-Raising Predicate by a negative element above the NRP. This test for Neg-Raising was first proposed by Kajita, as reported in Lakoff (1969). According to Lakoff there are certain NPIs that have the following characteristic: when in the complement of an embedding verb, they are licensed by negation above that verb ONLY IF that verb is a Neg-Raiser. Lakoff gives punctual *until* as an example.

2.1.1 Diagnostic NPIs

punctual *until*

Generally *until* combines felicitously with durational eventualities, such as states and activities. When combined with a punctual eventuality, *until* is unacceptable.

(46) Bill was sick until last Friday (state)

(47) Bill ran in a circle until 5:30 (activity)

(48) a. *Bill arrived until yesterday.

b. *Sue left until today.

However, when a punctual eventuality is embedded under certain negative expressions, *until* becomes felicitous again:

(49) a. Bill didn't arrive until yesterday.

b. No one left until yesterday.

Several hypotheses have been put forward to explain the compatibility of *until* and punctual eventualities as in (49). The most prominent are the following: (A) the negation of a punctual eventuality is durational, so if *until* scopes above negation it should become acceptable. Examples such as (50), where *until* scopes above negation overtly, suggest that this is a plausible analysis.

(50) Until Friday, Bill didn't leave.

(B) there are two *until*'s - one that combines with durational eventualities and another, an NPI, that combines with punctual eventualities. While (B) seems the less parsimonious theory, it is supported by that fact that negated *until* gives rise to implications not associated with the un-negated durational *until*. For example, (49a) strongly implies that Bill arrived yesterday. Compare this with (50) which weakly implicates that Bill left Friday. To better see the difference consider the following contrast:

(51) [Context: A: Was Bill still here on Saturday?
B: I don't know all I know is that]
a. Until Friday, he didn't leave.
b. #He didn't leave until Friday.

Karttunen (1974b) offers in addition striking contrasts such as the following, which pose a challenge to analysis (A) of negated *until*. While both (52a) and (52b), seem to imply that Nancy married in 1978, the difference between the two becomes apparent in (53).

(52) a. Nancy remained a spinster until 1978.
b. Nancy didn't get married until 1978.

(53) a. Nancy remained a spinster until she died.
b. # Nancy didn't get married until she died.

I find this last argument convincing (*pace* de Swart (1996)) and for that reason follow Karttunen and Lakoff in assuming that there is an NPI *until*.

Now we turn to the argument for Neg-Raising based on the licensing of NPI *until*. The crucial sentences involve an instance of *until* in an embedded clause that denotes a punctual eventuality. In some such case a negation above the embedding verb licenses the *until* and in other cases it does not. Lakoff (1969) claims that in order for the superordinate negation to license the lower *until* the intervening predicate has to be Neg-Raising.

(54) Non-Neg-Raisers

i. *Bill didn't **say** that Mary would leave until tomorrow[1]

ii. *Bill didn't **claim** that Mary would leave until tomorrow

(55) Neg-Raisers

i. Bill didn't **think** Mary would leave until tomorrow

ii. Bill doesn't **want** Mary to leave until tomorrow

Notice that this pattern of licensing is different from that of more familiar NPIs, such as *any* and *ever*, which are also licensed by negation above non-Neg-Raising predicates:

(56) i. Bill didn't claim that there was anything in the refrigerator.

ii. Bill didn't say that he had ever been to France.

(57) i. Bill didn't want there to be anything in the refrigerator.

ii. Bill didn't think Sue had ever been to France.

in years

Fortunately, *until* is not the only NPI that shows such sensitivity. Another, less controversial, class of NPIs that are licensed by a higher negation only if the intervening

[1]It is crucial in these examples to control for the scope of *until*. For example in (i) there are two possible scopes for *until*, (ia) and (ib):

(i) Bill didn't think Mary left until Friday

a. Bill didn't think [[Mary left] until Friday]

b. Bill didn't [[think Mary left] until Friday]

Whenever possible I will disambiguate the scope by means of tense clashes. For example, if the main clause is in the past and *until*'s complement is *tomorrow*, then the *until* clause can only be associated with a future oriented embedded clause.

(ii) Bill didn't think Mary would leave until tomorrow

a. Bill didn't think [[Mary would leave] until tomorrow]

b. # Bill didn't [[think Mary would leave] until tomorrow]

The infelicity of (iib) is analogous to that of (iii).

(iii) # Bill didn't think until tomorrow predicate is Neg-Raising is the class of *in* + indefinite temporal nominal NPIs, such as *in years*. In addition to requiring negation to be licensed, these NPIs require the perfect.

(58) *Bill has been here in years. (perfect, no negative)

(59) Bill hasn't been here in years. (perfect, negative)

(60) *Bill wasn't here in years. (no perfect, negative)

It is worth noting that these restrictions hold not only for *in* with a bare plural time expression but with any indefinite time expression.

(61) i. # I have seen her in some years.

ii. I haven't seen her in some years.

(62) i. # I have seen her in at least two years.

ii. I haven't seen her in at least two years.

(63) i. # I have seen her in a fortnight.

ii. I haven't seen her in a fortnight.

Note, actually, that *in* + a bare numeral time expression (or the indefinite article) does not require negation to be licensed. Instead, it can occur felicitously with an accomplishment or achievement, as pointed out by Dowty 1979.

(64) Bill ate seven cream pies in two days. (accomplishment)

(65) Sue found the solution to the problem in three months. (achievement)

Notice that these are crucially upper bounded readings of the numerals. If it took Bill three days to eat his seven cream pies, then (64) is false. When the numeral is modified with *at least*, which removes the upper bound, the sentences become ungrammatical again:

(66) *Bill ate seven cream pies in at least two days.

(67) *Sue found the solution to her problem in at least three months.

For this reason, in testing Neg-Raising I will use *in* with either bare time expressions or numerals modified with *at least*. The following examples show that *in years*-licensing shows the same sensitivity to Neg-Raising that punctual *until*-licensing does.

(68) Non-Neg-Raisers

i. *Bill didn't claim that Sue had visited Fred in (at least two) years.

ii. *Bill didn't say that Sue had visited Fred in (at least two) years.

(69) Neg-Raisers

i. Bill doesn't think that Fred has visited Sue in (at least two) years.

ii. Bill doesn't seem to have visited Sue in (at least two) years.

See Horn (1989) for a discussion of *in years* as a diagnostic for Neg-Raising.

either

Another NPI that fits this profile is additive *either*. See Rullman (2003) for a number of arguments that *either* is a NPI.

(70) (Mary didn't go to the party) SUE didn't go **either**.

In order to establish that licensing of *either* is sensitive to Neg-Raising, it will be useful to point out an unusual property that it shares with other additive particles, such as *too* and *also*. Heim (1992) observes that the presuppositions introduced by additive particles do not need to project through the usual channels. Normally, when a presuppositional element is contained in the complement of an attitude predicate, such as *think*, the attitude holder is presupposed to believe the presuppositions of its complement. For example,

(71) I stopped smoking.
presupposes: I used to smoke.

(72) My parents think I stopped smoking.
presupposes: My parents <u>believe</u> I used to smoke.

When the presupposition is introduced by an additive particle, however, this is not so. The presupposition is able to project as if attitude predicates were holes for presupposition. That is, the presupposition of the additive particle may be inherited unmodified by the sentence as a whole.

(73) $[\mathrm{I}]_F$ am also in bed.
presupposes: someone other than me is in bed.

(74) John: I'm already in bed.
Mary: My parents think $[\mathrm{I}]_F$ am also in bed (Heim, 1992)

Mary's response in (74) need not presuppose that her parents believe of someone else that they are in bed. Rather Mary can be taken to presuppose merely that someone other than her is in bed. This special property of additive particles facilitates the use of *either* to test for Neg-Raising. Thanks to the possible absence of the projected presupposition we can test for the scope of *either* relative to an embedding predicate rather easily. If *either* has to introduce a presupposition about the embedding predicate, then it must have scope over the predicate. If it need not, then it scopes below the predicate. For example,[2]

(75) (In a context where we know Bill mistakenly thinks that Sue is here)
A: Sue isn't here.
B: That's too bad. Bill doesn't think MARY is here either.

(76) A: Bill's really upset that Mary isn't coming.
B: I know. (He's so angry at Sue for talking her out of coming that) he doesn't want SUE to come either

In each of these cases, *either* need not introduce a presupposition about Bill's attitude. The presupposition introduced by *either* may be satisfied merely by the fact that Mary is not here (isn't coming).

Compare this with the behavior of non-Neg-Raising predicates:

(77) (In a context where Bill has been telling everyone Sue is here. Bill is an authority on Mary's whereabouts)

[2]Rullman (2003) gives similar examples demonstrating *either*'s sensitivity to Neg-Raising.

A: Sue isn't here. But I fear that Mary is.

B: Don't be so worried. #Bill didn't say that MARY was here either.

(78) a. A: Bill's really upset that Mary isn't coming.

b. B: I know. (He's so angry at Sue for talking her out of coming that) # he definitely doesn't hope SUE comes either.

Here the situation is different. In (77), B's response clearly presupposes that there is someone besides Mary that Bill didn't say was there. It cannot be taken to merely presuppose that someone besides Mary was (not) there. Similar comments hold for the non-Neg-Raising predicate *hope* in (78).

This suggests that the licensing of *either* is also a good test for Neg-Raising.

In the last three subsections, we have established three NPIs as being sensitive to Neg-Raising in long-distance licensing environments. Ideally, we would like to have some understanding of why this is so. In the next section, I will mount a detailed defense of the hypothesis that these three NPIs require Anti-Additive licensers to be licensed. In each of the three cases the hypothesis has been considered before and in some cases rejected. I will give responses to some of the objections and then show that under the presuppositional theory of Neg-Raising, NR predicates turn out to be Anti-Additive.

2.1.2 Anti-Additivity

In this section, I defend the hypothesis that the three NPIs just discussed are subject to a requirement that they occur in an environment that licenses Anti-Additive inferences. I defend this hypothesis both against an alternative hypothesis about their licensing (clausemate negation) as well as direct criticisms of the Anti-Additivity generalization.

Background: Licensing conditions on NPIs

The licensing of an NPI depends on the logical properties of the environment in which the NPI occurs. Ladusaw (1979) identified the valid inference from sets to subsets

(Downward Entailingness (DE-ness)) as a property necessary for licensing NPIs. In particular, Ladusaw states a condition according to which an NPI must occur in the scope of a downward entailing operator. (In the definitions below, I use '$\Rightarrow$' to stand for cross-categorial entailment).

(79) An NPI is licensed only if it occurs in the scope of an expression that denotes a Downward Entailing function.

(80) A function F is downward entailing iff for all A, B in the domain of F such that A $\Rightarrow$ B, F(B) $\Rightarrow$ F(A).

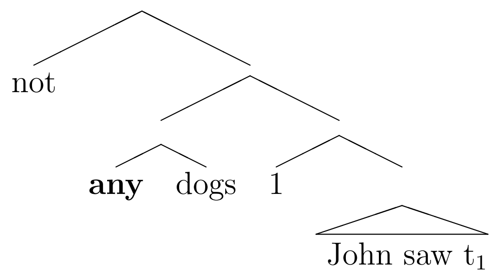
For reasons that I discuss in §2.1.6 I make use of a slightly different statement of the licensing conditions on NPIs. Rather than requiring NPIs to appear in the scope of an expression that denotes a DE function, I require that an NPI occurs in an environment that supports downward entailing inferences. Such a condition allows for a combination of expressions that do not themselves denote DE functions to create an environment that does support downward inferences. Furthermore, some subconstituents of the scope of an expression that denotes a DE function might fail to support downward inferences if the environment contains another expression that interferes with downward inferences. Crucial use of such principles of licensing has been made by Heim (1984), Zwarts (1996a) and Heim (2003) a.o.

(81) An NPI α is licensed in a sentence S only if there is a constituent β of S containing α such that β is Downward Entailing with respect to the position of α

(82) A constituent α is Downward Entailing with respect to the position of β ($[\![\beta]\!]$ $\in$ D_σ) iff the function $\lambda x.[\![\alpha[\beta/v_{\sigma,i}]]\!]^{g[v_{<\sigma,i>} \rightarrow x]}$ is Downward Entailing

(83) $\alpha[\beta/\gamma]$ is the result of replacing β with γ in α

An example, the occurrence of *any* is licensed in (84) since the entire sentence is Downward Entailing with respect to the position of *any*.

(84) 

(85) $\lambda x_{et,ett}.[\![\text{not } [\text{v}_{<et,ett,2>} \text{ dog}] \text{ 1 John saw } \text{t}_1]\!]^{g[v_{<et,ett,2>} \rightarrow x]}$

(86) $[\![$ at least 2 $]\!] \Rightarrow [\![$ any $]\!]$

(assuming $[\![$ any $]\!] = [\![$ some $]\!]$)

(87) ‘John didn’t see any dogs’ entails ‘John didn’t see at least two dogs’

(That is, (85) is a DE function)

In the remainder of this thesis, I will use a slightly more complex statement of such environment-related licensing principles. While the formulation is more complex it will ultimately make checking for NPI licensing simpler. The idea is simply this: in checking whether the environment that an NPI occurs in is Downward Entailing we do not need to pay attention to every expression that c-commands the NPI. Specifically, we can ignore any expression that c-commands the NPI but is taken by the NPI as an argument or is taken as an argument by the function that is the result of applying the NPI to another expression or ... etc. Simply put, expressions whose denotations are arguments of a function do no affect the logical properties of the environment in which the NPI occurs.

To achieve this simplification we must first define an auxiliary notion, F(unction)-projection.

(88) F(unction)-projection

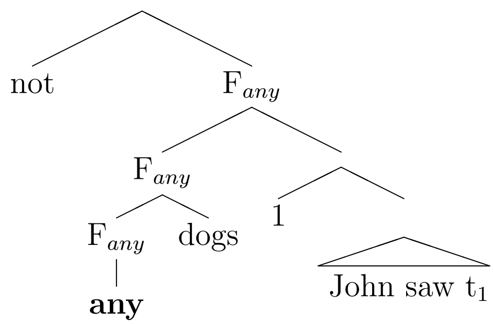
a. Every terminal node is an F-projection of itself.

b. If C is a branching node with daughters A, B, then C is an F-projection of A iff $[\![C]\!] = [\![A]\!]([\![B]\!])$ or B is a binding index.

c. F-projection* is the transitive closure of the F-projection relation

For example, the F-projections* of *any* are marked F_{any} in (89).

(89)



Now let's formulate a new principle for the licensing of NPIs based on this notion.

(90) An NPI α is licensed in a sentence S only if there is a constituent β containing α such that β is Downward Entailing with respect to the maximal F-projection* of α

Using this principle, the function that we have to check for DE-ness is much simpler: since the complement of *not* is the maximal F-projection* of *any*, the function to be checked for DE-ness is $\lambda p.[\![$ not $v_{<t,1>}$ $]\!]^{g[v_{<t,1>} \rightarrow p]}$, which is just $[\![$ not $]\!]$.

Now let's use these notions to formulate the licensing principle for the strict NPIs enumerated in §2.1.1. As mentioned above, an environment is downward entailing iff it licenses inferences from sets to subsets. For example,

(91) i. Not a single student read any books

ii. Not every student read any books

(92) LONG BOOK $\Rightarrow$ BOOK

(93) i. Not a single student read a book $\Rightarrow$ Not a single student read a long book

ii. Not every student read a book $\Rightarrow$ Not every student read a long book

The valid inferences in (93), show that not a single student (the negation of an existential) and not every student (the negation of a universal) create DE contexts. This explains why *any* is licensed in (91). The general lesson to be taken away from these examples is summarized schematically below.

(94) The environments NOT(SOME( )) and NOT(EVERY( )) are both DE

Zwarts (1996b) observes that DE-ness is not always sufficient to license an NPI. Some NPIs require environments that have logical properties in addition to DE-ness.

Zwarts offers a classification of negative strength that is based on a generalization of De Morgan's Laws:

(95) DeMorgan's Laws

$$\neg(X \wedge Y) \Leftrightarrow \neg X \vee \neg Y$$

$$\neg(X \vee Y) \Leftrightarrow \neg X \wedge \neg Y$$

These equivalences can be split up into four entailment relations and generalized so that functions other than negation can be tested to see which parts of DeMorgan's Laws they validate.

(96) Strengths of Negation (Zwarts 1998)

Antimorphic

Anti-additive

Downward Entailing

(i) $f(X) \vee f(Y) \Rightarrow f(X \wedge Y)$

(ii) $f(X \vee Y) \Rightarrow f(X) \wedge f(Y)$

(iii) $f(X) \wedge f(Y) \Rightarrow f(X \vee Y)$

(iv) $f(X \wedge Y) \Rightarrow f(X) \vee f(Y)$

DE functions validate at least (96i) and (96ii). An Anti-Additive function is one that in addition satisfies (96iii). An Antimorphic function validates all four entailments in (96). Essentially, only sentential negation qualifies as antimorphic.[3] More natural language expressions satisfy the criteria for being Anti-Additive. For example, *no student* creates an Anti-Additive environment; but *not every student* does not:

(97) a. Not a single student smokes and Not a single student drinks $\Rightarrow$
Not a single student smokes or drinks

b. Not every student smokes and Not every student drinks $\not\Rightarrow$
Not every student smokes or drinks

So, from this we learn the following:

[3] Zwarts (1996b) gives negated proper names as another example of an antimorphic function, e.g., *not John*. I think it is debatable that such operators exist in natural language.

(98) NOT(SOME( )) is ANTI-ADDITIVE

(99) NOT(EVERY( )) is not ANTI-ADDITIVE

This diffence in logical properties accounts for the difference in licensing of what Zwarts calls strong NPIs like *lift a finger*.

(100) Not a single student lifted a finger to help.

(101) *Not every student lifted a finger to help.

The principle (90) allows for the licensing of NPIs in the complements of both negated Neg-Raising predicates (NRPs) and negated non Neg-Raising predicates. This follows since the combination of negation and a universal quantifier creates a downward entailing environment. Our negated NRPs are stronger than negated universals since inferentially they behave as if negation were below the NRP. That is ¬NRP behaves like ALL¬. ALL¬ in contrast to ¬ALL is anti-additive.

(102) F is Anti-Additive iff F(A) $\wedge$ F(B) $\Leftrightarrow$ F(A$\vee$B)

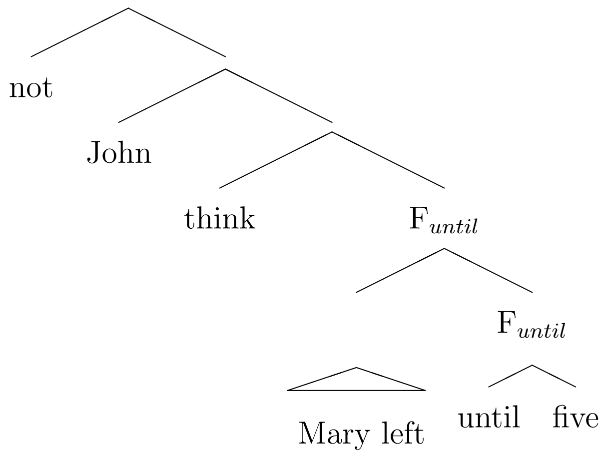
So we hypothesize that strict NPIs are allowed under negated NRPs because they are licensed in Anti-Additive environments.

(103) An strict NPI α is licensed in a sentence S if there is a constituent β containing α such that β is Anti-Additive with respect to the maximal F-projection* of α

(104) A constituent α is Anti-Additive with respect to constituent β ($[\![\beta]\!] \in$ D_σ) iff the function λx.$[\![\ \alpha[\beta/v_{\sigma,i}]\ ]\!]^{g[v_{<\sigma,i>} \to x]}$ is Anti-Additive.

Consider the NRP example (105) – the F-projections* of *until* are circled.

(105) John doesn't think Mary left until five.



So we need to test whether (105) is Anti-Additive with respect to the embedded clause (the maximal F-projection* of *until*). Note in this regard that the inference in (106) is intuitively valid.

(106) John doesn't think Mary left and John doesn't think Bill left.
$\Rightarrow$ John doesn't think Mary or Bill left

2.1.3 Neg-Raising and strong NPIs

The NPIs used to diagnose Neg-Raising also exhibit this licensing asymmetry: punctual *until* and in weeks require an Anti-Additive licenser.

(107) a. *Not every student arrived until 5 o'clock.
b. Not a single student arrived until 5 o'clock.

(108) a. *Not every student has visited Bill in (at least two) years.
b. Not a single student has visited Bill in (at least two) years.

Note that these examples pose a challenge to the clausemate condition on the licensing of *until/in years*. Here the NPIs appear to be clausemates with negation but are not licensed.

van der Wouden (1995) observes that (in Dutch) negated Neg-Raising predicates show the licensing capabilities of Anti-Additive functions. In other words, the negation of a NRP licenses NPIs like the negation of an existential.

(109) a. Bill doesn't think Sue lifted a finger to help

b. *Bill doesn't know that Sue lifted a finger to help

Van der Wouden stops short of giving a semantics for NRPs. The challenge, then, is to give a semantics that is universal but whose negation acts like the negation of an existential. The presuppositional account of Neg-Raising reconciles their universal semantics with the Anti-Additivity of their negations. Let's see how:

Under the presuppositional account, a negated Neg-Raising predicate is predicted to create an Anti-Additive context. Recall that I assume that Neg-Raising predicates have lexical entries of the form (42), where M is NRP's modal base:

(110) For any proposition P, and individual x,

$[\![$ NRP$]\!]$ (P)(x)

(i) presupposes: every M(x)-world is P-world or no M(x)-world is a P-world

(ii) asserts: every M(x)-world is a P-world

The crucial part of our story is what happens when you negate an NRP that carries an Excluded Middle presupposition:

(111) $[\![$ not NRP(P)(x)$]\!]$

(i) presupposes: every M(x)-world is P-world or no M(x)-world is a P-world

(ii) asserts: not every M(x)-world is a P-world

Under what conditions is (111) defined? It is a general property of negation that negated sentences carry the same presuppositions as their unnegated versions. So (111) is defined in the same cases as (110). The assertion of (111) is simply the negation of the universal assertion of (110).

Notice that the presupposition and assertion of (111) come together to entail the second disjunct of the presupposition:

(112) every M(x)-world is P-world

or

no M(x)-world is a P-world $\Rightarrow$ no M(x)-world is a P-world

not every M(x)-world is a P-world

So let's see what we predict about Anti-Additivity: is the entailment in (113) valid?

(113) not NRP(P)(x) and not NRP(Q)(x) $\Rightarrow$ not NRP(P$\vee$Q)(x)

(114) a. $[\![$ not NRP(P)(x)$]\!]$

1. presupposes: every M(x)-world is P-world or no M(x)-world is a P-world
2. asserts: not every M(x)-world is a P-world
3. Together (1) and (2) entail: no M(x)-world is a P-world

b. $[\![$ not NRP(Q)(x)$]\!]$

1. presupposes: every M(x)-world is Q-world or no M(x)-world is a Q-world
2. asserts: not every M(x)-world is a Q-world
3. Together (1) and (2) entail: no M(x)-world is a Q-world

c. $[\![$ not NRP(P$\vee$Q)(x)$]\!]$

1. presupposes: every M(x)-world is P$\vee$Q -world or no M(x)-world is a P$\vee$Q -world
2. asserts: not every M(x)-world is a P$\vee$Q -world

If (a) and (b) are both true, then the presupposition of (c) is satisfied. The reason is that (a) entails that no M-world is a P-world and (b) entails that no M-world is a Q-world, thus no M-world is a P$\vee$Q-world. The assertions of (a) and (b) entail that B is non-empty, it follows from this and the fact that no B world is a P$\vee$Q world that the assertion of (c) is true. So for arbitrary P, Q, the truth of (a) and (b) guarantees the truth of (c). Thus, negated NRPs create an Anti-Additive context.

This prediction is shown to be correct by the validity of the entailment in (115).

(115) Bill doesn't think Sue is here and Bill doesn't think Fred is here $\Rightarrow$
Bill doesn't think Sue is here or Fred is here

So, now we have an alternative account of the licensing of strong NPIs under NRPs. But it can also be shown that this approach is superior to the clausemate proposal.

Clausemate Condition

A hypothesis as to why these NPIs interact with Neg-Raising in this way is already present in Lakoff (1969). In fact, this interaction is pointed to as an argument in favor of the syntactic theory of Neg-Raising. Lakoff proposes that NPI *until* is required to be clausemates with negation. Under the syntactic theory this immediately accounts for the contrast in (54) and (55). A negation occurring above a Neg-Raising predicate can have been base-generated in the complement clause, as a clausemate with *until*. A negation above a non-Neg-Raising predicate cannot have such a source.

(116) α and β are clausmates iff
  a clause contains α iff it contains β

(117) Neg-Raiser
  Interpretive level: [John thinks [Mary not left until Friday]] $\Rightarrow$
  Surface: [John does not think [Mary ~~not~~ left until Friday]]

In the above derivation, *until Friday* and *not* are clausemates at the level of interpretation though on the surface they are separated by an intervening predicate *think*. If negation is clausemates with *until* under a non-Neg-Raising predicate, the two remain clausemates on the surface:

(118) non Neg-Raiser
  Interpretive level: [John claims [Mary not left until Friday]] $\Rightarrow$
  Surface: [John claims [Mary did not leave until Friday]]

In the next few sections, I will argue against the clausemate condition as the appropriate licensing condition for the NPIs *until*, *in years* and *either*.

Morphological Decomposition

One immediate hurdle that the clausemate story for NPI licensing faces is that it focuses too sharply on sentential negation. As Horn (1978) demonstrates, a variety of negative expressions trigger Neg-Raising and license strict NPIs in an embedded clause under a Neg-Raising Predicate. In this section, I'd like to focus on Negative Universal quantifiers, such as *no student*.

(119) a. No student thinks Mary left until Friday.

b. No student thinks Bill has been there in years.

c. (Mary won't come today) No student thinks BILL will come either.

This sentence demonstrates that the quantifier *no student* triggers Neg-Raising and licenses the strict NPIs. As it is, the subject of the predicate *think*, the negative quantifier itself cannot be clausemates with the strict NPIs at any level of representation. However, if the subject is morphologically decomposed into a negation and a different quantifier, it is possible that the negation could have originated in the embedded clause. But which quantifier? There is only one possibility if we want to get the meaning of the sentence right. The sentences in (119) are roughly equivalent to the sentences in (120), which have negation in the embedded clause.

(120) a. Every student thinks Mary didn't leave until Friday.

b. Every student thinks Bill hasn't been there in years.

c. (Mary won't come today) Every student think Bill won't come today either.

We can let these equivalences be our guide and conclude that underlyingly *no student* is composed of a universal quantifier and a negation:

(121) LF: [EVERY student thinks [Mary NOT leave until Friday]] ⇒
[EVERY student NOT t thinks [Mary ~~NOT~~ leave until Friday]] ⇒
[EVERY+NOT student thinks [Mary ~~NOT~~ leave until Friday]]
(EVERY + NOT spells out as *no*)

This would not be the only case in which a decomposition of negative quiantifiers into a universal and a negative has been proposed. Zanuttini 1991 gives such an analysis for N-words in Italian. According to Zanuttini an n-word such as *nessuno studente* 'no student' decomposes into a universal and a negation. This accounts for uses in which the n-word contributes a negation to a sentence:

(122) Nessuno studente e' arrivato.

No student has arrived. = Every student did not arrive.

In concord sentences, on the other hand, the n-words combine via a process of absorption, cf. May (1989). During the process of absorption, the negations of the n-words are factored out, so that only one negation enters into the interpretation of the sentence.

(123) Nessuno ha fatto niente

No one has done nothing

'No one did anything' = 'every thing is such that everyone didn't do it'

(124) Neg-Factorization

[[EVERY one]$_i$+NOT, [EVERY thing]$_j$+NOT][t$_i$ did t$_j$]]]

[[EVERY one]$_i$,[EVERY thing]$_j$]][NOT][t$_i$ did t$_j$]]

Despite this parallelism between Zanuttini's 1991 account and what is required for a syntactic theory of Neg-Raising, there is evidence that this is not the correct analysis of negative quantifiers. The problem is not the decomposition. Morphological decomposition of negative universal quantifiers is often proposed (see Kratzer (1995), Penka and von Stechow (2001)). The decomposition of negative quantifiers, however, generally splits the negative quantifier into a negation and a lower scope existential.

The choice is not arbitrary, despite the logical equivalence of $\neg\exists$ and $\forall\neg$. The chief evidence in favor of decomposition comes from cases in which the quantifier and negation take separate scope. In these cases, the negation scopes above some operator and the existential below it. Such split scope can be found also in English, cf. Potts (2000).

(125) Mary need fire no nurses.

a. There is no nurse that Mary needs to fire.

b. It is not the case that Mary needs to do some nurse-firing.

The sentence (125) has two readings corresponding to the paraphrases in (125a) and (125b). The former reading says that there is no nurse x such that Mary is required to fire x. The latter reading says, by contrast, that Mary is not required to engage in any nurse firing. The latter reading entails the former and it can be shown that this stronger reading is sometimes available for this sentence. It is possible to be required to engage in nurse-firing without there being any particular nurse that you are required to fire. It is not possible to be required to fire a nurse without being required to engage in nurse-firing.

(126) A: I'm really upset, the hosptial chief is making me fire nurses. At least she didn't specify anyone in particular.

B: Mary, you misread the memo. You need fire no nurses.

If (125) had only reading (125a), B's objection would make no sense: A already knows that there is no particular nurse x such that she needs to fire x. Here the only sensible available reading is the one corresponding to the paraphrase (125b).

(127) Surface: [need [you fire [NEG + SOME nurses]]] ⇒

Interpretation: [NEG need [you fire SOME nurse]]

This is strong evidence that negative quantifiers decompose into a negation and an existential quantifier. We might furthermore ask, however, whether Neg-Raising and Neg-splitting are compatible. If they are not, then it may be that the negative quantifiers involved in Neg-Raising receive a different representation in the grammar.

Here is an attempt to identify a reading of a sentence that is the result of both Neg-Raising and Neg-splitting. (128) demonstrates that a negative quantifier can split across a predicate that it c-commands on the surface.

(128) No student is allowed to enter.

a. There is no student x such that x is allowed to enter

b. It is not allowed for students to enter.

In this case, as well, the non-scope split reading (128) may be true while the scope split reading is false. It could be that the rules for entering specify that none of the people who happen to be students are allowed to enter, while permitting that students enter. This implies that in a world similar to our own in which the set of students differ from the actual set there could be a student that could enter.

Now consider (129), which features a negative quantifier above the predicate *allow* and a Neg-Raising predicate embedded below it. The question is can (129) have the reading (130).

(129) ?No student is allowed to think that Mary left until last Friday.

(130) It is required that every student think that Mary didn't leave until last Friday

We can safely assume that there is a non-Neg-split reading corresponding to (131). [M = Mary left until Friday]. Now if in addition the presupposition of the NRP *think* projects universally both throught the existential modal and existential quantifier, then we end up with the presupposition of (131) in (132). This presupposition (132) and the truth-conditions in (131) combine to entail (133).

(131) $\neg\exists$x [student$_w$(x) $\wedge$ $\exists$w'[wRw' $\wedge$ x thinks$_{w'}$ M]]

(132) $\forall$x[student$_w$(x) $\rightarrow$ $\forall$w'[wRw' $\rightarrow$ x thinks$_{w'}$ M or x thinks$_{w'}\neg$M]]

(133) $\forall$x[student$_w$(x) $\rightarrow$ $\forall$w'[wRw' $\rightarrow$ x thinks$_{w'}\neg$M]]

(134) 'Every student is required to think that Mary didn't leave until Friday'

Now we want to differentiate from this reading the one in which the subject splits across the matrix predicate *allow*.

(135) $\neg\exists$w'[wRw' $\wedge$ $\exists$x [student$_{w'}$(x) $\wedge$ x thinks$_{w'}$ M]][4]

If again the presupposition to the NRP projects universally, then the presupposition carried by the Neg-split reading is (136). This together with the truthconditions (135), this entails (137).

(136) $\forall$w'[wRw' $\rightarrow$ $\forall$x[student$_{w'}$(x)$\rightarrow$ x thinks$_{w'}$ M or x thinks$_{w'}\neg$M]]

(137) $\forall$w'[wRw' $\rightarrow$ $\forall$x[student$_{w'}$(x)$\rightarrow$ x thinks$_{w'}\neg$M]]

[4]I will not consider the *de re* reading, since it is equivalent to 131.

This reading, if available, results from the Logical Form of (129) in (138). Given such a reading, this LF is Anti-Additive with respect to the position of the most deeply embedded clause. Thus, we predict that punctual *until* is licensed in that environment.

(138) [not [allowed [SOME student [think that Mary left until Friday]]]]

To my ear, the sentence (129) is somewhat marginal. There are several reasons why this might be so. One salient possibility is that presupposition projection does not behave exactly as we have sketched it here. There are two dimensions in which we have over-simplified. First, it is controversial to claim the existential quantifiers project the presuppositions of their scope universally, cf. Heim (1982).

(139) A fat man pushed his bicycle
??presupposes: every fat man has a bicycle

Second, it is not clear that a presupposition in the complement of *allow* projects into the worlds of *allow*'s modal base (assuming that these are the best worlds according to the rules of the actual world).

(140) Bill is allowed to push his bicycle
??presupposes: Bill is required to have a bicycle

If we revise either of these assumptions (that existentials project universal presuppositions, that presuppositions of *allow*'s complement project as presuppositions about permissible worlds) then we no longer predicate that strict NPI's are licensed in the most deeply embedded clause.

2.1.4 Clausemate NPI licenser is not sufficient

A theory that says what is special about our strict NPIs is that they require a clausemate NPI-licenser cannot be exactly right, as we have seen. There are many cases in which an NPI-licenser that is clausemates with a strict NPI fails to license that NPI. As we have seen, strict NPIs require stronger negations, even when clausemates with a licenser. The examples below illustrate more fully the restrictions on strict NPIs.

They are licensed by sentential negation (a), and negative universals (b), which are Anti-Additive. They are not licensed by a negated universal (d), or a DE modified numeral (e). Licensing by *few* is more controversial, as I've indicated with a question mark. Consider, for example, the judgments of Hoeksema and Rullman on (144) and (145).

until

(141) i. John didn't arrive until five.
 ii. Not a single/No student arrived until five.
 iii. $^?$Few students arrived until five.
 iv. *Not every student arrived until five.
 v. *At most three students arrived until five.

in years

(142) a. John hasn't visited in years.
 b. Not a single/No student has visited in years.
 c. $^?$Few students have visited in years.
 d. *Not every student has visited in years.
 e. *At most three students have visited in years.

either

(143) (Bill doesn't like spaghetti)
 a. Sue doesn't like it either.
 b. Not a single/No student likes it either.
 c. $^?$Few students like it either.
 d. *Not every student likes it either.
 e. *At most three students like it either.

(144) He was one of the few dogs I'd met in years that I really liked.

(Sue Grafton, A is for Alibi, Hoeksema ms.)

(145) Few Americans have ever been to Spain. Few Canadians have either.

(Rullman (2003), p.345)

Instead of the clausemate condition, I maintain that these NPIs are subject to a licensing condition requiring a stronger negative licensing environment than other NPIs, namely an Anti-Additive licenser. The purported grammaticality of (144) and

(145) poses a challenge to this generalization since *few* is not Anti-Additive, neither on its cardinal nor on its proportional reading. [5]

Before moving on, it is worth noting another prima facie problem for the connection between Anti-Additivity and the licensing of strict NPIs. Licensing by *few* suggests that it is not <u>only</u> Anti-Additive functions that license strict NPIs. Now we will turn to some data that suggests that not <u>all</u> Anti-Additive functions license strict NPIs. This data shows the unacceptability of strict NPIs under *only*.

(146) *Only Bill left until yesterday.

(147) *Only Bill has visited Sue in years.

(148) (Only Bill likes pancakes) *Only Bill likes waffles, either.

This is surprising, if we take *only Bill* to denote an Anti-Additive function. In one direction, the Anti-Additive equivalence goes through uncontroversially.

(149) $F(A) \wedge F(B) \Rightarrow F(A \vee B)$

(150) Only Bill swims and Only Bill smokes ⇒ Only Bill swims or smokes

[5] One possibility is that Anti-Additivity is not the correct logical property required by strict NPIs; it could be that the line is drawn somewhere else between DE and AA. One candidate I would like to suggest is DE + Intolerance:

(i) A function F is intolerant if $F(A) \Rightarrow \neg F(\neg A)$

Of course, every function that is both DE and Intolerant is DE. But it is also the case that every (non-trivial) AA function is DE+Intolerant. Suppose a function F were both AA and not Intolerant. If F is not intolerant, there is some element A s.t. F(A) and $F(\neg A)$. Then by AA, $F(A \cup \neg A)$. So, since $A \cup \neg A = E$ (the top element in the domain), F(E). But now since every element in the domain is a subset of E and F is DE, $\forall X\ F(X)$, i.e., F is trivial.

This is helpful in this case, since *few* has two readings (Partee, 1989) and one of them is arguably DE and Intolerant:

(ii) Cardinal: $[\![\text{few}]\!] = \lambda A.\lambda B.\ |A \cap B| > n$ (n provided by context)

(iii) Proportional: $[\![\text{few}]\!] = \lambda A.\lambda B.\ |A \cap B| < k|A|$ (k a fraction provided by context)

The proportional reading is DE and Intolerant with respect to its scope if it is assumed that $k < \frac{1}{2}$ (assuming that $|A \cap B| + |A \cap \neg B| = |A|$). The cardinal reading is not Intolerant for any value of n. This may explain why judgments with *few* are variable.

In the other direction, matters are more complicated. Intuitively, the inference is not valid.

(151) F(A∨B) ⇒ F(A) and F(B)

(152) Only Bill swims or smokes ⇏ Only Bill swims and Only Bill smokes

On the basis of *only*'s ability to license *any* and *ever*, however, von Fintel (1999) argues that in judging the logical properties of expressions for the purposes of determining their NPI-licensing abilities we need to consider a kind of entailment that abstracts from the interference of presupposition. This is called Strawson Entailment. In assessing whether α Strawson entails β we take for granted the presuppositions of β. Applying this to (152), we see that the entailment does go through.

(153) Only Bill swims or smokes (+ Bill swims and Bill smokes) ⇒ Only Bill swims and Only Bill smokes

We return to a more detailed discussion of this data in §2.2.

2.1.5 Clausemate NPI licenser is not necessary

In this section, I will show that it is not necessary for a strict NPI to be clausemates with its licenser. What is crucial, I argue, is the semantic properties of the environment of the strict NPI. The crucial case then is a sentence in which a strict NPI is licensed in an environment that is Anti-Additive thanks to a negative operator located outside of the NPIs clause.

Guerzoni (2001) notes a contrast of the kind we are interested in, in the licensing of n-words in Italian, which she analyzes as NPIs (cf. Laka 1990, Ladusaw 1992). Rizzi (1982) famously noted the following subject object contrast in the licensing of n-words. While (154) may receive the single-negation, concord reading, (155) may only receive the double negation reading.

(154) Non pretendo che arrestino nessuno.
not demand-1Sg that arrest-3Pl-Subj noone
'I don't demand that they arrest anyone'

(155)	?Non	pretendo	che	nessuno	sia	arrestato.
	not	demand-1Sg	that	no-one	be-3Sg-Subj	arrested

'I don't demand that no one be arrested'

Guerzoni argues that this contrast arises because n-words must occur in an Anti-Additive environment. The sentential negation creates an Anti-Additive environment. However, any environment containing *pretendo* - a universal quantifier over worlds - will no longer be Anti-Additive. Recall that we have already seen that NOT(EVERY( )) is not Anti-Additive. So the environments that the n-words in (154) and (155) occupy at the surface are not Anti-Additive. So, in order to be licensed the n-words must move out of the scope of *pretendo*. This is how the subject-object asymmetry is derived: due to the ECP, or some similar constraint, movement of the subject out of the tensed embedded clause is prohibited. (The ECP explanation is Rizzi's, the motivation for movement is Guerzoni's).

Importantly, Guerzoni notes that the contrast exhibited in (154) and (155) disappears when the matrix predicate *pretendo* is replaced with an existential quantifier over worlds, e.g., *permetto* '(I) allow'. An n-word in embedded subject position may receive the existential, concord reading.

(156)	Non	permetteró	che	nessuno	studente	venga	punito
	not	permit-1Sg-Fut	that	no-one	student	come-3Sg-Subj	punished

'I will not allow any student to be punished'

Why is this so? Guerzoni argues that, because the combination of negation and an existential quantifier over worlds create an Anti-Additive environment, there is no need for an n-word in the embedded clause to move in order to be in an Anti-Additive environment at LF. Since no movement is required, the ECP is irrelevant and no subject-object asymmetry arises.

With this knowledge, let's look for a parallel in the licensing of strict NPIs in English (As we will see the parallel between Italian n-words and English strict NPIs is very striking.)

Under the Anti-Additive hypothesis, we predict that the combination of a negation with an existential quantifier ought to license a strict NPI. On the other hand, a theory that relies on a clausemate condition for explaining the distribution of strict NPIs predicts that strict NPIs under negated existential predicates should be acceptable only if the predicate is Neg-Raising. But Horn (1978) argues that no existential predicate is Neg-Raising, cf (157). So, we need not worry about this confound in our test.

(157) a. Bill is allowed to smoke and Bill is not allowed to smoke (Contradictory)
b. Bill is allowed to smoke and Bill is allowed not to smoke (Consistent)

Now in order for a contrast to arise between a negated existential predicate and a negated universal predicate, it must be the case that the NPIs in the latter case do not occupy an Anti-Additive environment. As we know, NOT(EVERY()) is not an Anti-Additive environment; but we also know that some NPIs such as Italian n-words can move to an Anti-Additive environment at LF. I assume that none of the adverbial NPIs that we are considering may undergo QR. This assumption appears justified by the ill-formedness of (158).

(158) a. *An applicant is not required to have left the country in at least two years
b. *An applicant doesn't have to have left the country in at least two years.

Now compare these unacceptable sentences with the following examples in which we have substituted the existential quantifiers over worlds, *allow* and *can*, for the universal quantifiers, *require* and *have to*.

(159) a. An applicant is not allowed to have left the country in at least two years
b. An applicant can't have left the country in at least two years.

Thus it appears that we can replicate Guerzoni's n-word contrast with strict NPIs in English. Before drawing such a sanguine conclusion, one difficulty should be noted. We have so far relied on an intuitive notion of clausehood for setting up our tests. In (158) and (159), I have used example in which negation and strict NPI are separated by a non-finite clause boundary. This is not an innocent oversight. Many researchers have identified finiteness as a relevant factor in determining whether a clause boundary interferes with NPI licensing (see, a.o., Giannakidou 1997). And it must be admitted that examples analogous to (159) involving finite clause boundaries are much degraded.

(160) *it is not certain that Bill has left the country in at least two years.

(161) ??it is not possible that Bill has left the country in at least two years.

So it appears that there may still be some room for a clausemate condition to apply in the licensing of strict NPIs. On the other hand, Guerzoni's original examples (156) do involve finite clause boundaries.

2.1.6 (Partial) Cyclicity

One classic argument given for the syntactic approach to NR was that the phenomena appeared to be cyclic. That is, the raising of negation over one Neg-Raising predicate is able to feed raising of negation over another Neg-Raising predicate. So given an unbroken chain of NRPs, one embedded under the other, we expect that a negation should be able to appear an unbounded distance away from the position in which it is interpreted, much as in the case of cyclic wh-movement. For example, (162a) can be taken to mean (162b).

(162) a. I don't imagine Bill thinks Mary wants Fred to go.

b. I imagine Bill thinks Mary wants Fred to not go.

Such facts led many to the hypothesis that negation underwent a series of local cyclic movements (Fillmore, 1963):

(163) [I _ imagine [Bill _ thinks [Mary _ want [Fred not to go]]]]

Horn and Morgan (reported in Horn (1972a)) point out a problem for this simple picture. As they show, the order of the predicates in the sentence determines whether or not "cyclic" Neg-Raising is possible. They consider minimal pairs such as (164). Both *think* and *want* allow for Neg-Raising.

(164) a. I don't think Bill wants Mary to leave.

b. I don't want Bill to think Mary left.

According to Horn and Morgan, cyclic Neg-Raising is possible in (164a) but not in (164b). That is, (164a) can be understood as equivalent to (165a), but (164b) cannot be understood as equivalent to (165b).

(165) a. I think Bill want Mary not to leave.

b. I want Bill to think Mary didn't leave.

They support this claim with evidence from NPI-licensing:

(166) a. I don't think Bill wants Mary to leave until tomorrow.

b. #I don't want Bill to think Mary left until yesterday.

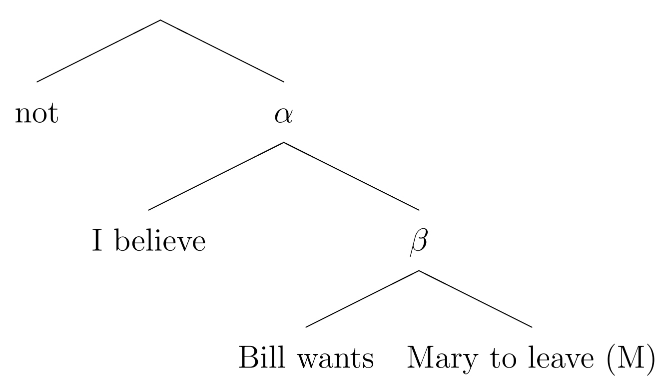
The strict NPI *until* is not licensed in (166), in which Neg-Raising predicate *think* is embedded under Neg-Raising predicate *want*. In the next section, we will show how the cyclicity of Neg-Raising follows also from the presuppositional analysis of Neg-Raising that we are pursuing. Furthermore, we will show that the asymmetry in cyclicity observed by Horn and Morgan follows from our analysis of Neg-Raising and independently justified principles of presupposition projection.

Explaining the contrast in (166)

The presuppositional approach to Neg-Raising offers an interesting explanation of this contrast. Consider how the presuppositional analysis captures the cyclicity of NR in (167), in which NRP *want* is embedded below NRP *think*.

(167) I don't think Bill wants Mary to leave.

(168)



(169) Presupp(β) = Bill want M or Bill wants $\neg$M

Presupp(α) = (i) I believe Bill wants M or I believe Bill does not want M

(ii) I believe Bill wants M or Bill wants $\neg$M

I have indicated the presuppositions of constituents α and β of the tree in (177) in (178). The entire structure (177) inherits the presuppositions of α. Presupposition (ii) of α is the Excluded Middle presupposition associated with *believe*. Presupposition (i), on the other hand, derives from projection of the presupposition of β. Now the assertion of (167) is (170).

(170) $\neg$I think Bill wants Mary to leave.

This combined with presupposition (ii) of α (178) gives us (171).

(171) I think $\neg$Bill wants Mary to leave.

This combined with presupposition (i) of α in (178)) entails that

(172) I think Bill wants Mary to leave.

If we try to use this reasoning when the predicates are in the reverse order we run into a problem. It is well-known that desire predicates differ from doxastic predicates in their presupposition projection properties (cf. Karttunen (1974a), Heim (1992)) . Doxastic predicates, on the one hand, assert that a proposition holds in some doxastic alternatives and presuppose that the presuppositions of its complement hold among those doxastic alternatives. Desire predicates, on the other hand, assert that the complement holds in some bouletic alternatives but presuppose that the presuppositions of their complement hold in the subject's doxastic alternatives. For example, (174) presupposes that Bill believes he has a cello and (175) presupposes not that Bill wants to have a cello, but that he believes he has one.

(173) Bill will sell his cello.

Presupposition: Bill has a cello.

(174) Bill thinks he will sell his cello.

Presuppositions: Bill thinks he has a cello.

(175) Bill wants to sell his cello.

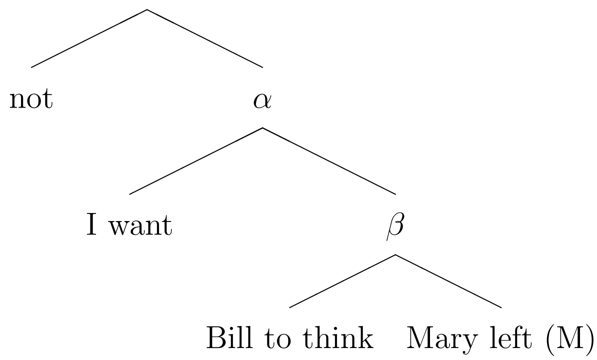
Presupposition: Bill thinks he has a cello

(#Bill wants to have a cello)

Knowing this, consider again the case in which *think* is embedded under *want*, repeated below as (176):

(176) I don't want Bill to think Mary left.

(177)



(178) Presupp(β) = Bill thinks M or Bill thinks ¬M

Presupp(α) = (i) I want Bill to think M or I want Bill not to think M

(ii) I believe Bill thinks M or Bill thinks ¬M

The assertion of (176) is (179).

(179) ¬I want Bill to think Mary left

This together with presupposition (ii) of α entails that

(180) I want ¬Bill to think Mary left

In the case of (167) we were able to use presupposition (i) of α to infer the final 'cyclic' step of Neg-Raising. In this case we cannot. I can believe that Bill believes M or that he believes not-M, want that it not to be the case that he believes M and still not want Bill to believe not-M.

Implementing this account formally encounters one technical difficulty. That difficulty is how to analyze the contribution of presuppositional constituents contained in other presuppositional constituents. This analysis requires that the presuppositions of the embedded predicates do not contribute to the meaning of the Excluded Middle presupposition of the predicates that embed them. More specifically, I propose that the presuppositions of the embedded item are "cancelled" within the Excluded Middle presupposition. They do contribute to the presuppositions of the larger constituent through projection. Let's see what I mean by this by spelling out some concrete lexical entries ['$D_{w,x}(w')$' abbreviates 'w′ is doxastically accessible to x from w']:

(181) $[\![$ believe $]\!]^w$(p)(x) is defined only if
(i) $\forall w'$ [$D_{w,x}(w') \rightarrow$ [$p(w') = 1$ or $p(w') = 0$]]
(projection of the presuppositions of the embedded clause)
(ii) $\forall w'$ [$D_{w,x}(w') \rightarrow p(w') = 1$] or $\forall w'$ [$D_{w,x}(w') \rightarrow p(w') \neq 1$]
when defined, $[\![$ believe $]\!]^w$(p)(x) = 1 iff $\forall w'$ [$D_{w,x}(w') \rightarrow p(w') = 1$]

In this definition, the crucial part is the consequent of the second disjunct in presupposition (ii). Here we have crucially written '$p(w')\neq 1$' rather than '$p(w')=0$'. This effectively cancels the presupposition of p within the Excluded Middle presupposition of $[\![$believe$]\!]$. In this particular lexical entry, the distinction does not ultimately make a difference for the definedness conditions. Clause (i) which projects the presuppositions of the embedded clause guarantees that p is true or false in each of the subjects belief worlds. So, if every belief world w is such that p is not true in w, then every belief world w is such that p is false in w. That is, we might just as well have written '$p(w')=0$'.

This decision to use '$p(w')\neq 1$' rather than '$p(w')=0$' in the Excluded Middle presupposition has a more dramaic effect in the lexical entry of *want*. The reason being, of course, that the projection clause of the definedness conditions does not match up with the excluded presupposition as it did in the lexical entry of *believe*. ['$B_{w,x}(w')$' abbreviates 'w′ is bouletically accessible to x from w']

(182) $[\![$ want $]\!]^w$(p)(x) is defined only if
(i) $\forall w'$ [$D_{w,x}(w') \rightarrow$ [$p(w') = 1$ or $p(w') = 0$]]
(projection of the presuppositions of the embedded clause)
(ii) $\forall w'$ [$B_{w,x}(w') \rightarrow p(w') = 1$] or $\forall w'$ [$B_{w,x}(w') \rightarrow p(w') \neq 1$]
when defined, $[\![$ want $]\!]^w$(p)(x) = 1 iff $\forall w'$ [$B_{w,x}(w') \rightarrow p(w') = 1$]

Clause (i) of the definedness conditions projects the presuppositions of the embedded clause, requiring that the subject of *want* believe them.

(183) $[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(p)(a))(c)$ is defined only if

(i) $\forall w'$ [$D_{w,c}(w') \rightarrow$ [$[\![\text{believe}]\!]^{w'}(p)(a) = 1$ or $[\![\text{believe}]\!]^{w'}(p)(a) = 0$]] iff

$\forall w'[D_{w,c}(w') \rightarrow [\forall w''[D_{w',a}(w'') \rightarrow p(w'') = 1\ ] \vee \forall w''[D_{w',a}(w'') \rightarrow p(w'')=0]]]$

(ii) $\forall w'[B_{w,x}(w') \rightarrow [\![\text{believe}]\!]^{w'}(p)(a)=1]$ or

$\forall w'[B_{w,x}(w') \rightarrow [\![\text{believe}]\!]^{w'}(p)(a) \neq 1]$ iff

$\forall w'[B_{w,c}(w') \rightarrow \forall w''[D_{w',a}(w'') \rightarrow p(w'') = 1]]$ or

$\forall w'[B_{w,c}(w') \rightarrow \neg\forall w''[D_{w',a}(w'') \rightarrow p(w'') = 1]\ ]$

when defined, $[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(p)(a))(c) = 1$ iff

$\forall w'$ [$B_{w,c}(w') \rightarrow [\![\text{believe}]\!]^{w'}(p)(a) = 1$] iff

$\forall w'[B_{w,c}(w') \rightarrow \forall w''[D_{w',a}(w'') \rightarrow p(w'') = 1]\ ]$

(184) Equivalences used in (183)

a. $[\![\text{believe}]\!]^{w'}(p)(a) = 1$ iff $\forall w''[D_{w',a}(w'') \rightarrow p(w'') = 1]$

b. $[\![\text{believe}]\!]^{w'}(p)(a) = 0$ iff $\forall w''[D_{w',a}(w'') \rightarrow p(w'') = 0]$

This is precisely the result we want. The negation of (183) does not entail that c wants a to believe that not p. It merely entails that c wants a to <u>not</u> believe that p. It furthermore presuppose that c believes that a either believes that p or that not p. However, without any further postulates about the relationship of belief worlds to desire worlds, this does not entail that c wants a to believe that not p. I t may be that practically we do assume beliefs constrain desire in this way. It is my hypothesis that this constraint is not imposed by the grammar.

Furthermore, given these proofs, it is simple to show that (176) does not contain a constituent that is Anti-Additive with respect to the most deeply embedded clause. Thus we correctly predict that strict NPIs are not licensed in (176).

(185) $[\![\text{c doesn't want a to believe that p}]\!]^w = 1$ iff

$[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(p)(a))(c) = 0$

(negation preserves presuppositions)

Now we show that (186ii) does not follow from (186i).

(186) i. $[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(p)(a))(c) = 0$ and

$[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(q)(a))(c) = 0$

ii. $[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(p \vee q)(a))(c) = 0$

To do so, we construct a simple model in which (i) holds and (ii) does not.

(187) a. B_c: $w \mapsto \{w_1, w_2\}$

b. D_a: $\begin{array}{ccc} w_1 & \mapsto & \{w_3, w_4\} \\ w_2 & \mapsto & \{w_5, w_6\} \end{array}$

c. $p(w_3) = p(w_5) = 1$; $p(w_4) = p(w_6) = 0$

d. $q(w_3) = q(w_5) = 0$; $q(w_4) = q(w_6) = 1$

e. $\forall u \in \{w_3, w_4, w_5, w_6\}\ p \vee q(u)=1$

In every one of c's bouletic alternatives there is a doxastic alternative for a in which p is false. Similarly for q. A quick inspection of (183) shows that this verifies (i). But (ii) does not hold. In fact, $[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(p \vee q)(a))(c) = 1$ in this model. In every one of c's bouletic alternatives, $p \vee q$ is true in every one of a's doxastic alternatives. This makes $[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(p \vee q)(a))(c)$ both defined and true.

Thus the environment $[\![\text{want}]\!]^w(\lambda u.[\![\text{believe}]\!]^u(\_\_\_)(a))(c)$ is not Anti-Additive since it fails the inference in (188).

(188) $F(A) \wedge F(B) \Rightarrow F(A \vee B)$

This contrast with the case in which *want* is embedded under *believe*. In that case, the inference in (188) does indeed go through.

(189) $[\![\text{believe}]\!]^w(\lambda u.[\![\text{want}]\!]^u(p)(a))(c)$ is defined only if

(i) $\forall w'\ [\ D_{w,c}(w') \rightarrow [\ [\![\text{want}]\!]^{w'}(p)(a) = 1$ or $[\![\text{want}]\!]^{w'}(p)(a) = 0\ ]\ ]$ iff

$\forall w'[D_{w,c}(w') \rightarrow [\ [\forall w''[B_{w',a}(w'') \rightarrow p(w'') = 1\ ] \vee \forall w''[B_{w',a}(w'') \rightarrow p(w'') \neq 1]$

$] \wedge \forall w''[D_{w',a}(w'') \rightarrow [p(w'') = 1 \vee\ p(w'') = 0\ ]\ ]]]$

(ii) $\forall w'[D_{w,c}(w') \rightarrow [\![\text{want}]\!]^{w'}(p)(a)=1] \vee \forall w'[D_{w,c}(w') \rightarrow [\![\text{want}]\!]^{w'}(p)(a) \neq 1]$ iff

$\forall w'[D_{w,c}(w') \rightarrow [\ \forall w''[B_{w',a}(w'') \rightarrow p(w'') = 1] \wedge$

$\forall w''[D_{w',a}(w'') \rightarrow [p(w'')=1 \text{ or } p(w'')=0]]]]$ or

$\forall w'[D_{w,c}(w') \rightarrow [\exists w''[D_{w,a}(w'') \wedge p(w'') \neq 1 \wedge p(w'') \neq 0] \vee$

$\exists w''[B_{w,a}(w'') \wedge p(w'')\neq 1]]]$

when defined, $[\![\text{believe}]\!]^{w}(\lambda u.[\![\text{want}]\!]^{u}(p)(a))(c) = 1$ iff

$\forall w'\ [\ D_{w,c}(w') \rightarrow [\![\text{want}]\!]^{w'}(p)(a) = 1\ ]$ iff

$\forall w'[D_{w,c}(w') \rightarrow [\forall w''[B_{w',a}(w'') \rightarrow p(w'') = 1] \wedge$

$\forall w''[D_{w'}(w'') \rightarrow [p(w'')=1 \text{ or } p(w'')=0]]]]$

(190) Equivalences used in (189)

$[\![\text{want}]\!]^{w'}(p)(a) = 1$ iff $\forall w''[B_{w',a}(w'') \rightarrow p(w'') = 1] \wedge$

$\forall w''[D_{w'}(w'') \rightarrow [p(w'')=1 \text{ or } p(w'')=0]]$

$[\![\text{want}]\!]^{w'}(p)(a) \neq 1$ iff $\exists w''[D_{w',a}(w'') \wedge p(w'')\neq 1 \wedge p(w'')\neq 0] \vee$

$\exists w''[B_{w',a}(w'') \wedge p(w'')\neq 1]$

$[\![\text{want}]\!]^{w'}(p)(a) = 0$ iff $\forall w''[B_{w',a}(w'') \rightarrow p(w'')\neq 1] \wedge$

$\forall w''[D_{w',a}(w'') \rightarrow [p(w'')=1 \text{ or } p(w'')=0]]$

(191) $[\![\text{believe}]\!]^{w}(\lambda u.[\![\text{want}]\!]^{u}(p)(a))(c) = 0$ iff

(i) $\forall w'[D_{w,c}(w') \rightarrow [[\forall w''[B_{w',a}(w'') \rightarrow p(w'') = 1\ ] \vee \forall w''[B_{w',a}(w'') \rightarrow p(w'')\neq 1]$

$\wedge\ \forall w''[D_{w',a}(w'') \rightarrow [p(w'') = 1 \vee p(w'') = 0\ ]]]]$

(ii) $\forall w'[D_{w,c}(w') \rightarrow [\ \forall w''[B_{w',a}(w'') \rightarrow p(w'') = 1] \wedge$

$\forall w''[D_{w',a}(w'') \rightarrow [p(w'')=1 \text{ or } p(w'')=0]]]]$ or

$\forall w'[D_{w,c}(w') \rightarrow [\exists w''[D_{w,a}(w'') \wedge p(w'')\neq 1 \wedge p(w'')\neq 0] \vee$

$\exists w''[B_{w,a}(w'') \wedge p(w'')\neq 1]]]$

(iii) $\neg\forall w'[D_{w,c}(w') \rightarrow [\forall w''[B_{w',a}(w'') \rightarrow p(w'') = 1] \wedge$

$\forall w''[D_{w'}(w'') \rightarrow [p(w'')=1 \text{ or } p(w'')=0]]]]$

We may simplify this as follows:

(192) (ii) and (iii) are equivalent to (iv):

(iv) $\forall w'[D_{w,c}(w') \rightarrow [\neg\forall w''[D_{w,a}(w'') \rightarrow p(w'')=1 \vee p(w'')=0] \vee$

$\neg\forall w''[B_{w,a}(w'') \rightarrow p(w'')=1]]]$

Notice now that (i) and (iv) entail (v):

(193) (v) $\forall w'[D_{w,c}(w') \rightarrow [\forall w''[B_{w',a}(w'') \rightarrow p(w'')\neq 1] \wedge$

$\forall w''[D_{w',a}(w'') \rightarrow [p(w'') = 1 \vee p(w'') = 0]]]]$

Notice now that if (v) holds for another proposition q, then (193) holds of p∨q as well. Why? well if in every one of c's doxastic alternatives p is false in every one of a's bouletic alternatives and the same holds of q, then in every one of c's doxastic alternatives p∨q is false in every one of a's bouletic alternatives. Furthermore if in every one of c's doxastic alternatives, p is either true or false in each of a's doxastic alternatives, and the same holds of q, then in every one of c's doxastic alternatives, p∨q is true or false in each of a's doxastic alternatives. These facts verify that $[\![\text{believe}]\!]^w(\lambda u.[\![\text{want}]\!]^u(p\vee q)(a))(c) = 0$. The inference in the other direction (194) is straightforward. The crucial step in our reasoning, what differentiated this case from the last, was the use of presupposition (i) to draw the inference in (193).

(194) $F(A\vee B) \Rightarrow F(A) \wedge F(B)$

Thus, the environment $[\![\text{believe}]\!]^w(\lambda u.[\![\text{want}]\!]^u(\_\_\_)(a))(c)$ is Anti-Additive.

Summary In this section, we have seen how a Zwarts approach to the distribution of strict NPIs, a Barstch/Heim approach to Neg-Raising, and some independently justified principles of presupposition projection dovetail neatly and predict an intricate contrast in the licensing of strict NPIs under multiple Neg-Raising Predicates.

2.2 Further Issues in the Licensing of Strict NPIs

In §2.1.4 we pointed out some data potentially problematic to our approach to the licensing of strict NPIs

(195) a. *Only John arrive until 5.

b. *Only John visited Marry in years.

c. (Only John likes pancakes.) *Only John likes waffles either.

The potential problem with these examples is that, under a certain perspective, *only John* can be argued to create an Anti-Additive environment. Under what perspective? von Fintel 1999 argues, on the basis of *only*+DP's ability to license NPIs such as *any* and *ever*, that *only* should be understood as Downward Entailing, despite the intuitive failure of such reasoning as (196).

(196) (i) Only John ate vegetables $\not\Rightarrow$ (ii) only John ate broccoli

The problem with the inference in (196) is that it is possible for (i) to be true and (ii) undefined (under Horn's 1969 analysis of *only*). This happens when John eats kale, no one else eats any vegetable and John eats no broccoli. Von Fintel proposes that we adjust our notion of entailment, so that *only*+DP does come out Downward Entailing. He propose that for the purposes of NPI-licensing use instead Strawson Downward Entailment:

(197) Cross-Categorial Entailment ($\Rightarrow$)

a. For p, q of type t: p $\Rightarrow$ q iff p = False or q = True.

b. For f, g of type $<\sigma,\tau>$: f $\Rightarrow$ g iff for all x of type σ: f(x) $\Rightarrow$ g(x).

(198) Strawson Downward Entailingness

A function f of type $<\sigma,\tau>$ is Strawson-DE iff for all x, y of type σ such that x $\Rightarrow$ y and f(x) is defined: f(y) $\Rightarrow$ f(x).

As von Fintel shows, *only*+DP indeed comes out Strawson-DE under this definition. This entails that *only*+DP satisfies the Right-to-Left implication in the definition of Anti-Additivity (if we understand '$\Leftrightarrow$' to stand for symmetric Strawson entailment).

(199) $F(A) \wedge F(B) \Leftrightarrow F(A \vee B)$

Furthermore, *only*+DP uncontroversially validates the Left-to-Right implication:

(200) Only John drinks and Only John smokes $\Rightarrow$ Only John drinks or smokes

So, from the perspective of Strawson Entailment, *only*+DP is Anti-Additive. But why the judgments in (195)? **I argue that it is strict entailment and not Strawson entailment that figures in the statement of the licensing condition on strict NPIs.** By way of supporting this generalization, I observe that two other environments that von Fintel identifies as Strawson Downward Entailing fail to license strict NPIs. Note that these constructions also validate the Left-to-Right implication of (199).

(201) Adversatives

Fred hasn't arrived. *Bill is sorry Sue has arrived either.

*Sue is sorry that Bill arrived until five

*Sue is sorry that Bill has visited John in years

(202) Antecedent of a Conditional

Fred isn't here. *If Bill is here either, then Mary is upset

*If Bill arrived until five, Mary was upset.

*If Sue has visited Bill in years, then Mary is upset.

(203) Bill is sorry Sue is here and Bill is sorry Fred is here ⇒ Bill is sorry Sue is here or Fred is here

(204) If Bill arrived at five, then Mary is upset and if Sue arrived at six, then Mary is upset ⇒ If Bill arrived at five or Sue arrived at six, then Mary is upset

So, these constructions would also count as Anti-Additive, if our underlying notion of entailment were Strawsonian. Let's refer to such functions as Strawson Anti-Additive(AA) and to functions that are AA on the standard notion of entailment strict-AA.

Note that there is one construction analyzed by von Fintel as (merely) Strawson Downward Entailing that defies this trend. This is the case of superlatives, which von Fintel 1999 assigns the semantics in (205).

(205) $[\![\text{the...–est}]\!](P)(Q)(\alpha)$ is defined only if $Q(\alpha)$=True

If defined, $[\![\text{the...–est}]\!](P)(Q)(\alpha)$=True iff

$$\forall x \neq \alpha\text{: } (Q(x) = \text{True} \rightarrow \iota dP(x)(d) < \iota dP(x)(d))$$

Under this analysis, superlatives turn out to be Strawson-DE. And indeed, in this case, strict NPIs are acceptable in a relative clause in the scope of the superlative morpheme. Furthermore, superlatives do intuitively validate the Left-to-Right direction of (199):

(206) Erin is the tallest girl in this class and Erin is the tallest girl in that class ⇒ Erin is the tallest girl in this class or that class.

So, superlatives do create Strawson-AA environments. And, actually, in this case, strict NPIs are licensed in relative clauses in the scope of a superlative morpheme.

(207) Superlatives

Erin is the tallest girl Fred has ever seen.

She is the tallest girl Bill has ever seen either.[6]

Erin is the tallest girl John has seen in years.

The tallest girl John had seen until Friday walked in the room.

If we wish to maintain that the licensing of strict NPIs requires a strict-AA environment and not merely a Strawson-AA environment, then we need a different semantics for superlatives: one in which superlatives create a strict-AA environment. First, let us recall the motivation for a Strawson-DE analysis of superlatives – superlatives do not intuitively validate downward inferences:

(208) (i) Emma is the tallest girl in her class $\not\Rightarrow$ (ii) Emma is the tallest girl in her class to have learned the alphabet

((i) can be true and (ii) undefined if Emma hasn't learned the alphabet)

While this is certainly true it does not necessarily tell us that superlatives do not create strict-AA environments. There is another instance of a construction that does not intuitively validate downward inferences and yet licenses strict NPIs: a negative universal quantifier + exceptive.

(209) No one but Bill left the party $\not\Rightarrow$ No one but Bill left the party early

[6] This judgment has been controversial with informants. To support the reported judgment, here are a handful of natural examples:

(i) Xine is by far the best DVD player I've used, and in fact the best media app I've ever used either.

`http://forums.kustompcs.co.uk/archive/index.php/t-7450.html`

(ii) Occasional recumbents pass, and several want to take a closer look at the first tadpole trike they have ever laid eyes on. For that matter, it is the only one I have ever seen either.

`http://www.qsl.net/ws8g/century.html`

(iii) A swimming dwarf- how rare! [...] I'd have to say this is the first swimming dwarf i have ever seen either!

`http://elfwood.lysator.liu.se/art/s/e/sebklement/swimdwarf.jpg.html`

(Conclusion does not follow if Bill did not leave early)

(210) Exceptives

a. No one but Bill likes pancakes.

No one but Bill likes WAFFLES either.[7]

b. No one but Bill has visited Sue in years.

c. No one but Bill left until yesterday.

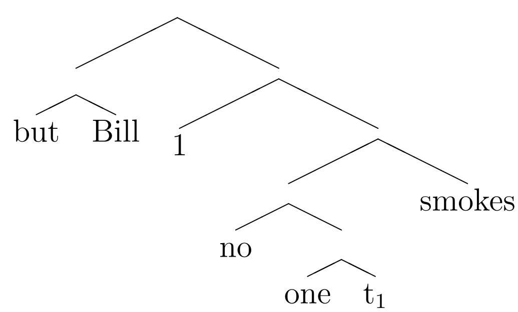
Gajewski (2004) offers an analysis of exceptive phrase which makes it possible to view sentences such as (210) as containing strict-AA environments embedded in larger environments that are not strict-AA. According to Keenan & Stavi (1986) and von Fintel (1993), the exceptive is interpreted as a consituent with the quantifier.

(211) $[\![\text{ no boy but Bill }]\!] = \lambda X. [\![\text{boy}]\!] \cap X = \{\text{Bill}\}$

This function is clearly not Downward Entailing. Gajewski proposes, for independent reasons, that the contribution of the exceptive phrase should be factored out of the quantifier, see Appendix A. Specifically, he proposes that the exceptive phrase QRs out of the DP taking sentential scope:

(212) No one but Bill smokes

(213)



This movement leaves behind a constituent [no one t_1] that creates a strict-AA environment in its scope. On the one hand, it is the presence of this strict-AA environment that licenses strict NPIs. On the other hand, the embedding of this environment in the scope of the exceptive phrase prevents downward inferences from being valid for the sentence as a whole.

[7]To my knowledge the difference in the abilities of *only*+DP and *no one but*+DP to license *either* was first noted in Nathan (1999). Thanks to Lance Nathan for bringing this to my attention.

Below, we will take this analysis as a guide for a tentative analysis of superlatives that creates strict-AA environments. Before doing so, I show in the next section that the generalization we are endorsing has validity beyond strict NPIs in English.

2.2.1 English strict NPIs and N-words in Italian and Spanish

In §2.1.5 we showed that English strict NPIs and Italian n-words showed a similar licensing asymmetry with respect to the negations of existential and universal quantifiers over worlds. A good deal of research has been done on the distribution of Romance n-words. Much of what has been discovered about their distribution overlaps with the distribution of strict NPIs in English. For example, it has been frequently proposed that Romance N-words require strong negative licensers, e.g., Anti-Additive operators (cf. Ladusaw (1992), Guerzoni (2001), a.o.). Consider (214a).

(214) a. Non-Anti-Additive

*Meno di tre studenti hanno mangiato niente.
Less than three students have eaten n-thing

b. Conditional

??Se Maria si accorgesse niente, sarebbe un problema.
If Mary noticed n-thing, it would be a problem

c. *Only*

??Solo Maria ha visto nessuno degli studenti.
Only Mary has seen n-one of the students

d. Adversatives

*Mi spiacerebbe che tu vedessi nessuno.
I would be sorry that you saw n-body

(Alonso-Ovalle and Guerzoni, 2003)

It has also been noted that this generalization is not adequate. Romance n-words are not licensed by Strawson-AA environments, such as the antecedents of conditionals, the scope of *only*+DP, and adversatives. This has led to a variety of proposals for accounting for the distribution, e.g., replacing Downward Entailingness with Non-Veridicality in the licensing conditions on NPIs (Giannakidou, 1997).

Less frequently noted is the fact that superlatives also license the existential concord reading of n-words in Italian and Spanish:

(215) É l'idea più stupida che abbia mai avuto nessuno
is the idea stupide that has ever had n-one
'It's the stupidest idea anyone ever had' (Acquaviva 1997)

(216) Es la ultima vez que te digo nada
is the last time that you I tell n-thing
'It's the last time that I tell you anything' (Herburger, 1997)

The starting point for an analysis is the simple semantics for superlatives proposed in Heim (1985) (and argued for recently in Sharvit and Stateva (2002)).

(217) $[\![$-est$]\!]$(K)(R)(x) is defined only if x∈K and for all y in K there is a degree d such that R(d)(y) = 1. Whenever defined, $[\![$-est$]\!]$(K)(R)(x) =1 iff there is a degree d such that {z∈K: R(d)(z) = 1}={x}

The argument 'K' is a comparison set provided by context, and 'R' is the relation between an individual and a degree denoted by the abstraction over the degree variable in the constituent containing the head noun and a gradable adjective:

(218) [the [-est$_K$ λd[tall(d) [$_\alpha$girl that I know]]]]

This semantics does not create a downward entailing environment in the position α in (218). Though it is Strawson-DE. Our goal is to take this semantics and change it so that it creates a strict-AA environment. What gets in the way of this property is clear. Consider again the intuitive failure of the Downward Entailing inference.

(219) Emma is the tallest girl in her class $\not\Rightarrow$ Emma is the tallest girl in her class to have learned the alphabet

We can see the superlatives statements as having two components. When we say that Emma is the tallest girl in her class we say (i) that Emma is taller than any girl in the class (that is not Emma) and (ii) that she is a girl in the class. It is (ii) that gets in the way of the inference in (219). If the conclusion did not require that Emma have learned the alphabet the inference would go through. Thus, I propose that we separate these two parts of the semantics. We want an operator at one level that contributes (i) and at higher level an operator that contributes (ii) and thereby masks the strict-AA property of the construction.

To achieve the first objective, we can amend the semantics of *-est* to be (220).

(220) $[\![\text{-est}]\!] = \lambda R.\lambda P.\lambda x.\exists d[\ R(d)(x){=}1 \text{ and } \neg\exists y[\ P(y)\ \&\ y{\neq}x\ \&\ R(d)(y) = 1]\ ]$

This denotation requires some comment. First, I have ignored the contribution of the context variable which is irrelevant for our purposes. Second, instead of feeding the superlative morpheme one relation between a degree and an individual, I feed it the relation denoted by a gradable adjective as well as the predicate of individuals denoted by the sister of the adjective. So, now superlatives have the structure in (221).

(221) [the [[-est tall(d)] [$_\alpha$girl that I know]]]

This denotation for *-est* does create a strict-AA environment α which is denotes an argument of the superlative morpheme. The problem with this denotation of course is that it leaves open the possibility that the tallest girl might be a boy, or an SUV. Something needs to guarantee that the tallest girl is a girl.

I will make two suggestions for how this might be accomplished. One possibility is that superlatives involve some unpronounced structure that contributes this inference. For example, we might take [[-est tall] girl that I know] to form an adjectival phrase that modifies an NP so that the entire structure looks like the following:

(222) [the [$_{ADJ}$ [-est tall(d)] [$_\alpha$girl that I know]] [$_{NP}$ ~~girl that I know~~]]

Here the adjectival phrase contributes a predicate true of individuals x taller than any girl that I know ($\neq$ x) and the noun phrase a predicate true of girls that I know. By Predicate Modification, we arrive at the predicate we want.

Such an analysis may find precedents in the matching analysis of relative clauses and the elliptical analysis of comparative quantifiers found in Hackl (2000).

Another possiblity is to draw an explicit analogy between superlatives and exceptives. The attentative reader will have noticed a striking similarity between the

Keenan & Stavi semantics for exceptives and the Heim (1985) semantics for superlatives.

(223) $[\![$ no boy but Bill $]\!] = \lambda X.\ [\![\text{boy}]\!] \cap X = \{\text{Bill}\}$

(224) $[\![\text{-est}]\!]$(K)(R)(x) is defined only if x∈K and for all y in K there is a degree d such that R(d)(y) = 1. Whenever defined, $[\![\text{-est}]\!]$(K)(R)(x) =1 iff there is a degree d such that {z∈K: R(d)(z) = 1}={x}

The truth-conditions of superlative constructions can be paraphrased with an exceptive:

(225) Emma is the tallest girl in the class

(226) There is a degree d such that no girl in the class but Emma is d-tall

If we use this paraphrase as a guide for the articulated syntax of superlative constructions, we come up with (227). The choice between (227a) and (227b) does not affect the truthconditions of the constructions, cf. (228). It may have some effect on the pragmatics.

(227) [the tallest girl in the class]

a. [the λx[∃d[[BUT x]$_1$[[NO girl in the class t$_1$]$_2$[t$_2$ d-tall]]]]]

b. [the λx[∃d[[BUT x]$_1$[[NO d-tall t$_1$]$_2$[t$_2$ girl in the class]]]]]

(228) No girl but Sue smokes iff No smoker but Sue is a girl

2.3 Remaining Issues

In this section, I discuss a few issues held over from previous sections. First, I discuss the question of whether there is independent evidence of the projection of Excluded Middle presuppositions. The results are mixed, raising questions about the nature of the explanation in this chapter. Next, I briefly address two more of the arguments given in favor of the syntactic theory of Neg-Raising in §1.4.2, showing that the presuppositional theory offers simple alternative analyses.

2.3.1 Neg-Raising and Presupposition

To this point in the chapter, we have used the Bartsch/Heim approach to Neg-Raising to explain some intricate facts about NPI-licensing. We have not yet stopped to ask whether Neg-Raising predicates pass certain basic tests for presupposition. The evidence turns out to be mixed, tending towards suggesting that Neg-Raising predicates are not presuppositional.

Basic Tests

One test for presupposition is truthvalue judgment. If Neg-Raising predicates carry an Excluded Middle presupposition, then we might expect to find the truth of (229) difficult to judge in a scenario in which we know that Sue has no opinion.

(229) Sue thinks Bill is here.

Most people, however, have no problem judging this sentence false in such a scenario. It is unclear though what we should make of this. von Fintel (2004) has argued that truthvalue judgments are not the most reliable indicators of the presence of a presupposition.

There are certain environments linguists use to diagnose the presence of a presupposition. The most common are the antecedents of conditionals, yes/no questions, and epistemic modals.

(230) If Bill thinks Sue is here, then he will leave.

(231) Does Bill think Sue is here?

(232) Bill might think Sue is here.

If *think* introduces the presupposition that its subject is opinionated about the truth or falsity of its complement, then we expect each of the sentences to imply that Bill has an opinion as to whether Sue is here. This does not seem to be the case.

Another important environment for diagnosing presuppositions is negation. We can only use this environment at pains of circularity. On the other hand, there does seem to be something marked about negating a Neg-Raising predicate without assuming the Excluded Middle. For example, to use the negation of *think* to convey that someone has no opinion about the embedded clause - as opposed to conveying that the subject thinks it is false - requires special marking. Stressing the negation or the predicate facilitates this reading.

(233) Bill DOESN'T think Sue is here. He has no opinion.

(234) Bill doesn't THINK Sue is here. He is unopinionated.

This special marking is similar to the marking generally required to cancel a presupposition:

(235) The King of France is NOT hiding in the closet. There is no King of France

So, the marking required to convey the non-NR reading of Neg-Raising predicates suggests that a presuppositional analysis is superior to a theory that posits a simple structural ambiguity between the Neg-Raised and non-Neg-Raised "readings".

There is a way of looking at one of the basic tests that suggests NRPs do presuppose the Excluded Middle. If you consider the negative answer to a yes/no question containing a NRP, it does imply the Neg Raised reading:

(236) Do you think Bill is here? No (=I think he isn't)

(237) Does Bill want to leave? No (=he wants to not leave)

(238) Should Bill leave? No (=he should not leave)

This test, however, is somewhat questionable lacking an analysis of "No". If "No" introduces a negation, then this case may not be independent of the simple cases with sentential negation.

2.3.2 No Factive Neg-Raisers

Our presuppositional analysis explains a gap in the class of NR predicates. As first noted by Kiparsky and Kiparsky 1970, there are no factive NR predicates. What does this mean? A predicate is considered to be NR if placing negation above the predicate yields a reading equivalent to placing negation below the predicate. So, as discussed above, (239) has a reading equivalent to (240).

(239) John doesn't believe Mary is here.

(240) John believes that Mary is not here.

There is no factive predicate that follows this pattern. For example, (241) has no reading equivalent to (242). While (242) presupposes that Mary is not here, (241) presupposes that Mary is here.

(241) Bill doesn't know that Mary is here.

(242) Bill knows that Mary is not here.

This fact follows straightforwardly from our proposal. If [x *knows* p] presupposes p, there is simply no presupposition that can be added to *know* that will yield the result that [not [x *knows* p]] presupposes ¬p. Why? As is well-known, the negation of a constituent inherits the presuppositions of that constituent. So, there is simply no way under the presuppositional analysis of Neg-Raising for a Neg-Raising factive, in the sense of Kiparsky and Kiparsky, to exist.

2.3.3 High NPIs

Recall from §1.4.2 that it is not possible for the negation of a Neg-Raising predicate P to license both an NPI above P and an NPI below P.

(243) * I didn't ever think that Bill would leave until tomorrow.

(244) a. I don't at all think that John will leave.

b. I don't think John will leave until next week.

c. * I don't at all think John will leave until next week.

I propose that this can also be explained under the presuppositional account of Neg-Raising. Recall our analysis of Neg-Raising when the negation is expressed by a negative universal quantifier such as *no one* or *never*.

(245) No one thinks Bill left (until yesterday).

In order for the environment of the embedded clause to be Anti-Additive, it is necessary that the Excluded Middle presupposition of *think* projects universally through the subject quantifier. Such a universal presupposition has been proposed by Heim (1982).

(246) Presupposition of (245)

Everyone either thinks Bill left or thinks Bill didn't leave

This presupposition together with (245) entails that

(247) Every one thinks Bill didn't leave

This inference is sufficient to make the environment of the embedded clause Anti-Additive. Suppose, however, that the Excluded Middle presupposition had projected existentially instead.

(248) Some one either thinks Bill left or thinks he didn't leave.

From this and the assertion, the most we can infer is that

(249) Some one thinks Bill didn't leave.

This is not sufficient to create an Anti-Additive environment since *someone thinks Bill didn't leave and someone thinks Sue didn't leave* does not entail that *someone thinks neither that Bill left nor the Sue left*.

Recall now that indefinites do not project universal presuppositions (see §2.1.3):

(250) A fat man pushed his bicycle Heim 1982

Now if we assume that indefinite NPIs like *ever* (and *at all*) are like normal indefinites in this regard, we see why the double licensing in (243) is not possible. The presupposition projected through the indefinite NPI and inherites by the higher negation is not strong enough to support Anti-Additive inferences in the lower clause. That is the presence of the higher NPI interferes with the licensing of the lower NPI.

2.4 Summary

In this chapter, we have seen how a presuppositional analysis provides an elegant account of NPI licensing and cyclicity in Neg-Raising. The account depends crucially on the projection of an Excluded Middle presupposition under negation, negative quantifiers and other Neg-Raising predicates. In the final section, however, we saw that the Excluded Middle presupposition does not seem to project in other typically diagostic environments such as the antecedent of conditionals. Finally, we showed how the presuppositional account can explain the lack of factive Neg-Raisers and why NPIs in a higher clause interfere with the licensing of strict NPIs in lower clauses.

Now we turn to consider an alternative analysis of Neg-Raising proposed by Horn. While attractive, Horn's analysis will be shown to have problems of its own, particularly in the area of NPI-licensing.

Chapter 3

R-implicature

In the last chapter, we have given a detailed analysis of NPI-licensing in the context of Neg-Raising predicates. Our chief foil in that chapter was the traditional movement analysis of Neg-Raising. In this chapter we shift gears to give extended attention to Horn's pragmatic theory of Neg-Raising. In §3.3.2, we look in detail at Horn's approach to NPI-licensing, suggesting that it is not built on firm foundations and cannot compete with the syntactic and semantic approaches in terms of predictive power.

Before advancing to NPI-licensing, we must take some time to fully explicate Horn's very different perspective on the nature of Neg-Raising. According to Horn, Neg-Raising derives from a very general Gricean pragmatic principle, which he calls **R**. Deriving the phenomenon from such general principles is certainly desirable. More than this, however, Horn argues that the details of the derivation explain certain gaps in the class of Neg-Raisers. According to Horn neither predicates that are too strong nor too weak may participate in Neg-Raising. He refers to this as the Mid-Scalar Generalization and argues that it follows from general restrictions on the application of **R**.

Into this discussion I bring the case of the negation of definite plurals (§3.4) and question whether they form a natural part of Horn's pragmatic picture. There is an alternative literature on definite plurals, dating back to Fodor 1970, that suggests treating definite plurals in terms of Excluded Middle presuppositions. Krifka 1996 suggests that definite plurals can indeed be brought under the aegis of the **R** principle. I will argue in favor of Fodor's approach - a position I develop in the next chapter, showing there is a significant correlation between obeying the Excluded Middle and having the characteristics of a definite plural.

3.1 Horn on Neg-Raising

According to Horn (1989), Neg-Raising, rather than being an idiosyncratic property of certain predicates, is an instance of"a fundamental grammatical, semantic and pragmatic phenomenon manifested across distinct, but systematically related, classes of predicates in generally and typologically diverse languages" (Horn (1989), p. 309). In this section I sketch and criticize Horn's approach to Neg-Raising. In addition, I suggest that Neg-Raising may belong to a class of semantic/pragmatic phenomena that are not obviously subsumed by the phenomena discussed by Horn.

3.1.1 R-implicature

The fundamental pragmatic phenomenon that comprises Neg-Raising, according to Horn, is R-implicature (Horn (1984), Horn (1993)). In this section, I describe the role that Horn ascribes to R-implicature in the pragmatics of natural language.

(251) The Maxims of Conversation

a. Quality: try to make your contribution one that is true

    1. Do not say what you believe to be false

    2. Do not say that for which you lack evidence

b. Quantity:

    1. Make you contribution as informative as is required (for the current purposes of the exchange)

    2. Do not make you contribution more informative than required

c. Relation: Be relevant

d. Manner: Be Perspicuous

1. Avoid obscurity of expression

2. Avoid ambiguity

3. Be brief (avoid unnecessary prolixity)

4. Be orderly

Grice (1975) proposes a set of rules that support a system of non-logical inference for natural language. The assumption that fellow speakers are (or should be) adhering to these rules allows one to draw inferences that do not logically follow from the content of their statements. The specific rules proposed by Grice are those in (251). Grice's maxims of conversation have been put to work to explain many phenomena, but the most worked out use has been in the explanation of scalar implicature (Horn, 1972b). Scalar implicatures are inferences based on the existence of scales and Quality and the first Quantity maxim (henceforth Q1). In very broad strokes, here is how the computation of scalar implicatures works: When a statement is made, it is made against the background of alternative statements, statements that could have been made instead. Which statements are taken as background is partly determined by the words used to make the statement and the scales to which they belong. For example, when someone utters (252) the alternative statements in (253) are invoked (at least), because *most* belongs to the same scale as *some* and *every*.

(252) Most students attended the meeting.

(253)
1. Some students attended the meeting.
2. All students attended the meeting.

(254) <some, ... , most, every>

When a statement is asserted against a background set of alternatives, the question is raised of why the speaker made the statement s/he did and not some other alternative. For example, given that Q1 enjoins us to make our statements as informative as required we can see immediately why the speaker of (252) chose it and not (253a). If the speaker is following the enjoinder of Quality in stating (252) to not state what s/he knows to be false then s/he must believe that most students attended the meeting. Given that (252) is more informative than (253a), if s/he had stated (253a) instead s/he would not have made the most informative s/he was in a position to make. Similarly, we can understand why s/he did not choose to utter (253b). Knowing that the speaker is enjoined to make the most informative possible given the requirements of the context, we can infer that s/he did not assert (253b) either because the extra information conveyed by (253b) was not required by the context, or else, asserting (253b) would have violated another maxim, namely Quality. That is, we can infer that (if the information would not have been more than required) the speaker either does not believe that (253b) is true or that s/he has no evidence for it. In a context where, in addition, we can safely assume that the speaker is informed about stronger alternatives, we infer that s/he believes that the stronger statement (253b) is false.

The nature of scalar implicature can be summed up in this way. When a statement is made, stronger alternatives are considered and when certain conditions hold we infer that those stronger alternatives are false. So, in the end, the utterer of a statement containing a weak scalar item ends up meaning something stronger than the content of their statement, specifically their statement plus the negation of stronger alternatives.

Horn suggests that there is another way to mean something more than what you say, not based on Quality and Q1. Specifically he proposes that the maxims of Relation and Quantity 2 (henceforth Q2) are set up in fundamental opposition to Q1. For example, you can mean more than you said by adhering to Q2 and not saying more than is required. That is, you can invite the inference to a strong statement by making a weak one in deference to Q2. To bring out the opposition of these principles Horn re-organizes Grice's system thus:

(255) Minding our Q's and R's

The Q Principle	The R Principle
(Hearer Oriented)	(Speaker Oriented)
Make your contribution SUFFICIENT:	Make your contribution NECESSARY:
Say as much as you can	Say no more than you must
(given both Quality and R)	(given Q)
Lower-bounding principle,	Upper bounding principle,
Inducing upper-bounding	inducing lower-bounding
Implicata	implicata
Quantity1 and Manner1,2	Quantity2 and Manner3,4

The fundamental work done by the R-principle is summed up by Horn thus: "The R principle is an upper-bounding law which may be (and systematically is) exploited to generate lower-bounding implicata: a speaker in saying '... P_i ...' implicates '... P_j ...' for some strong P_j stronger than P_i and/or representing a salient subcase of P_i." (Horn 1989, p. 195) So when an R-implicature is drawn, it is because a stronger alternative is invoked, as in the case of scalar (Q-)implicature. Unlike the case of scalar implicature, in the case of R-implicature, the hearer is invited to infer that the stronger case is true, not that it is false. Another difference between R-implicature and (at least scalar) Q-implicatures is that the stronger alternative may not be linked to the asserted statement by form. In the case of scalar implicatures there is a formal link between the original statement and the alternatives: the alternatives are generated by replacing scalar items with scalemates. In the case of R-implicatures, the stronger alternative to a statement P may simply represent a culturally salient/stereotypical subcase of P. This is illustrated in the examples below.

How do Q and R interact? When a statement invokes stronger alternatives, how do you know whether to infer that a stronger alternative is true or false? Horn suggests, following Atlas and Levinson 1981, that R-implicatures generally arise due to social and cultural factors that override the Q-principle. Chief among the social/cultural factors that motivate R-implicature is politeness; when it would not be polite to make a strong statement, a speaker makes a weaker statement the R principle invites the hearer to infer that the speaker, in fact, believes the stronger statement. So, for Horn, euphemism and indirect speech acts are the paradigmatic examples of R-implicature.

In many situations, it is regarded as impolite to issue a direct order to an interlocutor. So if it is necessary to have someone else perform an action, it is preferable to make the demand indirectly. For example, someone might issue the demand to have the salt at the dinner table by asking the question in (256)

(256) Can you pass me the salt?

While the question merely inquires about the ability of the addressee to pass the salt, it is taken via the R-principle to issue a demand to have the salt passed. This is so, according to the R-principle, because a salient subcase of the instances in which someone asks about another's ability to do X, is when they are asking as a precondition to demanding that the person do X.

Similar reasoning applies in the case of euphemisms. Due to cultural taboos it is impolite to refer to certain bodily functions, such as having a bowel movement. For this reason, less informative and thereby less offensive locutions are used for this purpose.

(257) "go to the bathroom" $\approx$ have a bowel movement

In this case, going to the bathroom is literally only the action of moving towards a room having a bath in it. Through the workings of the R-principle, this expression came to be used to refer to a salient/stereotypical subcase of when people go to bathrooms. In fact this meaning has now become lexicalized and refers directly to the action of having a bowel movement, even when it doesn't occur in the context of going to a bathroom.

Horn suggests that this kind of reasoning is pervasive in natural language and that, consequently, the R-principle has corollaries in domains other than euphemism and indirect speech acts. The most relevant for the study of Neg-Raising are its corollaries in the domain of negation:

(258)
1. Contrary negation tends to be maximized in natural language.
2. Subcontrary negation tends to be minimized in natural language.

(Horn (1989), p.330)

The terms contrary and subcontrary derive from Aristotle's study of the logical relations and his famous Square of Opposition. Contrariety is the relation that holds, for example, between a universal statement ($\forall$) and its inner negation ($\forall\neg$).

(259)
1. Every man is mortal
2. Every man is not mortal ($\approx$ No man is mortal)

This relation can be described more generally as the relation that holds between two statements that cannot simultaneously be true (but may be simultaneously false).

(260) Two statements are **contraries** iff they cannot be simultaneously true

So, in addition to the quantified examples in (259), the sentences in (261) count as contraries since it is not possible for a single object to be both (all) black and (all) white, while it is possible for an object to be neither white nor black.

(261)

a. X is black	c. X is tall
b. X is white	d. X is short

Subcontrariety on the other hand is the relation that holds, for example, between an existential statement ($\exists$) and its inner negation ($\exists\neg$).

(262) a. Some man is mortal

b. Some man is not mortal (= Not every man is mortal)

The logical relation between two such statements is characterized by the fact that they cannot be simultaneously false (but may be simultaneously true).

(263) Two statements are **subcontraries** iff they cannot be simultaneously false

Non-quantificational examples of subcontrariety are more difficult to find. Horn suggests this is so precisely because of the principles in (258). Let me now explicate how Horn thinks the principles in (258) affect natural language interpretation.

Suppose that an expression of natural language belongs to a contrary opposition. This expression may be, for example, a quantified statement or an antonymic predicate. Horn assumes that natural language negation semantically denotes logical negation. Logical negation applied to a statement or predicate yields its contradictory, not its contrary.

(264) Two statements are contradictories iff they can neither be simultaneously true nor simultaneously false

The proposition denoted by the contrary[1] is a subcase of the proposition denoted by the contradictory. Recall that the R principle invites the hearer to infer that a salient stronger subcase of a weak statement holds. Given this we can understand the principle (1) as a corollary of the R principle, assuming that the contrary of a statement is a salient/stereotypical subcase of its contradictory.

(265) The R principle is an upper-bounding law which may be (and systematically is) exploited to generate lower-bounding implicata: a speaker in saying '... P_i ...' implicates '... P_j ...' for some strong P_j stronger than P_i and/or representing a salient subcase of P_i.

(260) Contrary negation tends to be maximized in natural language.

How exactly does language maximize contrary negation? One could imagine a conspiracy on many levels of grammar that are directed to accomplish this maximization. Horn suggests, on the contrary, that this maximization is a pragmatic process (although its results can feed grammaticaliztion). According to Horn, in natural language there is a tendency towards polarization, a preference for binary opposition. This tendency shows itself through the evocation, whenever possible, of a disjunction that fills in the gap between a statement and its contrary, excluding the middle.

(266) The government is not good.
$\approx$ The government is bad

(267) The government is good or the government is bad.

Horn notes that "under the right conditions, then, a formally contradictory negation not-p will convey a contrary assertion q (Horn (1989) p.273). The right conditions

[1]Technically, we are not justified in referring to *the* contrary of a statement. The definition in (260) does not guarantee that every statement has a unique contrary. In fact, our definitions imply that any statement that entails the contradictory of a statement A is contraries with A. For example, X is red' is contrary to X is black' just as much as X is white'. Horn, however, identifies white' and black' as *polar* contraries since they are contraries and represent the endpoints of a scale.

will license the assumption of the middle-excluding disjunction and make use of the following basic principle:

(268) In a context licensing the pragmatic assumption p $\vee$ q, to assert not-p is to implicate q.

(Horn (1989), p.273)

As we will see this principle plays an important role in Horn's analysis of the Neg-Raising phenomenon.

To show how fundamental the pragmatic process behind Neg-Raising is, Horn shows how it affects two constructions other than Neg-Raising: affixal negation, sentence negation. In these two cases, Horn argues contrary negation tends to be maximized. This maximization, however, is subject to at least two constraints.

In the case of affixal negation, the maximization of contrary negation is shown in the meaning of negated forms such as *unhappy* and *unwise*, which convey not just the contradictories of *happy* and *wise*, but their contraries *sad* and *foolish*. That is (269a) and (269b) are roughly equivalent to (270a) and (270b), respectively.

(269) a. Bill is unhappy

b. That decision was unwise

(270) i. Bill is sad

ii. That decision was foolish.

Affixation with *un-* is only partially productive in English. Similarly the inference to the contrary in these cases is lexicalized. That is, the reasoning from the assumed Excluded Middle is not defeasible, as one would expect a pragmatic inference to be. Furthermore, the restrictions on the inference to the contrary are restrictions on lexicalization, not on possible meanings for complex words. For example, *un-*affixation is ungrammatical with the negative antonyms of words that do accept *un-*affixation.

(271) unhappy/*unsad, unwise/*unfoolish, unhealthy/*unsick uninteresting/*undull, untrue/*unfalse, unkind/*unrude

To Horn, this pattern reveals the roots of contrary-maximization in politeness. One says *unhappy* to avoid saying *sad*, the R-principle permits filling in the gap between lacking happiness and being sad. Similar motivation does not exist in the other direction, since *happy* is evaluatively positive there is no reason to avoid saying *happy* by using **unsad*.

Another constraint on *un-* affixation comes from the position of a predicate on its scale. According to Horn, only an unmarked positive predicate can be affixed with un-, a positive predicate associated with high values of the scale cannot.

(272) *unecstatic, *ungreat, *undelicious

Horn has no ready explanation for this restriction. He states the generalization about the distribution of *un-* thus:

(273) The stem to which a relatively nonproductive negative affix [such as, un- JG] can attach tends to be an UNMARKED, WEAK POSITIVE scalar value.

(Horn 1989, p. 286)

Horn demonstrates that similar forces are at work in the case of simple sentential negations of scalar predicates. For example, there is a tendency to interpret (274a) and (274b) as (275a) and (275b), respectively.

(274) a. Bill is not happy.

b. That decision was not wise.

(275) i. Bill is unhappy.

ii. That decision was unwise.

The tendency in this case, however, is defeasible - in contrast to the case of affixal negation. This is demonstrated by the following examples.

(276) a. Bill is not happy. In fact, he's unhappy

b. That decision was not wise. In fact, it was unwise.

(277) i. Bill is unhappy. #In fact, he's not happy.

ii. That decision was unwise. # in fact, it was not wise.

Nonetheless, the tendency exists and has been the subject of much investigation, as Horn documents. Chomsky 1970 suggests that the scope of negation may be at issue. Lyons 1977 suggests a pragmatic explanation much like Horn's.

Horn observes that just as affixal negation is unavailable for evaluatively negative predicates, a contrary reading is not forthcoming for the sentential negations of evaluatively negative predicates.

(278) a. Bill is not unhappy/sad.

b. That decision was not unwise/foolish.

(279) i. Bill is happy.

ii. That decision was wise.

We do not in general take someone who asserts (278a) or (278b) to be claiming (279a) or (279b), but something weaker the contradictories of (278a) and (278b). Once again this effect can be traced back to the fact that the maximization of contrary negation is rooted in politeness. Also parallel to the case of affixal negation is the fact that the sentential negations of high scalar value positive predicates do not tend to receive contrary readings.

(280) a. Bill is not ecstatic.

b. This cake is not delicious.

c. The performance was not great.

Horn states the generalization:

(281) A contrary reading is available for a sentential negation just in case the negated predicate is positive and relatively weak on its scale.

Horn proposes that a similar generalization holds about Neg-Raising, and thus we should view Neg-Raising as an instance of the same phenomenon.

(282) A contrary (NR) reading is available for a negated proposition-embedding higher predicate only when the negated predicate is positive and relatively weak (i.e., just above the midpoint on its scale).

(Horn 1989, p.338)

The next section is devoted to the examination of this claim, particularly the claim that the predicate must be relatively weak. Concerning the requirement that the predicate be positive little can be said since few proposition-embedding predicates could be labeled negative. Possibly relevant are cases such as (283).

(283) a. It is not unlikely that Bill left (*until 5).

b. It is not impossible that Bill left (*until 5).

Neg-Raising is not possible in these cases as shown by the unacceptability of the strong NPI until.

3.1.2 Mid-Scalar Generalization

Horn translates the requirement that a predicate be unmarked weak to the requirement that a proposition-embedding predicate be just above the midpoint on its scale. For example is not enough merely to be a positive proposition embedding predicate, such as *possible*.

(284) *possible* is not a Neg-Raiser

a. It's not possible that Bill left (*until 5)

b. It's possible that Bill didn't leave until 5

Horn argues that proposition embedding predicates such as *possible* do not permit Neg-Raising because they fall below the mid-point on the positive scale. But what property distinguishes predicates at or below the midpoint on a scale and those above the mid point? Horn claims it is the property of tolerance, cf. Löbner (1985).

(285) An operator P is tolerant if the conjunction $[P(p) \wedge P(\neg p)]$ is logically consistent

(286) An operator P is intolerant if the conjunction $[P(p) \wedge P(\neg p)]$ is logically inconsistent

A tolerant operator sits at or below the midpoint of a positive scale. Operators above the midpoint are intolerant. The notion of intolerance is an extension of the familiar

Law of Contradiction, which states that p $\wedge$ $\neg$p is logically inconsistent. Obeying the law of Contradiction is what distinguishes contradictory from contrary negation (they both obey the Law of the Excluded Middle).

The proposition-embedding predicate *possible* is tolerant, given that (287a) and (287b) are clearly consistent.

(287) a. It is possible that Bill left

b. It is possible that Bill didn't leave

This contrasts with the inconsistency of (288a) and (288b), sentences containing the Neg-Raising predicate *believe*.

(288) a. Bill believes that Mary is here

b. Bill believes that Mary is not here

One might quibble that (288a) and (288b) are not actually logically inconsistent. This, however, is true also in the case of the paradigmatic contraries in (289).

(289) a. No man is mortal

b. Every man is mortal

It has been argued that these sentences are also consistent; both can be true if there are no men. Strawson suggests that we can reconstruct Aristotle's oppositions if we always assume that the subject class is non-empty. We can appeal to this assumption also in the case of proposition-embedding predicates. Assume that every proposition-embedding predicate is judged relative to a modal base. For *believe* this is the set of the subject's doxastically accessible worlds. To say that this set is non-empty is precisely to say that the subjects beliefs are consistent. This assumption is clearly inconsistent with the propositions expressed by (289a) and (289b).

Horn proposes that the ban against low scalar Neg-Raisers has a parallel in the domain of affixal negation. Affixation of *un-* to deverbal adjective formed from *able* is fully productive. The negation however is not interpreted as internal negation.

(290) The stone is untouchable

a. It is not possible to touch the stone

b. It is possible not to touch the stone

This failure to convert external to internal negation is an instance of Horn's generalization that subcontrary negation is minimized in natural language. So, we have established a lower bound for the availability of Neg-Raising among the positive proposition-embedding predicates on a scale. Horn argues that there is also an upper bound, as in the cases of affixal negation and the negation of scalar predicates. That is Neg-Raising is not available for proposition-embedding predicates that are too high on their scales. Horn has a hypothesis here too as to what property defines what is too high on a scale of proposition-embedding predicates. He proposes that the property is implying the truth of the embedded proposition. This is the property that is sometimes referred to as veridicality.

(291) An operator P is veridical iff P(p) implies that p.

This idea can be illustrated with the predicate *certain*, which does not permit Neg-Raising.

(292) *certain* is not a Neg-Raiser

a. It is not certain that Mary left (*until 5)

b. It is certain that Mary didn't leave until 5

(293) *certain* is intolerant

i. It is certain Mary left

ii. It is certain Mary didn't leave

Horn claims that an assertion of *it is certain that p* virtually guarantees the truth of p, if it does not in fact entail it.

(294) a. It is certain that Mary is here.

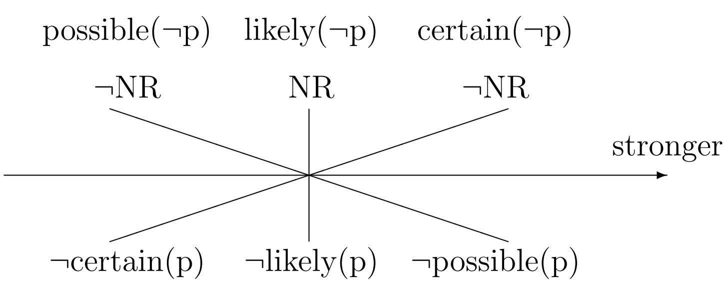
b. Mary is here.

Horn furthermore suggests that the reason that veridical predicates do not submit to Neg-Raising is that breakdowns in communication would result. There is simply too much functional difference between *it is not certain that p* and *it is certain that* *not p* for the former to be able to mean the latter. The latter sentence would after all imply that not p.

Horn goes on to argue that the ban on low scalar Neg-Raisers may follow from a similar communicative principle. He points out that just as *it is certain that p* implies p, *it is not possible that p* implies not-p. Horn concludes, "we must evidently inspect each pair of the form <P(p), ¬P(p)> and determine if an entailment [of p or ¬p JG] is derivable from either member; if so P is scratched from the roll of prospective neg-raisers. (Horn (1989), p. 326).

Another way to look at this ban on Neg-Raising at the scalar extremes is illustrated by the picture below. The external negation of possibility yields a very strong statement at the top of the negative scale. The internal negation on the other hand is very weak, at the bottom of the positive scale. Similarly, converting the external negation of necessity to internal negation takes you from one scalar endpoint to the other.

(295)



Only converting the negation of midscalar predicate from external to internal keeps you in the same quantitative range of the scale. In other words, the statement you get as a result of Neg-Raising doesn't differ in strength from the statement without Neg-Raising.

Horn summarizes his Mid-Scalar Hypothesis about the availability of Neg-Raising for predicates in the following chart:

(296)

be able	believe, think	know, realize
be possible	be likely	be clear, be obvious
	figure	be sure, certain
	seem, appear	be odd, significant
may, might	should, ought	must, have to
can, could	be supposed to	need, be necessary
allow, permit	be desirable	require
	want, suggest	order, demand

Given the generality of the principles invoked by Horn and the fact that the combination of sentential negation and a proposition-embedding predicate is fully productive, one would think that any predicate that meets the specifications Horn charts out would allow Neg-Raising. As Horn well knows, this is not the case. Within the class of Mid-Scalar predicates, which are Neg-Raisers and which are not is arbitrary:

(297)

Neg-Raiser	Non Neg-Raiser
suppose	guess
expect	anticipate
want/wish	desire
hoffen (German)	hope
souhaiter (Fr.)	esprer (Fr.)
xosev think' (Hebrew)	maamin believe' (Hebrew)

This is a serious problem for any general pragmatic theory of Neg-Raising. Horn, however, has a solution. He notes that pragmatic researchers have already shown that not all pragmatic processes are as productive as they ought to be. There are pragmatic phenomena whose derivations are clear (like scalar implicatures from Grice's maxims) but only show themselves in particular constructions. Limiting our sights to the realm of implicatures, there are, to use Grice's terminology, implicatures that are calculable but detachable.

The research to which Horn refers is that of Sadock (1972), Searle (1975) and Mogan (1978) on Indirect Speech acts. Above we briefly sketched how some Indirect

Speech Acts could be viewed as arising as a result of Horn's R-principle. In particular we saw that the question (298a) can be used to issue the command to pass the salt.

(298) a. Can you pass the salt?

b. Pass the salt.

What we did not consider at that point is that not every expression that is synonymous with (298a) can be so used. That is, the interpretation of the question expressed by (298a) as an indirect command is detachable. The basic observation is that (299a) is easily interpreted as an indirect command, (299b) is less easily interpreted this way, and (299c) even less so.

(299) a. Can you pass the salt?

b. Are you able to pass the salt?

c. Do you have the ability to pass the salt?

A similar phenomenon is illustrated in (300), where (300a) is conventionally used to wish luck in a performance whereas the nearly synonymous expressions in (300b) and (300c) can only be taken as expressing literal wishes to break a leg.

(300) a. Break a leg!

b. Fracture a tibia!

c. I hope you break a leg!

To use Horn's terminology, these phenomena show that pragmatic reasoning can become "short-circuited". That is, an implication that is calculated on the basis of general conversational principles can become conventionally associated with a construction that often triggers its calculation.

The proposal then is that this kind of pragmatic conventionalization also occurs in the domain of Neg-Raising. Whether a given combination of negation and a proposition embedding predicates invokes the middle-excluding disjunction that gives rise to Neg-Raising depends on whether the calculation of the implicature has become short-circuited for that combination.

We must imagine then that certain predicates are marked to trigger R-based reasoning. Though Horn does not propose such marking, we can propose to mark certain proposition-embedding predicates [+R].

(301) NRPs are marked with a feature that triggers the R-principle. For example, want[+R] desire[-R]

(302) Inclusion of +R feature on NRP invokes Excluded Middle Implicature.
X wants[+R] P $\Rightarrow$ X wants P or X wants $\neg$P

This gives us the rudiments of how one should represent Neg-Raising in the grammar. In Horn's theory much work is left to the pragmatics, but conventionalization also plays a role in accounting for the arbitrariness of the application of the pragmatic principles. A major question for the grammatical representation, so far left unanswered, is how it accounts for the licensing of strict NPIs. The licensing of an NPI clearly depends on the semantic representations associated with expressions in the environment of the NPI, cf. Ladusaw (1979), and many others.

Horn (1989) attempts to a demonstrate that negation expressed by a pragmatic convention like Neg-Raising is enough to license strict NPIs. This approach to licensing is similar to that employed by Linebarger (1980), Linebarger (1987), in that a negative implicature licenses an occurrence of an NPI. It differs slightly, as Horn shows, since not just any negative implicature will license a strict NPI; the implicature must be short-circuited.

First, as a baseline, recall that the strict NPI *until* is licensed by superordinate negation only when the intervening predicate is a Neg-Raiser:

(303) I don't think/*hope they'll hire you until you shave off your beard

Now notice the contrast between the two superficially similar constructions (304a) and (304b). While both give rise to a kind of negative implicature, only the former licenses *until*.

(304) a. I'll be damned if I'll hire you until you cut your hair.

b. *I'll be surprised if he hires you until you cut your hair.

(305) i. I won't hire you until you cut your hair.

ii. I expect he won't hire you until you cut your hair.

Horn claims that the negative implicature of (304a) is short-circuited and licenses strict NPIs. The negative implicature of (304b), on the other hand, is not conventionalized and thus is not sufficient to license a strict NPI.

Another pair that Horn uses to motivate this analysis is (306a) and (306b). Only the conventionalized rhetorical question in (306a) licenses a strict NPI.

(306) a. Why get married until you absolutely have to?

b. *Why are you getting married until you absolutely have to?

(307) You shouldn't get married until you have to.

According to Horn, only the conventionally rhetorical question in (306a) conventionally implicates the negative proposition (307).

3.2 Summary

In this section, I hope to have laid plain the logic behind Horn's appealing, principled account of Neg-Raising. The basis of the story is the very general R principle which enjoins a speaker to make his/her contribution necessary. Obeying this principle allows a speaker to say something weak but implicate a stronger salient alternative. Horn then proposes that in the case of negation, standard contradictory negation is strengthened by the R principle to the salient subcase of contrary negation. It is observed, however, that this strengthening does not always occur: the maximization of contrary negation is limited to cases in which an unmarked positive predicate is involved. Predicates that are too strong do not have their contradictories converted to contraries. Horn suggests that this may be because there is too much functional difference between the contradictory and contrary negations of a strong scalar predicate.

Horn then transposes these assumptions to the domain of Neg-Raising. Thus, he offers solution both to the problem of what Neg-Raising is (a strengthening implicature) and to the problem of which predicates under go Neg-Raising (only the unmarked positive ones). The solution does not fit exactly: not all positive predicates undergo Neg-Raising. In particular, weak (tolerant) positive predicates do not allow Neg-Raising. In these cases, however, converting an external to an internal negation yields a subcontrary, not a contrary. A subcontrary of a statement is weaker than its contradictory. So, Horn arrives at the Mid-Scalar generalization about which proposition-embedding predicates undergo Neg-Raising. This is still not quite right. Not all midscalar predicates undergo Neg-Raising. There are pairs of predicates that are comparable in their scalar qualities but differ in whether or not they allow Neg-Raising, e.g., *want* (Neg-Raising) and *desire* (not Neg-Raising). Consequently, Horn must allow the pragmatic reasoning we have just outlined to become a conventionalized part of the construction, applying to some predicates and not to others. He argues that this is not such an exotic option; it is already utilized in the analysis of Indirect Speech Acts. Finally, Horn argues that this conventionalization is responsible for the licensing of strict NPIs under negated Neg-Raising predicates.

3.3 Criticisms of Horn's account

Perhaps the most incisive criticism of Horn's approach is his own: "Given our current state of knowledge, it must be conceded that ascribing some phenomenon to the presence of an SCI [short-circuited implicature – JG] may amount more to labeling than explaining the phenomenon." (Horn 1989 p.350) Unfortunately, I cannot report that our knowledge of the grammar of short-circuited implicature is any greater at this later date. The questions of exactly how these pragmatic conventions are to be represented in the grammar and how that representation can affect NPI licensing have yet to be resolved.

Nevertheless I will attempt to state some concrete problems for this analysis

3.3.1 Generality

One question about R-based reasoning is how generally it applies, and if it does not apply in some particular case, why not. As I pointed out above, Horn looks at the specific domain of negation and asks what the salient subcase would be for strengthening. The answer he gives is the contrary. Presumably, the R principle ought to operate similarly in constructions that do not directly involve negation. A natural place to seek parallel behavior would be in constructions that are naturally paraphraseable in terms of negation. Do the effects of the R-principle show up here?

The case that I have in mind is conditionals. Consider what the R-principle predicts about how we should strengthen a conditional whose antecedent contains negation.

(308) If Bill thinks Mary is here he locks the door

As a guide we can use the paraphrase in terms of negation in (309).

(309) Either Bill doesn't think Mary is here or he locks the door

The key question is what does Bill do when he isn't sure where Mary is - when he can't rule out that she is here and he can't rule out that she's not. I think that (308) does not entail that he locks the door when he's not sure. The case with overt negation (309) on the other hand does seem to imply that he locks the door when he's unsure.

Paraphraseability by negation is not crucial here, it merely helps to focus our attention on what a plausible strengthening of a conditional under the R principle would be. Another perhaps more intuitive way to think of it is this: in the context of negation, a universal quantifier over accessible worlds is interpreted as an existential quantifier, yielding the stronger contrary reading. Interpreting a universal as an existential doesn't yield stronger readings only in the scope of negation; it leads to a stronger reading in any downward entailing environment. Conditional antecedents are downward entailing environments (cf. von Fintel 1999).

(310) $[\![\text{think}_{DE}]\!]$(w)(p)(x) = 1 iff $\exists$w' [wR$_x$w' & p(w') = 1]

(wR$_x$w' iff w' is compatible with what x believes in w)

(311) if Bill thinks Mary is here he locks the door =
if Bill thinks$_{DE}$ Mary is here he locks the door

(312) (308) is true in w iff every world w' maximally similar to w in which every world w" compatible with what Bill believes in w' is s.t. Mary is here in w" is s.t. Bill locks the door in w'

(313) (311) is true in w iff every world w' maximally similar to w in which some world w" compatible with what Bill believes in w' is s.t. Mary is here in w" is s.t. Bill locks the door in w'

The fact of the matter is that interpretation (313) is not available for sentence (308). Horn's theory makes us wonder why. In one sense there is a ready answer. No pragmatic conventionalization ever took place in the case of the conditional. Still if the general line of reasoning that Horn sketches for Neg-Raising doesn't surface in any other construction we might question the value of the general principles on which it is based.

Another issue arises here. That is how Horn's theory is to be cashed out in the grammar. It could be that, when we spell out how pragmatic conventionalization is represented, we will see that the conditional case is not a candidate for conventionalization. In the last section, I suggested implementing Horn's theory by marking certain proposition-embedding predicates for whether or not they trigger reasoning based on the R-principle. This marking could be considered analogous to marking an item for belonging to a scale that triggers Q-based reasoning. This seems to predict that they should give rise to the kind of reasoning just sketched for conditionals. It could be however that Horn would not endorse this kind of marking.

The line of research on which he bases the notion of short-circuiting draws an analogy between pragmatic conventionalization and idioms. Sadock (1972) refers to sentences such as *Can you pass the salt?* as speech act idioms. Perhaps we should take this more seriously. One frequently proposed condition on idiom interpretation is that the parts of an idiom must form a constituent at the level of interpretation (Chomsky (1995), Larson (1988)). It is sometimes proposed that VP adverbs interfere with

Neg-Raising (Heycock, 2003). One way to implement the interpretation of idioms is to assume that there are special interpretive rules for constituents dominating the pieces of an idiom:

(314) $[\![$ [kick [the bucket]] $]\!] = \lambda$x. x died

This sort of rule sits alongside other more general rules such as Functional Application. Of course, unless some principle of grammar chooses which rule applies (the idiomatic rule, or the more general rule) the definition of the interpretation function $[\![\cdot]\!]$ falls apart. That is, a function is no longer defined. One candidate for regulating rule application is the familiar Elsewhere Condition (Kiparsky 1982), which legislates in favor of the rule with the more specific conditions for application. This is not appopriate here, since we want to be able to speak of kicking buckets without always speaking of death.

Alternatively one might choose to abandon the assumption that the interpretation function is a function and simply allow the grammar to generate a set of meanings for expressions. Assuming that such a viewpoint could be coherently developed, the rule for Neg-Raising might look like the following:

(315) Special Interpretive Rule for Neg-Raising
$[\![$ [not [NRP p]] $]\!] = \lambda$x. $[\![$NRP$]\!]$ ($[\![$not$]\!]$($[\![$p$]\!]$))(x) = 1

An approach in terms of a constituent-based meaning convention faces difficulties similar to those faced by the syntactic theory of Neg-Raising. It is not only negation that triggers Neg-Raising but negative quantifiers (*nobody*) and adverbs (*never*) as well. The idiom theory would need to either state separate rules for each Neg-Raising trigger or decompose all other triggers into some element and a negation that could form a constituent with the VP headed by the Neg-Raising predicate. If we choose the former option we might expect to find predicates that Neg-Raise with sentential negation but not with negative quantifiers footnote on never-raising. If we choose the latter option we face the difficulties I have outlined for decomposition in §2.1.3.

3.3.2 NPI Licensing

As we have already suggested, any successful theory of Neg-Raising must explain their facilitation of strict NPI-licensing by a superordinate negation. Horn suggests that his theory passes this test. As proof, he offers cases where a conventionalized negative implicature (NI) licenses strict NPIs. His two examples are repeated below:

(316) I'll be damned if I'll hire you until you shave off your beard
Conventional NI: I won't hire you until you shave off your beard

(317) Why get married until you absolutely have to?
Conventional NI: You shouldn't get married until you absolutely have to

Horn further points out that a non-conventionalized negative implicature does not suffice to license strict NPIs

(318) I'll be surprised if he hires you
NI: I expect he won't hire you

(319) *I'll be surprised if he hires you until you shave off your beard

In order for this explanation to carry any force it is necessary to give independent evidence that the implicature is conventionalized in the case of *be damned* but not in the case of *be surprised*.

I do not know what evidence there is in favor this idea. Presumably the pragmatic reasoning that has been short-circuited in the case of *be damned* goes like this: I'll be damned if p. Being damned is an obviously undesirable state, perhaps the least desirable. So one can safely assume that I will do anything to prevent myself from being damned. If p happening is sufficient to damn me then I will do anything to stop p from coming about.

I think it's fair to say that we do not go through this chain of reasoning anymore. One piece of evidence in favor of short-circuiting is that the inference is detachable.

(320) a. I'll be committed to hellfire if I(ll) hire you.
b. I'll suffer eternal torture If I(ll) hire you.

While these might be used very indirectly to say that I won't hire you, the inference is far from automatic. And indeed, strict NPIs are not licensed in the paraphrases.

(321) a. *I'll be committed to hellfire if I(ll) hire you until you cut your hair

b. *I'll suffer eternal torture If I(ll) hire you until you cut your hair

So far so good for Horn's analysis. There is one possible problem as I see it. When looking at the NPI-licesing status of a construction we should not neglect its grammatical properties. It may very well be that the kind of reasoning I've sketched above is part of the history of this construction, but we need to ask what other, grammatical properties the constructions has. As it turns out, not only does the implicature disappear in paraphrases, it disappears in a very close structural analogue as well.

(322) a. I'll be damned if I'll hire you

b. #If I'll hire you, I'll be damned

c. If I hire you I'll be damned

To my ear, the fronted if-clause with the future tense sounds ungrammatical. When we replace the future with the present the construction improves but does not seem synonymous with (322a). Neither, more importantly, does it license NPIs.

(323) *If I hire you until you cut your hair, I'll be damned

One might object that fronting the if-clause disrupts the idiomatic interpretation. It is well known, however, that transformations don't affect the interpretation of an idiom. Chomsky 1993 studies examples in which wh-movement does not interfere with idiom interpretation.

(324) Bill took a picture = Bill created a picture by means of a camera

(325) How many pictures do you think Bill took?
can mean "How many pictures do you think Bill created with a camera?"

Larson posits a passive like transformation within VP that postposes an indirect object into a *to*-phrase. This transformation can break up the constituency of an idiom but does not prevent idiom interpretation:

(326) Bill [Phil [sent the showers]] ⇒
Bill [[sent [Phil t t]] to the showers]

(327) Bill sent Phil to the showers
"Bill took Phil out of the game"

Furthermore, this contrasts with the behavior of the other conditionals we've discussed. In these cases, the position of the *if*-clause appears to have no effect on meaning:

(328) a. I'll be surprised if he hires you

b. If he hires you I'll be surprised

(329) i. I'll suffer eternal torture if I hire you

ii. If I hire you, I'll suffer eternal torture

This property of the *be damned* construction suggests to me that it has progressed beyond carrying a pragmatic convention and in fact carries this implication as part of its conventional meaning. That is, I believe this construction is simply downward entailing.

(330) "I'll be damned if I'll run to the corner" entails
"I'll be damned if I'll run to the corner quickly"

In fact, the construction may even be Anti-Additive

(331) "I'll be damned if I'll hire Sue" and "I'll be damned if I'll hire Fred" entails
"I'll be damned If I'll hire Sue or I'll hire Fred"

A defender of short-circuiting might suggest that the position of the if-clause is crucial to the pragmatic convention, perhaps due to locality constraints on idiom interpretation.

(332) *be damned* selects an if-clause as a complement

(333) $[\![$ is damned$]\!]$(p)(x) = 1 iff x wants not-p
(x causes not p)
(x tries to cause not p)

Concerning Horn's second example, I have little to say. The problem of how NPIs are licensed in questions is still unsolved. Consequently, it is difficult to say how the special properties of this construction might contribute to the licensing of strict NPIs.

(334) Why get married until you absolutely have to?

(335) *Why would you get married until you absolutely have to?

3.4 A Separate Generalization

Horn analogizes Neg-Raising to other R-based effects such as Indirect Speech acts and euphemism. In this section I suggest Neg-Raising belongs to a different generalization. There are construction other than those discussed by Horn 1989 in which contrary negation is maximized. It is unclear whether these construction fit easily into Horn's class of R-based construction. They have been largely discussed in a literature separate from Horn's discussion of the R-principle (though not exclusively, cf. Krifka 1996). The paradigmatic example from this literature is definite plural noun phrases.

3.4.1 Definite Plural Noun Phrases

Fodor 1970

Janet Dean Fodor (1970) attempted to argue that definite noun phrases, both singular and plural, denoted universal quantifiers. In the course of this attempt she observed several ways in which definite plural noun phrases differed from garden variety universal quantifiers. One of these differences, relevant for our purposes, is the way definite plural noun phrases interact with negation. In particular, she observes that while (338) is odd, (339) is not.

(336) I saw the boys

(337) I saw all the boys

(338) I didn't see the boys but I did see some of them.

(339) I didn't see all the boys but I did see some of them.

She describes the difference by saying that "*all* can be used to contrast with *some*, but *the* cannot." (Fodor (1970), p.160) However, as she shows the difference between definite plurals and universals extends beyond the case of sentential negation. She observes that the sentence (340) is not easily judged false when some of the boys we met are orphans and other are not.

(340) The boys we met are orphans.

She draws an analogy between the reluctance to judge (340) false in this case, with Strawson's reluctance to judge (341) false in a scenario in which there is no King of France.

(341) The King of France is bald.

Strawson analyzes as (341) as carrying a presupposition that there is a King of France. When a presupposition of a sentence is false, the sentence is neither true nor false. This suggests to Fodor that a presupposition of definite plurals is at issue. This is supported by Fodor's observation that a negative answer to the question (342) implies (343); note that (344) is an odd answer to (342).

(342) Are the boys we met orphans? No.

(343) None of the boys we met are orphans.

(344) #No, some of them are.

Fodor (p.162): "It looks as though a simple definite noun phrase in the plural not only does not contrast with *some*, but does not even admit the possibility that the sentence might be true of some but not all things of the kind described." Consequently, Fodor ascribes to constructions involving definite plurals the presupposition that the predicate is true of all the individuals referred to by the definite plural or none of them (Fodor's all-or-none presupposition).

This presupposition explains another contrast noted by Fodor. Embedded under a desire predicate, a definite plural is odd if it is known that the embedded predicate already holds of some individuals referred to by the definite. For example, if it is known that John has some of the pictures, (345) is odd but (346) is not.

(345) John wants to have the pictures

(346) John wants to have all the pictures

The reason for the contrast is clear under Fodor's hypothesis. The sentence (347) presupposes that John has all of the pictures or none of them. This presupposition is inherited by the sentence (345) as a presupposition about John's beliefs (Karttunen 1973, Heim 1992).

(347) John has the pictures
Presupposes: John has all the pictures or John has none of the pictures

(348) John believes he has all the pictures or he has none of the pictures

Assuming that if John has some but not all of the pictures, then he believes he has some but not all the pictures, we derive that it is odd to use the definite plural in (345).

Fodor does not give a fully fleshed out analysis of definite plurals noun phrases. The groundwork for the representation of presuppositions and plurals had yet be laid. Once these pieces were in place, however, Fodor's thread was taken up again most notably in the work of Löbner (1985, 1987a, 1987b, 2000).

Löbner

The primary innovation in the analysis of definite plurals noun phrases was to provide a unified analysis of collective and distributive readings. In some cases (the distributive) when a predicate is applied to a definite plural the truth of the sentence depends directly on whether or not the predicate applies to the individuals referred to by the definite, in other cases (the collective) it does not but instead depends on some property the individuals have as a group. The solution to providing a common representation for definite plurals in both these readings is to assign them groups as a denotation (Link, 1983). There are several possible ways of representing groups. I will use sets, e.g.,(Schwarzschild, 1991).

(349) $[\![$ the boys $]\!] = \{$x: x is a boy$\}$

A predicate that contains groups as part of its meaning can give rise to a collective reading. An example of such a predicate is *heavy*, which can contain a group without containing any of its parts.

(350) $[\![$ heavy $]\!]$({the lamp, the box}) = 1 $[\![$ heavy $]\!]$(the lamp) = 0, $[\![$ heavy $]\!]$(the box) = 0

Distributive readings arise from the application of a distributive operator to a predicate (Link, 1983). The operator assembles groups out of the individuals in the predicate and adds these groups to the predicate.

(351) Dist(P)(x) = 1 iff P(x) = 1 or every y s.t. y $\in$ x P(y) = 1

Thus, a predicate P that is not itself true of a group may become true of that group in virtue of being true of all its parts. For example, consider the predicate *blond*, which does not apply directly to groups.

(352)
1. $[\![$ blond $]\!]$({Bill, Mary}) is undefined
2. $[\![$ blond $]\!]$(Bill) = 1; $[\![$ blond $]\!]$(Mary) = 1
3. Dist($[\![$ blond $]\!]$)(Bill, Mary) = 1

Against these assumption about plurality, Löbner presents a theory of Fodor's all-or-none presuppositions. According to Löbner, the all-or-none presuppositions noted by Fodor are but an instance of a presupposition that holds of all predications. He refers to this as the presupposition of indivisibility (the parts of the arguments can't be divided with respect to the predicate).

(353) Presupposition of Indivisibility
Whenever a predicate is applied to one of its arguments, it is true or false of the argument as a whole.

While this vaguely stated principle is intended to apply to all predications, its most dramatic intended effect is on a class of predications that Löbner refers to as summative. Summative predications are those in which the truth of a predicate applied to a complex argument depends on the truth of the predicate applied to its parts.

(354) Definition

A predication is summative with respect to a certain argument a iff:

It is true/false of a iff it is true/false of all parts of an admissible partition into proper parts of a.

This definition is meant to include distributive plural predications such as (355) but also predications such as (356). Löbner argues that in the latter case the inference that the entire flag is black follows from the fact that the predicate can apply to parts of its argument.

(355) The girls are blond.

(implies all the girls are blond)

(356) The flag is black.

(implies the entire flag is black)

Löbner guarantees that summative predications carry the presupposition of indivisibility by defining the operation of predicate summation as in (357).

(357) Definition

For any predication p with domain D(p)

Σp is a predication whose domain consists of all those groups of elements of D(p) for which p yields a uniform truth value (i.e., all homogeneous groups within the original domain). For any x in D(Σp), Σp(x) is true/false iff p(y) is true false for each y that belongs to x.

Löbner's operator Σ corresponds to the more common Distributive operator. So, in sum, Löbner implements Fodor's all-or-none hypothesis by adding a presupposition of indivisibility to the distributive operator involved in plural predications.

(358) $[\![\Sigma]\!] = \lambda \mathrm{P}.\ \lambda \mathrm{X}$: $\forall \mathrm{x}{\in}\mathrm{X}\ \mathrm{P(x)} = 1$ or $\forall \mathrm{x}{\in}\mathrm{X}\ \mathrm{P(x)} = 0$. $\forall \mathrm{x}{\in}\mathrm{X}\ \mathrm{P(x)} = 1$

Krifka 1996

Krifka proposes an alternative analysis of the interaction of definite plurals and negation one that he explicitly links to Horn's R-principle.

Total vs. Partial Predicates Krifka's point of departure is the distinction between total and partial predicates. These predicates show a marked difference in their interpretations with definite plurals noun phrases. Krifka offers (359) and (360) as examples.

(359) I returned to the house because I thought I left the windows open.

(360) But when I came back I found that the windows were closed.

The sentence (359) seems to imply merely that I thought I left some of the windows open. The sentence (360) on the other hand implies that all the windows were closed. This difference has been analyzed (Yoon, 1996) as a lexical distinction between *open* and *closed*, as well as many other pairs *dirty/clean*, *sick/healthy*.

(361) $[\![\text{ open }]\!] = \lambda\text{x}.\exists\text{y}[\text{ y}<\text{x and open(y) }]$

(362) $[\![\text{ closed }]\!] = \lambda\text{x}.\forall\text{y}[\text{ y}<\text{x and closed(y) }]$

Krifka does not endorse the lexical theory of total vs. partial interpretation. Instead, he thinks that whether a definite is interpreted existentially or universally is pragmatically conditioned. As evidence he offers the following scenario:

(363) The local bank has a safe that is accessible only through a hallway with three doors, all of which must be open to reach the safe:

a. I could reach the safe because the doors were open

b. I could not reach the doors because the doors were closed.

Here the preferred interpretations have switched. *Open* which favored an existential reading in (359) now favors the universal interpretation in (363a): I got in because all the doors were open. *Closed* which favors a universal reading in (360) now favors the existential interpretation in (363b): I couldn't get in because at least one of the doors was closed.

Plural Predication, Generally For Krifka, then, the quantificational force of a predication involving a definite plural is a matter of pragmatics. Certain predicates might favor one quantificational force over another but these are only tendencies, preferences. So in any given sentence, such as Krika's (364), either interpretation, existential or universal, is possible.

(364) The windows are made of security glass.

a. $\forall$ x[x $<$ the windows $\rightarrow$ x is made of security glass]

b. $\exists$x[x$<$the windows and x is made of security glass]

In the simple sentence (364), the universal interpretation (364a) is preferred. However, when negation is brought into the picture, the preference changes.

(365) The windows are not made of security glass.

a. $\neg\exists$x[x$<$the windows and x is made of security glass]

b. $\neg\forall$x[x$<$the windows $\rightarrow$ x is made of security glass]

Under negation, the preference is for the existential reading (365a). Krifka suspects that some general principle is at work here. In each case (364) and (365), the preferred interpretation is the logically stronger one. Thus Krifka proposes the following principles:

(366) If a predicate P applies to a sum individual x, grammar does not fix whether the predication is universal (...) or rather existential, except if there is explicit information that enforces one or the other interpretation.

(367) If grammar allows for a stronger or a weaker interpretation of a structure, choose the one that results in the stronger interpretation of the sentence, if consistent with general background assumptions!

Among the explicit information that forces one interpretation over another are floating quantifiers and the lexical preferences of total and partial predicates. Concerning the underspecification of quantificational force, Krifka says,

> The interpretation rule is in some sense a null hypothesis. Grammar has to specify the truthconditions for P(x) if x is an atomic individual. Furthermore, it is natural to assume that the truth of P(y), y being a sum individual, will somehow depend on whether P applies to the parts of y.

> Now, if nothing indicates any particular proportion to which P should apply to the parts of y, then the two natural extreme options are the universal interpretation and the existential interpretation.

According to Krifka, then, the parts of the structure in plural predication do not completely determine its interpretation. Instead, the interpretation is underspecified and the grammar steps in to make its best guess, guided on the one hand by the end points of the possibilities and on the other by the injunction to choose the strongest interpretation.

Krifka goes on to explicitly suggest a connection between the principle (367) and Horn's R-based implicatures. The parallel can be brought out more clearly with a shift in perspective on plural predication. Krifka assumes that the principle (367) determines the interpretation of a structure. Horn assumes, instead, that structures receive a determinate interpretation and the R-principle is used to calculate an implicature based on the assertion of the content of that interpretation. Can plural predication could be viewed in a similar light? All that is required, perhaps, is to pick a determinate interpretation for plural predication, universal or existential. Let's try.

Suppose we choose the universal interpretation. This automatically accounts for the universal interpretation of (364). When an existential interpretation is required under negation, we merely need to assume with Horn that contrary negation is maximized. Then the inference will be made from what is said, the negation of a universal, to the salient stronger subcase, the negation of an existential. Krifka might disagree with this choice, however, on the grounds that it does not allow for an existential interpretation with partial predicates.

If this is a valid objection, then it seems we must choose the existential interpretation. Now the universal interpretation of (364) must be seen as an R-based inference. Now we make the leap from an existential statement to the stronger salient subcase of the universal interpretation. This choice, however, would put Krifka in an awkward position since he argues that strengthening due to principle (367) must be cancelable.[2]

(368) The windows are not made of security glass.
Sure those two are, but this one is not.

If plural predication is basically existential, then the implication that none of the windows are made of security glass is part of the asserted content of (365), and thus cannot be cancelled.

In sum, neither option seems appropriate in this case. Thus, Krifka's account of plural predication cannot be directly unified with Horn's implicature based account of the maximization of contrary negation.

Before drawing this concusion, however, it seems to me that we must question the set up for the discussion as established by Krifka. In particular, we should question the relevance of the analysis of total and partial predicates to the general discussion of plural predication.

In addition to their apparent selection of particular quantificational force with plurals (total → universal, partial → existential), total and partial predicates show differences in their behavior when predicated of non-plural arguments. These differences are observed by Cruse 1986, who refers to pairs of a total and a partial predicate, like *clean* and *dirty*, as C-complementaries. Cruse notes that the total predicate attributes the complete lack of some property to its argument, whereas the corresponding partial predicates attributes the presence of that property to its argument. So, for example, x is clean says that x is free of dirtiness; x is dirty says that x has some dirtiness.

This distinction in meaning has a correlate in the admissibility of certain modifiers. The most straightforward is *almost*. While *almost* is completely acceptable with total predicates it is degraded with partial predicates:

(369) This glass is (dirty but) almost clean.

(370) #This glass is (clean but) almost dirty.

[2] This example is not Krifka's, he makes this argument for a parallel case involving donkey anaphora.

This restriction derives directly from the semantics Cruse attributes to C-complementaries. If a glass is dirty it has some degree of dirtiness, if that degree is particularly low, then it is almost clean it almost has no dirtiness. If a glass is clean, it has no dirtiness, so there is no way for it to be approaching dirtiness with out having some dirt, i.e., being dirty. The contrast in acceptability with *almost* can be compared with *almost*'s preferences among quantifiers.

(371) I have almost no food in my refrigerator.

(372) #I have almost some food in my refrigerator.

Given these properties of total and partial predicates in singular predications, it is puzzling that Krifka chooses to relate the properties they display with plurals to distributive predication. That is it is unclear to me whether the truthconditions of the examples Krifka gives, repeated here, truly depend on the application of the predicate to the atomic members of the plurality.

(373) I returned to the house because I thought I left the windows open.

(374) But when I came back I found that the windows were closed.

To be clear, distribution is a strategy for applying a predicate to a plurality based on the application of that predicate to the members of the plurality. It is standardly assumed (Link (1983), Schwarzschild (1991), a.o.) that distribution is (quasi-)universal. Krifka suggests that in some cases, such as partial predicates, the strategy employed is existential. What I would like to suggest is that Krifka's examples have nothing to do with distribution but are, in fact, examples of collective predication.

In fact, this is the most attractive analysis of Krifka's examples. It would be surprising if the facts about the distribution of *almost* in singular predication and the existential force of partial predicates applied to plurals had nothing to do with each other, as they appear to under Krifka's hypothesis. I propose that (375) says of the sum of windows that they have some degree of openness. This implies that at least one window is open. By contrast, (376) says that the sum of the windows has no degree of openness, implying that all the windows are closed.

(375) The windows are open.

(376) The windows are closed.

This hypothesis may also account for Krifka's scenario in which preferred quantificational forces are swapped, repeated below.

(377) The local bank has a safe that is accessible only through a hallway with three doors, all of which must be open to reach the safe:

  a. I could reach the safe because the doors were open

  b. I could not reach the doors because the doors were closed.

Assume that a sum of windows or doors is viewed as a single aperture. In the case of a house the sum of windows divides the space between inside and outside. If any single window is open then the spaces are no longer divided. However, in a scenario where the doors or windows are arranged serially, there is a degree of openness only if a path exists from the outside to the inside. This will only be the case if all the windows/doors are open. Thus, it is sufficient for one door to be closed for there to be no degree of openness.

This analysis explains Krifka's observations about differences in apparent quantificational force in plural predication while capturing the connection between these differences and asymmetries in singular predication involving the same predicates. A partial predicate attributes some degree of a property to an atomic entity and yields existential quantification in plural predications. A total predicate attributes complete absence of a property to an atomic entity and universal quantification in plural predications.

If the connection Krifka draws between the interpretation partial predicates and plural predication more generally is broken, then our objection to a basically universal semantics for the distributor disappears. We can assume then that (378) basically asserts that not every window is made of security glass and the inference that no window is made of security glass derives from application of the R-principle

(378) The windows are made of security glass.

The question that now faces this approach is why is universal distributive predication different in this regard from garden variety universal quantification. That is, why don't we take (379) to implicate that no window is made of security glass.

(379) Not every window is made of security glass.

On the contrary, (379) is generally taken to implicate that some window is made of security glass. That is, why does R win over Q in the case of distributive plural predication but Q wins over R in the case of universal quantification. Horn explicitly addresses the division of labor between the antinomic Q and R principles. He suggests that it is the conventions that underpin their use that decide which wins out. For example, use of the R principles is often rooted in cultural conventions of politeness, whereas the use of Q is more directly driven by the principles of rational exchange of information. He argues that politeness wins out. In this case, however, there appears to be no difference in the politeness status of these constructions . That is, why would it be more polite to use a negated definite plural than a negated universal quantifier?

In addition to suggesting that politeness may trump the Q-principle, Horn proposes another principle for legislating between the application of the principles Q and R, which he calls the Division of Pragmatic Labor. The Division of Pragmatic Labor applies to equivalent expressions, deciding that one expression will give rise to an R-based implicature and that the other will give rise to a Q-based implicature. According to this principle, between two equivalent expressions the briefer and/or more lexicalized expression gives rise to an R-based implicature. That is, the briefer and/or more lexicalized expression is associated with the stereotypical/salient subcase of the content associated with the two coextensive expressions. The less brief/lexicalized expression, on the other hand, gives rise to a Q-based implicature. So, the less brief/lexicalized expression is associated with the atypical/(antonym of salient) subcase of the content associated with two expressions. Marked form, marked message; unmarked form unmarked message.

The essential idea is that the markedness of an expression correlates with the markedness of the message conveyed (Horn, 1984). Horn offers the following examples as evidence of such a correlation.

(380)
1. Can you pass the salt?
2. Do you have the ability to pass the salt?

(381)
1. My brother went to jail
2. My brother went to the jail

(382)
1. Amanda killed the sheriff
2. Amanda caused the sheriff to die

In each of these cases, Horn argues that the briefer/more lexicalized (a) examples is associated with a stereotypical subcase of the content the (a) and (b) examples share. For example, in (a) the bare noun *jail* is used and the sentence implicates that the speakers brother was incarcerated. On the other hand, the less brief (b), which uses the definite description *the jail*, implicates that the speaker's brother merely paid the jail a visit. That is, the sentence implicates that the stereotypical case of being incarcerated does not hold.

A defender of Horn's theory could then attempt to argue that universal quantifiers and definite plurals differ in their length/lexicalization.

(383)
1. The students are blond.
2. Every student is blond.

On the basis of a syllable count, (383a) is shorter than (383b). Lexicalization is not at issue. (383a)'s edge in length, however, only applies with (383b), not with (384).

(384) All students are blond.

It may be, however, that (383a) and (384) are not synonymous. (384), in fact, is most easily interpreted as a generic statement. Given this, perhaps, (384) must be compared for length with the bare plural statement (385)

(385) Students are blond.

It is worth noting that the negations of (384) and (385) relate to each other in the same way as the negations of (383b) and (383a). The negation of *all* is interpreted as the negation of universal, the negation of the bare plural as a universal scoping over negation.

(386) 1. Not all students are blond. ($\neg$>A)

2. Students are not blond. (A>$\neg$)

Let us assume, then, that (384) is irrelevant.

(387) a. Not every student is blond

b. The students are not blond

This leaves us to say, perhaps, that the implicit quantification involved in definite plural predication is briefer than lexical quantification.

I think the evidence in favor of considering distributive definite plurals less marked (briefer/ more lexicalized) than universal quantifiers is equivocal and inconclusive. Without a solid reason for considering distributive definite plurals to be less marked than universal quantifiers , the Division of Pragmatic Labor does not apply. This leaves us resorting to lexically marking certain construction for application of the R-principle and others for application of the Q-principle. This is an unattractive option.

3.5 Summary

In this chapter, we have reviewed Horn's approach to Neg-Raising, highlighting its problems particularly in the area of NPI-licensing. Furthermore, we introduced the case of definite plurals and discussed the pros and cons of attempting to subsume definite plurals under Horn's theory, as suggested by Krifka 1996. An alternative theory in terms of Excluded Middle presuppositions was introduced. This theory of definite plurals is elaborated on in the next chapter and extended to other constructions that exhibit properties of definite plurals.

Chapter 4

Definite Plurality and the Excluded Middle

4.1 An Alternative Generalization

In this section I would like to suggest an alternative generalization to that offered by Horn. The generalization I argue for is that Excluded Middle presuppositions are associated with constructions that have the semantic structure of definite plurality. I will argue for this generalization by giving a sampling of constructions that have been argued to carry Excluded Middle presuppositions and then show that in each case an analysis of the construction in terms of definite plurality is plausible (and in most cases has already been proposed for independent reasons).

4.1.1 Bare Conditionals

Stalnaker has argued at length (see e.g. Stalnaker (1981)) that conditionals obey what he calls the law of Conditional Excluded Middle (CEM). The CEM is stated in (388)

(388) if P then Q or if P then $\neg$Q

Stalnker's semantics for conditionals guarantees that this statement holds. According to Stalnaker's semantics for conditionals the antecedent picks out a unique world.

Thus, given that in any particular world either Q or ¬Q holds, it follows that either if P then Q or if P then ¬Q holds. So, for example, in the case of indicative conditionals, the antecedent P picks out the world closest to the world of the context in which P holds. If Q holds in that world, the conditional is true; if it does not, the conditional is false. It is a presupposition of Stalnaker's semantics that there is a unique closest world in which the antecedent holds. If there is no unique closest world the conditional is undefined. As evidence for this presupposition, Stalnaker offers the counterfactual example in (389). The point of this example is that its antecedent does not pick out a unique closest world. Most of us have no reason to believe that it was any more likely for Verdi to be French than it was for Bizet to be Italian.

(389) #If Bizet and Verdi had been countrymen, Verdi would have been French.

It is worth noting at this point that not all conditionals with the same antecedent as (389) are infelicitous. In particular, the conditional (390) is fully acceptable.

(390) If Bizet and Verdi had been country men,
then Verdi would have been French or Bizet Italian

Stalnaker is fully aware of this wrinkle and proposes a solution. He suggests that van Fraasen's (1969) notion of supervaluation is relevant here. According to the supervaluational strategy of interpretation, if the denotation of a term is indeterminate, then the sentence containing it is (super-)true if it is true for all possible values of the term and (super-)false if it is false for all possible values of the terms. If it is true for some values and false for others, the sentence is undefined.

The conditional (390) is, therefore, (super-)true. There are two classes of candidates for world closest to the actual world in which Bizet and Verdi are countrymen the worlds in which both are French and the worlds in which both are Italian (worlds in which both are, say, German are too far away). In all the worlds in both classes, it is true that either Verdi is French or Bizet is Italian. Thus, the conditional is (super-)true.

The conditional (389) still comes out undefined, as desired. Since the candidates for closest world in which Verdi and Bizet are countrymen include worlds in which

Verdi is French and worlds in which he is Italian, the statement that he is French is true in the former and false in the latter. So, the conditional is undefined.

Another natural way of viewing the semantics of conditionals is to abandon Stalnaker's assumption that the antecedent picks out a unique world and assume the the antecedent picks out the set of closest worlds. [This requires us to make the Limit Assumption (See Lewis (1973) argument against)] Viewed from this perspective, we can see the truthvalue gap that Stalnaker analyzes in terms of supervaluation as a homogeneity presupposition associated with universal quantification over the set of closest worlds.

This is the perspective taken by von Fintel (1997), who also argues that bare conditionals carry Excluded Middle presuppositions. One argument von Fintel gives for this position is that bare conditionals pass Fodor's test for carrying an 'all-or-none' presupposition. For example, consider what we infer from negative answers to the questions in (391).

(391) a. Will this match light if I strike it? No

b. Would Bill have passed the test if he had studied? No

Take (391b). A positive answer to this question implies that in every closest world to our own in which Bill studied (instead of what he actually did) he passed the exam. It seems then that a negative answer should at most imply that if he had studied he might not have passed, i.e., there's some closest world to ours in which Bill studies in which he does not pass. What we naturally infer from the negative answer however is that he <u>would not</u> have passed even if he had studied, i.e., that no closest Bill-studying world is one in which he passes. Similarly for (391a).

4.1.2 Embedded Interrogatives

A case that is not often discussed in the same breath as these other construction is the embedded interrogative. To my knowledge the only person to draw a parallel between the interaction of definite plurals with negation and the interaction of embedded questions with negation is Krifka (1996).

Lahiri (2000) analyzes Quantification Variability Effect with embedded questions (first observed by Berman (1991)) as a case of qunatification over the answers to the question. For example the sentences (392) and (393) differ in their quantification over answers in accordance with the quantity adverb present in the sentence:

(392) Bill partly knows what they serve for breakfast at Tiffany's

(393) Bill mostly knows what they serve for breakfast at Tiffany's

(394) Lahiri's informal paraphrases of the meanings of (392) and (393)

a. some (x)[they serve x at Tiffany's][Bill knows they serve x at Tiffany's]

b. most (x)[they serve x at Tiffany's][Bill knows they serve x at Tiffany's]

Furthermore Lahiri notes that when no adverb is present, embedded questions are interpreted as default universal quantifiers over answers[1]:

(395) a. Mary knows who passed the exam

b. every (x)[x passed the exam] [Mary knows that x passed the exam]

Krifka observes that while (395a) implies that Mary knows all the answers to the embedded question, the negation of (395a), (396a), implies that Mary knows none of the answers.

(396) a. Mary doesn't know who passed the exam.

b. every(x)[x passed the exam] ¬[Mary knows that x passed the exam]

That is, the negation of a sentence with an embedded question and no adverb of quantity is interpreted as a contrary universal negative quantification over answers, not the contradictory negation of a universal. The analogy to definite plurals is plain.

Fodor's test also suggests that a homogeneity presupposition is involved here:

(397) Does Mary know which students passed the exam? No.

A negative answer to this question implies that Mary is clueless about the embedded question. That is, she is completely incapable of distinguishing an exam-passing student from an exam-failing one.

[1] Except in the puzzling case of 'mention-some' readings.

It is too general to say that all embedded questions display this behavior with respect to negation. The only relevant cases are those in which the sentence can be analyzed in terms of universal quantification over (true) answers. As Lahiri observes there are other classes of interrogative embedding predicates whose semantics interacts differently witht the question denoted by the embedded interrogative. First, there is the class of interrogative-embedding predicates that as Lahiri says relate individuals and questions irreducibly (2000 p.329, cf. Groenendijk & Stokhof 1984). The paradigm example in this case is *wonder*.

(398) Bill wonders who passed the exam

The semantics of *wonder* cannot be captured in terms of quantification over the answers to the embedded question. For example, an analysis like (399) fails to capture the meaning of (398).

(399) every(x)[x passed the exam][Bill wonders whether x passed the exam]

Consequently (401) does not capture the meaning of the negation of (400):

(400) Bill doesn't wonder who passed the exam

(401) every(x)[x passed the exam]¬[Bill wonders whether x passed the exam]

The issue here simply doesn't arise. Consequently, we will focus on predicates in the *know*-class, whose semantics is analyzable in terms of the answer to the embedded question. Other predicates in the *wonder*-class are *ask*, *inquire*, *investigate*.

Lahiri's *know-class*, on the other hand, includes predicates such as *forget*, *remember*, *tell*, *be certain*, *realize*, *agree (on)*. A salient property that all of these predicates share is that they all independently embed declarative clauses, in addition to interrogatives. For this reason, Lahiri argues that these predicates take propositions as arguments, both when they embed interrogatives and declarative clauses. Consequently, the properties these predicates display when they embed declarative clauses are relevant to the discussion of cases in which they embed interrogatives. There are certain distinctions among these predicates that are relevant to our discussion. For example, some of these predicates are factive (e.g., *know*) and some are not (e.g., *agree* *(on)*). Another important distinction noted by Lahiri is distributivity. For example *know* is distributive; whether or not *know* is true of a question depends on whether or not the subject knows each of the propositions that makes up the complete true answer to the question.

(402) Mary knows who came
Sue came and Bill came
So, Mary knows Sue came

The truth of a predication involving *know*-class predicate *surprise*, however, does not depend on *surprise* holding between the subject and each of the (atomic) true answers.

(403) a. Bill is surprised who came
b. Sue came and Fred came
c. #So, Bill is surprised Sue came

If Fred and Sue came, the sentence (403a) can be true because Bill is surprised by the combination, though not by either of the facts independently. So, (403a) cannot be paraphrased as (435).

(404) every(x)[x came] [Bill is surprised x came]

So, a predication involving *surprise* and its negation do not bear the same relation to each other that a predication involving *know* and its negation do.

(405) Bill is surprised which students came

(406) Bill isn't surprised which students came

The latter can be true if Sue and Fred came but it is surprising that Fred came, since it could still be unsurprising that combination of Sue and Fred came they usually are inseparable.

The (non-)distributivity of these predicates reveals itself also when they take declarative clauses as complements. Whereas (407i) entails both (407ii) and (407iii), (408i) entails neither (408ii) nor (408iii).

(407) i. Bill knows that Sue came and Fred left

ii. Bill knows that Sue came

iii. Bill knows that Fred left

(408) i. Bill is surprised that Sue came and Fred left

ii. Bill is surprised that Sue came

iii. Bill is surprised that Fred left

This non-distributivity will play a crucial rule in our hypothesis about the origins of the Excluded Middle.

4.1.3 Bare Plurals

Bare plurals in English also exhibit the property of quantificational variability. For example, in the scope of an adverb of quantification they take on the quantificational force of the adverb.

(409) Lions usually have manes
$\approx$ most lions have manes

When no adverb of quantification is present, bare plurals may receive two different interpretations: existential or quasi-universal generic. Many studies have contributed to our understanding of when bare plurals receive which interpretation (Kratzer 1995, Diesing1992, Chierchia 1995, a.o.). One particularly important generalization is that bare plural arguments of individual-level predicates receive the generic interpretation.

In his landmark study of bare plurals, Carlson (1977) notes that the negation of a generic bare plural statement is itself a generic statement. Carlson's examples involve bare plurals objects of the individual-level predicate *like*. (410) is a quasi-universal statement about wombats: Bill likes any normal/typical wombat.

(410) Bill likes wombats.

(411) Bill likes all wombats

The negation of (410), however, is not interpreted as the negation of a universal statement. Rather it is interpreted as the negation of an existential: there is no normal/typical wombat that Bill likes.

(412) Bill doesn't like wombats

(413) Bill doesn't like all wombats

von Fintel (1997) makes crucial use of this property of generic bare plurals in his analysis of *only* + a bare plural. The prejacent has universal force (cf. (414a) - all instances of the kind mammal give live birth), but the negated alternatives have the force of negated existentials, not negated universals, cf. (414b).

(414) Only mammals give live birth

a. All (typical, relevant) instance of the kind mammal give live birth

b. no instances of any other kind of animal give live birth

Also, the Fodor test corroborates the evidence from negation. A negative answer to a question involving a generic bare plural has the force of a negated existential:

(415) a. Do mammals lay eggs?

b. No. (No typical mammal lays eggs)

von Fintel (1997) gives a unified account of the Excluded Middle in generic bare plurals and bare conditionals. He does this by assuming that both constructions involve an implicit generic quantifier that is type general. So, the structure von Fintel assumes for these constructions are (416) and (417). In the case of conditionals, 'f' is a selection function in the sense of Stalnaker 1968 and Lewis 1973; in the case of the bare plurals, f is a function, analogous to the conditional selection function, that selects relevant, typical instances of a kind.

(416) Generic Bare Plural

a. Mammals give live birth

b. GEN(f)(x)[x is a mammal] [x gives live birth]

(417) Bare Conditional

a. If it rains, we play soccer

b. GEN(f)(w)[it rains in w] [we play soccer in w]

von Fintel (1997) proposes the following semantics for the operator GEN (p.37):

(418) For σ either e or s, for all p, q in $D_{<\sigma,t>}$, and worlds w:

$[\![$ gen $]\!]$ (f)(p)(q) is defined for w only if

(i) p is compatible with f(w) : there is x in f(w): p is true of x

(ii) $[\forall x \in f(w){:}\ p(x) \rightarrow q(x)] \vee [\forall x \in f(w){:}\ p(x) \rightarrow \neg q(x)]$

Where defined, $[\![gen]\!]$(f)(p)(q) is true in w iff $[\forall x \in f(w){:}\ p(x) \rightarrow q(x)]$

Relevant for our purposes is clause (ii). This is the presupposition that imposes the Excluded Middle on constructions involving the implicit generic operator. It is von Fintel's proposal that the source of Excluded Middle for generic bare plurals and bare conditionals is lexically stipulated. It follows then that definite plurals and embedded questions, which are not generic, obey the Excluded Middle for different reasons.

4.2 Definite Plurality

In this section, we turn to evidence that the four constructions just discussed are best analyzed as definite plurals:

4.2.1 Conditionals

Non-monotonic Universality

A simple semantics for conditionals in terms of universal quantification over possible worlds makes prediction about what we ought to be able to validly infer from a conditional. For example, since universal quantifiers are downward monotone with respect to their restrictors, we predict that strengthening the antecedent of a conditional ought to yield a logical consequence of the conditional. As Lewis points out, however, strengthening of the antecedent is not valid for conditionals. The classic example is (419)

(419) If the U.S. throws its nuclear weapons into the sea, there will be war, but if the U.S. and all other nuclear powers throw their nuclear weapons into the sea there will be peace.

Lewis (1973) goes on to point out that definite descriptions exhibit a similar failure of a strengthening inference. Under a Strawsonian referential analysis of definite descriptions we expect strengthening the restrictor to yield a valid consequence, so long as both descriptions satisfy their felicity conditions (cf. von Fintel (1999))

(420) The pig is grunting. But the floppy eared pig is not grunting. The spotted floppy-eared pig is grunting.

(421) (There is a unique pig)
The pig is grunting
(There is a unique floppy-eared pig)
Therefore, The floppy-eared pig is grunting

In addition to this failure of monotonicity, Schlenker (2004) points out other predictions of the universal quantification approach that are not borne out: contraposition, transitivity.

(422) If Goethe had survived the year 1832, he would be dead by now.
If Goethe were not dead by now, he would not have survived 1832.

(423) The professor is not Dean
The Dean is not a professor

(424) If Jones wins the election, Smith will retire. If Smith dies tomorrow, Jones will win the election. If Smith dies tomorrow he will retire.

(425) The students are vocal. The undergrads in Beijing are students.
The undergrads in Beijing are vocal.

Morphosyntax

In addition to these shared failures of expected logical properties, Schlenker notes with Bhatt and Pancheva (2001) an analogy between conditional clauses and correlative constructions. Bhatt and Pancheva (2001), for example, point out that in Marathi (an Indo-Aryan language) conditionals are fully structurally analogous to correlative constructions. A correlative construction consists of free relative clause adjoined to a clause that contains a pronominal element that picks up the reference of the free relative. In Marathi, the morpheme corresponding to *if* belongs to the same morphological paradigm as the relative pronouns used in individual-referring free relatives. Similarly, the pronoun corresponding to *then* in Marathi comes from the same paradigm as the pronouns that pick up the reference of individual-denoting free relatives.

(426) [dzar tyaane abhyaas kelaa] [tar to paas hoiil]
if he-erg studying do-Pst-3MSg then he pass be-Fut-3Sg
If he studies, he will pass'

In Jacobson's classic analysis (Jacobson (1995)), free relatives are taken to denote definite descriptions. The morphological composition of conditionals in Marathi suggests a correlative analysis of conditionals, which when combined with Jacobson's analysis of free relatives suggests an analysis of conditionals as definite descriptions of worlds.

4.2.2 Embedded Interrogatives

Lahiri (2000) gives an analysis of QVE with embedded questions in terms of mass quantification. This leads him to analyze the denotations of interrogative embedded under *know*-class predicates in terms of mass-like part-whole structures. One reason Lahiri opts for this analysis is that QVE with embedded questions is much more felicitous with adverbs of quantity (mostly, partly, to a large extent) than with adverbs of frequency (usually, sometimes, often) the latter familiar from the literature on QVE with indefinites.

(427) Adverb of Quantity

John mostly knows who was at the party last night

(428) Adverb of Frequency

John usually knows who was at the party last night

(429) Usually a student dislikes his first class

Lahiri's starting point for his semantics is a Hamblin semantics for interrogatives. Thus, to every interrogative, Lahiri assigns a set of propositions as its denotation, as in (430).

(430) $[\![$ who was at the party$]\!]$ =

{that a was at the party, that b was at the party, that c was at the party,...}

Lahiri then defines a relation that holds between a proposition and a question when that proposition constitutes an answer to the question:

(431) Ans(p, Q) if and only if $\exists$S $\in$ Pow(Q)[p = $\cap$S]

In other words a proposition counts as an answer to a question Q if p is the conjunction of a set of propositions belonging to the question Q.

Using these semantic notions he constructs a mass-like denotation for embedded interrogatives. The crucial next step is the use of Chierchia's (1993) question-to-answer shift rule.

(432) Question-to-answer shift

$\mathrm{Q} \rightarrow \sigma\mathrm{p}[\ \mathrm{Ans(p,Q)} \wedge \mathrm{C(p)}\ ]$

This assigns to a question the sum of answers to that question that satisfy a contextually identified property C. Though Lahiri wishes to assign mass structures to embedded questions (because of their interaction with adverbs of quantity), we might just as well view these structures as sums like those assigned to definite plurals. We can do this because the sums denoted by Lahiri's embedded interrogatives have atoms, as Lahiri notes. The atoms are the propositions that belong to the pre-type shifted denotation of the interrogative. Atomicity is often identified as the key difference between the denotations of mass terms and the denotation of definite plurals.

Lahiri notes that the part-quantification he assigns to QVE sentences is equivalent to simple quantification over propositions. So why introduce part-whole structures?

Collectivity Lahiri's first answer is that there are some interrogative-embedding predicates that are true or false of answer sums as wholes and not dependent on quantification over answer parts. The analogy here is to collective predication with definite plurals and mass terms. The predicates at issue here are those mentioned in §4.1.2, the non-distributive predicates Lahiri's *surprise*-class.

The argument goes as follows. The *surprise*-class predicates still belong to the *know*-class. That is, they are basically proposition-taking predicates. This is indicated by the fact that they all take declarative clauses as complements.

(433) a. It is surprising who came yesterday

b. It is surprising Bill came yesterday

(434) a. It is amazing which students Mary favors

b. It is amazing that Mary favors Fred

Consequently, in order to be interpreted the embedded interrogatives must either form the restriction of a covert quantifier over propositions or undergo question-to-answer shift. Lahiri argues that the former option is inadequate. The default covert quantifier is a universal. So a sentence like (433) would receive the truthconditions in (435), under covert quantification.

(435) $\forall$ p [p $\in$ $[\![$who came to the party $]\!]$ $\wedge$ $^{\vee}$p] [surprising(p)]

As we already know, however, this does not adequately represent the meaning of (433). (Though there may be a distributive reading like this) Instead an analysis in terms of question-to-answer shift is much more promising. Lahiri shows, in fact, how a mass analysis of the default universal quantifier can account for the meaning of sentences with non-distributive predicates. Demonstrating how his account works would take us too far afield; but we might imagine a dramatically simplified version which makes use of the sums via a sum-to-proposition shift rule:

(436) σp[F(p)] $\rightarrow \cap${p: F(p)} (We might implement this via Lahiri's Ψ operator which maps a sum to its atomic members: $\cap${p: p $\in \Psi(\sigma$p[F(p)])}.)

This would allow us to say that in a case where the answer to who came to the party is Bill and Sue, then (437) is just equivalent to (438), which does not imply, for example, that Fred is surprised that Bill came.

(437) Fred is surprised who came to the party

(438) Fred is surprised that Bill came and Sue came

For this version of Lahiri's analysis to be preferable to the quantificational analysis, there would have to be some principle banning a direct question to proposition rule.

(439) Q $\rightarrow \cap${p: Ans(p,Q) $\wedge$ C(p)} (or λw. $\forall$p [Ans(p,Q) $\wedge$ C(p) $\rightarrow$ p(w)=1])

I refer the reader to Lahiri (2000) §§2.2.5-2.2.6 for details on Lahiri's actual proposal.

This argument essentially depends on the existence of non-distributive proposition-taking predicates. Most sentence embedding predicates are analyzed as universal quantifiers over worlds. This predicts that they ought to distribute over conjunction.

(440) $[\![$FR(p)$]\!]^w$ = $\forall$ w' [wRw' $\rightarrow$ p(w)=1]

(441) FR(p$\wedge$ q) entails FR(p) and FR(q)

So these non-distributive predicates must not be represented in terms of universal quantification over worlds. They may involve a different quantificational force or they may be true of the set of worlds as a whole. For example, *surprise* might receive an analysis like (442).

(442) $[\![$ x is surprised that p $]\!]^t$ is defined only if p is true at t
when defined, is true iff (i) x believes p at t and (ii) at sometime before t, x believed that p would be false at t

Conditionals and Collectivity I bring this up since one question facing Schlenker's analysis of conditionals is whether it is ever necessary to interpret an if-clause, which he analyzes as a plural definite description, collectively. Schlenker says that as far as he is aware this does not happen (Schlenker (2004) fn.8).

Semidistributivity As we have seen, sentences containing definite plurals and a distributive predicate receive a universal interpretation. Consequently, we might think that when a sentence contains two definite plurals that it is interpreted as if it contained two universal quantifiers.

(443) a. The professors are blond.

b. Every professor is blond

(444) a. The professors admonished the students.

b. Every professor admonished every student.

Arguably, sentence (444a) does have a reading corresponding to (444b). However, it is well known that such sentences also have a weaker reading than this. Under this reading, (444a) entails that every professor admonished some student and every student was admonished by some professor. Scha (1981) illustrated the existence of this reading with sentences like (445). The reading of such sentences in terms of two universal quantifiers cannot possibly be true.

(445) The sides of square A are parallel to the sides of square B.

(446) Every side of square A is parallel to every side of square B.

(Geometrically impossible)

Though the reading corresponding to (446) is necessarily false, sentence (445) nevertheless may be judged true. The true reading is the one that is true iff that every side of square A is parallel to some side of square B and every side of square B is parallel to some side of square A. Such a reading is not available to a sentence containing two universal quantifiers, as shown by (447) and (448). Nor is such a reading available for a combination of a definite plural and a universal quantifier.

(447) The professors admonished every student.

(448) Every professor admonished the students.

[A pseudo-semidistrbutive reading is available for (447) through a collective reading of the subject the professors. Lahiri makes the same point about the collective reading of the subject in (449).

(449) The members of team A found out that the Leitches were the culprits

Lahiri (2002) p. 189

Separating genuine semidistributive readings from collective readings is crucial to Lahiri's account.]

(450) $^{**}P(x)(y) = 1$ iff $\forall\ x_1 < x\ \exists\ y_1 < y$ s.t. $P(x_1)(y_1)$ and

$\forall\ y_1 < y\ \exists\ x_1 < x$ s.t. $P(x_1)(y_1)$

With this background, now notice that embedded interrogatives give rise to semidistributive readings in combination with the subject of their embedding predicate. Lahiri gives (451) as an example:

(451) The witnesses know which Klansmen were at the rally.

As Lahiri notes, this can be true even if it is not the case that each witness knows of every Klansman who was at the rally that he was at the rally. It is enough if each witness knows part of the complete, true answer and that every part of the complete, true answer is known by some witness. For example, if Fred Sue and Mary are the witnesses and Bill, John and Simon were Klansmen at the rally then the sentence (451) can be judged true in a scenario like (452).

(452) Fred knows Bill and Simon were at the rally (but doesn't know John was)

Sue knows John was at the rally (but not Bill and Simon)

Mary knows Simon was at the rally (but not Bill and John)

I take this as evidence that embedded interrogatives can denote sums, or pluralities, of propositions. Lahiri draws a slightly different conclusion, exploration of which would take us too far afield. Basically, I propose that via question-to-answer shift the embedded question in (451) denotes (453), modeling pluralities as sets. [X.w.t.r. = that X was at the rally]

(453) {B.w.t.r., J.w.t.r., S.w.t.r.}

And the DP *the witnesses* denotes the plurality {Fred, Sue, Mary}. So, (451) has the truth conditions in (454).

(454) $[\![(451)]\!]$ is true iff $**[\![\text{knows}]\!](\{\text{B.w.t.r., J.w.t.r., S.w.t.r.}\})(\{\text{F, S, M}\}) = 1$
iff $\forall$ p$\in${B.w.t.r., J.w.t.r., S.w.t.r.} $\exists$ x$\in${F, S, M} s.t. x knows that p and
$\forall$ x$\in${F, S, M} $\exists$ p$\in${B.w.t.r., J.w.t.r., S.w.t.r.} s.t. x knows that p

Conditionals and Semidistributivity If we follow Schwarzschild (1996) and Lahiri (2002) in believing that semidistributivity is generally available when two plurals are arguments of the same relation, then we should expect to find semidistributivity involving other non-canonical definite plurals, such as conditionals under Schlenker's account. That the precondition for a semidistributive reading, namely two plurals filling argument slots of the same relation, can be met with a conditional clause can be seen in the logical form Schlenker assigns to conditionals:

(455) If John is sick, Mary is unhappy

(456) unhappy(Mary, t_0{local}, [ι_{w_0} w_i sick(J,t_0{local}, w_i)]{LOCAL})

Ignoring information about tense and mood, irrelevant for our purposes:

(457) unhappy(Mary, [ι_{w_0} w_i sick(J, w_i)])[2]

Here we see that the subject of the consequent and the sum of worlds denoted by the conditional clauses are both arguments of the main predicate of the consequent, in this case unhappy.

Now imagine a similar case in which the subject of the consequent denotes a plurality. Now we would expect a semidistributive reading to be possible via application of the **-operator to the predicate of the consequent.

(458) If John is sick, the students are unhappy

Assuming that the students are Mary, Fred and Sue, this conditional has the logical structure in (459) and truth conditions in (460).

(459) **unhappy({M, F, S}, {w': w_0Rw' and sick(J, w')})

[2]'[$\iota_{w'}$ w ϕ]' refers to the set of worlds accessible from w′ that satisfy the property ϕ.

(460) $[\![(458)]\!] = 1$ iff $\forall$ w" $\in$ {w': w_0Rw' and sick(J, w')} $\exists$ x $\in$ {M,F,S} unhappy(x, w") and

$\forall$ x $\in$ {M,F,S} $\exists$ w" $\in$ {w': w_0Rw' and sick(J, w')} unhappy(x, w")

We might paraphrase this predicted reading in the following way: if John is sick, one of the students is certainly unhappy and each of the students might be unhappy (if John is sick). As far as I can tell, such a reading is not possible for (458). So, as in the case of collective readings, the evidence for the plurality of embedded questions does not extend neatly to conditional clauses.

4.2.3 Bare Plurals

The position that (at least generic) bare plurals denote sums of individuals like definite plurals is already well established in the literature. The specific position concerning bare plurals that I will endorse is the one put forward in Chierchia 1998. According to Chierchia's Neo-Carlsonian theory, bare plurals denote kinds unambiguously. The crucial question is what is a kind. Carlson (1977) originally proposed that kinds were special individuals that are related to the object that instantiate them through a realization relation R. Chierchia takes a more reductive position, identifying kinds with a function from worlds to the sum of the instances of the kind in that world:

> It seems natural to identify a kind in any given world (or situation) with the totality of its instances. Thus, the dog-kind in our world can be identified with the totality of dogs, the scattered entity that comprises all dogs, or the fusion of all dogs around. In our framework this entity is modeled by the set of dogs.
>
> (Chierchia (1998) p. 349)

It is instructive to examine Chierchia's definition of the down-operator that maps properties to kinds in order to see what he has in mind for the structure of kinds.

(461) ιX = the largest member of X, if there is one (else, undefined)

(462) For any property P and world s

$^\cap$P = λs. ιP_s is in K undefined otherwise

where P_s is the extension of P in s

[and K is the set of kinds - JRG]

(Chierchia (1998) p.351)

Morphosyntax The idea of treating bare plurals as terms referring to sums of individuals is supported by the fact that in many language generic plurals are not in fact bare. Rather, in languages such as French and Italian the noun phrases corresponding to English bare plurals are plural definite descriptions. Chierchia (1998) gives the examples such as (463) and (464) to illustrate this difference between English and Italian.

(463) a. Dogs love to play

b. *Cani amano giocare

c. I Cani amano giocare

(464) a. Dogs are rare in these parts

b. *Cani sono rari in queste parti

c. I cani sono rari in queste parti

This perhaps suggests that bare plurals are sum-denoting definite plurals in English. This view could be implemented by assuming that bare plurals in English have a covert determiner that maps a predicate to the sum of its instance, if that sum denotes a kind. The Italian definite determiner serves this function as well as the function of mapping a predicate to the sum of its members that are salient in a context (the function served by English *the*).

In the next few sections, I ask whether the arguments for the plurality of conditional clauses and embedded interrogatives apply to bare plurals.

Collectivity First, do bare plurals exhibit collective readings? Assuming Chierchia's analysis of kinds, I think that the uncontroversial answer is yes. Any instance of irreducible kind predication involving a bare plural can now be viewed as a collective predication of the sum of the instances of a kind. Common examples of kind-predication with bare plurals are in (465).

(465) Pandas will become extinct soon
Potatoes were introduced into Ireland by the end of the 17th century
Rhinos are common

Semidistributivity Next, we ask whether generic bare plurals give rise to semidistributive readings when they are the arguments of a predicate that takes another plural argument. Here are some attempts to formulate relevant examples. In each case, the main predicate is individual level (Carlson, Kratzer) with respect to the subject. This forces the generic reading of the bare plural. The object position is filled by a definite description or conjunction of individuals.

(466) a. Children love the books of Dr. Seuss
b. Children love *The Cat in the Hat* and *The Lorax*

(467) a. Linguists trained by Ken Hale know (how to speak) the native languages of Australia
b. Linguists trained by Ken hale know Warlpiri and Lardil

In these examples, a double distributive reading is the most prominent. That is, these sentences could be naturally paraphrased with two universal quantifiers. For example, I would find it difficult to judge (466b) true, if a large number of typical children, while loving *The Cat in the Hat*, hated *The Lorax*. Similarly, (467) strongly implicates that every linguist trained by Ken Hale knows both Warlpiri and Lardil.

(468) Every (typical) child loves every book by Doctor Seuss

(469) Every typical student of Ken Hale's knows every native language of Australia.

There are slightly different sentences, however, in which a semidistributive reading is available. The crucial property of these sentences is that it is impossible for an individual part of their subjects to bear the relation denoted by the main predicate to two different individual parts of their objects. If only a double distributive reading were available, these sentences ought to be trivially false.

(470) Blitvans have the (three) eye-colors common in Blatvia

(471) Blitvans have blue eyes and brown eyes

(472) $\forall$ x [x is Blitvan $\rightarrow \exists$ y [y is eyecolor in Blatvia $\wedge$ x has y] $\wedge$
$\forall$ x [x eyecolor in Blatvia $\rightarrow \exists$ y [y is Blitvan $\wedge$ y has x]

For comparison, consider analogous examples in which the generic bare plural is replaced by a garden variety quantifier over individuals.

(473) #Every Blitvan has the (three) eye-colors common in Blatvia

(474) #Every Blitvan has blue eyes and brown eyes

Concerning the contrast between (466) and (470), Winter (2001) has observed a similar effect in the conjunction of predicates. (475), whose main predicate is the conjunction of two incompatible predicates, allows the interpretation (476). (477), on the other hand, whose main predicate conjoins compatible predicates, only allows the interpretation in (478a), not (478b).

(475) The ducks are swimming and flying

(476) Some ducks are swimming, the others are flying

(477) The ducks are flying and quacking

(478) a. Each of the ducks is flying and quacking.
b. Some ducks are flying, the others are quacking.

On the basis of this effect, Winter argues against an analysis in terms of semidistributivity and in favor of a more general theory making use of the Strongest Meaning Hypothesis. While this indeed raises questions about the simple approach that I will adopt, it does not affect the point that generic bare plurals are to be interpreted as definite plurals, since this is a precondition also for the operation of the Strongest Meaning Hypothesis.

I tentatively conclude on the basis of these examples that semidistributive readings are available to generic bare plurals.

4.3 Summary

In this section, we have seen evidence for a correlation between obeying the Excluded Middle and showing signs of definite plurality. Among the properties that we have seen that diagnose definite plurality are non-monotonicity, cross-linguistic morphosyntax, non-distributivity and semidistributivity. While none of the constructions examined exhibits each of these properties, each exhibits some of them. In the next chapter, I give a formal account of the semantic structure that I believe underlies each of these constructions. Furthermore, I offer an explanation of why this semantic structure gives rise to Excluded Middle presuppositions.

Chapter 5

Explaining the Excluded Middle

5.1 A General Theory of the Excluded Middle

In this section, I put forward an explicit implementation of the idea that the four construction discussed in chapter 4 are to be analyzed as definite plurals and that the Excluded Middle is an effect of distributive plural predication.

First, I define the domains of the model for the semantics. Definitions (479) and (480) are standard.

(479) Domain of semantic objects

D_e = the set of atomic individuals

$D_t = \{0, 1\}$

D_s = the set of atomic worlds

(480) For every pair of types σ, τ there exists the domain of object

$D_{<\sigma,\tau>} = {D_\sigma}^{D_\tau}$

More controversially, I assign to every type a domain of the pluralities of that type. Pluralities of type σ are taken to be non-empty sets of the atomic elements of D_σ.

(481) For every domain D_σ there exists the domain

$D_{\{\sigma\}} = \mathrm{Pow}(D_\sigma) - \emptyset$

I follow Quine and Schwarzschild in identifying individuals with the singleton sets containing them. Crucial for our purposes are the domains in (482), the pluralities of individuals, worlds and propositions.

(482) $D_{\{e\}}$, $D_{\{s\}}$ and $D_{\{<s,t>\}}$

Assumptions about selection functions:

I posit a class of functions that subsumes the English definite determiner, which I will call the class of selection functions. These functions map functions to subsets of their characteristic sets. I illustrate with the definite determiner in English

(483) $[\![$ the $]\!] = \lambda f_{<e,t>}$:$\exists$y f(y) =1. {x: f(x) = 1 and x is salient}

This definite determiner maps a predicate whose extension is non-empty to the set of salient objects that satisfy the predicate. I follow Sauerland (2004) in his interpretation of number morphology in DPs. According to Sauerland, above each D head there is a ϕ-head where the semantically interpreted singular/plural feature resides. The singular feature introduces the presupposition that the individual denoted by the DP is a member of D_e. The plural feature is semantically vacuous, but its use is circumscribed by a version Heim's principle of Maximize Presupposition: whenever the presuppositions of the singular feature are satisfied it must be used instead of the plural. Singular/plural morphology on N heads is uninterpretable and must be licensed by the number feature in the ϕ-head.

So when the characteristic set of the nominal predicate in a definite DP contains only one individual, the ϕ-head above the definite DP must host the singular feature and, consequently, singular morphology must appear on N. If, on the other hand, the characteristic set is not a singleton the ϕ head hosts the plural feature and plural morphology appears on N.

(484) if $[\![$the$]\!]$($[\![$student$]\!]$) = {Bill}, then the LF of the student must be
[SG [the [$\text{student}_{[uSG]}$]]]

(485) if $[\![$the$]\!]$($[\![$student$]\!]$) = {Bill, Fred}, the LF of the students is
[PL [the [$\text{student}_{[uPL]}$]]]

The structure of selection functions more generally is given in (486)

(486) $[\![$ s.f. $]\!] = \lambda f_{<\sigma,t>}$. {x: f(x)=1 and C(x)}

(= λ $f_{<\sigma,t>}.\lambda y_{\tau}$. {x: f(x)=1 and R(x,y)})[1]

The selection function serves to select a subset of f's by means of C (or R) and forms a plurality of the selected subset. The type of the selection function and the content of C(or R) is what differentiates one construction from another. Below I list the selection functions that belong to the four constructions under discussion.

I follow Schlenker in assuming that definites involve a selection function parallel to that in conditionals. The function selects individuals from a predicate on the basis of their salience in the context.

(487) Definite Plural DPs (*the*):

function selects individuals based on salience

(von Heusinger, Schlenker)

λ $f_{<e,t>}$: $\exists$ y f(y) =1. {x: f(x) = 1 and x is salient}

One might object that selection on the basis of salience is redundant since domain selection accomplishes a similar kind of selection for all quantifiers. Schlenker notes that definites exhibit an ability to winnow down the domain of quantification beyond what is done by domain selection in standard quantifiers. He credits Peter Svenonius (p.c.) for pointing this out:

(488) [There are ten girls and ten boys in the class. Three girls raise their hands. Talking to the speaker, I say:]

a. Wait, the girls have a question!

b. Wait, the three girls have a question!

c. ?Wait, the girls each have a question!

d. #Wait, every girl has a question!

e. #Wait, all girls have a question!

[1] I include the relational version for cases in which selection is done relative to another object as in modal accessibility relations.

f. #Wait, all the girls have a question!

g. #Wait each of the girls has a question!

A similar point is made by Larson and Segal (1995) in their discussion of incomplete descriptions:

(489) Suppose Boris enters the hall twice through different doors, each time leaving the door open. Suppose there are five doors in the room, and consider [(i)] versus [(ii)]:

(i) The door is open

(ii) Every door is open

Whereas Natasha could say something true by uttering [(i)] in this situation, it certainly doesn't seem that she could do so by uttering [(ii)]. But if the mechanisms governing domains are the same in the two cases, then this is mysterious. If the domain of quantification can shrink small enough to exclude all doors but the one Boris left open, which thereby allows Natasha to say something true with [(i)] one would expect the same possibility in the case of Natasha's saying [(ii)].

(Larson and Segal 1995, p. 333)

The use of selection functions in conditionals is well-established.

(490) Conditional Clauses (*if*):

function selects worlds based on comparative similarity

(Stalnaker 1968, Lewis 1973)

λ $f_{<s,t>}$: $\exists$ y f(y) =1.λ u_s.{v: f(v) = 1 and uRv}

'uRv' = 'v is at least as similar to u as any other world'

It is common to analyze generic bare plurals in terms of universal quantification over typical/normal cases (and there are well-known problems with this analysis, cf. Carlson (1977)). von Fintel (1997) proposes the use of a selection function (analogous to those used in conditionals) in picking out the typical/normal cases. Under our analysis, the quantificational force is factored out as a contribution of the distributivity operator. The function of

(491) Generic Bare Plurals (null in English, definite determiner in Italian)
function selects individuals based on typicality/normality
(von Fintel 1997)

λ $f_{<e,t>}$: $\exists$ y f(y) =1. {x: f(x) = 1 and x is a typical f}

A function that quantifies over a subset of the propositions in the set denoted by an interrogative (or forms their intersection) is also common. Heim (1994) calls such a function Ans1. We have already discussed the Chierchia/Lahiri version of this function (see (431) above).

(492) Embedded Interrogatives (Ans)
function selects a set of propositions
(Heim, Chierchia, Lahiri)

λ $f_{<<s,t>,t>}$: $\exists$ y f(y) =1.{p: f(p) = 1 and C(p)}

'C(p)' is often $^{\vee}$p, though not always

The cornerstone of this analysis of the interaction of these expressions with negation is the denotation of the distributivity operator. For this purpose, I adopt a denotation for the distributivity operator along the lines of Löbner (2000) and Schwarzschild (1994).

(493) Generalized Distributivity Operator (*)
$\lambda f_{<\sigma,t>}.\lambda x_\sigma$: $\forall$y$\in$ x f(x) = 1 or $\forall$ y $\in$ x f(x) = 0. $\forall$ y $\in$ x f(x) = 1

(494) [s.f.$_{<<\sigma,t>,\{\sigma\}>}$[P1$_{<\sigma,t>}$]] [*$_{<<\sigma,t>,<\{\sigma\},t>>}$ [P2$_{<\sigma,t>}$]]

(495) $[\![$ (494) $]\!]$ is defined only if
$\forall$x [x $\in$ $[\![$s.f.$]\!]$($[\![$P1$]\!]$) $\rightarrow$ $[\![$P2$]\!]$(x) =1] or $\forall$x [x $\in$ $[\![$s.f.$]\!]$($[\![$P1$]\!]$) $\rightarrow$ $[\![$P2$]\!]$(x) =0]
when defined, $[\![$ (494) $]\!]$ = 1 iff $\forall$x [x $\in$ $[\![$s.f.$]\!]$($[\![$P1$]\!]$) $\rightarrow$ $[\![$P2$]\!]$(x) =1]

(496) $[\![$ not (494) $]\!]$ is defined only if

$\forall x$ [x $\in$ $[\![$s.f.$]\!]$($[\![$P1$]\!]$) $\rightarrow$ $[\![$P2$]\!]$(x) =1] or $\forall x$ [x $\in$ $[\![$s.f.$]\!]$($[\![$P1$]\!]$) $\rightarrow$ $[\![$P2$]\!]$(x) =0]

when defined, $[\![$ not (494) $]\!]$ = 1 iff $\neg\forall x$ [x $\in$ $[\![$s.f.$]\!]$($[\![$P1$]\!]$) $\rightarrow$ $[\![$P2$]\!]$(x) =1]

(497) The Semidistributivity Operator

$\lambda R_{<\sigma,<\tau,t>>}.\lambda x_{\{\sigma\}}.\lambda y_{\{\tau\}}.\ \forall z\ [\ z \in y \rightarrow \exists w\ [\ w \in x \wedge R(z)(w) = 1]$ and

$\forall\ z\ [\ z \in x \rightarrow \exists w\ [\ w \in y \wedge R(w)(z) = 1]$

An interesting question arises with respect to the semidistributivity operator: does it also carry an excluded middle presupposition? What would that presupposition look like? We can take as our guide the truthconditions of sentences in which a predicate with two plural arguments is negated. For example, (498) permits a semidistributive reading:

(498) The men danced with the women

Under what conditions is (498) false? Analogously, what are the truthconditions of (499)?

(499) The men didn't dance with the women

Addressing the second question, it seems that (499) can only be true if no man danced with any woman. It does not seem sufficient to verify (499) to point to a man who didn't dance with any of the women to or to a woman who didn't dance with any of the men. Similarly, a negative answer to the question in (500) implies that no man danced with any woman.

(500) Did the men dance with the women? No

There are many possible Logical Forms for a sentence with two plurals such as (498). First of all, there exists the possibility that the sentence contains two distributive operators, one for each plural. (501) and (502) are cases in which negation does not have the widest scope. Both of these entail that no man danced with any woman.

(501) [the men $^*\lambda x$ [not [the women $^*\lambda y$ [x dance y]]]]

(502) [the men $^*\lambda x$ [the women $^*\lambda y$ [not [x dance y]]]]

It is crucial to examine the case in which negation has widest scope, however. Consider the question answer pair in (500). Here it cannot be the case that the negation implied by the negative answer has scope under the definite NPs.[2] For example,

(503) Did someone empty the dishwasher?

No ($\neq$ someone didn't)

The relevant structure is (504).

(504) $[_\delta$ not $[_\gamma$ the men $*[$ 1 $[_\beta$ the women $*[_\alpha$ 2 t_1 dance t_2]]]]

As it turns out, even a sentence with this structure entails that no man danced with any woman. (d.w. = danced with)

(505) a. $[\![\alpha]\!]^g = \lambda y_e.$ g(1) d.w. y

b. $*([\![\alpha]\!]^g) = \lambda z_{\{e\}}$: $\forall$y $\in$ z g(1) d.w. y or $\forall$y$\in$ z g(1) $\neg$d.w. y. $\forall$y$\in$ z g(1) d.w. y

c. $[\![$the women$]\!]$ = {x: x is a woman}

d. $[\![\beta]\!] = *([\![\alpha]\!]^g)([\![$the women$]\!])$

$[\![\beta]\!]$ is defined only if $\forall$y[y is a woman $\rightarrow$ g(1) d.w. y] or $\forall$y[y is a woman $\rightarrow$ g(1) $\neg$ d.w. y]

when defined is true iff $\forall$y[y is a woman $\rightarrow$ g(1) d.w. y]

e. $[\![1\ \beta]\!]^g = \lambda x$: $\beta \in$dom$([\![\ ]\!]^{g[x/1]})$. $[\![\beta]\!]^{g[x/1]}$

= λx:$\forall$y[y is a woman $\rightarrow$ x d.w. y] or $\forall$y[y is a woman $\rightarrow$ x $\neg$d.w. y].

$\forall$y[y is a woman $\rightarrow$ x d.w. y]

f. $*[\![1\ \beta]\!]^g = \lambda z_{\{e\}}$:$\forall$u$\in$z $[\![1\ \beta]\!]$(u) = 1 or $\forall$u$\in$z $[\![1\ \beta]\!]$(u) = 0. $\forall$u$\in$z $[\![1\ \beta]\!]$(u)=1

= $\lambda z_{\{e\}}$: $\forall$u$\in$z $\forall$y[y is a woman $\rightarrow$ x d.w. y] or $\forall$u$\in$z $\forall$y[y is a woman $\rightarrow$ x $\neg$ d.w. y]. $\forall$u$\in$z $\forall$y[y is a woman $\rightarrow$ x d.w. y]

g. $[\![\ \gamma\ ]\!] = [\![\ 1\ \beta\ ]\!]([\![$ the men $]\!])$ is defined only if

$\forall$u[u is man $\rightarrow$ $\forall$y[y is a woman $\rightarrow$ x d.w. y]] or $\forall$u[u is a man $\rightarrow$ $\forall$y[y is a woman $\rightarrow$ x $\neg$ d.w. y]

[2] My use of the notion of scope is somewhat vague here, but could be made precise in terms of the scope of negation introduced by the yes/no question operator. Guerzoni 2001 argues that even may scope over this negation, but notes that nothing else is able to.

when defined is true iff $\forall$ u[u is man $\rightarrow$ $\forall$ y [y is a woman $\rightarrow$ x d.w. y]]

h. $[\![\delta]\!]$ is defined only if $[\![\gamma]\!]$ is defined

when defined is true iff $\neg\forall$ u[u is man $\rightarrow$ $\forall$ y [y is a woman $\rightarrow$ x d.w. y]]

Thus, whenever $[\![\delta]\!]$ =1 it must be the case that no man danced with any woman.

Sub-proof:

$\forall$u$\in$z $[\![1\ \beta]\!]$(u) = 0 iff $\forall$u$\in$z $\forall$ y [y is a woman $\rightarrow$ x $\neg$ d.w. y]

For arbitrary u (in z) assume $[\![1\ \beta]\!]$(u) = 0.

But $[\![1\ \beta]\!]$(u) is defined only if $\forall$y[y is a woman $\rightarrow$ x d.w. y] or $\forall$y[y is a woman $\rightarrow$ x $\neg$ d.w. y]

So, $\forall$y[y is a woman $\rightarrow$ x d.w. y] or $\forall$y[y is a woman $\rightarrow$ x $\neg$ d.w. y]

Furthermore, when defined $[\![1\ \beta]\!]$(u) =1 iff $\forall$ y [y is a woman $\rightarrow$ x d.w. y]

Since $[\![1\ \beta]\!]$(u) = 0, we conclude that $\neg\forall$y[y is a woman $\rightarrow$ x d.w. y]

Therefore it follows that $\forall$y[y is a woman $\rightarrow$ x d.w. y] or $\forall$y[y is a woman $\rightarrow$ x $\neg$ d.w. y] and $\neg\forall$y[y is a woman $\rightarrow$ x d.w. y], which entails that $\forall$y[y is a woman $\rightarrow$ x $\neg$ d.w. y]

The proof that $[\![1\ \beta]\!]$(u) = 0 if $\forall$y[y is a woman $\rightarrow$ x $\neg$ d.w. y] is obvious.

Schwarzschild (1994) defines a generalization of the distributivity operator that includes a generalization of the excluded middle presupposition. To understand his definition we need to take a brief detour to understand the system he uses (and why he uses it).

First let's examine a general definition of the distributivity operator given by Sauerland (1998) (following Krifka (1986)). This definition does not incorporate an excluded middle presupposition into the meaning of the distributivity operator.

(506) For F of type $< e, < e, ..., < e, t > ... >$, $\star$F is the function such that:

a. $\forall$ x_1,..., x_n: if F(x_1)...(x_n) = 1, then $\star$F(x_1)...(x_n) = 1

b. $\forall$ x_1,..., x_n, y_1,..., y_n:

if $\star$F(x_1)...(x_n) = 1 and $\star$F(y_1)...(y_n) = 1, then

$\star$F($x_1 \oplus y_1$)...($x_n \oplus y_n$) = 1

c. For any function F' that satisfies a. and b. :

$\forall$ x_1,..., x_n: if $\star$F(x_1)...(x_n) = 1, then F'(x_1)...(x_n) = 1

This definition only picks out a unique $\star$F for each F when we limit our attention to the set of total functions. Partial functions, however, are crucially the focus of our investigation. The reason this definition fails to pick out a unique function when we consider the set of partial functions is that there may be multiple functions that agree on the tuples that they map to 1, while disagreeing on which remaining tuples they map to 0 and which they are undefined for.

Let's begin to accommodate partial functions by amending a special case of Sauerland's definition, namely the case of functions of type <e,t>. In the following alteration, I have added a clause for tuples the function maps to 0 fully parallel to the clauses for the tuples the function maps to 1.

(507) For F of type $< e, t >$, $\star$F is the function such that:

a. $\forall$ x: if F(x) = 1, then $\star$F(x) = 1 and if F(x) = 0, then $\star$F(x) = 0

b. $\forall$ x, y: if $\star$F(x) = 1 and $\star$F(y) = 1, then $\star$F(x$\oplus$y) = 1 and

if $\star$F(x) = 0 and $\star$F(y) = 0, then $\star$F(x$\oplus$y) = 0

c. For any function F' that satisfies a. and b. :

$\forall$ x: if $\star$F(x) = 1, then F'(x) = 1 and if $\star$F(x) = 0, then F'(x) = 0

Interestingly, the result is a function with a built-in excluded middle presupposition (this result is already derived in Schwarzschild (1994)). The attempt to extend this symmetric definition to relations of higher types fails, however:

(508) For F of type $< e, < e, ..., < e, t > ... >$, $\star$F is the function such that:

a. $\forall$ x_1,..., x_n: if F(x_1)...(x_n) = 1, then $\star$F(x_1)...(x_n) = 1 and if F(x_1)...(x_n) = 0, then $\star$F(x_1)...(x_n) = 0

b. $\forall$ x_1,..., x_n, y_1,..., y_n:

if $\star$F(x_1)...(x_n) = 1 and $\star$F(y_1)...(y_n) = 1, then $\star$F($x_1 \oplus y_1$)...($x_n \oplus y_n$) = 1 and

if $\star$F(x_1)...(x_n) = 0 and $\star$F(y_1)...(y_n) = 0, then $\star$F($x_1 \oplus y_1$)...($x_n \oplus y_n$) = 0

c. For any function F' that satisfies a. and b. :

$\forall$ x_1,..., x_n: if $\star F(x_1)...(x_n) = 1$, then $F'(x_1)...(x_n) = 1$ and $\forall$ x_1,..., x_n: if $\star F(x_1)...(x_n) = 0$, then $F'(x_1)...(x_n) = 0$

This definition fails because it does not define a function. Consider the following relation:

(509) $[\![$ admires $]\!]$:
<Bill, Mary> → 1
<Fred, Sue> → 1
<Bill, Sue> → 0
<Fred, Mary> → 0

By clause a. of (508), $\star[\![$ admires $]\!]$ maps each of the pairs in (509) to the same value as $[\![$ admire $]\!]$. Furthermore, by clause b. it maps the pair of sums <Bill⊕Fred, Mary⊕Sue> both to 1 and 0; <Bill, Mary> and <Fred, Sue> map to 1, therefore <Bill⊕Fred, Mary⊕Sue> does as well, <Bill, Sue> and <Fred, Mary> map to 0, therefore <Bill⊕Fred, Mary⊕Sue> does as well. A function, however, cannot map one argument to two distinct values.

(510) $\star[\![$ admires $]\!]$:
<Bill⊕Fred, Mary⊕Sue> → 1
<Bill⊕Fred, Mary⊕Sue> → 0

To accommodate such a relationship between a predicate denotation and an individual (or tuple), we would have to abandon the attempt to assign predicates functions as their denotations. Such a move is made by Schwarzschild (1994), following Cooper (1983), who assigns to each predicate a pair of extensions, the positive and negative extension. In such a system it is possible for the positive extension and negative extension to overlap.

Here is a very brief sketch of Schwarzschild's system: Every predicate is assigned an ordered pair of intensions. If R is an ordered pair, then R_+ will stand for the first member and R_- will stand for the second. For example, a predicate such as walk is assigned a positive extension $[\![\text{walk}]\!]_{w,+}$ containing the individuals in the domain of walk that walk in w and a negative extension $[\![\text{walk}]\!]_{w,-}$ containing the individuals in the domain that do not.

(511) Function Application

If c is in the domain of the one-place predicate F:

$[\![\text{F(c)}]\!]_{+} = 1$ iff $[\![\text{c}]\!] \in [\![\text{F}]\!]_{+}$, 0 otherwise

$[\![\text{F(c)}]\!]_{-} = 1$ iff $[\![\text{c}]\!] \in [\![\text{F}]\!]_{-}$, 0 otherwise

(512) A sentence S is true w.r.t. w if $[\![\text{S}]\!]_{w,+} = 1$, false otherwise.

(513) A sentence S is an instance of presupposition failure w.r.t. w if $[\![\text{S}]\!]_{w,+} = [\![\text{S}]\!]_{w,-}$

(514) Q presupposes p iff p is true in all worlds w in which $\text{Q(w)}_{+} = \text{Q(w)}_{-}$

(515) Sentence Negation

$[\![\text{not-S}]\!]_{w,+} = 1$ iff $[\![\text{S}]\!]_{w,+} = 0$

$[\![\text{not-S}]\!]_{w,-} = 1$ iff $[\![\text{S}]\!]_{w,-} = 0$

(516) A sentence S is an instance of presupposition denial w.r.t. w iff $[\![\text{S}]\!]_{w,+} = [\![\text{S}]\!]_{w,-} =1$

(517) Cumulativity Operator

For a one world predicate F and world w,

$[\![\text{*F}]\!]_{w,+} = \text{CLOSE}([\![\text{F}]\!]_{w,+})$

$[\![\text{*F}]\!]_{w,-} = \text{CLOSE}([\![\text{F}]\!]_{w,-})$

$\text{CLOSE(A)} = \{\text{X} \in \text{D} : \text{X} \in \text{A or X} \subseteq \text{A}\}$

So, (if we limit our attention to only predicates that contain no pluralities) the cumulation of a one-place predicate P is true of the sets of individuals of which P is true and false of sets of individuals of which P is false. *P is undefined for sets containing some individuals in $[\![\text{P}]\!]_{+}$ and some in $[\![\text{P}]\!]_{-}$. This is how Schwarzschild implements the excluded middle presupposition.

Now we are ready to see how his definition generalizes to two-place predicates. Just as a one-place predicate is assigned two sets in each world, its positive and negative extensions, a relation is assigned a positive extension containing tuples of which it is true and a negative extension containing the tuples of which it is false. Cumulation extends to two-place relations as follows:

(518) Two-Place Cumulativity Operator

For a transitive verb T, world w and valence # (=+ or -, JRG), $[\![{*}T]\!]_{w,\#}=$ CLOSE$_2$($[\![T]\!]_{w,\#}$)

CLOSE$_2$(R) = $\{\langle X,Y\rangle \in D \times D: \exists K \subseteq R\ [\ (\{X\} = K_l \text{ or } X = K_l) \text{ and } (\{Y\} = K_r \text{ or } Y = K_r)]\}$

Where $K_l = \{x: \exists y\ \langle x,y\rangle \in K\}$ and $K_r = \{x: \exists y\ \langle y,x\rangle \in K\}$

Here it is possible for the positive and negative extension of the cumulated relation to overlap. That is, there may be $\langle X,Y\rangle$ such that $[\![R]\!]_+(X,Y) = 1$ and $[\![R]\!]_-(X,Y) = 1$. (This is not possible in one-place cumulation if the uncumulated predicate only contains individuals.) By definition (513), this is a presupposition failure.

So, for valence #, $\langle X,Y\rangle \in [\![{*}R]\!]_\#$ iff $\forall\ x \in X\ \exists\ y \in Y\ [\ \langle x,y\rangle \in [\![R]\!]_\#\ ]$ and $\forall\ y \in Y\ \exists\ x \in X\ [\langle x,y\rangle \in [\![R]\!]_\#]$. And, furthermore, *R(X,Y) is true and not a presupposition failure iff $\langle X,Y\rangle \in [\![{*}R]\!]_+$ and $\langle X,Y\rangle \notin [\![{*}R]\!]_-$. This means that (498) is true and not a presupposition failure if each of the men danced with a woman and each of the women was danced with by a man and either some man danced with all the women or someone was danced with by all the men. It is false and not a presupposition failure iff $\langle X,Y\rangle \in [\![{*}R]\!]_-$ and $\langle X,Y\rangle \notin [\![{*}R]\!]_+$. That is, for every man there is a woman he didn't dance with and for every woman there's a man she didn't dance with and either there's a man who danced with no woman or there's a woman who danced with no man.

(498) The men danced with the women

Schwarzschild correctly notes that this predicts (519) to be a false presupposition failure if Daniel is related to Nicole and not Kathleen.

(519) Daniel is related to Nicole and Kathleen.

The reason is that $[\![\text{related to}]\!]_+(\{D\}, \{N,K\}) = 0$ and $[\![\text{related to}]\!]_-(\{D\},\{N,K\}) = 0$. This is a welcome result. However, (520) is also predicted to be a presupposition failure if Daniel is related to Nicole (and not Kathleen) and Martin is related to Kathleen (and not Nicole), since there is no one related to each of the others.

(520) Daniel and Martin are related to Nicole and Kathleen.

On the other hand, if Daniel is in addition related to Kathleen, this sentence is no longer a presupposition failure, and is simply true. This predicted distinction has no basis in intuition, as far as I can tell.

I leave the integration of the Excluded Middle with semidistributivity as an open puzzle.

5.1.1 A New Analysis for Neg-Raising Predicates

In chapter one we saw arguments in favor of an analysis of Neg-Raising predicates in terms of an Excluded Middle presupposition. In that chapter, we assumed that the excluded middle presupposition was a lexical property of certain predicates. In this chapter and the last, we have proposed a single semantic structure underlying several other constructions that carry an excluded middle presupposition, namely definite plurality. Under this analysis, quantificational force and the excluded middle presupposition are factored out as the contribution of a distributivity operator. In this section I briefly suggest treating Neg-Raising predicates in the same manner.

Consider how this would work for the NRP *think*. Instead of treating this predicate as a generalized quantifier over worlds, we treat it as denoting a plurality of worlds.

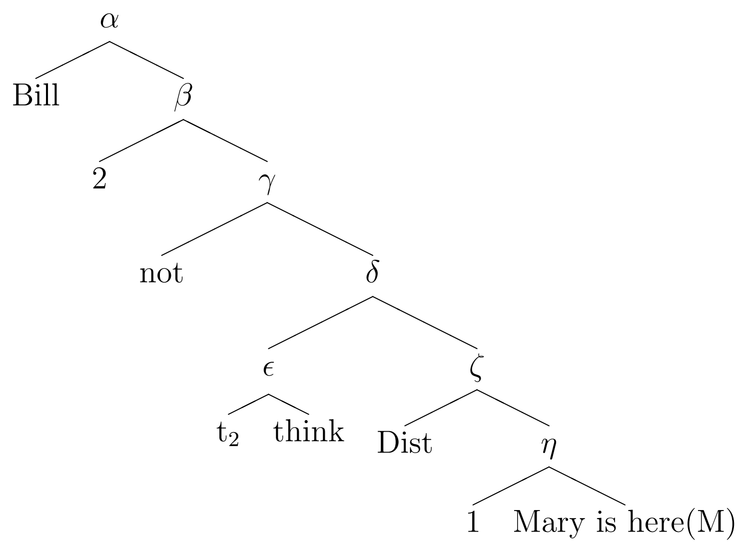
(521) Standard Analysis:
$[\![\text{think}]\!]^w = \lambda p_{st}.\lambda x.\forall w'[w' \text{ is compatible with x's beliefs in w} \rightarrow p(w')=1]$

(522) New Proposal:
$[\![\text{think}]\!]^w = \lambda x.\{w': w' \text{ is compatible with x's beliefs in w}\}$

Given (522), the structure for a sentence with an NRP is as in (523).

(523) a. Bill doesn't think Mary is here.

b.



(524) i. ⟦Dist⟧ = $\lambda f_{st}.\lambda F{:}F{\subseteq}D_s$: $\forall x[x{\in}F \rightarrow f(x){=}1] \vee \forall x[x{\in}F \rightarrow f(x){=}0]$. $\forall x[x \in F \rightarrow f(x){=}1]$

ii. $[\![\ \epsilon\ ]\!]^{w,g}[2{\rightarrow}\text{Bill}] = \{$w': w' is compatible with Bill's beliefs in w$\}$

iii. $[\![\ \zeta\ ]\!]^{w} = \lambda F{:}F{\subseteq}D_s\ \&\ \forall x[x \in F \rightarrow [\![\ \eta]\!]\ (x){=}1] \vee \forall x[x{\in}F \rightarrow [\![\ \eta]\!]\ (x){=}0]$. $\forall x[x{\in}F \rightarrow [\![\ \eta]\!]\ (x){=}1]$

iv. $[\![\ \delta\ ]\!]$ w is defined only if
$\forall x[x{\in}\{$w': w' is compatible with Bill's beliefs in w$\} \rightarrow [\![\ \eta]\!]\ (x){=}1] \vee$
$\forall x[x{\in}\{$w': w' is compatible with Bill's beliefs in w$\} \rightarrow [\![\ \eta]\!]\ (x){=}0]$
(i.e., if Bill thinks Mary is here or he thinks she is not)

v. When defined $[\![\ \delta\ ]\!]$ w is true iff
$\forall x[x{\in}\{$w': w' is compatible with Bill's beliefs in w$\} \rightarrow [\![\ \eta]\!]\ (x){=}1]$
(i.e., Bill thinks Mary is here)

(525) a. Bill doesn't believe Mary is here.

b. Assertion: $\forall$w[w is compatible with Bill's beliefs $\rightarrow$ Mary is here in w]

c. Presupp.: $\forall$w[w is compatible with Bill's beliefs $\rightarrow$ Mary is here in w] $\vee$
$\forall$w[w is compatible with Bill's beliefs $\rightarrow$ Mary is here in w]

Again, together the assertion and presupposition of (525) entail (526).

(526) Bill believes that Mary is not here.

Thus we have demostrated how Neg-Raising could be brought under this generalization. This view however, comes with a number of further commitments. For example, since the embedded clause takes the verb and its subject as an argument we are committed to the view that the subject of Neg-Raising predicates is introduced lower than the 'complement' clause. This makes a Neg-Raising predicates a kind of unaccusative. Do Neg-Raising predicates satisfy diagnostics for unaccusativity? The answer appears to be no. For example, *ne*-cliticization from the post verbal subject of a Neg-Raising verb in Italian is not possible.

(527) *ce ne stanno pensando tre
so of-them are3.pl thinking three
Three of them think so

Diagnostics for Definite Plurality The question also arises as to whether Neg-Raising predicates show any of the signs of definite plurality used to diagnose this quality in other constructions. Non-monotonicity is difficult to assess, since the restriction of the definite is implicit in the case of Neg-Raising predicates and cannot be directly manipulated. To my knowledge, there is no morphosyntactic evidence cross-linguistically in favor of treating Neg-Raising predicates as definite plurals. Neither, as far as I know, are there non-distributive propositions that would take the definite description of worlds as an argument as a whole. Finally, even though a Neg-Raising predicate would be a co-argument of the embedded predicate along with the predicates other arguments, it does not give rise to semidistributive readings in combination with these arguments. Recall that conditionals, which under Schlenker's analysis also denote definite plurals of worlds, do not exhibit either of these two latter properties either.

Mid-Scalar Generalization revisited

This analysis does, however, give us a new perspective on the Mid-Scalar Generalization. For example, we can see automatically that if a predicate denoted a definite plural of worlds that it would not sit at or below the midpoint of a scale. Why? Recall that Horn defined 'midpoint or below' in terms of the property of Tolerance. Distributive definite plurals never give rise to Tolerant predications. That is, (528) and (529) are incompatible.

(528) The girls are blond

(529) The girls are not blond

Furthermore, it makes sense under this theory that Neg-Raising predicates are not the strongest predicates of their kind. Why? It has often been pointed out that definite plurals are weaker than universal quantifiers. Link 1983 observes that definite plurals admit exceptions. For example, (530) can be true even if some of the relevant girls do not jump in the lake.

(530) The girls jumped in the lake.

(531) Every girl jumped in the lake.

The same is not true of (531) which requires that all relevant girls have jumped in the water.

Consequently, any sentence-embedding predicate that does not denote a definite plural of worlds but a standard universal quantifier over worlds will be stronger than the corresponding Neg-Raising predicate assuming they are based on the same accessibility relation. Possible examples of this are the pairs *believe* and *be certain*, or *should* and *must*.

(532) a. $[\![\text{x believes}]\!]^{u} = \{\text{w: w is compatible with x's beliefs in u}\}$

b. $[\![\text{is certain}]\!] = \lambda\text{p}.\lambda\text{x}.\forall\text{w}[\text{ w is compatible with x's beliefs in u} \rightarrow \text{p(w)=1}$

For an analysis of definites' tolerance of exceptions compatible with the present account see Brisson (2003).

5.2 Plural Predication and the Excluded Middle

5.2.1 Motivating question

In this section, I speculate on why distributive operators carry excluded middle presuppositions. There is no apparent reason for a universal quantifier to carry such a presupposition; nominal and adverbial universal quantifiers do not. I speculate that the excluded middle presupposition is the grammaticalized by-product of a pragmatic repair strategy. The pragmatic repair strategy I have in mind is a strategy for judging semantically undefined sentences either true or false. This strategy, I hypothesize, underlies the semantics of the distributive operator, which can be seen as mediating the application of a predicate of individuals to an object that is not in its domain, a plurality. In other words, distributivity is a repair of a presupposition failure in predication.

5.2.2 Repairing Presupposition Failure

This section is written in the spirit of von Fintel (2004) and Yablo (2004), who have recently pioneered research into non-catastrophic presupposition failure, to borrow Yablo's term. The position I take on the problem of presupposition failure is pithily stated by Yablo, "The claim will be that there is no such problem more like an opportunity of which natural language takes extensive advantage." In other words, speakers can use semantic undefinedness to a variety of communicative purposes, an undefined sentence is not a dead end. The thesis underlying von Fintel's and Yablo's work is summarized in (533).

(533) a. There are semantic 1's and 0's and there are the pragmatic categories TRUE and FALSE these two sets are not in direct correspondence.

b. Sentences that receive neither a 1 nor a 0 may be FALSE (von Fintel) or even TRUE (Yablo).

I will make use of this assumption in my analysis of distributive plural predication.

An example of Yablo's 2004 will prove instructive for our purposes:

> If in Oxford, I declared, "The Waynflete Professor of Logic is older than I am" it would be natural to describe the situation by saying that I had confused the titles of two professors [the Waynflete Professor of Metaphysics and Wykeham Professor of Logic], but whichever one I meant, what I said about him was true (Strawson 1954, 227)

> This becomes radical failure of the uniqueness presupposition if we suppose that Strawson in confusing the titles had confused the individuals too, so that his remark was no more directed at the one than the other. Does the failure remain non-catastrophic? I think it does. Strawson's remark seems factually incorrect if the Waynflete and Wykeham Professors are both younger than him and correct (or anyway correct-er) if both are older. (Yablo (2004), 7)

What Yablo notes here is a kind of excluded middle in the evaluation of the presupposition failure (534).

(534) The Waynflete Professor of logic older than I am.

This sentence is judged true if on every way of understanding the sentence it is true and judged false if on every way of understanding it, it is false. We can add to this description of its correctness that the sentence cannot be judged true or false if it is true on some understandings and false on others. (Yablo himself would not agree, see Yablo (2004) fn. 11)

Another important aspect of this example is just what is wrong with it. It fails because it intends to refer to a unique individual but instead picks out two[3,4]. Despite

[3] Another way of looking at it might suggest that it picks out none, making it a failure of the existence presupposition. To correct for this Yablo also gives the example below,

(i) The Professor of Philosophy at St Andrews is older than me

which he claims is true even if there are two such professors so long as both are taller than me and false if neither is.

[4] An example from literature in which a failure of uniqueness presupposition is arguably interpreted universally:

> "You see?" said Norrell, grimly. "The spell will not allow us to move too far from one another. It has gripped me too. I dare say there was some regrettable impreciseness in the fairy's magic. He has been careless. I dare say he named you as the English magician - or some such vague term. Consequently, his spell - meant only for you - now entraps any English magician who stumbles into it!"

Susanna Clarke, 2004, *Jonathan Strange and Mr. Norrell* this failure, the sentence is judged true (false) by narrowing our view to each one of the two candidate referents and noting that the sentence is true (false) in each of these cases.

5.2.3 Goal

I now make use of this kind of reasoning to give an analysis of distributive plural predication as a repair of a non-catastrophic presupposition failure. I hypothesize that at some point natural language lacked a distributive operator. On the one hand, language possessed the ability to refer to pluralities, a capability needed for collective predication. On the other hand, there existed predicates that, because of their meanings, applied only to atomic individuals. Application of predicates of this sort to pluralities resulted in semantic undefinedness. In other words, there was a sortal mismatch between a predicate of individuals and a plurality-denoting argument.

This is where, I propose, the pragmatic repair strategy explored by von Fintel and Yablo kicks in. Before showing how this repair strategy works, I make my assumptions about the state of the pre-distributivity operator language clear.

(535) Assumptions:

1. There are predicates that apply to pluralities and predicates that do not.
2. There is no distributivity operator for VPs.
3. Assume a version of Sauerland's theory of number for DPs:
    (a) 'the' maps a predicate to its characteristic set
    (b) SG is presuppositional, checks to see if DP denotes an singleton
    (c) PL is vacuous, competes with SG via Maximize Presuppositions

(536) $[\![\text{the}]\!] = \lambda f_{<e,t>}.\{x\text{: } f(x) = 1\ \}$

(537) $[\![\text{SG}]\!] = \lambda x \in D_e.\ x$

(538) Maximize Presupposition Heim (1991)

Make your contribution presuppose as much as possible

Corollary: When there is a choice between SG and PL, use SG

Refer to the domain of atomic individuals as D_e, the domain of pluralities $D_{\{e\}}$ = Pow(D_e) - $\emptyset$. We might alternatively take the set of minimal elements in $D_{\{e\}}$ as our set of atoms AT.

So for example,

(539) $[\![$ blond $]\!] = \lambda x \in D_e$. x is blond

(540) $[\![$ gather $]\!] = \lambda x \in D_{\{e\}}$. x gathered

(541) a. the boy = [SG [the [boy]]] = $[\![$ boy $]\!]$ (when defined, i.e., when $[\![$ boy $]\!]$ is a singleton)

b. the boys = [PL [the [boy]]] = $[\![$ boy $]\!]$ (infelicitous if $[\![$ boy $]\!]$ is a singleton, Max Presupp. says should have used singular)

Yablo's case: Let's now take these assumptions and show how we would account for the intuitions expressed by Yablo. First, how do we evaluate (542) if there are two boys?

(542) The boy is blond

Let's begin by first going through the case von Fintel (2004) uses to motivate his algorithm:

(543) The King of France is sitting in that chair.

Why do we say this false, knowing there is no King of France (i.e., that this is a presupposition failure)? According to von Fintel, we may judge this sentence flase because we can reason like this:

(544) Say I'm mistaken and there is a King of France, he still isn't sitting in that chair.

In other words, we rehabilitate the presupposition (assuming its truth) and evaluate the sentence based on known properties of other entities, in this case the chair. Merely assuming the existence of a king of France does not affect the properties of the chair I see before me; it is still empty, the sentence is false.

(545) and (546) show how von Fintel formalizes this strategy for rejecting a sentence. (545) says that we may judge a sentence false if once we have revised our beliefs to accommodate the truth of the presupposition π of ϕ, the semantic value of the sentence is 0. (546) gives some details as to how one goes about revising his/her beliefs in the face of evidence contradicting their prior beliefs.

(545) Rejection

Reject a sentence ϕ as FALSE with respect to a body of information D iff for all worlds w compatible with rev_π(D): $[\![\phi]\!]$(w) = 0.

(546) Conversational revision [instructions for forming rev_π(D)]

Remove $\neg\pi$ from D.

Remove any proposition from D that is incompatible with π.

Remove any proposition from D that was in D just because $\neg\pi$ was in D, unless it could be shown to be true by examining the intrinsic properties of contextually salient entity without at the same time showing that π is false.

Add π to D.

Close under logical consequence.

Von Fintel does not relativize acceptance in this way, but Yablo does. So, following Yablo, one might extend von Fintel's analysis symmetrically:

(547) <u>Fintel-Acceptance</u>

Accept a sentence ϕ as TRUE with respect to a body of information D iff for all worlds w compatible with rev_π(D): $[\![\phi]\!]$(w) = 1.

Let's consider, informally first, how such a story might be extended to the failure of a uniqueness presupposition like (542). Here's one idea about we might reason that (542) is true:

(548) Say I'm mistaken and there's really only one boy. I don't know exactly who that would be, but it must be one of the people I think is a boy and they are all blond, so it's true.

Let's use Fintel-Acceptance to look at the Yablo case (542).

(549) $[\![$ [SG [the boy]] is blond $]\!]$

(550) $\pi = [\![$ the boy $]\!]$ is a singleton

To apply Fintel-acceptance, we must remove from our body of information the proposition that $[\![$boy$]\!]$ is not a singleton and anything incompatible with $[\![$boy$]\!]$ being a singleton. We then add π; so our revised body of information must entail there is one boy. But who will this boy be? Let's assume that the body of information does not decide.

Now revision is intended to yield a body of information that is maximally similar to the original but entails π. It seems plausible to assume the following (Assume I know who the boys are, i.e., there is some set of individuals **boy** s.t. $\forall$w$\in$D[$[\![$the boy$]\!]$(w) = **boy**]):

(551) a. $\forall$w$\in$rev$_\pi$(D), $[\![$the boy$]\!]$(w) $\subseteq$ **boy**
(No matter how we choose our one boy, we pick someone we believe to be a boy)
b. $\exists$X [$\forall$w$\in$rev$_\pi$(D)[$[\![$blond$]\!]$(w) = X] & $\forall$w$\in$D[$[\![$blond$]\!]$(w) = X]]
(Changing our mind about who is a boy doesn't make us change our mind about who is blond)

(552) The assumption that the body of information is unhelpful in picking out a particular boy:
$\forall$x [x$\in$ **boy** $\rightarrow$ $\exists$w$\in$rev$_\pi$(D)[$[\![$the boy$]\!]$(w) = x]]

So now we accept this sentence as true iff for all worlds compatible with rev(D) $[\![$blond$]\!]$(w)($[\![$the boy$]\!]$(w)) = 1. Once again, this will be the case iff every actual boy is actually blond.

Similar reasoning shows that (4) is FALSE iff no actual boy is actually blond.

5.2.4 Plurals

Now let us turn to the case that interests us most: a definite plural argument and predicate of individuals.

(553) The boys are blond

(554) [PL [the boys]] are blond

According to the principle Maximize Presuppositions, $[\![$ the boys $]\!]$ must not denote a singleton. But given that it denotes a plurality, it is not in the domain of $[\![$blond$]\!]$, which is limited to singletons/individuals. Can we apply von Fintel's algorithm to this case? Yes, but Sauerland's theory of number is crucial. The feature PL is vacuous, it does not carry a lexical presupposition requiring its argument to be a plurality. Consequently, it is possible for there to be worlds in which [PL [the boys]] denotes a singleton. This means that in the process of belief revision, we can move to worlds in which [PL [the boys]] denotes a singleton and not face semantic undefinedness. Rather, the sentence (553) can be mapped to the value 1 in such worlds so long as the individual in the singleton is blond in that world.

In other words, the fact that PL enforces plurality through a pragmatic principle that applies to assertions and not a semantic presupposition gives us just the space we need to use von Fintel's belief revision algorithm to judge such sortal mismatches true or false.

The reasoning that we go through in this case is substantially similar to the reasoning we go through in the case of the failure of a uniqueness presupposition:

(555) $[\![$ the boys are blond$]\!]$ = $[\![$blond$]\!]$($[\![$the boys$]\!]$)

(556) $[\![$blond$]\!]$($[\![$the boys$]\!]$) is defined only if $[\![$the boys$]\!] \in$ dom($[\![$blond$]\!]$) that is, only if $[\![$ the boys$]\!]$ is a singleton
(Assume I know who the boys are, i.e., there is some set of individuals boys s.t. $\forall$w$\in$D[$[\![$the boys$]\!]$(w) = **boys**])

(557) a. $\forall$w$\in$rev$_\pi$(D), $[\![$the boys$]\!]$(w) $\subseteq$ **boys**
(No matter how we choose our one boy, we pick someone we believe to be a

boy)

b. $\exists X\ [\forall w \in rev_\pi(D)[\ [\![blond]\!](w) = X]\ \&\ \forall w \in D[\ [\![blond]\!](w) = X]\ ]$

(Changing our mind about who is a boy doesn't make us change our mind about who is blond)

(558) The assumption that the body of information is unhelpful in picking out a particular boy:

$\forall x\ [\ x \in \textbf{boys} \rightarrow \exists w \in rev_\pi(D)[\ [\![the\ boys]\!](w) = \{x\}]\ ]$

So now we accept this sentence as true iff for all worlds compatible with rev(D) $[\![blond]\!](w)([\![the\ boys]\!](w)) = 1$. Once again, this will be the case iff every actual boy is actually blond.

Similar reasoning shows that (4) is FALSE iff no actual boy is actually blond.

5.2.5 Singular vs. Plural

Now, unfortunately, we have derived that (542) and (553) should be judged true in the same situations. This is not a desirable result. We must answer the question of why singular definites do not have the same meaning as definite plurals.

What von Fintel 2004 describes is an algorithm that hearer can use to ascribe truth or falsity to a sentence that is semantically undefined. I propose that in the case of the definite plural, the speaker exploits the algorithm intending for the hearer to use the algorithm to evaluate the truth of the sentence. I argue that in the case of the singular this is not possible.

On what grounds might we differentiate (542) and (553) in a situation in which there is more than one boy.

(559) SG the boy is blond

(560) PL the boys are blond

The sentences carry exactly the same presupposition, namely (561).

(561) $[\![$ the boy(s) $]\!]$ is a singleton

However, in (542), this presupposition has two sources: the feature SG and the sortal restriction of *blond*. We should pin the difference in meaning between these sentences on this difference. We might suppose that a speaker who uses (542) knowing that there is more than one boy has done something irreparably uncooperative. How could we formulate this? In the case of the feature SG there is a salient alternative expression which would not impose the offending presupposition, namely PL. So, a cooperative speaker should use PL if she/he knoes there is more than one boy. There is no such alternative for the predicate *blond*. Recall we are considering a hypothetical stage of the language in which there is no distributivity operator. In some sense, then, the presupposition failure in (553) is irreducible – we are doing the best we can. In (542), this is not the case, if the speaker knows there is more than one boy she/he could do better – namely, use the feature PL. Of course, in this case, this choice does not change the presuppositions of the sentence as a whole, since the predicate introduces the same presupposition as the feature singular. So, we must assume that this principle of cooperation applies down to the level of every constituent of an asserted sentence.

(562) A speaker S asserting sentence α containing constituent [ϕ χ] (whose presuppositions are inherited by α) is being <u>fatally uncooperative</u> if S does not believe the presuppositions of [ϕ χ] and there is an expression ψ in competition with ϕ such that S believes the presuppositions of [ψ χ].

Naturally, I suppose SG and PL to be in competition in the relevant sense. The crucial assumption now is that there is no expression in such competition with *blond*. This allows a speaker to use sentence (553) while believing there is more than one boy, without being fatally uncooperative.

So, a speaker can use (542) in a scenario in which there is more than one boy and say something that is judged to be true by the hearer, so long as he believes there is only one boy.

(542) The boy is blond

In other words, even if the hearer judges this sentence true on the basis of the fact that every boy in the situation is blond, the speaker cannot meant to have convey that every boy is blond. He must believe there is only one boy.

This restriction does not hold of the sentence with the plural subject. In this case, the speaker does violate the principle that a speaker should believe the presuppositions of his/her assertion, but he/she avoids being fatally uncooperative in the sense of (562).

(563) The boys are blond

This leaves open the possibility that the speaker might use (563) to mean that every boy is blond. This is possible since the speaker can use (563) believing there is more than one boy and intending the hearer to use von Fintel's algorithm for assigning the otherwise undefined sentence a truth value.

Appendix A

Moving Exceptions

A.1 Introduction

It is commonly held that a connected exceptive phrase (EP) and its associated quantifier, such as *but John* and *no one* in (564), form a constituent at LF that denotes a generalized quantifier. Von Fintel 1993 proposes that EPs are Det-modifiers (i.e., of type $<< et, ett >, < et, ett >>$); Moltmann 1995 that they are DP-modifiers (i.e., of type ¡ett, ett¿); Keenan & Stavi 1986 that Det + EP is a discontinuous determiner. In this paper, I present arguments against such "constituent" analyses. Instead I propose a movement analysis in which EPs are base-generated as sisters to NP, but denote generalized quantifiers over properties (i.e., are of type $<< et, t >, t >$). The type mismatch between the EP and its sister NP (of type $< e, t >$) forces the EP to raise out of the DP in which it originates and take scope at a node of type t. I show that this articulation of the structure of exceptive constructions solves a puzzle in their semantics.

(564) No one but John ate the herring.

The Puzzle: NPI *any*

A major goal in the analysis of exceptive constructions has been to provide a unified account of their truthconditions. Von Fintel 1993 took a major step in this direction by proposing a semantics for exceptive constructions which gives the correct truth-conditions compositionally for EPs with both positive and negative universal quantifiers. His account, however, predicts that trivial truthconditions, and therefore anomaly, result whenever an EP co-occurs with a non-universal quantifier. This is generally a welcome prediction (see (565)); but it is problematic in the case of NPI *any*.

(565) All/None/*Many/*Three/*Some/*Most of the students but Bill came

The thesis that NPI any denotes an existential determiner is well-supported (see Ladusaw 1979, Kadmon & Landman 1993 a.o.). Yet NPI *any* and EPs do co-occur felicitously, see (566). The puzzle of their compatibility is best illustrated by examples in which an any is associated with an EP in an environment that has been used to argue for the existential (and against a universal) analysis of any (cf. Ladusaw 1979), such as in there-insertion sentences (566c) and in the scope of non-anti-additive NPI licensers such as few (566d).

(566) a. Bill didn't see anyone but Mary

b. No one saw anyone but Mary

c. There isn't anyone in the room but Bill.

d. Few boys talked to any girl but Sue.

The puzzle then is how can we maintain von Fintel's successful account and still assign the correct truthconditions to sentences in which an EP is associated with NPI any?

In section 2, I review von Fintel's semantics for exceptive constructions and show precisely why it runs into this problems. In section 3, I show how a simple amendment to von Fintel's account that is compatible with the movement of EPs can solve the puzzle.

A.1.1 Von Fintel 1993

Truthconditions

The basic truthconditional facts that von Fintel 1993 accounts for are the following:

(567) every student but John complained is true iff

- i. John is a student, and
- ii. John did not complain, and
- iii. every student who is not John complained

(568) no student but John complained is true iff

- i. John is a student, and
- ii. John complained, and
- iii. no student who is not John complained

Von Fintel treats each of the implications listed in (567) and (568) as a truthconditional entailments, since none of them is cancellable in the manner of an implicature or a presupposition.

Before looking at von Fintel's analysis, it will be useful to introduce some terminology. Let's refer to a determiner D together with its restrictor A and scope P as a quantification. Given a quantification Q (=D(A)(P)), let's call a set C an exception set relative to Q if P belongs to D(A-C).

The first attempt that von Fintel makes at spelling out the truthconditions for exceptive constructions is (569). These truthconditions encode the basic idea that the complement of but denotes an exception set relative to the quantification in which the exceptive occurs, as defined above. An example is given in (570).

(569) $[\underset{\uparrow}{\mathrm{D}}\ \underset{\uparrow}{\mathrm{A}}\ \underset{\uparrow}{[[\mathrm{but}]]}\ \ \underset{\uparrow}{\mathrm{C}}]\ \ \underset{\uparrow}{\mathrm{P}} = \mathrm{True} \Leftrightarrow \mathrm{P} \in \mathrm{D(A–C)}$

(570) no boy but John ran is True iff $\{x{:}x \text{ ran}\} \in [\![\text{no}]\!](\{x{:}x \text{ is a boy}\}–\{\text{John}\})$

Von Fintel points out, however, that (569) does not assign exceptive constructions strong enough truthconditions. Specifically, it does not account for the implications (a-b) in (567) and (568). Because of this weakness in truthconditions, (569) predicts counterintuitive entailments such as (571):

(571) no student but John complained ?⇒?

no student but John and Mary complained

For this reason, von Fintel rejects (569) as inadequate. He proposes, however, to keep domain subtraction as part of the meaning but to add another clause that will account for the additional implications. What he adds is the statement that the complement of but denotes the least exception set relative to it associated quantification. That is, the set denoted by the complement of but is a subset of every exception set of the quantification. This is formalized in the following way:

(572) $$[\text{D A } [\![\text{but}]\!] \text{ C}]\text{ P} = \text{True} \Leftrightarrow \underbrace{\text{P} \in \text{D(A–C)}}_{\text{Domain Subtraction}} \;\&\; \underbrace{\forall \text{S(P} \in \text{D(A–S)} \rightarrow \text{C} \subseteq \text{S)}}_{\text{Unique Minimality}}$$

When this schema is applied to the determiners every and no we get (573) and (574) as special cases of (572).

(573) $$[\![\text{every}]\!]\text{A}[\![\text{but}]\!]\text{C P=True} \Leftrightarrow \text{P} \in [\![\text{every}]\!]\text{(A–C) \& } \forall\text{S(P} \in [\![\text{every}]\!]\text{(A–S)} \rightarrow \text{C} \subseteq \text{S)}$$
$$\Leftrightarrow \text{A} \cap \text{P}' \subseteq \text{C \& } \forall \text{S(A} \cap \text{P}' \subseteq \text{S} \rightarrow \text{C} \subseteq \text{S)}$$
$$\Leftrightarrow \text{A} \cap \text{P}' \subseteq \text{C \& C} \subseteq \text{A} \cap \text{P}'$$
$$\Leftrightarrow \text{A} \cap \text{P}' = \text{C}$$

(574) $$[\![\text{no}]\!] \text{ A } [\![\text{but}]\!] \text{ C P = True} \Leftrightarrow \text{P} \in [\![\text{no}]\!]\text{(A–C) \& } \forall \text{S(P} \in [\![\text{no}]\!]\text{(A–S)} \rightarrow \text{C} \subseteq \text{S)}$$
$$\Leftrightarrow \text{P} \cap \text{A} \subseteq \text{C \& } \forall\text{S(P} \cap \text{A} \subseteq \text{S} \rightarrow \text{C} \subseteq \text{S)}$$
$$\Leftrightarrow \text{P} \cap \text{A} \subseteq \text{C \& C} \subseteq \text{P} \cap \text{A}$$
$$\Leftrightarrow \text{A} \cap \text{P} = \text{C}$$

These truthconditions straighforwardly yield the implications listed in (567) and (568), and bar entailments like (571).

Ungrammaticality via "immediate falsity"

Let us turn now to how the truthconditions in (572) can account for the co-occurrence restrictions on EPs. Von Fintel's idea is that the "immediate falsity" at which one arrives by combining EP with the wrong kind of quantifier results in ungrammaticality. The truthconditions in (572) require there to be a smallest set such that its subtraction from the restrictor of D makes the quantification true. On the one hand, universal quantifiers like every and no guarantee the existence of such least exception sets. Left upward monotone (↑mon) quantifiers, on the other hand, never have minimal exception sets. That is, whenever a left upward monotone determiner is substituted for D in the schema of (572), the right hand side of the equivalence is false, no matter what A,C and P are. Given that D is left upward monotone and that $P \in D(A \cap C)$, we know that for any $S \subseteq C$, e.g. the empty set, $P \in D(A \cap S)$ since $A \cap C \subseteq A \cap S$. For this entire class of quantifiers, then, it is predicted that combination with an EP yields trivial falsity. And from triviality in truthconditions we make the leap to ungrammaticality. This idea has a precedent in Barwise & Cooper's 1981 analysis of there existential constructions.

Aside from some difficulties with proportional quantifiers (see Moltmann 1995, Lappin 1996), this analysis is successful in picking out the positive and negative universal quantifiers as a class – the class of quantifiers that have least exception sets. In the next section, I outline the two challenges to this approach.

Problems for von Fintel 1993

Compatibility with NPI any

NPI *any* is an existential determiner In the literature, there are two main positions on the semantics of NPI any. Under one theory, that of Quine 1960 and Lasnik 1975, NPI *any* denotes a universal determiner that must take wide scope over negation. Under the other theory, that of Ladusaw 1979, Linebarger 1980 and Carlson 1981, NPI *any* denotes an existential determiner that must occur in the scope of a downward entailing operator. Advocates of the latter position have given a number of arguments that show NPI *any* must to receive a narrow scope existential reading in some contexts. One such context is in the associate position of a there insertion sentence.

(575) There aren't any students in the room.

(576) a. *There aren't all the students in the room

b. *There isn't every student in the room.

(577) There must be someone in John's house. (Heim 1987)

(575) illustrates the acceptability of a DP headed by any as the associate of there. (576) shows that universal quantifiers are generally unacceptable as there-associates. (577) makes the further point that there-associates are generally restricted to narrowest scope: (577) does not have the reading that there is some particular person who is required to be in John's house. If any were always required to take wide scope over negation we would expect (575) to be ungrammatical since any would be limited to narrowest scope.

Another argument that any is existential is based on the readings of sentences in which NPI any occurs in the scope of non-anti-additive DE quantifiers like few.

(578) An operator f is anti-additive iff $f(A \cup B) = f(A) \cap f(B)$

(579) *few* NP is not anti-additive since it is possible for (i) to be true and (ii) false.
(i) few students smoke and few students drink
(ii) few students smoke or drink

Consider (580). This sentence has two readings. On one reading, it says that few students are such that they read every book on the list. But on another reading that is at least marginally available, it says that every book on the list is such that it was read by few students.

(580) Few students read every book on the list.

The second reading corresponds to the LF in which the universal object takes scope over the subject. However, if we turn now to (581) in which we replace the universal with NPI any we see that such a reading is unavailable.

(581) Few students read any book on the list.

Consider a scenario in which there are twenty students in the class and twenty books on the reading list and each student read exactly one book on the list and no two students read the same book. If (581) had a reading corresponding to the second reading of (580), we would expect it to be judged true since every book was read by few students, in this case just one. (581), however, is clearly false in this scenario, since all the students read a book from the list.

NPI any and EPs are compatible The data in (566), repeated here as (582), clearly show that EPs felicitously co-occur with quantifiers headed by NPI any. Take special note that EPs are compatible with NPI any in those environments that argue most strongly for an existential analysis of that determiner, see (582c,d).

(582)
a. Bill didn't see anyone but Mary
b. No one saw anyone but Mary
c. There isn't anyone in the room but Bill.
d. Few boys talked to any girl but Sue.

These sentences present a difficult challenge to von Fintel's theory. If any is an existential determiner, then it is $\uparrow$mon. As discussed above, under von Fintel's semantics, all $\uparrow$mon determiners yield trivial truthconditions when combined with an EP. So, von Fintel predicts incorrect (in fact trivial) truthconditions for such cases and furthermore predicts them to be ungrammatical.

A.1.2 A Movement Approach to Exceptives

In section 3.1, I present an account of the truthconditions of sentences in which an EP co-occurs with NPI any. In essence, I retain the semantics for exceptive construction proposed by von Fintel 1993. I propose to change, however, the compositional implementation of the semantics. This change allows EPs to take scope at LF outside of the quantifier in which they originate. This is shown to yield correct truthconditions for sentences involving NPI any and to provide an account of the NPI licensing properties of [no A but C].

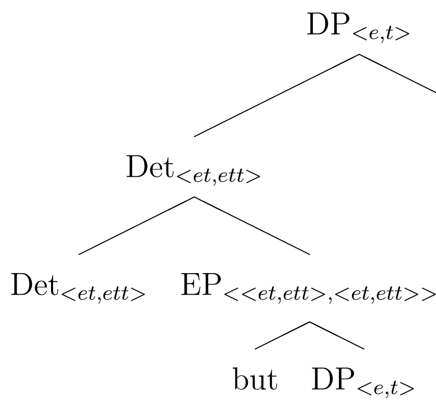
EPs Move at LF: A Revision of von Fintel 1993

Recall the problem confronting von Fintel's analysis. Von Fintel assigns the meaning schema in (572) to sentences containing exceptive phrases. He implements this compositionally by assigning but the denotation in (26b), resulting in structures of the form (37).

(572) $[\text{D A } [\![\text{but}]\!] \text{C}]\text{P=True} \Leftrightarrow \text{D}(\text{A}\cap\text{C}')(\text{P})=1 \ \& \ \forall\text{S}(\text{D}(\text{A}\cap\text{S}')(\text{P})=1 \rightarrow \text{C}\subseteq\text{S})$

(583) $[\![\text{but}]\!] = \lambda\text{C}_{et}.\lambda\text{D}_{<et,ett>}.\lambda\text{A}_{et}.\lambda\text{P}_{et}.\text{D}(\text{A}\cap\text{C}')(\text{P})=1\&\forall\text{S}(\text{D}(\text{A}\cap\text{S}')(\text{P})=1 \rightarrow \text{C}\subseteq\text{S})$

(584)



Applying (572) to an ↑mon determiner results in immediate falsity. Thus, if NPI any denotes an existenital determiner, then [any NP but DP] ought to be ungrammatical due to immediate falsity. Such a combination of NPI any and but is not ungrammatical, e.g., (585). Not only are such sentences grammatical but they also have clear entailments that a theory of exceptives should account for:

(585) John didn't see any student but Mary is true iff

i. Mary is a student

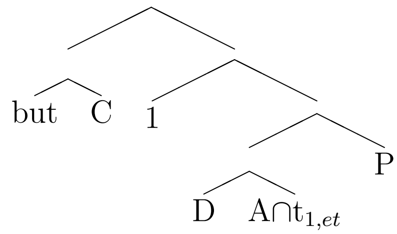
ii. John saw Mary

iii. John didn't see a student who was not Mary

I suggest that von Fintel's analysis can capture these entailments, if we give it a different compositional structure. If we abandon the structure in (584), we see that there are other ways to slice the meaning pie of (572) compositionally. Consider, for example, a structure in which the EP has moved out of its NP-modifier position, adjoined to a constituent of type t, and left behind a trace of type $< e, t >$:

(586) [[but John] 1[[no student $t_{1,et}$] smokes]]
C D A P

(587)



Such a structure provides us with all the necessary pieces to arrive at the meaning von Fintel proposes. To see how, consider again von Fintel's meaning schema:

(572) D A ⟦but⟧ C P = True $\Leftrightarrow$ $D(A \cap C')(P)=1$ & $\forall S(D(A \cap S')(P)=1 \rightarrow C \subseteq S)$

Notice that ⟦but⟧ need not take D, A, C and P individually as arguments. The only one of these that is crucially singled out in the truthconditions is C, the exception set. The others, D, A and P, are referred to in both conjuncts as a unit, namely the function $\lambda X.D(A \cap X)(P)=1$: in the first conjunct, it is stated that $D(A \cap C')(P)=1$; in the second conjunct, $D(A \cap S')(P)=1$ is used in the antecedent of a conditional, where S is a variable bound be a higher quantifier. The structure in (587) provides us with exactly these two pieces: C is simply the denotation of the complement of but, $\lambda X.D(A \cap X')(P)=1$ is the denotation of the sister of the moved EP:

(588) ⟦1 [D [A $t_{1,et}$] P] ⟧ = $\lambda X_{et}.D(A \cap X)(P)=1$

Now we are ready to give a new denotation to but that puts these pieces together.

(589) ⟦but⟧= $\lambda C_{et}.\lambda F_{<et,t>}.F(C')=1$ & $\forall S(F(S')=1 \rightarrow C \subseteq S)$

Let's now quickly check that this denotation does in fact give us back the semantics von Fintel proposes for exceptives constructions of the form in (588).

(590) a. ⟦but C⟧ = $\lambda F_{<et,t>}.F(C')=1$ & $\forall S(F(S')=1 \rightarrow C \subseteq S)$

b. ⟦but C⟧(⟦1[D [A $t_{1,et}$] P]⟧) =1 iff ⟦but C⟧($[\lambda X_{et}.D(A \cap X)(P)=1]$)=1 iff $[\lambda X_{et}.D(A \cap X)(P)=1](C')=1$ & $\forall S([\lambda X_{et}.D(A \cap X)(P)=1](S')=1 \rightarrow C \subseteq S)$ iff $D(A \cap C')(P)=1$ & $\forall S(D(A \cap S')(P)=1 \rightarrow C \subseteq S)$

We have now shown that for structures of the form (588) we derive the same truth-conditions as von Fintel's. The power of our new analysis, though, is that the EPs need not appear in a configuration of the form in (588). Since the EP is of type $<< et, t >, t >$, it cannot be interpreted in its base position as an NP-modifier, where it sister is of type $< e, t >$. So, the EP QRs to a node of type t, leaving and abstracting over a trace of type $< e, t >$. Nothing we have said so far dictates which node of type t the EP adjoins to. It might QR to a node that contains just the DP from which it moved and a predicate, as in (588), or it might QR to a higher node of type t, taking other operators into its scope.

This freedom in landing sites for a QR'd EP gives us the flexibility we need to capture the truthconditions of examples with NPI any. Let's consider (585) as a concrete example. Under von Fintel's 1993 proposed denotation for but, the EP was constrained to take the deteminer any as an argument, resulting in trivial truthconditions. We showed above that we make the same predictions as von Fintel 1993 when the EP takes scope as in (585). So, we predict that an LF is ill-formed in which the EP [but Mary] takes scope just above the existential NPI any student and below sentential negation.

(591) *LF of (585) = [not [but Mary 1 [any [student t_1] 2 [John saw t_2]]]]

Here the constituent that is sister to sentential negation is of the same form as (588) and therefore induces ungrammaticality because of trivial truthconditions. We do not yet predict, however, that (585) is ungrammatical; other LFs are available. The other scope option that is available to [but Mary] is the root node. This is a scope option that is not of the form (588): there is another operator in the scope of [but Mary], namely negation. Here then we make different predictions from von Fintel 1993.

(592) Well-formed LF for (585): [[but Mary] 1[not [[any student t_1] 2[John see t_2]]]]

The reason that this LF is well-formed is intuitively clear. The denotation of the sister of [but Mary] is equivalent to the denotation of a constituent that does have the form (588): namely one in which D = *no*.

(593) [1[not [[any [student t_1]
2[John saw t_2]]]]] = [[1[[no [student t_1]] 2[John saw t_2]]]]

Given this equivalence we can fall back on von Fintel's analysis to determine the truthconditions of (585). For the sentence with no, von Fintel makes the following predictions:

(594) John saw no student but Mary is True iff {x: x is a student}∩{x: John saw x}={Mary}

These are exactly the truthconditions we want for (585), since they account for the entailments in (585i-iii).

Beyond *not ... any*

This analysis is equally successful with cases that cannot be so easily reduced to structures for which von Fintel 1993 makes correct predictions. One such case is (595).

(595) No student read any book but War & Peace is true iff

a. W&P is a book, and

b. Some student read W&P, and

c. No student read a book that wasn't W&P

Similarly to (585), the LF of (595) in which the EP scopes below the NPI licenser is ill-formed due to its trivial truthconditions. That is, (596) is filtered out at LF.

(596) *LF of (595): [no student 1[but W&P 2 [any [book t_2] 3[t_1 see t_3]]]]

The other scope possibility for [but W&P], that is above [no student], does yield a well formed LF:

(597) Well-formed LF for (595): [but W&P 1[no student 2[any [book t_1] 3[t_2 see t_3]]]]

To show that this LF is indeed well-formed, I offer a sketch of the calculation of its truthconditions.

(598) Sketch of predicted truthconditions for (595):

Domain Subtraction (see (572)) says that taking Bill out of the set of men makes the statement 'No woman saw any man' true. Unique Minimality (see (572)) says that any way of taking individuals out of the set of men that does not include taking out Bill makes the statement false. Why would that be? It must be because some woman saw Bill.

This is an encouraging result. By making a slight change to von Fintel 1993, we have managed to assign intuitively correct truthconditions to sentences in which NPI any co-occurs with an EP.

Before turning to the solution for our second puzzle, let's calculate the truthconditions for one last example, (581).

(581) Few students read any book but Ethan Frome.

Once again, in order for this sentence to avoid trivial truthconditions, the EP must scope over the c-commanding DE operator few students. So we predict that (581) will be true iff, if we subtract EF from the set of books, its true that few students read any books and furthermore any set of books that we subtract from the books that makes it true that few students read any books will contain EF. This entails that the number of students that read EF is not few. If it were few, there would be a set not containing EF that could be subtracted from the set of books that would make it true that few students read any book, namely the complement of {EF} in the set of books. The truth conditions are expressed formally in (599), suppose few is contextually resolved to mean less than 3:

(599) $|\{$x: x is a student and $\exists$y: y is a book and y is not EF and x read y$\}| < 3$ & $\forall$S($|\{$x: x is a student and $\exists$y: y is a book and y$\in$S and x read y$\}| < 3 \rightarrow$ $\{$EF$\}\subseteq$ S)

Here's a scenario in which we predict (581) to be true:

(600) There's a class with 10 students. The reading list for the class has four books on it: Ethan Frome EF, The Age of Innocence AI, House of Mirth HM &

Glimpses of the Moon GM. Bill read the AI and EF, Sue read HM and EF and Sara read EF. No other students read any other books.

Here's a scenario in which we predict (580) to be false:

(601) Same class and reading list as in (600). Fred read AI. Sue read AI. Bill read HM. Sara read HM. Ted read GM. Mary read GM. Everyone read EF.

Indeed, (581) is false in this scenario since nearly everyone read a book that wasn't Ethan Frome. This scenario is of interest since we would predict that (581) should be true in (601) if we analyzed it with a wide scope universal, i.e., as being equivalent to (602):

(602) Every book but Ethan Frome is such that few students read it.

Bibliography

Alonso-Ovalle, Luis, and Elena Guerzoni. 2003. Double negatives, negative concord and metalinguistic negation. In *Proceedings of CLS 38*.

Bartsch, Renate. 1973. “Negative Transportation” gibt es nicht. *Linguistische Berichte* 27.

Berman, Stephen. 1991. On the semantics and logical form of wh-clauses. Doctoral Dissertation, UMass Amherst.

Bhatt, Rajesh, and Roumyana Pancheva. 2001. Syntax of conditionals. Ms. UTexas and USC.

Brisson, Christine. 2003. Plurals, all, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26:129–184.

Carlson, Gregory. 1977. Refernce to kinds in English. Doctoral Dissertation, UMass Amherst.

Chierchia, Gennaro. 1998. Refernce to kinds across languages. *Natural Language Semantics* .

Chomsky, Noam. 1995. A minimalist program for linguistic theory. In *The minimalist program*. MIT Press.

Cooper, Robin. 1983. *Quantification and syntactic theory*. Dordrecht: Reidel.

Fillmore, Charles. 1963. The position of embedding transformations in grammar. *Word* 19:208–231.

von Fintel, Kai. 1993. Exceptive constructions. *Natural Language Semantics* .

von Fintel, Kai. 1997. Bare plurals, bare conditionals and 'only'. *Journal of Semantics* 14:1–56.

von Fintel, Kai. 1999. NPI licensing, strawson entailment and context dependency. *Journal of Semantics* 16:97–148.

von Fintel, Kai. 2004. Would you believe it? the King of France is back! (presuppositions and truth-value intuitions). In *Descriptions and beyond*, ed. A. Bezuidenhout and M. Reimer. Oxford University Press.

Fodor, Janet Dean. 1970. The linguistics description of opaque contexts. Doctoral Dissertation, MIT.

Gajewski, Jon. 2004. Raising exceptions. In *Proceedings of WCCFL 23*. Cascadilla Press.

Giannakidou, Anastasia. 1997. The landscape of polarity items. Doctoral Dissertation, University of Groningen.

Grice, J. Paul. 1975. Logic and conversation. In *Syntax and semantics: Speech acts*, ed. P. Cole and J. Morgan. Academic Press.

Guerzoni, Elena. 2001. Italian n-words as NPIs. Incontro di Grammatica Generativa.

Hackl, Martin. 2000. Comparative quantifiers. Doctoral Dissertation, MIT.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertation, UMass, Amherst.

Heim, Irene. 1984. A note on negative polarity and downward entailingness. In *Proceedings of NELS 14*.

Heim, Irene. 1985. Notes on comparatives and related matters. Ms. University of Texas, Austin.

Heim, Irene. 1991. Articles and definiteness. MIT.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.

Heim, Irene. 1994. Interrogative semantics and karttunen's semantics for know. In *Proceedings of IATL 1*, ed. Buchalla and Mittwoch. Hebrew University.

Heim, Irene. 2000. Degree operators and scope. In *Proceedings of SALT X*.

Heim, Irene. 2003. On quantifiers and NPIs in comparative clauses. Ms. MIT.

Herburger, Elena. 1997. The negative concord puzzle revisited. *Natural Language Semantics* .

Heycock, Caroline. 2003. On the interaction of adjectival modifiers and relative clauses. Ms. Edinburgh.

Hoeksema, Jack. 1996. In days, weeks, months, years, ages: A class of temporal negative polarity items. Ms. Groningen.

Horn, Laurence. 1972a. Negative transportation: Unsafe at any speed? In *Proceedings of CLS 7*, 120–133.

Horn, Laurence. 1972b. On the semantic properties of logical operators in english. Doctoral Dissertation, UCLA.

Horn, Laurence. 1978. Remarks on neg-raising. In *Pragmatics*, ed. Peter Cole, volume 9 of *Syntax and Semantics*, 129–220. Academic Press: New York.

Horn, Laurence. 1984. Toward a new taxonomy for pragmatic inference: Q-based and r-based implicature. In *Meaning, form and use in context: Linguistic applications*, ed. Deborah Schiffrin, chapter Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature, 11–42. Georgetown University Press.

Horn, Laurence. 1989. *A natural history of negation*. Univeristy of Chicago Press.

Horn, Laurence. 1993. Economy and redundancy in a dualistic model of natural language. In *1993 yearbook of the linguistic association of finland*, ed. S. Shore and M. Vilkuna, 33–72. SKY.

Horn, Laurence, and Samuel Bayer. 1984. Short-circuited implicature: A negative contribution. *Linguistics and Philosophy* .

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In *Quantification in natural language*. Kluwer.

Karttunen, Lauri. 1974a. Presupposition and linguistic context. *Theoretical Linguistics* 1:181–193.

Karttunen, Lauri. 1974b. Until. In *Proceedings of CLS 10*.

Kratzer, Angelika. 1995. Stage and individual level predicates. In *The generic book*, ed. Francis Pelletier Greg N. Carlson. Univeristy of Chicago Press.

Krifka, Manfred. 1986. Nominalreferenz und zeitkonstitution. zur semantik von massentermen, pluraltermen und aspektklassen. Doctoral Dissertation, University of Munich.

Krifka, Manfred. 1996. Pragmatic strengthening in plural predications and donkey sentences. In *Proceedings of SALT IV*, ed. T. Galloway and J. Spence. CLC Publications.

Ladusaw, William. 1979. Polarity as inherent scope relations. Doctoral Dissertation, University of Texas at Austin.

Ladusaw, William. 1992. Expressing negation. In *Proceedings of SALT I*, ed. M. Harvey C. Barker, D. Dowty and L. Santelman. Cornell University Press.

Lahiri, Utpal. 2000. Lexical selection and quantificational variability in embedded interrogatives. *Linguistics and Philosophy* .

Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford University Press.

Lakoff, Robin. 1969. A syntactic argument for negative transportation. In *Proceedings of CLS 5*, 149–157.

Landau, Idan. to appear. Chain resolution in Hebrew V(P)-fronting. *Syntax* .

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* .

Larson, Richard, and Gabriel Segal. 1995. *Knowledge of meaning*. MIT Press.

Lewis, David. 1973. *Counterfactuals*. Harvard University Press.

Linebarger, Marcia. 1980. The grammar of negative polarity. Doctoral Dissertation, MIT.

Linebarger, Marcia. 1987. Negative polarity and grammaical representation. *Linguistics and Philosophy* 325–387.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In *Meaning, use and interpretation of language*, ed. R. Bauerle et. al., 302–323. Walter de Gruyter.

Löbner, Sebastian. 1985. Definites. *Journal of Semantics* 4:279–326.

Löbner, Sebastian. 2000. Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences. *Linguistics and Philosophy* 23:213–308.

May, Robert. 1989. Interpreting logical forms. *Linguistics and Philosophy* 12:387–435.

Mogan, Jerry. 1978. Two types of convention in indirect speech acts. In *Syntax and semantics: Pragmatics*, ed. P. Cole, 261–280. Academic Press.

Nathan, Lance. 1999. Either: Negative polarity meets focus sensitivity. Ms. MIT.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Number 43 in Linguistic Inquiry monographs. MIT Press.

Partee, Barbara. 1989. Many quantifiers. In *Proceedings of the 5th escol*, ed. J. Powers and K. de Jong, 383–402. Ohio State University.

Penka, Doris, and Arnim von Stechow. 2001. Negativen indefinia unter modalverben. In *Modalitat und modalverben im deutschen*, ed. Reimar Muller and Marga Reis. Sonderheft 9 edition.

Potts, Chris. 2000. When even *no*'s neg is splitsville. URL `http://ling.ucsc.edu/Jorge/potts.html`.

Prince, Ellen. 1976. The syntax and semantics of neg-raising, with evidence from french. *Language* 52:404–426.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Foris Dordrecht.

Ross, John Robert. 1973. Slifting. In *The formal analysis of natural languages: Proceedings of the first international conference*, ed. M.P. Schutzenburger M.Gross, M. Halle, 133–169. Mouton: The Hague.

Rullman, Hotze. 2003. Additive particles and polarity. *Journal of Semantics* 20:329–401.

Sadock, Jerrold. 1972. Speech act idioms. In *Cls 8*, 329–339.

Sauerland, Uli. 1998. Plurals, derived predicates and reciprocals. In *The interpretive tract*. MITWPL.

Sauerland, Uli. 2004. A new semantics of number. In *Proceedings of SALT*. CLC Publications.

Scha, Remko. 1981. Distributive, collective and cumulative quantification. In *Formal methods in the study of language*, ed. T. Jansen J. Groenendijk and M. Stokhof. Amsterdam Mathematical Center.

Schlenker, Philippe. 2004. Conditionals as definite descriptions (a referntial analysis). *Research on Language and Computation* 417–462.

Schwarzschild, Roger. 1991. On the meaning of definite plural noun phrases. Doctoral Dissertation, UMass Amherst.

Schwarzschild, Roger. 1994. Plurals, presuppositions and the sources of distributivity. *Natural Language Semantics* 2:201–248.

Schwarzschild, Roger. 1996. *Pluralities*. Kluwer: Dordrecht.

Searle, John. 1975. Indirect speech acts. In *Syntax and semantics: Speech acts*, ed. P. Cole and J. Morgan, 59–82. Academic Press.

Sharvit, Yael, and Penka Stateva. 2002. Superlative expressions, context and focus. *Linguistics and Philosophy* 25:453–504.

Stalnaker, Robert. 1981. A defense of conditional excluded middle. In *Ifs: conditionals, belief, decision, chance and time*, ed. R. Stalnaker W. Harper and G. Pearce, 87–104. Dordrecht:Reidel.

de Swart, Henriette. 1996. Meaning and use of *not ... until*. *Journal of Semantics* 13:221–263.

Uribe-Etxebarria, Miriam. 1994. Negation and the LF structural conditions on NPI licensing. In *Proceedings of WECOL XXIV*.

Winter, Yoad. 2001. Plural predication and the strongest meaning hypothesis. *Journal of Semantics* .

van der Wouden, Ton. 1995. A problem with the semantics of neg-raising.

Yablo, Steven. 2004. Non-catastrophic presupposition failure. MIT.

Yoon, Y. 1996. Total and partial predicates and weak and strong interpretations. *Natural Language Semantics* 217–236.

Zwarts, Frans. 1996a. Facets of negation. In *Quantifiers, logic and language*, ed. J. van der Does and J. van Eijck, 385–421. CSLI Publications: Stanford.

Zwarts, Frans. 1996b. Three types of polarity. In *Plural quantification*, ed. F. Hamm and E. Hinrichs. Kluwer: Dordrecht.