‘Only’ and Association with Negative Antonyms

by

Sam Al Khatib

B.Sc. Computer Science, University of British Columbia, 2001
M.A. Linguistics, Simon Fraser University, 2008

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 2013

Author: ....
Department of Linguistics and Philosophy
August 20th, 2013

Certified by: ....
Irene Heim
Professor of Linguistics
Thesis Supervisor

Certified by: ....
Martin Hackl
Associate Professor of Linguistics
Thesis Supervisor

Accepted by: ....
David Pesetsky
Head of the Department of Linguistics and Philosophy

Abstract

A problem is detected in how *only*, traditionally viewed, associates with Negative Quantifiers (NQs) like *few*, [*at most n*], and [*less than n*]. The predicted meanings, which negate the stronger alternatives of the relevant NQs, is shown to be incorrect. The attested meanings are shown to be truth-conditionally equivalent to the result of removing *only*, hence giving the illusion that *only* is vacuous. It is further observed that, given an NQ, sentences where *only* appears to be vacuous become ungrammatical when the NQ is replaced with its positive counterpart. I argue for a view that relates the two phenomena: the ungrammaticality of the *only*-positive cases explains the unavailability of the predicted meanings in *only*-negative cases. The attested readings for the negatives, where *only* appears to be vacuous, is derived from LFs that feature a silent existential quantifier above the NQ.

Thesis supervisor: Irene Heim
Title: Professor of Linguistics

Thesis supervisor: Martin Hackl
Title: Associate Professor of Linguistics

Acknowledgements

I am incalculably indebted to my committee members, Irene Heim, Martin Hackl, and Robert Stalnaker, for their help and support. Their contribution to what this thesis became is beyond description. To have discussed this material with them, and to have benefited from their knowledge, articulateness, and generosity over these past years, is one of the great fortunes of my life.

This thesis grew out of a short squib that I wrote (hastily and badly) for a seminar that Rick Nouwen taught at MIT in the fall of 2011. Without his help, it would have never come this far. Since then, the project benefited from discussions with Gennaro Chierchia, Kai von Fintel, Danny Fox, Natasha Ivlieva, Sasha Podobryaev, Yasutada Sudo, Guillaume Thomas, and Wataru Uegaki. I thank them all for their invaluable input.

I thank also my instructors for many inspiring and enjoyable lectures: Adam Albright, Michel DeGraff, Edward Flemming, Kai von Fintel, Danny Fox, Martin Hackl, Irene Heim, Sabine Iatridou, Michael Kenstowicz, Van McGee, Rick Nouwen, David Pesetsky, Norvin Richards, Robert Stalnaker, Donca Steriade, and Ken Wexler.

Finally, my friends in Cambridge. In them I found a family, and sad though it is to part after five unforgettable years, I am comforted knowing that, wherever we are, we will always be family.

And finally finally, my biological family. To them I dedicate this thesis, and to my grandfather.

Contents

Chapter 1

Introduction and Background

1.1 Introduction

This work concerns the behavior of the exclusive particle *only* and negative quantifiers like *few* and *at most*. I discuss what I claim to be an incorrect prediction given the standard semantics of particle, and the semantics of negative quantifiers. The general signature of the data I consider is the following: for a pair of antonyms $\langle P,N\rangle$—P the positive and N the negative—sentences of the form [*only* $[\cdots N_{\mathrm{F}}\cdots]$] do not have the predicted reading where all of N's non-weaker alternatives are negated. Instead, the detected reading is one where the excluded alternatives seem to be ordered non-logically with respect to the prejacent, in a sense that will be made clear later. The study begins with an empirical discussion, Chapter 2, which concludes with a generalization that I later attempt to capture. I focus on three groups of negative quantifiers: unmodified/intensified gradable quantifiers (Chapter 4), the negative modified numeral *at most* (Chapter 5), and comparative negative constructions (Chapter 6). I use the abbreviation NQ when I want to refer to all three.

1.2 *only*

Since Horn (1969) semanticists have generally taken the particle *only* to have two components: an assertoric and a presuppositional. The components are best illustrated if the particle is assumed to operate on entire sentences, as in the LF in (1). I follow common usage and refer to *only*'s sentential argument as its 'prejacent'.[1]

[1]The practice of treating *only* as a sentential operator is fairly common. The reason for it is to abstract away from the technical complications that arise in more compositional treatments of the particle, where the

(1) John only spoke to some$_{\mathrm{F}}$ of the students.

LF: [*only* [John spoke to some$_{\mathrm{F}}$ of the students]]

On Horn's theory, a sentence like (1) presupposes that its prejacent is true. Its other component, the assertoric, strengthens the prejacent by adding that all its *alternative* sentences are false, provided that they are not logical consequences of the prejacent itself (Rooth 1992).[2] Alternatives to a sentence S are those sentences that result from replacing the focused element in S with its formal/contextual alternatives (to be discussed later). The two components are formalized in (2).

(2) ***Only* – version 1:**

Given a set of alternatives A and a sentence S,

$[\![\text{only}_A\ S]\!]$ is defined only if $[\![S]\!] = 1$, and if defined

$[\![\text{only}_A\ S]\!]$ is true iff for all $S' \in A$, if $S \not\models S'$, then $[\![S']\!] = 0$

I will later argue for revisions to (2). Let us now see how the current definition applies to (1), where the focused element is the determiner *some*. Assume for now that *some* has the determiner *all* as its formal alternative. Then the set A in [*only*$_A$ [John spoke to some$_{\mathrm{F}}$ students]] will be:

(3) $A = \{[\text{John spoke to some of the students}], [\text{John spoke to all of the students}]\}$

Given A, the result of applying (2) to (1) is:

(4) $[\![(1)]\!]$ is defined only if $[\![\text{John spoke to some of the students}]\!] = 1$, and if defined,

$[\![(1)]\!] = 1$ iff for all $S' \in (3)$, if $S \not\models S'$, then $[\![S']\!] = 0$, i.e.

iff $[\![\text{John spoke to all of the students}]\!] = 0$

(1) is therefore defined only if John spoke to some of the students, and if defined, (1) is true iff John did not speak to all of them. For the remainder of this work I will focus on the assertoric component of the particle. Let me point out, however, that the inference to the prejacent, which is assumed in Horn (1969) to be a presupposition, has been the topic of much dialog over the past few decades. Horn (1996), for example, argues for a weaker 'existential' presupposition, where the particle is said to be defined only if at least *one* of

flexibility of its distribution is mirrored in its semantics (see Rooth 1985).

[2]Other treatments of *only*, particularly its assertoric component, include Groenendijk and Stokhof (1984), Kratzer (1991), Bonomi and Casalegno (1993), and van Rooij and Schulz (2007). To my knowledge the differences do not affect the findings of this work.

the prejacent's alternatives is true. Ippolito (2006, 2008) weakens the condition further to an 'implicational' presupposition. To her, *only* is defined provided that, if one of the prejacent's alternatives is true, then the prejacent itself is true. Note that this presupposition is met even if nothing in the set of alternatives A is true. Ippolito argues that this is desired, given the suspendability of the prejacent in cases like (5).

(5) Only John$_{\mathrm{F}}$ knows how to draw, and maybe not even he can.

Other analyses classify the prejacent as an entailment of *only* (Atlas 1993), and as a conversational implicature of it (McCawley 1981,van Rooij and Schulz 2007).

Because my concern is primarily with the particle's assertoric component, I will adopt Horn's (1969) analysis of the prejacent, and refer the interested reader to the works cited above (see also Beaver and Clark 2008 for a review). The rest of this section is dedicated to two issues: the generation of alternatives (Section 1.2.1), and the problem of disjunction in association with *only* (Section 1.2.2). In the former section I present Katzir's (2007) theory of alternatives, and in the latter I argue for a revision of the semantics of *only* that uses the notion of Innocent Exclusion.

1.2.1 The generation of alternatives

On Rooth's (1985) theory of focus, an expression qualifies as an alternative to another if the two are of the same semantic type. This characterization correctly groups together e.g. the determiners *some* and *all*, numerals, proper names, etc. However, building on observations that date back to Kroch (1972), Fox and Katzir (2011) argue that Rooth's characterization requires further restriction. Take again our example (1).

(1) John only spoke to some$_{\mathrm{F}}$ of the students.

The danger of not constraining Rooth's recipe is that for any prejacent S and any alternative $S' \in A$, there can be another alternative S'' whose meaning is equivalent to $S \wedge \neg S'$. In our example, where S is '*John spoke to some$_F$ of the students*', the availability of the *all*-alternative S' (in (6a)) comes with the availability of S'' (6b).

(6) a. S' = John spoke to all of the students.
b. S'' = John spoke to some but not all of the students.

Neither S' nor S'' is a logical consequence of (1)'s prejacent, and both are generated by replacing the focused determiner *some* with another expression of the same type.[3] If *only*

[3] The reader may recall that the generalized rules of negation and conjunction in Rooth and Partee (1982) are capable of generating the complex determiner [*some but not all*] from the conjoinable *some* and *all*.

negates (excludes) the two alternatives, we get the result in (7).

(7) $[\![(1)]\!]$ is defined only if $[\![$John spoke to some of the students$]\!] = 1$, and
if defined, $[\![(1)]\!] = 1$ iff $[\![(6a)]\!] = [\![(6b)]\!] = 0$,
i.e. iff John did not speak to all of the students ($[\![S']\!] = 0$), and
John either spoke to *none* or to *all* of the students ($[\![S'']\!] = 0$).
i.e. iff John spoke to none of the students.

The exclusions in (7) jointly contradict the prejacent. So, if the prejacent is presupposed, this will make (1) contradictory.

(7) is an example of what is sometimes called *the symmetry problem*.[4] The problem is that for any sentence S and any alternative S' that entails S, there is a sentence $S'' = S \wedge \neg S'$ that also entails S. In this case, S' and S'' are said to be *symmetric* with respect to S, where:

(8) Two sentences S_1, S_2 are symmetric with respect to S iff
a. $[\![S_1]\!] \wedge [\![S_2]\!] \vDash \bot$.
b. $[\![S_1]\!] \vee [\![S_2]\!] = [\![S]\!]$.

The key property that symmetric alternatives have is (8b). If both S_1 and S_2 are negated, the resulting proposition is equivalent, by De Morgan's law, to the negation of their disjunction: $(\neg[\![S_1]\!] \wedge \neg[\![S_2]\!]) \equiv \neg([\![S_1]\!] \vee [\![S_2]\!])$. But by (8b), the disjunction $[\![S_1]\!] \vee [\![S_2]\!]$ is equivalent to $[\![S]\!]$, which means that negating S_1 and S_2 amounts to negating S.

This has important consequences for Rooth's characterization of alternatives: we saw that having a stronger alternative S' to S gives us another stronger alternative $S'' = (S \wedge \neg S')$ (e.g. *all* and *some but not all*). This makes S' and S'' symmetric (see (9)), and so their negation (under *only*) is predicted to contradict the prejacent.

(9) $S' \wedge S'' = S' \wedge (S \wedge \neg S') = \bot$
$S' \vee S'' = S' \vee (S \wedge \neg S') = (S' \vee S) \wedge (S' \vee \neg S') = (S' \vee S) = S$

The goal of any approach to the symmetry problem is to remove one of the symmetric pair from the set of alternatives. In the case of (1), *only* is intuitively felt to negate the *all*-alternative, not the symmetric *some-but-not-all*. The question is in what sense could the former be an alternative that would not hold of the latter. Much of the literature on Scalar Implicatures, where the symmetry problem arises in the same way, bypasses it by assuming *lexically*-defined scales of alternatives. The determiner *some*, for example, is

[4] The term is attributed by Fox (2007a) and Katzir (2007) to class notes by Kai von Fintel and Irene Heim.

lexically associated with the determiner *all*, and not with *some-but-not-all*. Associations like this are known as Horn Scales (Horn 1972, 1989).

More recently, Katzir (2007), and, in the context of association with focus, Fox and Katzir (2011), proposed that alternatives to an expression S are generated by any of three operations: (i) lexical substitution, (ii) structural-simplification, and (iii) contextual replacement.[5] By (i) we generate *all* as an alternative to *some*, but we cannot generate *some-but-not-all* from either (i) or (ii).[6]

To see the advantage of (ii), consider the disjunctive associate in (10).

(10) To pass my course, you only need to [write a paper or take the final exam]$_{\mathrm{F}}$.

The addressee of (10) is not required to write a paper, nor is s/he required to take the final exam. In order to derive this meaning, the disjuncts in (11a,b) must make it into the prejacent's set of alternatives, so that the two (stronger) requirements may be negated by *only*.

(11) a. You need to [write a paper].
b. You need to [take the final exam].

A relatively liberal theory of alternatives, like Rooth's, is capable of generating (11a,b); to Rooth the substitution is legal because the (bracketed) replacements in (11) are both of the same semantic type as the focus associate in (10). But the previous discussion showed that more needs to be added to Rooth's characterization, if the symmetry problem is to be overcome. In the current case, resorting to a strictly lexical approach, e.g. Horn Scales, is not obviously an option, for what lexical connection is there between *or* and whatever disjuncts it happens to accompany? In Sauerland (2004), the connection is made through two binary operators $\boxed{\text{L}}$ and $\boxed{\text{R}}$, which are assumed to be scale-mates to *or*:

(12) $\boxed{\text{L}}, \boxed{\text{R}} \in \text{ALT}(or)$

(13) $S_1 \boxed{\text{L}} S_2 = S_1$

(14) $S_1 \boxed{\text{R}} S_2 = S_2$

With (12-14), we generate the following alternatives to (10)'s prejacent:

[5] Structural simplification results from replacing a node with one of its subconstituents. See Fox and Katzir for details.

[6] Note crucially that (iii) can generate symmetric alternatives as long as they are made salient in context. Judgements indicate that even in these cases, the negation of the contextual alternative is less clearly felt to be an SI. See Fox and Katzir, footnote 16.

(15) ALT('need to write a paper or take the exam')

$= \{$'need to write a paper $\boxed{\text{L}}$ take the exam'

'need to write a paper $\boxed{\text{R}}$ take the exam'$\}$

$= \{$'need to write a paper',

'need to take the exam'$\}$

While Sauerland's idea successfully produces the correct alternatives, the adoption of (12-14), as he notes, is "more of a technical trick than a real solution". It is here that the advantage of Katzir's condition (ii) becomes apparent: simplifying a disjunction produces its individual disjuncts as alternatives, and there is no need to refer to any arbitrary associations from the lexicon.

Appealing though it is, Katzir's theory is not without problems (see e.g. Swanson 2010 for a reply).[7] I will nonetheless continue to use it in the remainder of the study.

1.2.2 Disjunction, the assertion of *only*, and Innocent Exclusion

In this section I revise the semantics of *only* in (2), in light of challenges that are posed by disjunctive focus associates. The revision incorporates Fox's (2007a) condition of Innocent Exclusion, a notion that will be of key importance throughout this work.

Let us take a scenario where three types of dessert are salient: cake (p), ice cream (q), and cookies (r). Consider now the sentence in (16).

(16) You are only allowed to have [cake or ice cream]$_{\text{F}}$.

(16) forbids its hearer from having cookies, but gives him/her permission to have either cake or ice cream.[8] Our treatment of *only* as a sentential operator gives us the LF in (17), schematized as (18).

(17) $only_A$ [you are allowed to have [cake or ice cream]$_{\text{F}}$]

(18) $only_A \Diamond(p \vee q)_{\text{F}}$

On Fox and Katzir's theory of alternatives, the set A will include the result of simplifying the focused constituent $\Diamond(p \vee q)_{\text{F}}$, along with the result of replacing it with whatever else is salient, e.g. $\Diamond r$ (cookies):

[7] See also Matsumoto (1995) for a characterization of alternatives that relies on monotonicity

[8] (16) also licenses the so-called Free Choice permission inference. The inference has no bearing on the conclusions of this section, but see Alxatib (In progress) for a discussion of FC under *only*.

(19) $A = \{\Diamond p, \Diamond q, \Diamond r\}$

Since none of the elements of A are logical consequences of the prejacent in (18), *only*, on the entry in (20), will negate them all. This is shown in (21).

(20) (*repeated from* (2)) $[\![\text{only}_A\ S]\!]$ is defined only if $[\![S]\!] = 1$, and if defined
$[\![\text{only}_A\ S]\!]$ is true iff for all $S' \in A$, if $S \not\models S'$, then $[\![S']\!] = 0$

(21) $[\![only_A \Diamond(p \vee q)]\!]$ is defined only if $[\![\Diamond(p \vee q)]\!] = 1$, and if defined
$[\![only_A \Diamond(p \vee q)]\!] = 1$ iff $[\![\Diamond p]\!] = [\![\Diamond q]\!] = [\![\Diamond r]\!] = 0$

The exclusions in (21) negate the prejacent of (18), and so the prediction is that sentences like (16) prohibit p (cake) and prohibit q (ice cream), which is incorrect.

The problem that leads to the result in (21) is that some elements of A can *individually* be negated consistently with the prejacent, e.g. the alternative $\Diamond p$, but when they are negated together with some other elements of A, e.g. $\Diamond q$, the result contradicts the prejacent.[9] So, if the prejacent is held true, at most *one* of these alternatives can be negated (excluded): $\Diamond(p \vee q)$ and $\neg\Diamond p$ are consistent, and $\Diamond(p \vee q)$ and $\neg\Diamond q$ are consistent, but all three are not. Let us then define a Consistent Set of Excludables (CSE), given a sentence S and a set of alternatives A as follows:

(22) A set M is a CSE given S and A (CSE$(S)(A)$) iff $M \subseteq A$ and $\bigwedge\{\neg[\![S']\!] : S' \in M\} \wedge [\![S]\!] \not\models \bot$

According to (22), a set M is a CSE given a sentence S and alternatives A iff M is a subset of A, and if the negation of every member of M is consistent with S. If $S = \Diamond(p \vee q)$ and $A = \{\Diamond p, \Diamond q, \Diamond r\}$, then the sets in (23) are CSE$(S)(A)$, and the sets in (24) are not.

(23) CSE$(S)(A)$: $\{\Diamond p\}$, $\{\Diamond q\}$, $\{\Diamond r\}$, $\{\Diamond p, \Diamond r\}$, $\{\Diamond q, \Diamond r\}$

(24) Not CSE$(S)(A)$: $\{\Diamond p, \Diamond q\}$, $\{\Diamond p, \Diamond q, \Diamond r\}$

Now let a set M be a *Maximal* Consistent Set of Excludables, MCSE, given S and A, iff M is a CSE$(S)(A)$, and no superset of M is a CSE$(S)(A)$:

[9] Note that $\Diamond p$ and $\Diamond q$ satisfy the second of the subdefinitions of symmetry: $S_1 \vee S_2 = S$ (their conjunction is not contradictory, so they do not satisfy the first). But in the earlier discussion we were looking for ways to admit only one of the two symmetric sentences into the set of alternatives. Here we have reasons to admit both, as discussed in Section 1.2.1, so while the symmetry problem arises here as it did in the earlier discussion, the solution then was to prefer one of the two alternatives, while here the goal is to remove them both.

(25) A set M is a MCSE given S and A iff M is a CSE$(S)(A)$, and there is no M' such that $M \subset M'$ and M' is a CSE$(S)(A)$

The first three (singleton) sets in (23) are not MCSE$(S)(A)$, because each of them can be expanded to a superset that is a CSE$(S)(A)$. However, the remaining two sets in (23) are MCSE, because every subset of A that includes them is not a CSE. Therefore,

(26) MCSE$(S)(A)$: $\{\Diamond p, \Diamond r\}$, $\{\Diamond q, \Diamond r\}$

(27) Not MCSE$(S)(A)$: $\{\Diamond p\}$, $\{\Diamond q\}$, $\{\Diamond r\}$, $\{\Diamond p, \Diamond q\}$, $\{\Diamond p, \Diamond q, \Diamond r\}$

We now define the set of Innocently-Excludable alternatives (IE) given sentence S and alternatives A as the intersection of all the MCSEs given S and A:

(28) $\text{IE}(S)(A) = \bigcap \mathbf{MC}(S)(A)$, where,

(29) $\mathbf{MC}(S)(A) = \{M : M \text{ is an MCSE(S)(A)}\}$

In our example, IE$(S)(A)$ is the intersection of the MCSEs $\{\Diamond p, \Diamond r\}$ and $\{\Diamond q \Diamond r\}$, which is the set $\{\Diamond r\}$. To put it briefly, what IE delivers is the set of alternatives whose exclusion is consistent with the prejacent, no matter what other alternatives are excluded. If an alternative S' does not have this property, i.e. if its negation is inconsistent with the prejacent, or if it conspires to block the exclusion of some other alternative, then S' is non-innocently excludable given S, A.

Let us first note that the set IE$(S)(A)$ can only include members of A that are non-weaker than S. If it included a weaker member S', then it must be the case that $\neg S'$ belongs to all the CSEs of S given A, which in turn means that $\neg S'$ is consistent with S. But this cannot be true; S' is by assumption a logical consequence of S, so its negation $\neg S'$ must entail $\neg S$. Therefore no CSE$(S)(A)$ can include a weaker alternative of S.

So, in light of (16) and the utility of IE, I restrict the reach of the assertoric component of *only* to IE-alternatives. My semi-official definition of *only* is shown in (30).

(30) ***Only* – IE version (semi-official):**

Given a set of alternatives A and a sentence S,

$[\![\text{only}_A\ S]\!]$ is defined only if $[\![S]\!] = 1$, and if defined

$[\![\text{only}_A\ S]\!]$ is true iff for all $S' \in \text{IE}(S)(A)$, $[\![S']\!] = 0$

Because IE can only include non-weaker alternatives, the revision in (30) does not add any exclusions to those of the initial version in (2). The revision therefore restricts *only* to exclude just those non-weaker alternatives that can be negated along with any others, and still be consistent with the prejacent.

1.2.3 The non-vacuity constraint

A final fact about *only* that we need to mention is that it cannot be vacuous. This is evident from the oddness of examples like (31).

(31) a. #John only saw [every student]$_F$.
     b. Of Mary and Sue, #John only saw [Mary and Sue]$_F$.
        ✓John only saw Mary$_F$.
        ✓John only saw Sue$_F$.

At first sight this shows that a revision of *only* is needed where the particle is banned from appearing vacuously. Ultimately I will argue this, but not on the basis of (31). To see what IE$(S)(A)$ alternatives there are for e.g. (31b), we need to determine the set of MCSEs: what alternatives (from among *Mary* and *Sue*) can be negated consistently with the prejacent? The answer is none. So there are no CSEs that can be constructed for (31b). Therefore the set of MCSEs is empty. IE$(S)(A)$, by definition, is the grand intersection of all MCSEs, but since there aren't any, IE$(S)(A)$ will contain every sentence in the language.[10] *Only* therefore negates everything, which makes (31b) contradictory, hence unacceptable.

We may think, then, that no revision is needed, since examples like (31b) are already ruled out, owing to their contradictory truth-conditions. But let us again consider (31a), and this time suppose that the determiner *no* is made salient before the sentence is uttered:

(32) A: Did John see every student, or did he see no student(s)?
     B: #He only saw [every student]$_F$.

The consideration above do not apply to the odd (32). Here we do have a salient alternative that can be consistently negated with the prejacent: the negation of *John saw no student(s)* is consistent with (in fact it follows from) the proposition that he saw every student. Therefore, we can construct a CSE containing just this alternative, and since nothing else is salient, the set is maximal.

(33) $S =$ *John saw every student*

[10]This feature of IE is pointed out in Gajewski (in press): if the operator $\bigcap$ is defined as in (1)

(1) $\bigcap \mathbf{P} = \{x : \forall P(P \in \mathbf{P} \rightarrow x \in P)\}$

then the intersection of an empty $\mathbf{P}$ will contain every x in the domain, because the membership condition of the intersection is trivially satisfied.

(34) $A = \{[\textit{John saw every student}], [\textit{John saw no student}]\}$

(35) $\text{IE}(33)(34) = \{\textit{John saw no student}\}$

(32) is therefore defined only if John saw every student, and, as is already implied from the prejacent, true iff John saw some student. So far, nothing in what we've assumed about *only* rules out the sentence. Nonetheless its unacceptability is clear, and its oddness evidently results from the vacuity of *only*'s contribution. We therefore need to formulate a yet finer definition of the particle, as follows.

(36) ***Only* – IE version + nonvacuity:**

Given a set of alternatives A and a sentence S,

$[\![\text{only}_A\ S]\!]$ is defined only if $[\![S]\!] = 1$, and $\exists S' \in \text{IE}(S)(A).([\![S]\!] \not\models \neg[\![S']\!])$. If defined,

$[\![\text{only}_A\ S]\!]$ is true iff for all $S' \in \text{IE}(S)(A)$, $[\![S']\!] = 0$

(36) adds that *only* is defined provided that some IE-alternative is such that its negation does not follow from the meaning of the prejacent. The condition is falsified if every excludable alternative is already excluded from the meaning of the prejacent, in which case *only* is vacuous.

1.3 Summary

In Section 1.2 I developed a non-vacuous, IE-sensitive semantic definition of *only*. I assume this entry for the particle in the remainder of this study (though I will add further assumptions in Chapter 4).

(36) ***Only* – IE version + nonvacuity:**

Given a set of alternatives A and a sentence S,

$[\![\text{only}_A\ S]\!]$ is defined only if $[\![S]\!] = 1$, and $\exists S' \in \text{IE}(S)(A).([\![S]\!] \not\models \neg[\![S']\!])$. If defined,

$[\![\text{only}_A\ S]\!]$ is true iff for all $S' \in \text{IE}(S)(A)$, $[\![S']\!] = 0$

In the next chapter I will apply (36) to Negative Quantifiers, and show that the definition makes incorrect predictions given standard assumptions about alternatives. The general empirical picture, as I will show in the next chapter, is that *only* is intuitively felt to be vacuous, which contradicts the definition in (36). I will ultimately argue that the vacuity is illusory, and that prejacents containing NQ associates have a different form at LF from sentences where *only* is absent.

Chapter 2

The empirical picture, and the *only*-P/N Generalization

In this chapter I present my empirical claims about association between *only* and Negative Quantifiers (NQs). I show that the facts are problematic for the analysis of *only* that was developed in Chapter 1, where the particle was taken to negate those alternatives that are Innocently-excludable with respect to the prejacent. When I show my findings, I will focus on a subset of these IE-alternatives, namely the stronger ones, and show that *only* fails to negate them in every case. I will also focus on unmodified/intensified quantifiers—e.g. [(*very*) *few*]—as the representative NQs, and discuss other NQs later on.

The chapter begins with the assumption that an NQ like [(*very*) *few*] has the determiner *no* as a formal alternative. This assumption, which will be reconsidered later, is derived rather naturally from the scalar implicatures that [(*very*) *few*] gives rise to: *few* standardly licenses an existence inference, which on simple accounts is generated (non-semantically) by negating what is thought to be its stronger formal alternative, the determiner *no*. I will show that maintaining this assumption presents a problem for the analysis of *only* from Chapter 1: by its definition, the particle is predicted to *semantically* negate the stronger alternative *no*, a prediction that I will show is incorrect. Crucially, I will argue that even after we revise our assumptions about *few* and its alternatives—in Chapter 3 where we will analyze the term as gradable predicate rather than a determiner—the problem with association with *only* does not disappear: whatever alternative is negated in generating the existence implicature for *few*, that same alternative is predicted to be negated by *only*, thus generating the existence inference in the particle's assertoric component.

2.1 The data

Consider (37) and (38). Here the NQs *few* and *rarely* are placed in association with *only*.

(37) John only spoke to [(very) few]$_F$ students.

(38) John only [rarely]$_F$ spoke to students.

What, intuitively, is the contribution of *only* in these examples? And what is its contribution predicted to be according to the analysis developed in Chapter 1? The main part of this chapter is devoted to showing that these two questions have conflicting answers: in sentences like (37) *only* does not add any truth conditions to what already follows from the prejacent. A brief comparison between (37) and its *only*-less variant in (39) suggests this.[1]

(39) John spoke to (very) few students.

In later sections I will test this claim more rigorously, and conclude that *only* does indeed make no assertoric contribution in sentences like (37).

Now we turn to the second question: what is the contribution of *only* predicted to be? In each of (37,38), *only* presupposes its prejacent, and asserts the negation of the prejacent's stronger alternatives. If we assume that *no* is an alternative to *few*, and if replacing *few* with *no* produces a stronger sentence, then associating *few* with *only* should semantically negate the stronger *no*-alternative. This should then introduce the existence inference in the semantics of the sentence:

(37) John only spoke to [(very) few]$_F$ students.

a. *Presupposes*: John spoke to very few students.

b. *Asserts*: ¬(John spoke to *no* students),

i.e. John spoke to *some* students.

(38) John only [rarely]$_F$ spoke to students.

a. *Presupposes*: John rarely spoke to students.

b. *Asserts*: ¬(John *never* spoke to the students)

i.e. John *sometimes* spoke to the students.

[1]To my knowledge, this fact about *only* and *few* was only noticed in Beck (2012). I will come back to her account in Chapter 7 (Section 7.4).

Before we go through the tests that disconfirm these predictions, we ask first what motivation there is for thinking that *no*/*never* are alternatives to *few*/*rarely*. Moreover, we want to know in what sense the former are stronger than the latter. I address these questions by taking a short detour to SIs, where I show reasons—based on how the relevant SIs are intuitively understood—to assume that *no*, or something like it, is indeed an alternative *few*, and likewise for *never* and *rarely*.

2.1.1 *no* as an alternative to *few*: evidence from Scalar Implicatures

Consider the *only*-less examples in (40,41).

(40) John spoke to (very) few students.

(41) John rarely spoke to students.

By default, (40) is understood to say that John spoke very few *but some* students. Similarly, (41) says that John spoke to students on a small number of occasions, but that there were at least *some* such occasions. These two inferences, to *some* in the case of *few* and to *sometimes* in the case of *rarely*, are on standard accounts thought to be scalar implicatures of (40,41), rather than part of their semantic content.

Implicatures are thought to arise from conversational considerations (see Davis 2013 for a recent survey). They differ from semantic inferences in their cancellability, and in their weakness under downward monotone operators, e.g. under negation, in the antecedent of a conditional, the restrictor of a universal quantifier, etc. Compare the (a) and (b) examples in (42) and (43).

(42) a. John spoke to some of the students. #In fact he spoke to none.
 b. John spoke to some of the students. In fact he spoke to all.

(43) a. John spoke to few of the students. #In fact he spoke to many.
 b. John spoke to few of the students. In fact he spoke to none.

(42a,43a) are contradictory, and they are predicted to be if the truth conditions of *some* are taken to require existence (e.g. of students that were spoken to), and if those of *few* are taken to require falling below some standard of “fewness”. By contrast, the coherence of (42b,43b) suggests that *some* is semantically compatible with *all*, and *few* is with *none*. So even though *some* is by default understood to imply *not all*, and *few* is understood to imply *not none* (i.e. *some*), the inferences cannot be said to arise from the semantics of the determiners, given the acceptable continuations in the (b) examples.

Contemporary literature has it that these inferences (SIs) are pragmatic (a line of research that originated in Grice 1975, and also Ducrot 1973). The idea is that, if the speaker believes the stronger alternative (*all* in the case of *some*, and *none* in the case of *few*), and is moreover inclined to maximize his/her informativity, s/he would have uttered the stronger alternative. So, when a speaker utters *some*/*few*, it is inferred that s/he does not believe the stronger alternatives *all*/*none* (or that s/he is ignorant about their truth). This inference can be suspended in the right circumstances, however, as (42b,43b) show.

The suspendability of the inference to *some* (in the case of *few*) can also be seen when *few* is embedded in a DE environment, e.g. the antecedent of the conditional in (44).

(44) If John scores few goals this season, he'll have a tough time advancing to the major league.

If *some* (the negation of *no*) were part of the *semantic* meaning of *few*, we would expect (44) to be equivalent to (45).

(45) If John scores few but some goals this season, he'll have a tough time advancing to the major league.

However, (44) and (45) are intuitively different: in the former the consequent holds if John does not score any goals, while in the latter it does not. This indicates that in (44), the inference to *some*—the putative SI—does not contribute to the meaning of the antecedent, unlike in (45) where *some* is explicitly added. A similar paradigm is constructed for *rarely* in (46).

(46) a. John rarely speaks to students. In fact he never does.
b. John sometimes speaks to students. #In fact he never does.
c. If John rarely shows up to practice, he'll have a tough time during the game.
$\neq$ If John rarely but sometimes shows up to practice, he'll have a tough time during the game.

Conclusion of detour, and look-ahead

We have seen that, though *few* licenses the existence inference in simple cases, the inference does not come from its semantic meaning. Two pieces of evidence were used to illustrate this: the acceptable [*few*... in fact *no*] continuation, (43b), and the weakness of the existence inference in DE environments, e.g. (44). We took this to show that any situation where the *no*-alternative is true is one where the *few*-alternative is also true. The *no*-sentence is therefore semantically stronger than the *few*-sentence, and it is by negating the former alternative that *few*'s existence implicature is derived.

Now I move to association with *only*. The plot is this: we postulated *no* as an alternative to *few* in order to derive the latter's existence implicature. If association with *only* is sensitive to the same alternatives as those for SI calculation, then we expect *no* to take part in the meaning of *few* when it associates with *only*.

The more general point is that *whatever alternative* is responsible for generating the existence implicature for *few*, that same alternative is predicted to semantically generate the existence inference when *only* takes *few* as associate. When we look at the semantics of *few* in more detail—in Chapter 3, where the term is treated as a gradable adjective—we will look at other alternatives that might help derive the existence implicature. The important point now is that, once these alternatives are found, their negation is predicted to form part of *only*'s composition with *few*.[2]

I will now show evidence that this is wrong.

2.1.2 Associating *few* with *only*

Consider again (37) and (38).

(37) John only spoke to [(very) few]$_F$ students.

a. *Presupposes*: John spoke to very few students.

b. *Asserts*: ¬(John spoke to *no* students),
i.e. John spoke to *some* students.

(38) John only [rarely]$_F$ spoke to students.

a. *Presupposes*: John rarely spoke to students.

b. *Asserts*: ¬(John *never* spoke to the students)
i.e. John *sometimes* spoke to the students.

Having seen reasons to treat *no* as alternative to *few*, we now see why the existence inference is predicted to be part of the semantics of (37): because *no* entails *few*, it should be negated by *only*'s assertoric component. In summarize my claim against this prediction in (47), a generalization that I will revise in Section 2.2.

[2] The correspondence between SIs and the semantics of of association with *only* is discussed in Fox (2003), Fox and Hackl (2006). See specifically Fox's Only-Implicature-Generalization (OIG), which draws the parallel between the two phenomena. The case of NQ association with *only* may be seen as a potential exception to the generalization, though, as I will argue throughout this work, the apparent exception follows from other, independently motivated factors.

(47) **The *only*-NQ Empirical Claim**

For any of the relevant NQs, the semantic meaning of [*only* $\cdots$ NQ_F $\cdots$] does not semantically negate the NQ's stronger alternatives.

Once the empirical findings are demonstrated, I return to *few* and qualify my assumptions about its semantics. We will see that the revisions do not directly affect the significance of the data, but they will provide an important step in locating the cause of (47).

2.1.3 Data Illustrations

This section consists of three illustrations of the robustness of (47). In the first two I demonstrate the weakness of the existence inference, which is incorrectly predicted to be semantic.

First Illustration. If we imagine that a bet is made that (37) is true, and if the world turns out to verify the stronger *no* alternative, i.e. that John spoke to no students at all, then the bet, according to judgements, is won (likewise for (38), in a world where John never spoke to students).

(37) John only spoke to $[\text{very few}]_F$ students. True/Bet won when *none*

(38) John only rarely_F spoke to students. True/Bet won when *never*

On the other hand, a win could not be claimed if the bet is made on either of (48,49), and if the stronger alternatives to the prejacents — 'John spoke to all' and 'John always spoke' — turn out to be true.

(48) John only spoke to $[\text{some}]_F$ students.

(49) John only $[\text{sometimes}]_F$ spoke to students.

We might explain the judgements regarding (48,49) as follows: in these sentences *only* makes a semantic contribution to the content of the bet, specifically, the exclusion of the stronger alternatives (*all* and *always*). The wager, then, concerns the truth of these stronger alternatives: if they are false, then the *only*-sentence is true, and the bet is won; if they are true, the *only*-sentence is false, and the bet is lost.[3]

But if this same reasoning is applied to (37,38), we predict the (unattested) parallel intuition, that the bet is lost in situations where their stronger alternatives (*no*/*never*) are true.

[3] See Beaver and Clark (2008), Ch. 9, for experimental tests on cases when the prejacent itself fails.

The difference between *some*/*sometimes* and *few*/*rarely* becomes clearer when we compare the sentences with their *only*-less variants. We see here that the introduction of *only* changes the content of the bet in the case of *some*/*sometimes*, but not in the case of *few*/*rarely*.

(37′) John spoke to very few students. True/won when *none*
(same as (37) – unexpected)

(38′) John rarely spoke to students. True/won when *never*
(same as (38) – unexpected)

(48′) John spoke to some students. *True*/won when *all*
(different from (48) – expected)

(49′) John sometimes spoke to students. *True*/won when *always*
(different from (49) – expected)

In (37′,38′,48′,49′), the negation of the stronger alternatives—*no*, *never*, *all*, and *always*—is not felt to be an obligatory part of the content of the respective bets. Approaches that treat these inferences as non-semantic (e.g. as conversational implicatures) thus gain support from these intuitions. But while the intuitions are undone through association with *only* in (48,49), where the relevant inferences become *semantic* and thus part of the content of the bet, the expected parallel is not found in (37,38).

Second Illustration. Here I build on the intuition elicited in the previous illustration, and show more evidence that [*only very few*] does not semantically negate the stronger alternative *no*. I show that the inference to *some* in [*only very few*] can be explicitly cancelled, and that it disappears in Downward Entailing (DE) contexts. This is unexpectedly similar to analogous sentences where *only* is removed.

Consider the contrast between (50) and (51).

(50) John only spoke to some$_{\mathrm{F}}$ of the students. #In fact he spoke to all of them.

(51) John only spoke to [very few]$_{\mathrm{F}}$ students. In fact he spoke to none.

The oddness of the continuation in (50) is expected on the analysis of *only* from Chapter 1: the particle semantically negates the stronger *all*-alternative, so the continuation in which that alternative is asserted is odd. Now, when *few* is associated with *only*, we predict the *no*-continuation to be equally odd, since its negation should be contributed semantically by *only*. We find, however, that the continuation is acceptable.

In continuing to build the paradigm, let us turn to sentences where the relevant constructions are placed in a DE environment, e.g. the restrictor of a universal quantifier. As discussed earlier, Scalar Implicatures (by default) disappear in Downward Entailing contexts. Though (52a) has the strengthened meaning in (52b), embedding the scalar item *some* in the restrictor of a universal quantifier, as in (53a), fails to license the parallel inference in (53b).

(52) a. John did some of his homework

b. Strengthened Meaning = John did *some but not all* of his homework

(53) a. Everyone who does some of his homework will pass the course.

b. Strengthened meaning $\neq$ Everyone who does some but not all of his homework will pass.

At present there is no need to explain these facts. For our purposes it is enough to observe that the implicature ordinarily associated with e.g. (52a) is not felt to arise when the scalar item (*some*) is itself placed in the restrictor of a universal quantifier, as in (53a). This implies, correctly, that the argument of the quantifier, [will pass the course], holds of those who did all of their homework. In other words, stronger satisfaction conditions (conditions satisfying 'all' in the case of 'some') under the restrictor of a universal quantifier, but less so when *some* appears unembedded.

We will shortly examine these facts with *few*. Now let's note that adding *only* inside the restrictor of (53), as in (54), forces the some-but-not-all reading, and blocks the 'all' truth-conditions from satisfying the restrictor clause. This is expected, because the strengthening of 'some' to 'some but not all' is in this case contributed *semantically* by *only*, rather than pragmatically.

(54) Everyone who does some of his homework will pass the course. Everyone who does only some$_F$ of his homework will not ace it.

In the first restrictor of (54) the speaker intends to include those who do some/all of the homework: they all pass. In the second restrictor, *only* blocks students who do *all* of their homework, and the restrictor is limited to just those who do part of it. The speaker says of those that—though they pass, given the first sentence—they will not ace the course.

Let us now turn to *very few*. The prediction should be similar to the case of *some*: when *few* is embedded in the restrictor of a universal quantifier, the implicature that negates *no* should disappear. We predict, then, that the restrictor is satisfied even when the stronger *no*-alternative is true. But importantly, we predict that this strengthening (the inference from *few* to *some*) reappear in the semantics when *only* is added to the restrictor, because the inference will be contributed by the particle. The example is shown in (55).

(55) Everyone who submits very few assignments won't ace the course. #Everyone who submits only [very few]$_{\mathrm{F}}$ assignments will pass it.

The first half of (55) says that those who submit few *or no* assignments will not ace the course. Those who submit none are included because the restrictor does not semantically keep them out. In the second half, the restrictor is predicted to refer to just those who submit few *but at least some* assignments. The sentence says that they will pass the course. The contrast between the two restrictors should be (roughly) the same as the contrast in (54): there, the instructor says that submitting some homework is enough to pass it, but failing to submit it all leads to not acing it. Here, the announcement is the same: submitting a small number of assignments (few) leads to not acing the course, but submitting few but not-none—i.e. few but some—is enough to pass it. Why, then, is (55) odd? Intuitively, the addition of *only* in the second restrictor does not refine the range covered by the first. This indicates that both restrictors refer to the same group of students.

It seems, then, that placing *few*, or [*very few*], in association with *only* does not give rise to a meaning that excludes the stronger 'no' alternative. It was instead shown that '*only* [*very few*]$_{\mathrm{F}}$' is in fact compatible with scenarios where the 'no' alternative is true.

Third Illustration. Suppose a conversation is taking place between A and B, both executives in a company. A and B are discussing the recent behavior of John, an employee.

(56) A: It looks like we're gonna have to let John go. He's been slacking off recently, and over the past two weeks he's shown up late to every meeting.

B: Well, let's be fair. He only showed up late to some$_{\mathrm{F}}$ meetings.

(57) A: It looks like we're gonna have to let John go. He's been slacking off recently, and over the past two weeks he submitted none of his reports.

B: Well, let's be fair. #He only submitted [very few]$_{\mathrm{F}}$ of his reports.

B's defense of John is legitimate in (56), but not in (57). If B's sentences had the predicted meanings, given the proposed semantics for *only* in Chapter 1, both (56) and (57) should be equally good: in both cases, B weakens A's claim by negating the stronger alternative, *all* for *some*, and *no* for *few*, and in both cases the negations should serve to make John's position seem better than A makes it out to be. Yet, only the continuation in (56) is felicitous. The same (incorrect) prediction applies to *rarely*.

(58) A: It looks like we're gonna have to let John go. He's been slacking off recently, and over the past while he always came in late.

B: Well, let's be fair. He only sometimes$_{\mathrm{F}}$ came in late.

(59) A: It looks like we're gonna have to let John go. He's been slacking off recently, and over the past while he never came in on time.

B: Well, let's be fair. #He only rarely$_{\mathrm{F}}$ came in on time.

Extension to other NQs

In order to examine the behavior of other NQs (comparatives and *at most*), we first need to make assumptions about their alternatives. What alternatives does a comparative like [*less than* 3], or a modified numeral like [*at most* 3], have? Naturally, the numeral contained in either expression can be replaced with another, giving us [*less than* 2], for example, as an alternative to [*less than* 3], and [*at most* 2] as an alternative to [*at most* 3]. In sentences like (60,61), substitution with smaller numerals generates logically stronger alternatives:

(60) It takes less than 5 hours to get to New York from here.

It takes less than 4 hours to get to New York from here. (*Stronger than* (60)).

(61) It takes at most 5 hours to get to New York from here.

It takes at most 4 hours to get to New York from here. (*Stronger than* (61)).

This means that, if the NQ [*less than n*]/[*at most n*] is associated with *only*, the resulting semantics should negate the stronger alternatives, i.e. those containing lower numerals.

In (62,63) I show web examples where (*less than*)-comparatives (*at most*) constructions are associated with *only*.

(62) a. Jay-Z goes off about people saying he only owns less than one percent of the Brooklyn Nets

b. It only takes less than one hour to finish the DVD copying.

(63) a. However, I use enough crocking so that the depth of medium will only be at most 4 to 5 inches (10 to 12.5 cm).

b. Go with smaller groups, you can usually only go through with at most 6 people at a time.

Observe first that, just like in the case of *few*, the contribution of *only* is these examples appears to be vacuous, since their meanings do not change if the particle is removed:

(64) Jay-Z owns less than half of the Nets $\approx$ Jay-Z only owns less than half of the Nets.

(65) The depth of medium will be at most 4 to 5 inches $\approx$ The depth of medium will only be at most 4 to 5 inches.

I now run through the three tests used earlier. The result will confirm the intuition that *only* does not negate its prejacent's stronger alternatives.

Illustration 1. Bets Like in the case of *few*, a bet on (66) is intuitively won if John meets no students. This is not predicted if *only* negates alternatives where the numeral is replaced with a smaller one. The same holds of (67).[4]

(66) a. ?John only met [less than five]$_{\mathrm{F}}$ students $\approx$
b. John met less than five students

(67) a. ?John only met [at most five]$_{\mathrm{F}}$ students $\approx$
b. John met at most five students

Illustration 2. Here I revisit the scenario of the instructor, who may felicitously utter (68), repeated from (54), but neither one of (69) and (70).

(68) Everyone who does some of his homework will pass the course. Everyone who does only some$_{\mathrm{F}}$ of his homework will not ace it.

(69) Everyone who turns in less than 3 assignments will not ace the course. #Everyone who only turns in [less than 3]$_{\mathrm{F}}$ assignments will pass.

(70) Everyone who turns in at most 3 of his assignments will not ace the course. #Everyone who only turns in [at most 3]$_{\mathrm{F}}$ of his assignments will pass.

(69) and (70) are both predicted to be acceptable, if in each case *only* is predicted to negate the stronger alternatives, i.e. those containing lower numerals. In (69), for example, the instructor may be understood to say that whoever turns in 0-2 assignments will not ace the course, but whoever turns in exactly 2 assignments (less than 3 but not less than 2) will pass. But the sentence does not have this meaning, which suggests that the predicted exclusion (negation) introduced by *only* is not attested.

[4](66) and (67) are not perfectly acceptable, but as far as our discussion is concerned, the goodness of the earlier examples (62) and (63) is enough to make the empirical point. I use (66) and (67) because they are simpler, and leave it to future work to study the factors that distinguish them from the more acceptable sentences in (62,63).

Illustration 3. Finally, consider the scenario of the company executives A and B:

(71) A: It looks like we're gonna have to let John go. He's been slacking off recently, and over the past two weeks he submitted none of his reports.

B: Well, let's be fair. #He only submitted [less than 5]$_F$ of his reports.

B′: Well, let's be fair. #He only submitted [at most 5]$_F$ of his reports.

The logic here is the same as in the case of *few*. When B uses *only*, his utterance should exclude alternatives containing lower numerals: John only submitted [at most 5]$_F$ reports should mean that he did not submit at most 4, i.e. that he submitted at most 5, but more than 4 (= exactly 5). This should contrast with A's statement, but B's reply is still felt to be infelicitous.

Conclusion of Data

The discussions above were aimed at highlighting an unexpected empirical finding: that treating *only* as an exclusive of (logically) stronger alternatives falls short of accounting for the behavior of NQ associates. In the next section I propose the empirical generalization that I draw from the data.

2.2 The Empirical Generalization

There is a sense in which the paradigms presented above are incomplete: the comparisons were drawn between NQs and some other upward monotone determiner/adverb, specifically *some* and *sometimes*. A fairer comparison, and a more revealing one as we will see, would have included the positive counterpart of each of the examined NQs. The comparison is insightful because associating *only* with any of these positive counterparts results in generally ungrammatical constructions.

Unmodified/intensified

(72) a. John only spoke to [(very) few]$_F$ students.

b. *John only spoke to [(very) many]$_F$ students.

(73) a. John only rarely$_F$ spoke to students.

b. *John only frequently$_F$ spoke to students.

Comparatives

(74) a. (?)John only brought [less than one]$_F$ liter of beer.

b. *John only brought [more than one]$_F$ liter of beer.

<u>*At least/At most*</u>

(75) a. (?)John only brought [at most one]$_F$ liter of beer.

b. *John only brought [at least one]$_F$ liter of beer.

Given these findings, I propose the generalization in (76).

(76) <u>**Only P/N Generalization (*o*PN):**</u>

Given a pair of antonyms $\langle P,N\rangle$, if P is dispreferred/ungrammatical as (part of) *only*'s focus associate, in *[*only* [$_S \cdots$ [$\cdots P \cdots$]$_F \cdots$]], then [*only* [$_S \cdots$ [$\cdots N \cdots$]$_F$ $\cdots$]] can only have a non-logical reading.

(77) [*only* ϕ] is non-logical if the alternatives excluded by *only* do not include any *stronger* alternatives to ϕ[5]

The *o*PN Generalization is unidirectional. It can be equivalently formulated as (78), the contrapositive of (76): when association with an NQ is ambiguous between the predicted logical reading and the non-logical reading, association with the positive counterpart is grammatical.

(78) <u>Contrapositive *o*PN:</u>

Given a pair of antonyms $\langle P,N\rangle$, if [*only* [$_S \cdots$ [$\cdots N \cdots$]$_F \cdots$]] can be interpreted on its logical reading, then [*only* [$_S \cdots$ [$\cdots P \cdots$]$_F \cdots$]] is grammatical.

A case that falls outside of the domain of (76), owing to the generalization's unidirectionality, is when P is grammatical in [*only* [$_S \cdots$ [$\cdots P \cdots$]$_F \cdots$]] (failing the antecedent of (76)/satisfying the consequent of (78)). Here *o*PN is still true even if we find that [*only* [$_S$ $\cdots$ [$\cdots N \cdots$]$_F \cdots$]] can only be interpreted non-logically. We will later find examples of this kind, and relate their behavior to other properties that are known about them.

[5]Reader familiar with the literature on *only* may be reminded here of what is known as *only*'s 'scalar' reading, where the particle's alternatives are not ordered on a logical scale, e.g. military ranks or poker hands. I will discuss this reading briefly in Chapter 4, but for the majority of this study I refrain from using it, in an effort to capture as much data as possible from *only*'s logical reading. We will discuss possible limitations in Chapter 7.

2.3 Overview of Chapters 4-7

Let me now summarize the plan for the remainder of the study. The two main questions that will be addressed are:

Given a pair of antonyms $\langle P, N \rangle$,

(79) What blocks the logical readings of [*only* $[_S \cdots [\cdots P \cdots]_F \cdots]]$ and [*only* $[_S \cdots [\cdots N \cdots]_F \cdots]]$.

(80) What allows the non-logical reading [*only* $[_S \cdots [\cdots N \cdots]_F \cdots]]$:

a. What assumptions make the non-logical reading available for the negative case [*only* $[_S \cdots [\cdots N \cdots]_F \cdots]]$?

b. Why do these assumptions not predict a non-logical reading for [*only* $[_S \cdots [\cdots P \cdots]_F \cdots]]$?

Chapters 4-6 concern (79), and will be divided by type of construction. In Chapter 4 I look at cases where *only* associates with *few* in its bare form (which I assume includes the unpronounced positive morpheme POS), and when it co-occurs with the intensifier *very*. The proposal there is to my knowledge novel, but draws on von Stechow's (2009) treatment of temporal adverbs, and Zeevat's (2008) analysis of *only*. In Chapter 5 I consider the possibility of applying Büring's (2008) and Schwarz's (2011) analysis of *at least*/*at most*, but I argue that more is needed in order to account for the facts in full. I suggest that the findings from Chapter 4, for POS and *very*, are applicable to *at least* and *at most*. In Chapter 6 I discuss comparatives, where I build primarily on the findings of Fox and Hackl (2006) and Heim (2006).

I turn my attention to (80) in Chapter 7, where I argue that the non-logical reading of *only*-NQ arises from LFs where the NQ appears under an existential quantifier. The presence of the quantifier changes the monotonicity of the prejacent, so that its stronger alternatives are those that contain *greater* degrees. When these LFs appear under *only*, the nonlogical reading arises. Analogous LFs that contain the positive antonym will not be different from LFs without the existential quantifier, so whatever blocks the latter from associating with *only* will also block the former.

Chapter 3

many, *few*, and their alternatives

In this chapter I outline some basic assumptions about the semantics of *many* and *few*. The plan is to provide sufficient background for later chapters, where the analysis is complemented with further assumptions, depending on what particular construction is considered. I develop a system where *many* is treated as a gradable adjective, and derive its determiner-like semantics by adding a silent existential quantifier to the phrase that contains it. The move is not favored over other competing analyses (e.g. Choice Functions or type-shifting mechanism), but is chosen for its technical simplicity. I then turn to *few*, which I take to be a suppletive form consisting of an antonymizing component ANT together with *many* itself. I show that the move is needed in order to derive the correct semantics for the term, and that it is motivated by the scopal behavior of *few* in modal environments.

3.1 *many*

3.1.1 *many* as a determiner

Hackl (2000) takes *many* to denote a "parametrized" determiner, i.e. a function from degrees to relations between sets of individuals.

(81) $[\![many]\!] = \lambda d.\lambda P.\lambda Q.\exists x(|x| \geq d \,\&\, P(x) = Q(x) = 1)$[1]

The saturation of the degree argument in (81) creates a semantic object of the same type as the determiners *some* and *every*. One reason for adding the degree argument to *many*

[1] I sloppily use the set cardinality operator on individuals in this entry. One would more accurately use a measure function μ that is defined for (possibly complex) individuals. Given an input x, $\mu(x)$ returns the maximal size of x.

is the morphological/semantic similarity that it bears to other gradable expressions, e.g. adjectives like *tall* and *wide*. In all these cases, the gradable term can take a comparative or an equative form (82,83), can appear with a direct measure phrase (84), or appear with the intensifier *very* (85).

(82) a. John is taller than Mary.
 b. The desk is wider than the table.
 c. John ate more cookies than Bill.

(83) a. John is as tall as Mary.
 b. The desk is as wide as the table.
 c. John ate as many cookies as Bill.

(84) a. I had no idea that John was that-tall.
 b. I had no idea what the table would be that-wide.
 c. I had no idea that John ate that-many cookies.[2]

(85) a. John is (very) tall.
 b. The desk is (very) wide.
 c. John ate (very) many cookies.

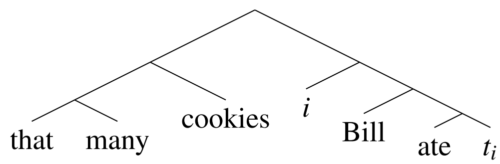
Much research has been dedicated to the unification of these paradigms. A contemporary view has it that each of these gradable terms denotes a function from degrees to some other type of object, e.g. a determiner $\langle et, \langle et, t \rangle \rangle$, as in the case of *many*, or a property of individuals $\langle e, t \rangle$, as in the case of the adjectives *tall* and *wide*. The given gradable term may merge with a degree-referring expression, like a measure phrase, or merge with an expression of a higher semantic type, e.g. a quantifier over degrees, and compose in the semantics with the trace of the expression once it raises to an interpretable position. Examples of this latter type include the comparative morpheme and the modifier *very*, among others. We will examine these constructions in detail in later chapters.

Let us see first how Hackl's account of *many* composes with a direct measure phrase. Suppose that Bill ate 12 cookies yesterday, and in this context someone utters (86).

[2]More obvious examples of measure phrases are those that refer to real measurements, e.g. *John is 6-ft tall*. But these measure phrases do not behave in exactly the same way with the determiner *many*, e.g. *John ate 3(*-many) cookies*. This oddness seems to be arbitrary, however, since in many languages numerals can only combine with nouns through classifiers, which are taken by some to have 'many'-like denotations. See e.g. Krifka (1995a) on classifiers in Mandarin Chinese, and recently Wilhelm (2008) on Dëne Sųłiné, and Bale and Coon (2012) on Mi'gmaq and Chol.

(86) (I heard that) he ate that-many cookies today too.

The phrase [*that-many cookies*] denotes a generalized quantifier, and so it cannot be interpreted in its base position as the argument of *ate*. At LF, shown in (87), the quantifier is interpreted at a higher structural position where it binds its individual-type trace.

(87)


The lexical entries for the relevant items are provided in (88), and the semantic composition is detailed in (89).[3,4]

(88) $[\![\text{that}]\!]^g = 12$

$[\![\textit{many}]\!] = \lambda d.\lambda P.\lambda Q.\exists x(|x| \geq d \ \&\ P(x) = Q(x) = 1)$

(89) $[\![(87)]\!] = [\![\text{that many cookies}]\!]([\![i \text{ [John ate } t_i]]\!])$

$= [\![\text{that many}]\!]([\![\text{cookies}]\!])([\![i \text{ [John ate } t_i]]\!])$

$= [\![\text{many}]\!]([\![\text{that}]\!])([\![\text{cookies}]\!])([\![i \text{ [John ate } t_i]]\!])$

$= \exists x(|x| \geq [\![\text{that}]\!] \ \&\ [\![\text{cookies}]\!](x) = 1 \ \&\ [\![i \text{ [John ate } t_i]]\!](x) = 1)$

$= \exists x(|x| \geq 12 \ \&\ [\![\text{cookies}]\!](x) = 1 \ \&\ [\![\text{ate}]\!](x)([\![\text{John}]\!]) = 1)$

3.1.2 *many* as an adjective

Another view of *many*, the one I ultimately adopt, classifies the term as an adjective rather than a determiner proper, in light of what seems in (90) to be adjective-like distribution (taken from Solt 2009, 2013).[5]

(90) a. John's friends are many.

b. The many students who attended enjoyed the lecture.

As a predicate of individuals, *many* may be defined as in (91).

[3]The denotation of *many* is reminiscent of Barwise and Cooper's (1981) set-theoretic semantics.

[4]The indexical *that* is here evaluated via the contextually-supplied assignment g.

[5]Proponents of this view include Hoeksema (1983), Partee (1989), Winter (1997), and Kennedy and McNally (2005a; 2005b).

(91) $[\![many]\!] = \lambda d\,.\,\lambda x\,.\,|x| \geq d$

On the entry in (91), a sentence like (90a) composes semantically in the same way that it would if, instead of *many*, the sentence featured an adjective like *tall*:

(92) John is tall.

(93) $[\![\text{tall}]\!] = \lambda d\,.\,\lambda x\,.\,\textsc{height}(x) \geq d$

In (90) and (92) the gradable terms appear unmodified. Let's assume that in these cases the degree argument is saturated with a silent degree-pronoun that refers to some contextually-supplied standard.[6] John is (d)-tall iff his height measures up to degree d or more, and similarly John's friends are (d)-many iff the unique (complex/plural) individual that is John's friends measures up to d in size/cardinality, where d is some implicitly understood value.

(94) $[\![\text{tall}]\!](d)([\![\text{John}]\!]) = 1$ iff $\textsc{height}([\![\text{John}]\!]) \geq d$

(95) $[\![\text{many}]\!](d)([\![\text{John's friends}]\!]) = 1$ iff $|[\![\text{John's friends}]\!]| \geq d$[7]

Importantly, note that "d-manyhood", just like other adjectives, can compose with noun phrases by Predicate Modification and produce another property. The denotation of [d-many cookies], for example, is the set of (plural) individuals which are cookies, and whose size/cardinality is greater than or equal to d:

(96)
$$\begin{aligned}[\![[d\text{-many}]_{\langle e,t\rangle}\ [\text{cookies}]_{\langle e,t\rangle}]\!] &= \lambda x_e\,.\,[\![d\text{-many}]\!](x) = [\![\text{cookies}]\!](x) = 1 \\ &= \lambda x_e\,.\,|x| \geq d\ \&\ [\![\text{cookies}]\!](x) = 1\end{aligned}$$

How does *many* get its determiner-like meaning on this analysis? On Hackl's approach, *many* is lexically specified as a determiner. Once it composes with its first predicate argument, the result QRs to a clausal position and leaves a trace over individuals behind. But if *many* is instead treated as an adjective, then how can it compose in an LF like (97), where it appears as a complement to a transitive verb?

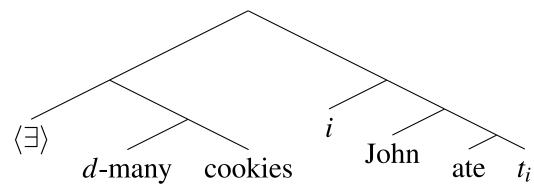
(97)
$$[\ \text{John}\ [\ \text{ate}_{\langle e,et\rangle}\ [_{\langle e,t\rangle}\ d\text{-many}\ \ \text{cookies}\ ]\ ]\ ]$$

[6]We get back to this in the next chapter.

[7]The property of being d-many is here assumed to apply to plural individuals only, unlike *tall* which applies to singulars.

There are a number of ways to resolve the mismatch in (97): one is to apply a choice function variable to the resulting predicate [d-many cookies], and existentially quantify over the function at a higher level (see e.g. Reinhart 1997, Winter 1997, Abels and Martí 2010). Another is to use existential closure (e.g. in the style of Lewis 1975), where the phrase [d-many cookies] is itself taken to introduce a free variable, and the variable is bound by the higher closure operation (Kamp 1981, Heim 1982, Diesing 1992). Yet another option involves type-shifting the adjectival meaning to a quantifier meaning, e.g. de Swart (2001). And finally, one may stipulate that, in indefinite phrases like [many cookies], an unpronounced determiner composes with the predicate and produces an existential quantifier over individuals. The result then QRs to a clausal position where it can be interpreted, as shown in (98).

(98)



(99) $[\![\langle\exists\rangle]\!] = \lambda P.\lambda Q.\exists x(P(x) = Q(x) = 1)$

(100) $[\![(98)]\!] = 1$ iff $\exists x(|x| \geq d \;\&\; [\![\text{cookies}]\!](x) = 1 \;\&\; [\![\text{ate}]\!](x)([\![\text{John}]\!]) = 1)$

I choose the final option because it is the simplest (technically) of the four. I should note, however, that I do not claim that the approach is capable of fully handling the indefinite-like behavior of *many* (or numerals generally). My goal is have a working system where *many*—and the other constructions that will later be investigated—can be predicated of individuals, and also give rise to determiner-like meanings. As I will show in later chapters, treating *many* as an adjective is an important ingredient in deriving the detected readings of *only*-NQ associations.

3.2 *few*

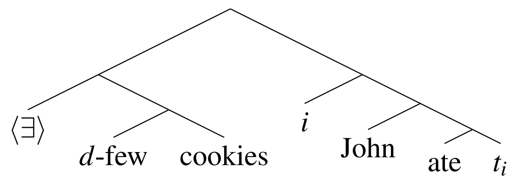
The most obvious way of extending our analysis of adjective-*many* to *few* is to assign *few* the entry in (101).

(101) $[\![\text{few}]\!] = \lambda d.\lambda x.|x| \leq d$

The adjective pairs a degree d and an individual x if x's size is d or less. If we now attempt to derive a determiner-like semantics for *few*, in e.g. (102), using the same existential quantifier for *many*, we generate the LF in (103).

(102) John ate d-few cookies.

(103)



(104) $[\![(103)]\!] = 1$ iff $\exists x(|x| \leq d$ & $[\![\text{cookies}]\!](x) = 1$ & $[\![\text{ate}]\!](x)([\![\text{John}]\!]) = 1)$

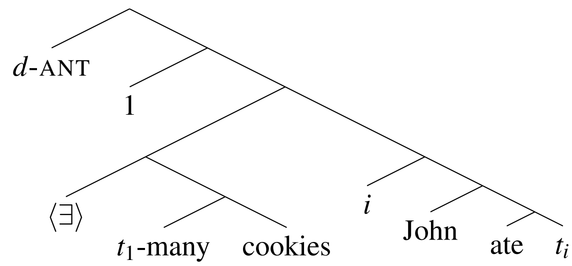
For the sake of discussion, continue to assume that d is saturated by a silent degree-pronoun that refers to a standard, in this case a standard of 'fewness'. (102) is true iff there is a plurality of size d or less, and the plurality is a plurality of cookies, and is eaten by John. The problem with these truth conditions is that they are too weak. Suppose d is 2. Then the sentence requires that there be a pair of cookies (or a single cookie) that John ate. But this holds even if John ate a dozen cookies, because if he ate 12, then there exists a collection of two cookies (or less) that he ate.

This problem is known in the literature as van Benthem's problem (after van Benthem 1986). It arises in this context in any approach that attempts to derive determiner-like meanings from adjectival *many* and *few* in a uniform way: if the determiner-*many* results from introducing an existential quantifier above the adjective *many*, then the analogous LF in the case of *few* will generate the problematic truth conditions seen in (104).

At least two issues are involved in addressing van Benthem's problem. First, if the adjectival approach to *many* and *few* is maintained, and if some version of existential closure is used in deriving the quantificational reading of *many*, then how can the correct reading of *few* be generated? Second, how can the problematic LFs, e.g. the one in (103), be blocked? If nothing rules out semantic structures where *few* is outscoped by an existential quantifier, then any sentence containing *few* should in principle have this meaning at least as an available reading, which is incorrect.

My answer to the first issue follows the recent tradition of decomposing d-*few* into an antonymizing component along with *many* itself: d-*few* = (d-ANT)-*many*. The unit (d-ANT) fills the degree slot that *many* takes as its argument, but it denotes a generalized quantifier over degrees, type $\langle dt, t\rangle$, that raises at LF to a clausal position. d-ANT takes a set of degrees D as its argument, and returns True iff d does not belong D. The LF is shown in (105).

(105)



(106) $[\![\text{ANT}]\!] = \lambda d.\lambda D.D(d) = 0$

(107) $[\![(105)]\!] = [\![\text{ANT}]\!](d)(\lambda d' [\![[\langle\exists\rangle\ t_1\text{-many cookies}]\ [i \text{ John ate } t_i]]\!]^{d'/1})$
$= 1 \text{ iff } [\![[\langle\exists\rangle\ t_1\text{-many cookies}]\ [i \text{ John ate } t_i]]\!]^{d/1} = 0$
$= 1 \text{ iff } \neg\exists x([\![\text{cookies}]\!](x){=}1\ \&\ [\![\text{ate}]\!](x)([\![\text{John}]\!]){=}1\ \&\ |x|{\geq}d)$
$= 1 \text{ iff } \forall x(([\![\text{cookies}]\!](x){=}1\ \&\ [\![\text{ate}]\!](x)([\![\text{John}]\!]){=}1) \rightarrow |x|{<}d)$

According to (107), the LF in (105) is true iff there is no collection of cookies that John ate, whose size is greater than or equal to d. This holds if John ate less than d-many cookies, e.g. less than whatever the standard of fewness is.[8]

Support for this 'syntactic' decomposition of *few* comes from the findings of e.g. de Swart (2000), who—building on observations that were initially made in Bech (1955) on German scope-splitting—points out that downward entailing determiners like *few* show scope ambiguity in modal contexts. (108) is an example.

(108) a. <u>You need to come up with few objections</u> if you don't want to seem like a party-pooper.
(Every possible world where you do not seem like a party-pooper is one where you come up with a small number of objections: $\Box >$ *few*)

b. This proposal is skimpy, and <u>you need to come up with few objections</u> to prove it wrong.
(The number of objections that you need to come up with is small: *few* $> \Box$)

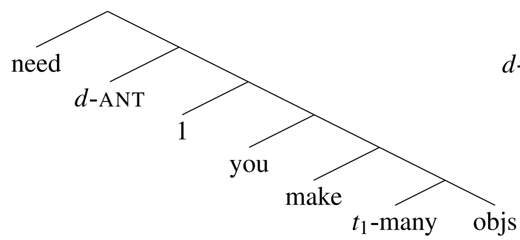
Considering that *come up with* is a creation predicate, there can be no *de re* reading of the DP [*few objections*]. Nevertheless, the sentence is ambiguous: on one reading it is

[8]There are other proposals on how to overcome van Benthem's problem. In de Swart (2001) an operation of Universal Closure is introduced alongside Existential Closure. In Diesing (1992) and Landman (2004) an operation of maximization strengthens the truth conditions, so that [*John ate few cookies*] is true iff some small collection of cookies are eaten by John, and *all such cookie collections* are subparts of this collection (see also Herburger 2000).

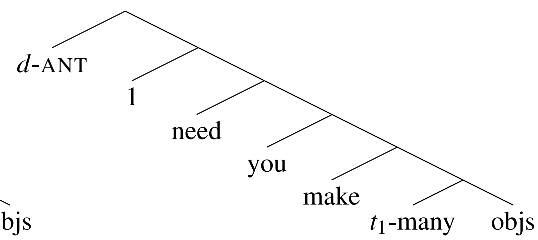
interpreted according to its surface representation, i.e. true in situations where the goal requires this: that a small number of objections be made. On the other reading, the sentence is true iff the number of objections that *need* to be made is a small.

On the decompositional approach to *few*, as in e.g. Heim (2006), the ambiguity of (108) results from the availability of multiple (clausal) landing sites for the raising antonymizer. The LFs responsible for the two readings are shown in (109,110).

(109)



(110)



The composition of (109), which produces the surface reading, is partially shown in (111). The inverse reading is composed, also partially, in (112).[9]

(111)
$$\begin{aligned} [\![(109)]\!] &= [\![\text{need}]\!]([\![\text{ANT}]\!](d)(\lambda d'[\![\text{you make } t_1\text{-many objections}]\!]^{d'/1})) \\ &= 1 \text{ iff } \forall w([\![\text{you make } t_1\text{-many objections}]\!]^{w,d/1} = 0) \\ &= 1 \text{ iff } \forall w(\text{you make less than } d\text{-many objections}) \end{aligned}$$

(112)
$$\begin{aligned} [\![(110)]\!] &= [\![\text{ANT}]\!](d)(\lambda d'[\![\text{need}]\!]([\![\text{you make } t_1\text{-many objections}]\!]^{d'/1})) \\ &= 1 \text{ iff } [\![\text{need}]\!]([\![\text{you make } t_1\text{-many objections}]\!]^{d/1}) = 0 \\ &= 1 \text{ iff } \neg\forall w([\![\text{you make } t_1\text{-many objections}]\!]^{w,d/1} = 1) \end{aligned}$$

Side remark on decomposition and antonymy

Despite the appeal of the decompositional analysis of antonymy, e.g. between *few* and *many*, the approach does not find empirical support in all cases of positive-negative correspondence. Here I merely wish to point to one example where no ambiguity is detected in modal contexts. Assume, following our current line of thought, that the adverb *slow* is composed of ANT together with *fast*, just like *few* is composed of ANT+*many*. Consider now (113), based on Heim (2008).

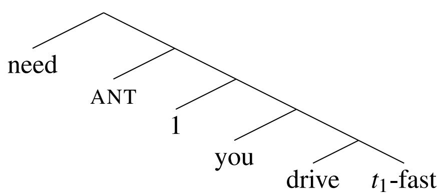
[9] I abstract away from the semantics of creation predicates. See e.g. von Stechow (2001).

(113) a. This highway is heavily patrolled, so you will need to drive slow if you don't want to get a ticket.

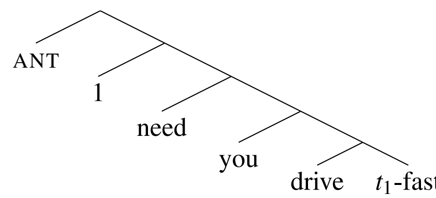
b. Mary lives far, so she will need to drive fast to get here. You live close, so #you will need to drive slow.

On an unconstrained decompositional account of antonyms, we expect that both LFs (114,115) be available for the underlined sentences in (113a,b).

(114)



(115)



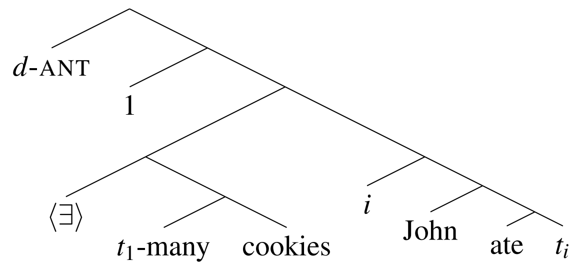
This means that two sets of truth conditions should be available: according to one, (114), it is required that the greatest speed you reach fall below some standard. This is appropriate for (113a), where failing the requirement will get you a speeding ticket. On the other reading, (115), the greatest speed that you *need* to reach falls below some standard, which leaves it open whether you are *allowed* to exceed that standard. This is the reading intended in (113b). But the unacceptability of the discourse suggests that the reading is unavailable. (115) must therefore not be a possible LF for (113b).

The scope of the so-called 'split'-reading, and the mechanisms behind it, are the subject of ongoing work. See e.g. Rullman (1995), de Swart (2000), Hackl (2000), Büring (2007a,b), Heim (2008), Penka (2011), to name a few. I have little to add to the findings of these works. Clearly the decompositional theory of antonymy must be constrained, but there is support for it at least in some cases, e.g. for the determiners *few* and *little*. Since the majority of this study concerns these determiners, I will continue to assume they are decomposed at LF.

Returning to van Benthem's problem

Having now derived the correct reading of *few*, using a decompositional treatment of *few*, the question is whether we expect the grammar to generate LFs where $\exists$ outscopes *few*. At first sight, one may think that the answer is no. To see why, consider (116) again, repeated from (105).

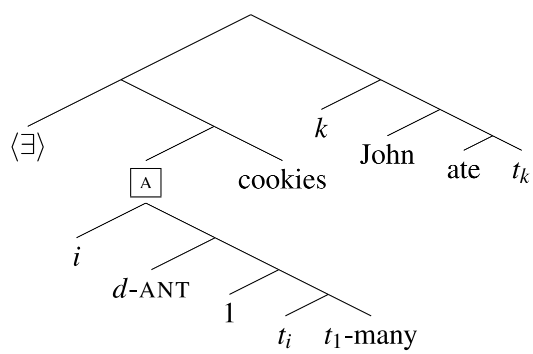
(116)



In (116), the trace t_1 can only be bound if the antonymizer raises to a position that c-commands it. Since d-ANT requires a clausal landing site, it can be only interpreted above $\langle\exists\rangle$, since none of the positions between $\langle\exists\rangle$ and t_1, where binding is possible, are clausal. It might seem, then, that the only way to decompose *few* is to raise its antonymizer above the existential quantifier, thereby preventing the derivation of the problematic van Benthem LFs. The proposal thus wins on two counts: it creates the correct reading of *few*, and by its nature it makes the unwanted LFs underivable.

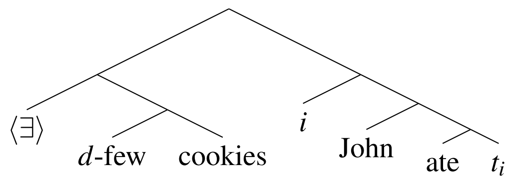
However, this is not quite true. There are, it turns out, other ways of embedding *few* under $\langle\exists\rangle$, but as I will argue in Chapter 7, the LFs are needed in order to derive the correct readings of *only*-NQ associations. Consider (117).

(117)



In (117), node $\boxed{\text{A}}$ denotes the property of having size less than d. The node is derived by merging a silent wh-like operator with d-*few*, i.e. with (d-ANT)-many. The antonymizer QRs to a clausal position and binds the trace in t_1-many, and the wh-operator raises to a higher node and introduces an index i that binds its trace. The result is a property of individuals that holds of a given x iff x's size is not greater or equal to d, which we continue to assume is supplied contextually. The composition of $\boxed{\text{A}}$ with *cookies* denotes the property of being 'less than d-many cookies', which is then taken as the first argument of $\langle\exists\rangle$. The resulting truth conditions require that there be a plurality of cookies eaten by John whose size is less than d. This is equivalent to (118), repeated from (103).

(118)



I will get back to these LFs in Chapter 7. Though I argue that the LFs are needed, in order to account for the seemingly non-logical reading of *only* with NQs, I will offer ways to constrain their distribution, in answer to van Benthem's problem.

3.3 *few* in association with *only*

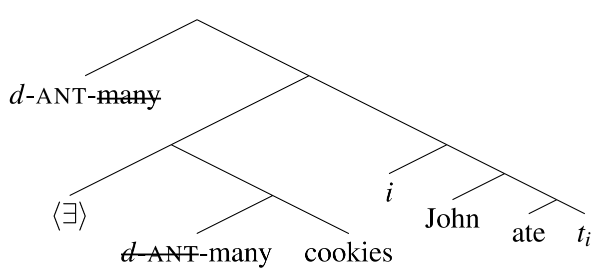
Having laid out my assumptions about association with *only*, and about the decomposition of antonyms, I now discuss how the two phenomena interact. The sentences that I will draw attention to in the coming chapters are sentences where *only* takes a NQ as its focus associate, e.g. (119).

(119) John only ate [(very) few]$_{\mathrm{F}}$ cookies.

How is (119) is interpreted in the semantics? So far I assume that (a) antonyms are decomposed at LF, and, following the bulk of the literature on association with focus, (b) focus-marking is used to determine the alternatives that *only* operates on. If we now have a decomposing associate, what happens to its focus-marking?

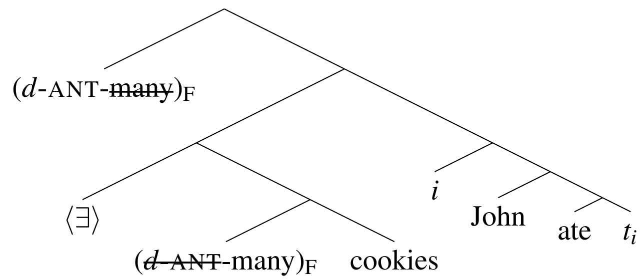
If decomposition is interpreted as an instance of movement, and if we analyze movement according to Chomsky's (1995) copy-theory, or as internal merge, then in LFs that feature decomposition we expect to see the "decomposed" constituent appear fully, in multiple syntactic positions. Our LF containing *few* is repeated below. Uninterpreted components are struck-out.

(120)



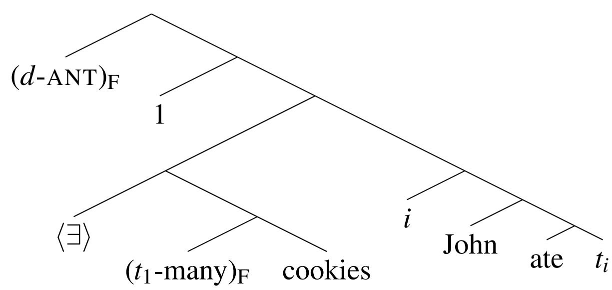
If the decomposed (i.e. copied) element is focus-marked, then it is natural for its copies to bear focus marking as well. So whenever we find *few* in focus, we expect to see focus marking on each copy of the ‘decomposing’ cluster in the movement chain:

(121)



Now if we abstract away from how traces are created in movement chains, e.g. via trace conversion (Fox 2002), we get the focus-marking seen in (122).

(122)



If a structure like (122) is taken as *only*’s prejacent, its alternatives will be those in which the focused elements are replaced with their own alternatives. On Katzir’s approach, we may replace, by lexical substitution, the degree d with other degree names. Some of the alternatives that result from this substitution will have a lower degree d' in place of d, and this makes them logically-stronger than the prejacent.[10] One possible way of deriving the existence inference, as an implicature of *few*, is by negating these stronger alternatives that feature a lower degree d'. If the existence SI is derived this way, then the same inference will be derived semantically when (122) is associated with *only*, since the particle will negate the same alternative that gives rise to the implicature in the absence of *only*. As I argued in Chapter (2), while the SI is indeed present, it does not come up semantically under *only*.

Another kind of alternative that Katzir’s theory produces is that where ANT is removed. The utility of this will be demonstrated in Chapter 7. Let us for now see how the alternative

[10] On the current analysis, eating ‘3-few’ cookies is the same as not eating 3 or more cookies. If we replace 3 with 2, the resulting alternative ‘2-few’ is stronger, because not eating 2 or more cookies logically entails not eating 3 or more. This is why *few* becomes stronger if its degree argument is replaced with a lower one: ‘fewness’ increases as quantity, or ‘manyhood’, decreases.

is generated: the focused structure (d-ANT) can be simplified, i.e. replaced with a sub-constituent of itself, giving us the degree variable d in place of (d-ANT). Effectively, this produces the positive antonym *many* as an alternative to the negative; the result of removing ANT, [1 [d-~~ANT~~ [2 ϕ]]], has the same semantic value as [2 ϕ].

3.4 Chapter summary

We analyzed *many* as a property of individuals, and derived its quantificational reading by merging it with a silent existential quantifier. The analysis was extended to *few*, whose scopal behavior was taken to suggest a decompositional account, one where the term is analyzed as a complex expression consisting of ANT and *many*. The quantificational reading of *few* was derived by interpreting its antonymizing component above the (silent) existential quantifier, but it was shown that there are LFs where the quantifier outscopes ANT, thus giving rise to weak readings that are traditionally thought to be problematic (van Benthem's problem). These LFs, I will show, are crucial in deriving the puzzling readings of *few* that were surveyed in Chapter 2, though as I will argue in Chapter 7, their distribution must be constrained. Finally, I offered a brief discussion of how alternatives are derived on a decompositional account of *few*.

Chapter 4

Mirativity, *only*, and the positive morpheme

In this chapter I focus on how *only* interacts with bare gradable forms like *few*, and intensified forms like *very few*. I argue that *only*, particularly its "mirative" component, is incompatible with the semantic contribution of the positive morpheme POS and the intensifier *very*. As a start, I remind the reader of what the *o*PN generalization claims: if the positive part in a pair of antonym is unacceptable/dispreferred as *only*'s associate, then the negative part can only have the non-logical reading. The positive of *few* is *many*, and as we already observed, *many* is unacceptable as *only*'s associate:

(123) #John only spoke to [(very) many]$_F$ students.

Intuitively, (123) is bad because *only*, on the one hand, suggests that John did not meet a large number of students, while [(very) many] says the opposite. In order make this intuition clearer, I use the notion of mirativity, and attribute it not to the determiner *many*, but to *very*, and in the absence of *very*, to an unpronounced morpheme POS that is thought to accompany the determiner in its bare form. I argue that POS/*very* are semantically incompatible with a licensing condition on the use of *only*, and it is this incompatibility that blocks the logical reading for both *many* and *few*.

4.1 The concept of mirativity, and definitional problems

The term "mirative" was introduced into contemporary linguistics by DeLancey (1997), who attributes it to Jacobsen (1964), and cites precursors in earlier studies of the "admirative" in the Balkan languages (Friedman 1986). Later work, like Friedman (2003), credit yet earlier sources for using the term "admirative" in grammatical descriptions, e.g. Dozon (1879). Both terms are intended as labels for markers of unexpected information, and since DeLancey, a wide variety of languages have been claimed to feature it in their morphology (see Aikhenvald 2012 for a review).

To my knowledge, the first application of the term to English particles came in Zeevat (2008). Zeevat's study focuses on *only*, but claims that similar reference to expectation/surprise is made by the particles *already*, *still*, and *even*:

> In English, one can find the markers *even*, *still*, *already* and *only* that seem to be mirative [...] In all four cases, they are specialised mirative markers, they express surprise at the large size of a quantity (even), surprise at the small size of a quantity (only), surprise at the early time of some event or the advent of some state (already) or at the long continuation of a state (still). Surprise would be a question of conflict with an expectation.
>
> (Zeevat, 2008, p.1)

The temporal particles *already* and *still* have seen similar treatments in earlier work. Löbner (1989), for example, cites (and criticizes) two approaches where expectation/surprise is thought to figure in the semantics of the particles:

> Several authors, e.g., Steube (1980) and Hoepelman and Rohrer (1981) include modal conditions for the time after t_e [evaluation time] in the truth-conditions. Hoepelman and Rohrer, e.g., make the expectation of the speaker that the state p does not obtain at the time t_e part of the truth-conditions of both $schon(t_e, p)$ [*already*] and $noch(t_e, p)$ [*still*]. This assumption, however, cannot be maintained in view of simple counterexamples such as
>
> (19) *Wie ich erwartet hatte, war das Licht schon/noch an.*
> As I expected, the light was already/still on.
>
> (Löbner, 1989, p.10)

Löbner is right to question accounts that refer *strictly* to the expectations of the hearer/speaker, and his example does indeed pose a problem to them. Of course, this applies equally to Zeevat's characterization of *only* (among the other particles) as a mirative item, since there too the goodness of sentences like (124) seems problematic.

(124) As everyone/I expected, only John$_\mathrm{F}$ showed up.

In fact, similar concerns arise in the literature on gradable quantifiers and adjectives. In Keenan and Stavi (1986), K&S, the determiner *many* is taken to hold of "significant" numbers, where speakers might have "almost any random reason" to decide on what qualifies as significant. This intuition is made more precise in Fernando and Kamp (1996), F&K, where reference to expectation is formalized and added to the determiner's semantic entry: a quantity/number counts as "many" iff it exceeds expectation. But here it is also easy to construct counterexamples:

(125) As everyone/I expected, (very) many people showed up.[1]

I will come back to these counterexamples shortly. Now I want to recast K&S's and F&K's observations in terms of what is often referred to as the positive morpheme, or POS. It is clear that K&S/F&K are concerned with *many* in its unmodified form (and when it is intensified with *very*), for whatever intuition there is regarding surprise or excess cannot be attributed to the modified versions in (126).

(126) a. Last year about twenty students enrolled in the program, and this year we will have about that-many also.

b. John drank as many beers as Bill.

In contemporary work, the vagueness of bare gradable terms like *many* and *few* is pinned on a covert morpheme, POS. The move is made to overcome the difficulty of composing the non-vague/non-mirative meanings of comparative/equative forms, from the overtly simpler, yet vague/mirative, bare form (the problem is addressed e.g. in Kamp 1975 and Klein 1980). And so it is postulated that *many*, understood in its ordinary use as a vague determiner, is composed of POS together with a precise, non-mirative version of *many*, and it is the result of this composition that the familiar vague meaning arises. When POS is replaced, say, with the comparative morpheme, the result is the precise/non-mirative comparative form (see Chapter 6).

My reason for this brief mention of POS is to generalize K&S/F&K's observations to all vague adjectives/quantifiers. The idea that a vague term, through the work of POS, holds of entities/quantities that stand out in some way, relative to some interest of the interlocutors, is discussed at length in Fara (2000). There it is also argued that expectation/surprise need not obtain in order for the bare vague term (or POS) to be licensed: a *tall* man need not be taller than expected, nor does there need to be an unexpectedly large number of, say, beer bottles in order for one to claim that *many* beers are there. Still it is acknowledged that

[1] The intensifier *very* is not discussed in either F&K or K&S, but their claims about "standing out" or "exceeding expectation" apply equally well to sentences that include it.

there ought to be some implicit interest or purpose that the interlocutors have in mind, so that expectation, instead of being understood broadly, is made relative to a specific person, perhaps, or a specific state of affairs, or a set of rules, etc. John might count as *tall* if he is taller than one would expect for someone his age, or for someone from among his family members, or from a group of candidates auditioning for the role of Napoleon in a film.

By now I hope to have pointed to a property that is shared by the temporal particles *already*, *still*, the exclusive *only*, and the positive morpheme. Mirativity, or something like it, does seem to be part of the semantics of these items, but because of the counterexamples that we've seen, it cannot be defined solely in terms of expectation. Löbner's complaint against the proposals of Steube (1980) and Hoepelman and Rohrer (1981), and the parallel complaint that can be leveled against Zeevat's treatment of *only*, and F&K's of POS/*very*, do not tell against modalized analyses of these items in general, but instead show that a modalized approach should be one where the modal base (e.g. what is compatible with expectations, norms, preferences, etc.) is made more flexible. So for example, when (127) is uttered, (Löbner's (19)),

(127) As everyone/I expected, the light was already/still on.

the use of *already* should not be taken to signal contrariness to everyone's/my expectations, but rather to what perhaps *ought* to be the case, given considerations of energy conservation, or saving on electric bills, etc. The variability of these construals, and the apparent ease with which they are accommodated, may be the reason why Keenan and Stavi allow "almost any random reason" to be good enough for a number/quantity to count as *many*.

What is striking, however, is that despite this flexibility, there seems to be a good degree of overlap between the modal bases that these items refer to. This comes as a surprise, because we are unable to decide on a precise definition of mirativity for any single one of these lexical items. And so we are forced to allow the modal bases to vary in ways that we have so far left unconstrained. Given this, one would expect it to be generally possible for the different items to co-occur, and for each to refer to a potentially different set of expectations, norms, or preferences. Interestingly, this is not the case.

(128) By 8 o'clock John was already there. #Mary came late too.

(128) is based on examples from von Stechow (2006), where (as far as I know) the first link is made in the literature between the temporal adverbs *already*/*still* and the positive morpheme.[2] We can find the relevant intuition from von Stechow in his paraphrases of the

[2]The additive *too* does not appear in von Stechow's examples. I thank Irene Heim for suggesting its use.

meaning of *already*: given a state of affairs ϕ, to say [*already* ϕ] is to say that it is early for ϕ. Note crucially the use of the adjective *early* in its bare form: [*already* ϕ] means ϕ is POS-early, i.e. that ϕ took place earlier than expected/desired/etc. So the norm referred to in the use of *already* is the same as the one that POS refers to in POS-*early*. In the second half of (128), the additive particle *too* adds the presupposition that some salient individual other than Mary came (POS-)late.[3] In this context, only John is salient, so the presupposition is that John came (POS-)late. But by the meaning of *already* (using its paraphrase from von Stechow), John arrived POS-early, so he couldn't possibly have come POS-late. This makes the utterance infelicitous.

Intuitive though this account may be, it remains surprising that the instance of POS in the second sentence of (128) invokes the same set of expectations/norms as *already*; if it didn't, the oddness of the utterance would not fall out, because it could be understood to say, for example, that John was there earlier than he *should* have been, but later than *expected*. The "earlier than *should*" licenses the use of *already* in the first half, and the "later than expected" satisfies the presupposition of additive *too* in the second.

Since this section is dedicated to the study of *only* and its interaction with POS, let me present the same conclusion drawn from (128), but with the slightly different (129).

(129) There were only fifty_F people at the party last year, #and again this year many people came.

Note crucially the improvement in (130), where *only* is removed.

(130) There were fifty people at the party last year, and again this year many people came.

The analysis of (129) is similar to that of (128): in this case, *again* presupposes that an earlier (salient) event was one where POS-*many* people came, and the only other salient event is the one where only fifty people came. But it seems that placing *fifty* in focus with *only* disqualifies the quantity from counting as *many*, hence the oddness. We might restate the conflict (in the spirit of von Stechow) as follows: for any quantity referring expression Q, [*only* Q_F] conveys that Q is not enough for what is needed/expected/etc. By contrast, *many* says of Q that it exceeds what is needed/expected/etc. In the next two sections, I will argue that *only* is for this reason incompatible with POS/*very*, regardless of the polarity of the scalar item that POS/*very* appears with. I claim that this incompatibility blocks the association of *only* with POS/*very*-P and POS/*very*-N in any pair of antonyms $\langle P, N \rangle$. The result is that the logical reading of both associations will not be available.

[3]See Beaver and Clark (2008), Section 3.4, and Beaver and Zeevat (2007).

4.2 Mirativity and the semantics of POS and *very*

One of von Stechow's (2006,2009) important contributions is his analysis of the positive morpheme. Since Cresswell (1976) semantic accounts of bare gradable adjectives depended on contextually understood standards: John is tall iff John's height reaches (or possibly exceeds) the standard of tallness. Among the puzzles that came with this analysis are

(i) why in the case of negative antonyms, like *short*, the requirements are turned around, so that John counts as short iff his height falls *below* the standard of shortness, and

(ii) why the standard for a negative term is always lower, on the relevant scale, than the standard of its positive counterpart.[4]

For Cresswell, the answer to (i) comes from the lexicon. Each gradable term comes with a scale, i.e. an ordered triple consisting of a dimension (e.g. height), a standard s, and an ordering relation R (e.g. $\geq$ in the case of tall, and $\leq$ in the case of its antonym). But this account falls short of answering *why* the relation is reversed between a gradable term and its antonym.

These two questions are answered neatly on von Stechow's proposal. POS maps a set of degree to True iff it includes in it the "neutral" set of degrees **N**.

(131) $[\![\text{POS}]\!]^{\mathbf{N}} = \lambda D_{\langle d,t\rangle} . \mathbf{N} \subset D$

Because of its semantic type, POS can only be interpreted at a clausal node, so it must QR at LF from its base position, leaving a trace of type d behind. So a sentence like (132a) is interpreted by composing the LF in (132b).

(132) a. John is POS-tall

b. [POS [1 [John t_1-tall]]]

Recall from Chapter 3 our assumption that *tall* denotes a relation between degrees and individuals, as in (133).

[4] A similar question to (i) arises in the context of comparative constructions (see Chapter 6). The answer provided by Heim's (2006) subset-semantics will be shown to lend support to her proposal.

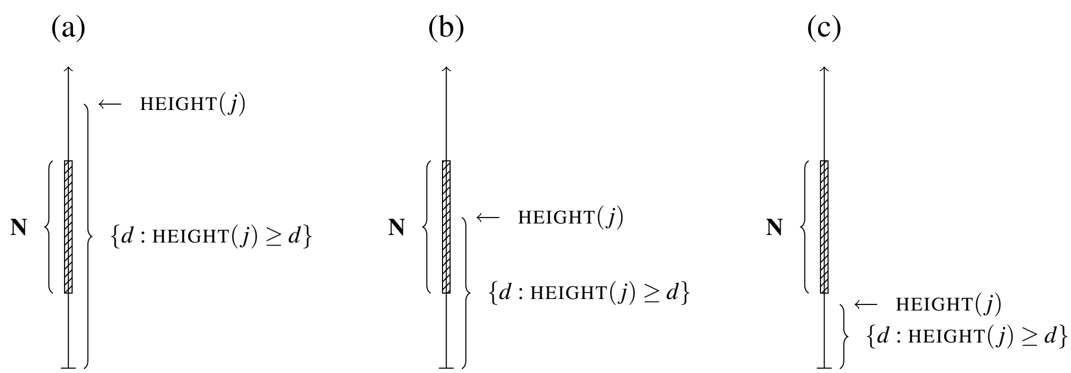
(133) $[\![\text{tall}]\!] = \lambda d\,.\,\lambda x\,.\,\textsc{height}(x) \geq d$

So the denotation of POS's sister node is the set of degrees that John's height measures up to. This will include John's height, and every degree that falls below it.

(134) $\lambda d\,.\,\textsc{height}(j) \geq d$

(132) will then be true iff **N** is properly included in (134). To von Stechow, and likewise to Heim (2006), **N** is an interval consisting of neutral degrees. In the case of height, **N**'s maximal element is the greatest neutral/normal height, and its minimal element is the smallest. So John counts as POS-tall if the set of height degrees he reaches includes the neutral interval **N**, as in (135a), and does not count as POS-tall otherwise (135b,c).

(135)



$[\![(132)]\!] = 1$ $[\![(132)]\!] = 0$ $[\![(132)]\!] = 0$

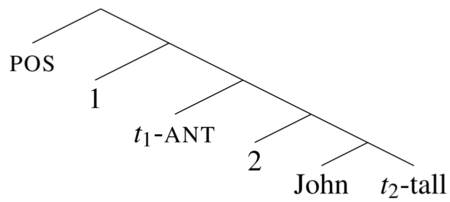
Continuing now with our treatment of negative antonyms as suppletives, consisting of ANT, (136), together with the positive antonym, we get *short* = ANT-*tall*.[5]

(136) $[\![\textsc{ant}]\!] = \lambda d.\lambda D.D(d) = 0$

Once ANT is combined with its first degree argument, it produces a GQ over degrees, so like POS, d-ANT/POS-ANT must be interpreted at the clausal level. The LF for (137a) is shown in (137b).

(137) a. John is POS-short = John is POS-ANT-tall

[5] The split-scope argument that was presented for *few* does not extend to *short*, as noted in Chapter 3. Here I use the adjectives to illustrate how POS interacts with ANT. The same is applicable to *few*.

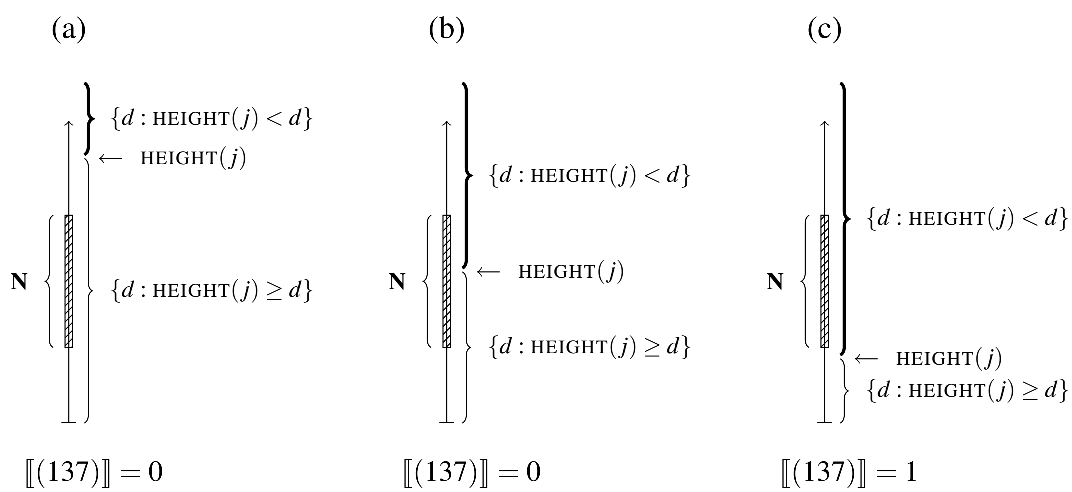
b. 

(137b) is composed as follows:

(138) $$\begin{aligned}[\![(137\text{b})]\!]^{\mathbf{N}} &= [\![\text{POS}]\!]^{\mathbf{N}}(\lambda d.[\![t_1\text{-ANT 2 John } t_2\text{-tall}]\!]^{d/1})\\ &= [\![\text{POS}]\!]^{\mathbf{N}}(\lambda d.[\![\text{ANT}]\!](d)(\lambda d'[\![\text{John } t_2\text{-tall}]\!]^{d'/2}))\\ &= [\![\text{POS}]\!]^{\mathbf{N}}(\lambda d.[\lambda d'[\![\text{John } t_2\text{-tall}]\!]^{d'/2}](d) = 0)\\ &= [\![\text{POS}]\!]^{\mathbf{N}}(\lambda d.[\![\text{John } t_2\text{-tall}]\!]^{d/2} = 0)\\ &= [\![\text{POS}]\!]^{\mathbf{N}}(\lambda d.\text{HEIGHT}(j) < d)\\ &= 1 \text{ iff } \mathbf{N} \subset \{d : \text{HEIGHT}(j) < d\}\end{aligned}$$

The truth conditions that result in (138) refer to two sets, the first is **N**, our set of normal/neutral degrees, and the second is the set of degrees that John's height falls short of. POS, by its definition, will demand that the former be included in the latter, i.e. John's height must be short of all of **N**. The only way to satisfy this requirement is for John's maximal height to fall below the lower end of **N**. The conditions are depicted in (139).

(139)



From this analysis of antonyms we derive both answers to questions (i) and (ii): regarding (i), in the positive case the set inclusion requirement can only be fulfilled if the maximal degree (of height in our example) extends above the higher end of **N**, i.e. above the positive standard (of tallness in our example). In the negative case, the inclusion requirement, together with the negation embedded in the denotation ANT, will make POS hold of an antonym only if its positive counterpart falls short of the lower tip of **N**, because in this case **N** will be a subinterval of those degrees that falsify the positive. So, in the positive case (e.g. *tall*), the greatest height has to be *above* the positive standard (e.g. of tallness), and in the negative case (e.g. *short*), the greatest height has to fall *below* the negative standard (e.g. of shortness). As long as **N** is not empty, the positive standard (**N**'s upper end) will always be above the negative standard (**N**'s lower end). This answers (ii).

Both von Stechow and Heim conceive of **N** as a set of degrees. But in order to link these degrees to mirativity i.e. to what might be expected or what might count as normal, **N** must be defined as a property of *propositions*/states-of-affairs. Once this is done, it will be possible to relate the different 'mirative' lexical items (e.g. *only*, POS, *already*, etc.) to each other. I therefore take **N** to be a set of expected/neutral propositions, and redefine POS as in (140).

(140) $[\![\text{POS}\ \phi]\!]^{w} = 1 \text{ iff } \{d : \mathbf{N}(\lambda w'.[\![\phi]\!]^{w'}(d){=}1)\} \subset [\![\phi]\!]^{w}$

where **N** holds of a proposition ψ iff ψ is true in at least one world w' that is compatible with expectation (i.e. that belongs to the set of worlds Exp).

(141) $\mathbf{N}(\psi) \text{ iff } \exists w'(w' \in \text{Exp} \ \&\ \psi(w') = 1)$

According to the syncategorematic definition in (140), POS holds (in w) of a set of degrees ϕ iff ϕ (in w) properly contains those degrees that it would contain in ordinary/neutral worlds. Take for example (142).

(142) [POS [1 [John is t_1-tall]]]

The semantics of (142) is composed in (143):

(143)
$$
\begin{aligned}
[\![(142)]\!] &= 1 \text{ iff } \{d : \mathbf{N}(\lambda w'.[\![1 \text{ [John is } t_1\text{-tall]]}]\!]^{w'}(d){=}1)\} \\
&\qquad \subset [\![1 \text{ [John is } t_1\text{-tall]]}]\!]^{w} \\
&= 1 \text{ iff } \{d : \mathbf{N}(\lambda w'.[\![\text{John is } t_1\text{-tall}]\!]^{w',1/d}{=}1)\} \\
&\qquad \subset [\![1 \text{ [John is } t_1\text{-tall]]}]\!]^{w}
\end{aligned}
$$

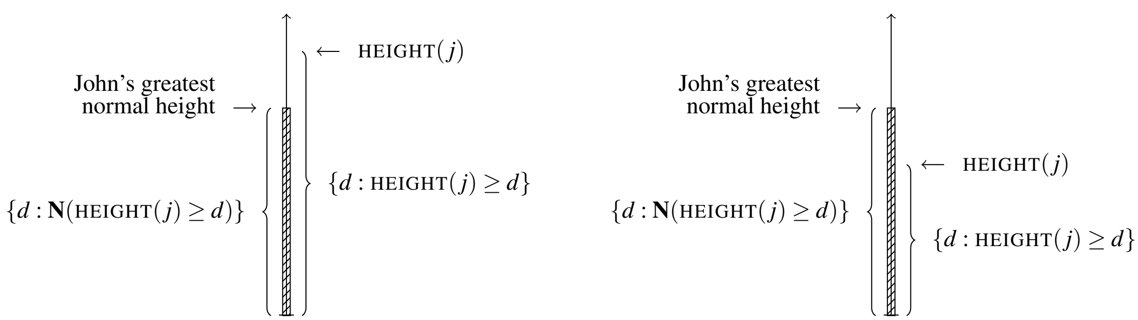
The truth conditions in (143) compares two sets of degrees: the degrees d such that John's height is greater or equal to d in the actual world, and the degrees d such that it is ordinary that John's height is greater or equal to d, that is, those degrees d for which some world w

that is compatible with expectation is a world where John's height is greater or equal to d. (From now on I occasionally omit abstraction over worlds from **N**'s argument, in order to simplify the notation).

There are two important changes that are introduced by (140). First, the definition shifts the extent of von Stechow's/Heim's neutral degrees. To von Stechow and Heim, **N** stands somewhere within the height continuum (in the case of *tall* and *short*). Now **N** is defined as a function of propositions, and POS refers to degrees that satisfy $\mathbf{N}(\phi)$ for some proposition ϕ. In (142), ϕ is the proposition that John is d-tall, so POS refers to those degrees d for which **N** maps 'John is d-tall' to true, i.e. for which there is a world w that is compatible with what is expected, and John's height is at least d in w. If it is compatible with expectation that John's height reach 170 cm, then **N** will hold of $\textsc{height}(john) \geq d$ for every d at 170 cm *and below*, because there is a world in Exp where John's height is greater or equal to 170, 169, 168, etc. (144) demonstrates the new truth conditions (note that the grey zone now stretches down to the bottom of the height scale).

(144)

(a) (b)



$[\![$John is POS-tall$]\!] = 1$ $[\![$John is POS-tall$]\!] = 0$

The new zone **N** does not affect the truth conditions that were derived on von Stechow's and Heim's proposals (compare (144) with (135)); the requirements in both entries to POS amount to the requirement that John's actual height exceed his greatest expected height.

Let us now look at the case of the negative antonym *short* (=[ANT-*tall*]). It will be seen that here too the revision in (140) is innocuous. The sentence 'John is POS-short' has the LF in (145), repeated from (137).

(145) [POS 1 [t_1-ANT [2 [John is t_2-tall]]]]

On the new entry for POS, the truth conditions for (145) are as follows:

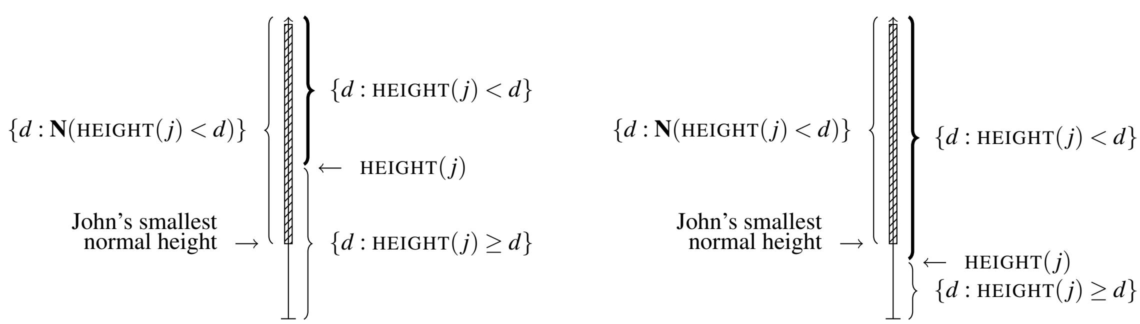
(146) $[\![\text{POS } 1 \text{ [}t_1\text{-ANT [2 [John is } t_2\text{-tall]]]}]\!] = 1$ iff

$$\{d : \mathbf{N}([\![t_1\text{-ANT [2 [John is } t_2\text{-tall]]]}]\!]^{d/1})\}$$
$$\subset \{d : [\![t_1\text{-ANT [2 [John is } t_2\text{-tall]]]}]\!]^{d/1}\}$$

The condition in (146) demands proper inclusion between two sets of degrees. The first is the set of degrees *d* for which it is normal that John's height does not reach *d*. The second is the set of degrees *d* that John's actual height does not reach. To see what degrees belong to the first set, assume a range of heights that John is expected to have, e.g. 160-170 cm. If it is normal for John to have a maximal height of 170 cm, then it is normal for him to *not* reach 171 cm, 172 cm, etc. By the same reasoning, if it is normal for John to have a maximal height of 165 cm, then it is also normal for him to not reach 166, 167, etc. Finally, if it is normal for John to stand at 160 cm, then it is normal for him to not reach 161, 162, etc. The first set in (146) is therefore the interval whose lower bound lies just above the smallest normal height for John, i.e. just above 160 cm, and ranges up indefinitely. The lower end of this interval is the same as the lower end of von Stechow's/Heim's **N**, but its upper end is infinitely high. This is depicted in (147).

(147)

(a) (b)



$[\![\text{John is POS-short}]\!] = 0$ $[\![\text{John is POS-short}]\!] = 1$

The set-inclusion requirement imposed by POS can only be satisfied if John's height falls short of all the degrees of height that he is expected to fall short of. This can only be true if John's maximal height falls below the lowest normal height, which is the same result we get from von Stechow's/Heim's entry for POS.

The second change introduced by (140), my redefintion of POS, is that the sentence 'John is POS-tall' is true iff John's height includes those degrees of height that *he* would be expected to have. Technically, this is different from the earlier definition of POS, because the new truth conditions refer to what is expected of *John*'s height, rather than some neutral set of degrees. If the subject in our sentence was *Mary* instead of *John*, the argument of **N** would change accordingly, and the truth conditions will compare Mary's degrees of height to the degrees of height that *she* is expected to have.

I say that this difference is technical because it is not clear what its real consequences are. One might think that, since the neutral interval is now a function of POS's argument, we expect the requirements of POS to vary depending on the clause it appears in. But there are reasons to be skeptical about this. The original definition of POS is not committed to a rigid extension of **N**, because the interval is provided contextually, and it is not obvious that contextual parameters should resist the influence of the utterances that they help interpret. If this is right, we conclude that the original definition of POS does not commit us to a more restrictive characterization of **N** than the revision in (140). For similar reasons, namely the interference of context, we cannot conclude that the revised definition of POS *forces* covariance with its input just because its input plays a role in the definition of **N**. So the claim that Mary is POS-tall, and the claim that John is POS-tall, do not *have to* refer to different expectations, because context (including the interests and purposes of the interlocutors) may allow total overlap between the degrees d that satisfy **N**(John is d-tall), and the degrees d that satisfy **N**(Mary is d-tall).

The semantics of *very*

Heim takes *very* to refer to a super-interval of **N**, call it $\mathbf{N}^+$. The intensifier's truth conditions are otherwise the same as those of POS:

(148) $[\![\text{very}]\!]^{\mathbf{N}} = \lambda D\,.\,\mathbf{N}^+ \subset D$, where $\mathbf{N} \subset \mathbf{N}^+$

From (148) it follows correctly that whenever [John is very tall] is true, [John is POS-tall] is also true. The conditions of the first sentence require that his degrees of height include the interval $\mathbf{N}^+$. If they do, then his degrees of height must also include **N**, because they include $\mathbf{N}^+$, and $\mathbf{N}^+$ includes **N**.

An intuitive way of viewing $\mathbf{N}^+$ on our approach, where **N** (and $\mathbf{N}^+$ like it) hold of propositions, is to imagine $\mathbf{N}^+$ as an expansion of what is normal/likely. Suppose we add more worlds to Exp, the set of worlds that are compatible with expectations. Let us call the result Exp^+. Suppose now that $\mathbf{N}^+$ is defined just like **N**, but with reference to

Exp^+ instead of Exp. This relaxes the conditions on what counts as expected, because adding more worlds to Exp allows more propositions to count as **N**. So if $\mathbf{N}^+$ contains more propositions of the form [John is d-tall] than **N**, then the set of degrees d that satisfy $\mathbf{N}^+$(John is d-tall) will contain in it the set of ds that satisfy **N**(John is d-tall). This, then, preserves the relation between Heim's **N** and $\mathbf{N}^+$, and *very*-tall (for example) is correctly predicted to be stronger than POS-tall.

(149) $[\![\mathit{very}\ \phi]\!]^w = 1$ iff $\{d : \mathbf{N}^+(\lambda w'.[\![\phi]\!]^{w'}(d){=}1)\} \subset [\![\phi]\!]^w$

4.3 Mirativity and the semantics of *only*

In this section I propose a formulation of the mirative component of *only*. Before I do, I want to note that Zeevat, though the first to connect mirativity to the semantics of *only*, was not the first to study the inference. An earlier attempt at formulating similar ideas is found in Klinedinst (2005), who calls the relevant component the 'scalar presupposition' of *only*. I will stick to Zeevat's terminology, in keeping with my strategy of linking the inference to the one licensed by POS.

To Klinedinst, *only*'s scalar presupposition says that the prejacent is 'low' on the contextually salient scale (I will come back to this idea shortly). Zeevat, on the other hand, chooses to discuss specific examples where *only* quantifies over individuals. In (150) for example,

(150) Yesterday, only Ronald went shopping

the mirative presupposition is that another (salient) individual, e.g. Susan, is expected to have gone shopping also. This is represented in (151), where r is Ronald, s Susan, and S the property of having gone shopping.[6]

(151) $[\![(150)]\!]$ is defined only if $exp(S(r+s))$. If defined, $[\![(150)]\!]$ is true iff $\neg S(s)$

Klinedinst, having noted the generality of the mirative component, and also the variability of the type of scale that *only* can operate on, offers a broader definition of the presupposition:

(152) **Klinedinst's Scalar Presupposition of *only***:
Given a scale σ, [$only_\sigma$ S] is defined only if S is low on σ.

[6](150) and (151) correspond to Zeevat's (11) and (12).

(153) S is low on σ iff there are sufficiently many alternatives $S' \in \text{ALT}(S)$ such that $S' >_{\sigma} S$.

The lowness that Klinedinst has in mind can be connected to Zeevat's criterion of expectation: if Ronald and Susan are expected to go shopping, then the alternative sentence [*Ronald and Susan went shopping*] must be relevant, and so it must appear in σ. Since on this scale the prejacent is lower than the conjunctive alternative, being asymmetrically entailed by it, we may decide that the prejacent is indeed low, provided that the existence of this one higher alternative satisfies the requirement that 'sufficiently many' of them exist.

Another difference between the two accounts is reference to scales. Klinedinst, like others before him, notes that *only* can have a non-logical reading, where the excluded alternatives to the prejacent are not ranked higher logically, but in some other relevant sense (see among others Jacobs 1983 and Bonomi and Casalegno 1993. See also Riester 2006). (154) is an example.

(154) A: I got an ace. What do you have?
B: I only got [a jack]$_{\text{F}}$.

No hand in poker can consist of a single card, and so the *logical* reading of B's reply in (154) is always false. But this is not what B is understood to say. *Only* is here intended to signal that B has nothing *better* than a jack, so σ here is an ordering of cards/hands by value. Now, Klinedinst's version of the mirative presupposition will require that σ contain sufficiently many alternatives to *jack* on σ, and since σ is a ranking of value, the requirement will be satisfied by the higher alternatives *queen*, *king*, and *ace*, etc. Note that the presupposition will not be satisfied if a good hand is combined with *only*, e.g.

(155) A: I got an ace. What do you have?
B: #I only got a [full house]$_{\text{F}}$.

A full house is in most circumstances a good hand, and in those circumstances it is odd to associate the hand with *only*, as B does in (155).

Zeevat's specific formulation is not equipped to handle cases of non-logical scales, nor even cases where the scale is logical, but where *only* does not quantify over individuals. We may, however, generalize Zeevat's definition as in (156).

(156) ***Only*'s mirative presupposition – All scales**:
Given a scale σ, $[\![only_{\sigma}\ S]\!]$ is defined only if there is an alternative $S' \in \text{ALT}(S)$ such that $S' >_{\sigma} S$ and $\mathbf{N}(\lambda w [\![S']\!]^{w})$.

Applying (156) to B's utterance in (155) yields the requirement that some hand of greater value than a full house be neutral/expected, which in ordinary circumstances is not met.

Since we will be primarily concerned with the scale of logical strength, we will focus on the instantiation of (156) seen in (157). (I use the turnstile $\Vdash$ for *asymmetric entailment*).

(157) ***Only*'s mirative presupposition – Logical scale**:
$[\![only_{\Vdash}\ S]\!]$ is defined only if there is an alternative $S' \in \mathrm{ALT}(S)$ such that $[\![S']\!] \Vdash [\![S]\!]$ and $\mathbf{N}(\lambda w [\![S']\!]^w)$.

A useful consequence of (157) is that neutrality will also be attributed to the prejacent itself. This is because the **N**-alternative S' *entails* the prejacent. So if it is neutral/expected for the stronger alternative to hold, then it must be neutral/expected for its logical consequences, e.g. the prejacent, to also hold.

(158) Consequence of (157)
$[\![only\ S]\!]$ is defined only if $\mathbf{N}(\lambda w [\![S]\!]^w)$.

Let us illustrate (157) by observing the oddness of (159).

(159) #John only has six_{F} children.

In ordinary contexts, (159) is infelicitous, and this is felt to come from combining *only* with what seems to be a higher-than-ordinary number (of children).[7] Assuming that numerals have 'at-least'-like semantics, and assuming a logical ordering, (157) will require that there be a neutral stronger alternative to (159)'s prejacent. This means that it should be neutral

[7]There are two examples from existing texts that I want to point out. I find them interesting because they both seem to explicitly isolate *only*'s mirativity. The first example is a dialog from J. D. Salinger's *A Perfect Day for Bananafish*, which takes place between the protagonist and a child that he befriends:

(1) - Did the tigers run all around that tree?
- I thought they'd never stop. I never saw so many tigers.
- There were only six.
- *Only* six? Do you call that *only*?

The second example is from the foreword to Schlesinger and Kinzer (1999), written by Richard Nuccio:

(2) The "good news" of the IOB report was its conclusion that CIA's local employees had killed no US citizens, "only" Guatemalans.

The mirativity of *only* is marked clearly by the use of quotation marks, which is intended by the author to indicate disagreement with the implication of *only*, namely that its prejacent is low/insufficient.

for John to have 7 or 8 or 9 (etc) kids, which is not true. The oddness of (159) is also derivable from (158), which requires neutrality of the prejacent itself. The requirement is not met, and (159) is correctly ruled out.

4.3.1 Embedded mirativity and *only*

In the previous section I illustrated the mirative component of *only* by associating it with a focus, where the resulting prejacent was felt to be too 'high'. With this condition in the particle's semantics, we were able to derive the oddness of sentences like (159).

I will now examine the behavior of the mirative component when association takes place across some other operators. I will show that some predicates can improve examples like (159), but some others cannot. The relevance of this will be clear in Section 4.4, where I discuss how *only* and POS interact. In this section, my goal is to show embedding environments that, as I will say, "plug" mirativity: these environments have the property that, where *only* cannot take ϕ as its associate due to a mirativity violation, e.g. in (159), embedding ϕ under the mirativity-plugging environment makes association with *only* acceptable. The examples I look at are the verbs *say*/*claim*, and antecedents of conditionals/restrictors of bare plurals. Throughout the discussion, I will informally talk about (159) as a case where the focus is 'high', but I want to remind the reader that more formal details lie behind this, along the lines discussed above.

Embedding verbs

Consider first the verbs *say*/*claim* and *know*. The acceptability of (160) and (161) shows that the mirative component is not violated when any of these verbs intervenes between *only* and its 'high' associate.

(160) Mary only said/claimed that John has six$_{\mathrm{F}}$ children.

(161) Mary only knows that John has six$_{\mathrm{F}}$ children.

Compare these judgements with (162) and (163).

(162) #You only need to have six$_{\mathrm{F}}$ children (in order to qualify for this tax break).

(163) ??You are only allowed to bring six$_{\mathrm{F}}$ cartons of cigarettes into the country.

Let us schematize (roughly) these judgements as in (164,165). The asterisks that indicate unacceptability mark mirativity violations.

(164) $*[\textit{only}\ [\cdots\phi_{\mathrm{F}}\cdots]] \rightarrow *[\textit{only need}\ [\cdots\phi_{\mathrm{F}}\cdots]]$

$$\rightarrow *[\textit{only allowed } [\cdots \phi_{\mathrm{F}} \cdots]]$$

(165) $$*[\textit{only } [\cdots \phi_{\mathrm{F}} \cdots]] \not\rightarrow *[\textit{only say } [\cdots \phi_{\mathrm{F}} \cdots]]$$
$$\not\rightarrow *[\textit{only claim } [\cdots \phi_{\mathrm{F}} \cdots]]$$
$$\not\rightarrow *[\textit{only know } [\cdots \phi_{\mathrm{F}} \cdots]]$$

The claim in (164) is that, when a prejacent *S* violates the mirative presupposition of *only*, embedding *S* under *need* or *allowed* also violates the presupposition. In contrast, embedding *S* under *say/know* seemingly resolves the conflict.

We may attempt to explain (164) by first determining the content of the mirative presupposition. Let us use OP in place of the predicates/modals that appear in the examples.

(166) $[\![\textit{only } \textsc{op}(S)]\!]$ is defined only if $\mathbf{N}(\lambda w' [\![\textsc{op}(S')]\!]^{w'})$

Let us take the case of *need*: the presupposition says that there is a world that is compatible with expectations where you need to have 6 children. Intuitively, the requirement seems excessive, because worlds in which the requirement is met, e.g. worlds where a given person has 6 children, do not fit with what we take to be ordinary circumstances. That is, the set of worlds that are compatible with expectations, Exp, does not include worlds where the given person has 6 children. A requirement is therefore expected/ordinary only if its content is expected/ordinary:[8]

(167) $\mathbf{N}(\textit{need}(\phi)) \rightarrow \mathbf{N}(\phi)$

We therefore derive the oddness of (162) by applying the mirative requirement to the prejacent $\textit{need}(\phi)$, and derive from (167) that ϕ itself has to be ordinary:

(168) $[\![\textit{only}$ need [you have 6_{F} kids]$]\!]$ is defined only if $\mathbf{N}(\lambda w' [\![$need [you have 6 kids]$]\!]^{w'})$,
and, by (167), only if $\mathbf{N}(\lambda w' [\![$you have 6 kids$]\!]^{w'})$

The same rationale does not carry over to cases where OP is a verb like *claim/say* or *know*. If OP $=$ *claim*, then it should be expected for Mary to claim that John has 6 children. This does not necessarily conflict with the number of children that we might *expect* John to have, because unlike laws, the content of Mary's claims need not conform to neutral/ordinary circumstances.

[8]The same can be repeated for *allowed*. Note, crucially, that the explanation refers to the modal base that the operators *need/allowed* quantify over: if *need* did not quantify over 'reasonable' deontic worlds, i.e. worlds that are compatible with what is ordinarily attainable in the evaluation world, then it does not follow from $\mathbf{N}(\textit{need}(\phi))$ that $\mathbf{N}(\phi)$. I leave it to future work to investigate how changing the modal base affects mirativity.

(169) $\mathbf{N}(\textit{say/claim}(\phi)) \not\rightarrow \mathbf{N}(\phi)$,

From this the acceptability of (160) follows. We will soon see that this difference between *need* and *say* has consequences on the acceptability of *many* as *only*'s associate. The verb *say*, which by default plugs mirativity, can acceptably embed *many* as an associate of a higher *only*. This contrasts significantly with *need*:

(170) Mary only said that [(very) many]$_\mathrm{F}$ students failed.

(171) ??John only needs to have aced [(very) many]$_\mathrm{F}$ courses to get into the program.

The mirativity of *only* with conditionals and bare plurals

Because the mirativity of *only* is assumed to apply to full sentences, and because it is assumed to require an **N**-alternative that is stronger than the prejacent, we expect to see a difference when the unacceptably 'high' focus associate is placed in a DE/non-monotone environment, e.g. (172,173).

(172) You can only apply for this tax break if you have six$_\mathrm{F}$ children.

(173) Only parents of six$_\mathrm{F}$ can apply for this tax break.

Earlier we saw that 6 was too 'high' to be compatible with *only*, e.g. in (159).

(159) #John only has six$_\mathrm{F}$ kids.

Here we see that the associate becomes acceptable inside the antecedent of a conditional (172), or the restrictor of a bare plural generic (BP, (173)). I will attempt to capture this contrast by looking at how the monotonicity of these environments interacts with *only*'s mirative presupposition.

I assume that in conditionals and BPs, like in other constructions, *only* takes scope above the entire sentence.[9] The rough LF for e.g. the conditional (172) is by this assumption the structure in (174), and for the BP in (173) it is (175).

(174) *only* [[if you have six$_\mathrm{F}$ children] [you can apply for this tax break]]

(175) *only* [GEN [parents of six$_\mathrm{F}$] [can apply for this tax break]]

An intuitive check on (174/175) shows that the prejacents satisfy what we think are ordinary circumstances. In (175), for example, *only* requires that the following be expected/ordinary: that parents of six can apply for the tax break.

[9]See von Fintel (1997) for a detailed analysis of these constructions.

The alternatives to the prejacents in (174,175) are those where the associate is replaced with other numerals. *only*'s mirative component requires that some stronger alternative S' be expected, i.e. $\mathbf{N}([\![S']\!])$, and owing to the monotonicity of the expressions, stronger alternatives are generated by replacing the focus with (locally) *weaker* alternatives—lower numerals in this case. The requirement is therefore that one of the alternatives in (176/177) be neutral, and by the consequence in (158), repeated, this gives us (178).[10]

(176) [if you have 5/4/... children] [you can apply for this tax break]

(177) [GEN [parents of 5/4/...] [can apply for this tax break]

(158) $[\![$*only S*$]\!]$ is defined only if $\mathbf{N}([\![S]\!])$.

(178) $[\![$(174)$]\!]$ is defined only if $\mathbf{N}([\![$[if you have six children]
[you can apply for this tax break]$]\!])$.

There may be some confusion regarding the alleged neutrality in (178). The claim might make it seem neutral to *require* having 6 kids in order to apply for the tax break, and given the oddness of (162)—recall earlier discussion—the requirement is not neutral.

(162) #You only need to have 6_{F} kids in order to apply for the tax break.

But this is incorrect. The conditional prejacent, which is claimed to be neutral, does not make any requirements. Rather, the conditional issues *permission* to apply for the tax break if you (the applicant) have 6 children. Seen differently, if the rules required that applicants have, say, 3 or more children, then the prejacent conditional is still true, but the prejacent in (162) is obviously false.

The conclusion is that, when *only*'s mirative component conflicts with an associate ϕ, it does not follow that it will conflict with the conditional where ϕ appears in the antecedent of a conditional, nor when ϕ appears in the restrictor of a BP.

(179) $*[\textit{only}\ [\cdots\phi_{\mathrm{F}}\cdots]] \not\rightarrow *[\textit{only}\ [\text{if}\ [\cdots\phi_{\mathrm{F}}\cdots]\ \text{then}\ \psi]]$
$\not\rightarrow *[\textit{only}\ [\text{GEN}\ [\cdots\phi_{\mathrm{F}}\cdots]\ \psi]]$

The reasons for (179) is that, as we've seen above, the property **N** can hold of a conditional without holding of its antecedent (same for BPs): given a conditional/BP that satisfies **N**, the antecedent/restrictor of the conditional/BP need not satisfy **N**:

(180) $\mathbf{N}([\![\text{if}]\!](\phi)(\psi)) \not\rightarrow \mathbf{N}(\phi)$

[10]This assumes that antecedents of conditionals (and restrictors of BPs) are downward entailing, which is disputed. See Bennett (2003).

$$\mathbf{N}([\![\text{GEN}]\!](\phi)(\psi)) \not\rightarrow \mathbf{N}(\phi)$$

A natural question to ask now is whether we can find conditionals/BPs that violate mirativity. In simple cases we found prejacents that conflict with *only* because they were too high, which we interpreted as too strong. If it is possible for *only* to take a full conditional/BP as its prejacent, as I am assuming, then in principle we should find unacceptable *only-if* constructions, and *only*-BPs, whose oddness derives from the conditional's (or BP's) excessive strength. This is exactly what we find: strong conditionals/BPs are those that contain weak antecedents/restrictors, and as is shown in (181,182), combining these constructions with *only* produces odd sentences:

(181) #John will only accept the job if it offers (at least) \$20k$_{\mathrm{F}}$ per year.
only [[if it offers \$20k$_{\mathrm{F}}$] [John will accept the job]]

(182) #Only jobs that pay (at least) \$20k$_{\mathrm{F}}$ are good enough for John.
only [GEN [jobs that pay \$20k$_{\mathrm{F}}$] [are good enough for John]]

The mirative component of *only* requires that some stronger alternative to (e.g.) the prejacent in (181) be expected.

(183) $[\![(181)]\!]$ is defined only if there is an alternative $S' \in \mathrm{ALT}(184)$ such that $\mathbf{N}([\![S']\!])$ and $[\![S']\!] \Vdash [\![(184)]\!]$.

(184) [if it offers \$20k$_{\mathrm{F}}$] [John will accept the job]

And by the consequence in (158), repeated below, we get (185).

(158) $[\![\textit{only } S]\!]$ is defined only if $\mathbf{N}([\![S]\!])$.

(185) $[\![(181)]\!]$ is defined only if $\mathbf{N}([\![(184)]\!])$.

But (184) is not neutral, so the conditional is correctly predicted to be odd. Note that here we have non-neutral conditionals that contain neutral antecedents; to see the neutrality of the antecedent, recall that from $\mathbf{N}([\![\text{the job offers \$45k}]\!])$ it follows (logically) that $\mathbf{N}([\![\text{the job offers \$20k}]\!])$.

The findings relating to conditionals/BPs is that excessively strong prejacents, whose strength renders them incompatible with *only*, can appear as antecedents/restrictors of conditionals/BPs that are compatible with *only*. This, I will show, makes it possible to have *many* appear grammatically as *only*'s associate when *many* is placed in either of these environments:

(186) I will only go if [(very) *many*]$_{\mathrm{F}}$ people go.

4.3.2 Section summary

I proposed a formulation of *only*'s mirative component, largely following the proposals of Zeevat and Klinedinst, and showed cases where it is obviated. In the next section I will propose that the mirative condition conflicts with the meanings of POS and *very*. This, I will argue, is the reason why the logical reading of *only*-*many*$_{\mathrm{F}}$ and *only*-*few*$_{\mathrm{F}}$ is unavailable. Specifically, I argue that *only*, by its mirative component, requires neutrality of its prejacent. In contrast, POS holds of a set of degrees only if it extends outside of what is neutral. This makes the two incompatible.

An important prediction of the analysis is that, when an intervening operator OP 'plugs' neutrality, i.e. has the property in (187), *only* and POS become compatible again.

(187) $\mathbf{N}(\textsc{op}(\phi)) \not\rightarrow \mathbf{N}(\phi)$

The details of how this prediction is derived will be shown in Section 4.4. The data that (I claim) support this prediction are summarized in (188-194). I add here cases where *only*-[(*very*) *few*] restores its logical reading: where *only* is compatible with *many*, *only* with *few* is interpreted on a logical scale, i.e. where *few*'s stronger alternatives are negated.

(188) #John only needs to have aced [(very) many]$_{\mathrm{F}}$ courses to get into the program.

(189) John only needs to have aced [(very) few]$_{\mathrm{F}}$ courses to get into the program. (*non-logical exclusion*).

(190) #John only needs to have failed [(very) few]$_{\mathrm{F}}$ courses to get into the program. (*logical exclusion unavailable*, intended, but unattested meaning: John does not need to have failed *no* courses).

(191) Mary only said that [(very) many]$_{\mathrm{F}}$ students failed.

(192) Mary only said that [(very) few]$_{\mathrm{F}}$ students passed. (*logical exclusion*: Mary did not say that *no* students passed.)

(193) I will only go if [(very) *many*]$_{\mathrm{F}}$ people go.

(194) I will only go if [(very) *few*]$_{\mathrm{F}}$ people go. (*reading is logical – negates all non-weaker conditionals*: e.g. I will not go if *many*$_{\mathrm{F}}$ people go).[11]

[11]Note that (194) is (correctly) predicted to not negate weaker conditionals, i.e. those where *few* is replaced with its locally-stronger alternative *no*. In other words, the sentence does not entail the negation of $[\![$I will go if no people go$]\!]$.

4.4 *only* $[\cdots \text{POS}_F \cdots]$ is contradictory

In this section I examine cases of associating the positive morpheme with *only*. The findings that we want to capture are the following:

(195) (i) *[*only* $[\cdots \text{POS}_F \cdots]$]
(ii) *[*only* [*need* $[\cdots \text{POS}_F \cdots]$]]
*[*only* [*allowed* $[\cdots \text{POS}_F \cdots]$]]
(iii) ✓[*only* [*say* $[\cdots \text{POS}_F \cdots]$]]
✓[*only* [*claim* $[\cdots \text{POS}_F \cdots]$]]
(iv) ✓[*only* [*if* $[\cdots \text{POS}_F \cdots]$ *then* $\cdots$]]
✓[*only* [GEN $[_{NP} \cdots \text{POS}_F \cdots]$ VP]]

My theoretical assumptions are the following:

(196) (i) Movement/decomposition at LF preserves focus marking (See Section 3.3).
(ii) Unless specified otherwise, focus alternatives must be of the same semantic type as the focus.
(iii) Alternatives may be generated by structural simplification: S' is an alternative to S if S' is a subconstituent of S (Katzir 2007, Fox and Katzir 2011).
(iv) The focus alternatives to POS/*very* are degree names of semantic type d.
(v) *only* has a mirative presupposition.

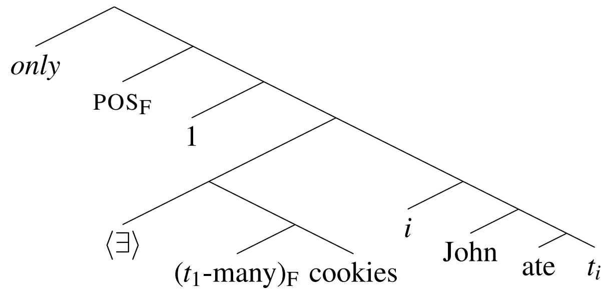
Before I show how the analysis works, I want to mention two other attempts at capturing the *only*-POS incompatibility, without resorting to *only*'s mirative component. Each of these attempts relies on a subset of the assumptions in (196). The first drops assumptions (iv) and (v): POS/*very* are assumed to have no alternatives at all, and the mirativity of *only* is not made use of. The second attempt drops (v), and instead treats POS/*very* as comparatives. Both analyses are discussed in the appendix to this chapter. The discussion shows that they fail to account for (195) in full: the first approach fails (195iii,iv), where *many* is acceptable as *only*'s associate, and the second incorrectly predicts *need* to allow *only*-POS association.

4.4.1 *only*, POS, and mirativity

Consider (197), and its LF in (198).

(197) #John only ate (POS-*many*)$_F$ cookies.

(198)



From (196iv) we get as alternatives to the prejacent every LF where POS is replaced with a degree name d.

(196iv) The focus alternatives to POS/*very* are degree names of semantic type d.

Below I repeat the definitions of *only*'s mirative presupposition, (157), together with the definition of POS, (140).

(157) ***Only*'s miratative presupposition – Logical scale**:
$[\![only\ S]\!]$ is defined only if there is an alternative $S' \in \text{ALT}(S)$ such that $[\![S']\!] \Vdash [\![S]\!]$ and $\mathbf{N}(\lambda w'[\![S']\!]^{w'})$.

(158) Consequence of (157)
$[\![only\ S]\!]$ is defined only if $\mathbf{N}(\lambda w[\![S]\!]^{w})$.

(140) $[\![\text{POS}\ \phi]\!]^w = 1$ iff $\{d : \mathbf{N}(\lambda w'.[\![\phi]\!]^{w'}(d){=}1)\} \subset [\![\phi]\!]^w$

(141) $\mathbf{N}(\psi)$ iff $\exists w(w \in \text{Exp}\ \&\ \psi(w) = 1)$

Let us now see how an associate containing POS, as in (198), interacts with *only*'s mirative component. By (158),

(199) $[\![only\ [\text{POS}_\text{F}\ \phi]]\!]$ is defined only if $\mathbf{N}(\lambda w[\![\text{POS}\ \phi]\!]^{w})$, i.e.
only if $\exists w(w \in \text{Exp}\,\&\,[\![\text{POS}\ \phi]\!]^{w}{=}1)$

The condition in (199) requires that [POS ϕ] be true in some world that is compatible with expectation. But POS, by its semantics, requires that its argument exceed expectation. The derivation is continued below:

(200) ... defined only if $\exists w(w \in \text{Exp}\,\&\,\{d : \mathbf{N}(\lambda w'.[\![\phi]\!]^{w'}(d){=}1)\} \subset [\![\phi]\!]^{w})$, i.e.
only if $\exists w(w \in \text{Exp}\,\&\,\{d : \exists w'(w' \in \text{Exp}\,\&\,[\![\phi]\!]^{w'}(d){=}1)\} \subset [\![\phi]\!]^{w})$

only if $\exists w(w \in \text{Exp}\,\&\, [\![\phi]\!]^w \subset [\![\phi]\!]^w)$

The sentence is defined only if, in some world w that is compatible with expectations, ϕ properly includes any degree that it maps to True in any world that is compatible with expectations. But w is itself compatible with expectations, so in order for the requirement to be met, ϕ in w must *properly* include the degrees that it maps to True in w. In other words, ϕ must be a proper subset of itself. (199) is therefore undefined.

From this it follows from the mirative presupposition is incompatible with POS, if we assume that **N** is uniform. (Recall that the uniformity of **N** was empirically motivated in Section 4.1). The same inconsistency arises when *very* is associated with *only*:

(201) $[\![$*only* [*very*$_F$ ϕ]$]\!]$ is defined only if $\mathbf{N}(\lambda w[\![\textit{very}\ \phi]\!]^w)$,
and since $[\![\textit{very}\ \phi]\!]$ entails $[\![\text{POS}\ \phi]\!]$,
$[\![$*only* [*very*$_F$ ϕ]$]\!]$ is defined only if $\mathbf{N}(\lambda w[\![\text{POS}\ \phi]\!]^w)$, ... (same as (199))

Embedding POS

Let us now turn to cases where *only* associates with POS/*very* across the environments discussed in Section 4.3.1. We will see that the generalizations above will lead to different predictions for the different operators. If mirativity permeates through the embedding OP, as in the case of *need*/*allowed*, we predict ungrammaticality for the positive, and unavailability of the logical reading of the negative:

(202) $[\![$*only* [OP(POS$_F$ ϕ)]$]\!]$ is defined only if $\mathbf{N}(\lambda w[\![\text{OP}(\text{POS}\ \phi)]\!]^w)$,

If OP, like *need*, is such that $\mathbf{N}(\lambda w[\![\text{OP}(\psi)]\!]^w) \to \mathbf{N}(\lambda w[\![\psi]\!]^w)$, then,

(203) $[\![$*only* [OP(POS$_F$ ϕ)]$]\!]$ is defined only if $\mathbf{N}(\lambda w[\![\text{POS}\ \phi]\!]^w)$

Because OP, e.g. *need*, allows the move from $\mathbf{N}(\text{OP}(\phi))$ to $\mathbf{N}(\phi)$, then from the mirative requirement in (202) we get the requirement that $\mathbf{N}(\lambda w[\![\text{POS}\ \phi]\!]^w)$. But this is the same inconsistent conclusion derived in (199). We predict, then, that an embedding predicate OP which fails to plug mirativity cannot embed an acceptable POS$_F$: by its semantic/pragmatic properties, when OP embeds a proposition ϕ that counts as 'high'/'excessive', then the proposition OP(ϕ) will also count as 'high'/'excessive'. So, when ϕ is too high to be compatible with *only*, then OP(ϕ) will be also. By the derivation in (199), POS violates the mirative requirement of *only*, because $[\![\text{POS}\ \phi]\!]$ is too high to qualify as **N**. This means that $[\![\text{OP}(\text{POS}\ \phi)]\!]$ can never have the property **N** either, which makes it unassociable with *only* for the same reason.

(162) #You only need to have six$_{\mathrm{F}}$ children (in order to qualify for this tax break).

(163) ??You are only allowed to bring in six$_{\mathrm{F}}$ cartons of cigarettes.

(188) #John only needs to ace [(very) many]$_{\mathrm{F}}$ courses to get into the program.

(189) John only needs to ace [(very) few]$_{\mathrm{F}}$ courses to get into the program. (*non-logical exclusion*).

(190) #John only needs to fail [(very) few]$_{\mathrm{F}}$ courses to get into the program. (*logical exclusion unavailable*: the sentence cannot mean that it is not needed that John fail *no* courses).

If, on the other hand, OP plugs mirativity, (that is, if $\mathbf{N}(\text{OP}(\phi)) \not\rightarrow \mathbf{N}(\phi)$), e.g. *say/claim*, or the restrictor of a conditional/BP, then we do not derive the inconsistency:

(204) $[\![only\ [claim(\text{POS}_{\mathrm{F}}\ \phi)]]\!]$ is defined only if $\mathbf{N}(\lambda w[\![claim(\text{POS}\ \phi)]\!]^w)$

But $\mathbf{N}(\lambda w[\![claim(\psi)]\!]^w) \not\rightarrow \mathbf{N}(\lambda w[\![\psi]\!]^w)$,

So $[\![only\ [claim(\text{POS}_{\mathrm{F}}\ \phi)]]\!]$ does not require $\mathbf{N}(\lambda w[\![\text{POS}\ \phi]\!]^w)$

From this it follows that, in these cases, *only*-POS may be read logically, as is indeed the case.

(191) Mary only said that [(very) many]$_{\mathrm{F}}$ students failed.

(192) Mary only said that [(very) few]$_{\mathrm{F}}$ students passed.

(193) I will only go if [(very) *many*]$_{\mathrm{F}}$ people go.

(194) I will only go if [(very) *few*]$_{\mathrm{F}}$ people go. (*reading is logical – negates all non-weaker conditionals*: e.g. I will not go if *many*$_{\mathrm{F}}$ people go; *does not negate weaker conditionals* *I will not go if no people go.)

4.5 Direct measure phrases and evaluativity

We have so far seen cases that, by and large, seem to support the *o*PN: when association with POS-*many* is ungrammatical, POS-*few* can only have a non-logical reading; when association with POS-*many* is grammatical, e.g. across the verbs *say/claim*, and into the restrictor of a conditional/BP, POS-*few* is read logically.

I will now show a case where *only* is associable with the positive, but where association with the negative favors a non-logical reading. This is not a counterexample to the *o*PN, because the generalization makes claims about *un*associable positives and their negative counterparts. The case is nonetheless important to consider

In (205) *many* takes a (direct) measure phrase, and associates acceptably with *only*.

(205) If we only have [that-many]$_{\mathrm{F}}$ attendants, the course might be cancelled.

Interestingly we find that in the negative case, (206), association with *only* results in a non-logical reading.

(206) If we only have [that few]$_{\mathrm{F}}$ attendants, the course might be cancelled.

Let us sharpen the judgement further. In (207) the speaker talks about a recent meeting in which s/he was in charge of the headcount, and of ensuring whether enough people were there to get the meeting started:

(207) We needed *n* participants to have quorum, and after the head count it turned out that we could start the meeting. #Thank god we only had [that-few]$_{\mathrm{F}}$ attendants.

(207) is odd. What would the sentence mean if it could have a logical reading? The speaker is grateful that in the meeting they had *that*-few participants, and not fewer. S/he should therefore be understood to say that s/he is grateful that there were enough people for quorum, so the meeting could get started. Nevertheless the sentence is odd.

Building on Bierwisch (1989), Rett (2008) has recently argued that even in the presence of direct measure phrases, negative antonyms are "evaluative". Empirically this means that a sentence like '*that-few people arrived*' licenses the inference that '*few people arrived*', which, on our theoretical set-up, means that the sentences entails 'POS-*few people arrived*'.

(208) $[\![\text{that-}\mathit{few}\ \phi]\!] \vDash [\![\text{POS-}\mathit{few}\ \phi]\!]$

If we take (208) as a lesson from Rett, then we can see why (206) is ungrammatical on its logical reading:

(209) $[\![\mathit{only}\ [[\text{that-few}]_{\mathrm{F}}\ \phi]]\!]$ is defined only if $\mathbf{N}(\lambda w[\![[\text{that-few}]_{\mathrm{F}}\ \phi]\!]^{w})$,

and since $[\![\text{that-}\mathit{few}\ \phi]\!] \vDash [\![\text{POS-}\mathit{few}\ \phi]\!]$,

(210) $[\![\mathit{only}\ [[\text{that-few}]_{\mathrm{F}}\ \phi]]\!]$ is defined only if $\mathbf{N}([\![[\text{POS-few}]_{\mathrm{F}}\ \phi]\!])$,

which is false by (199).

Rett's analysis relies generally on a blocking effect: the extra "evaluative" inference (the inference to POS) in the case of *few* is triggered in order to save the construction from having the same meaning as its positive counterpart.[12] Whether Rett's account is the correct one is less important for our analysis than her empirical generalization: the claim that POS is entailed by direct measure phrases (in the negative case) gives us the prediction that only the non-logical reading is available in sentences like (206).

4.6 Chapter Summary

In this chapter I addressed the question of why *only* cannot grammatically associate with *many*, and why *only* does not give rise to the predicted logical reading when it associates with (*very*) *few*. I proposed that the two questions have the same answer, which is that the positive morpheme (or *very*) is semantically incompatible with the mirative presupposition of *only*: POS/*very* entails a kind of 'excession', while *only* demands neutrality. In support of this claim, I showed environments where *only* can take an excessive associate, e.g. under the verb *say* and under the antecedent/restrictor of a conditional/BP. I showed that in these cases, POS/*very* do not enter into the same conflict with *only* that they do in other environments. This, I argued, predicts correctly that *many* and *few* retain their acceptability, and predicted logical readings, when they appear as associates to *only*.

4.7 Appendix: two failed attempts at capturing the *only*-POS/*very* incompatibility

Attempt 1: $\text{ALT}(\text{POS}) = \emptyset$. If no alternatives are generated for POS/*very*, we predict that both *only-P* and *only-N* be infelicitous, because the set of focus alternatives will be empty. Obviously this fails to account for cases where association between *only* and POS is acceptable, e.g. under *say* and across conditionals/BPs.

Attempt 2: POS as a comparative. Another possibility is to use F&H's account of comparatives (see Chapter 6 for a detailed review). Readers familiar with the literature on the comparative may have already noticed some similarity in the semantic accounts of POS on the one hand, and *-er* on the other. My official take on both morphemes is that they require a subsethood relation to hold (following Heim 2006 on the comparative, and von Stechow

[12] Rett considers many other gradable constructions. For comparatives she observes that *fewer* and *more* are semantically not equivalent, and so the blocking effect is not triggered. The antonym is therefore predicted not to be evaluative, correctly: '*John ate fewer cookies than Bill*' does not imply that John ate *few* cookies. See Rett for details, and see also Breakstone (2012) for a reply.

on POS). Given the similarity, why not treat POS essentially as a comparative, and explain its (in)compatibility with *only* in the same way? To see what this may involve, let us use the simple entry for POS in (211). (This discussion presupposes some familiarity with Fox and Hackl 2006 – see Chapter 6).

(211) $[\![\text{POS}_{\mathbf{s}}\text{-many}]\!] = \lambda D.\max(D) > \mathbf{s}$

In (211) I assume that the standard **s** is supplied in the logical form of the sentence. I do not wish to commit to this view here, nor do I want to attribute it to any of the researchers who worked on the topic. What I want to do is try to extend F&H's account of the comparative to POS, and in doing so, I want to make POS (and its alternatives) as similar as possible to the comparative (and its alternatives). This way, we will be able to see with some certainty whether F&H's density-based account can be made use of with POS. In the end I will argue that it can't.

Let us add the following assumption about the alternatives to POS.

(196iv′) $\text{ALT}(\text{POS}_{\mathbf{s}}) = \{\text{POS}_d : d \in D\}$

The parallel to the comparative can now be seen: 'John ate more than 3 cookies' is true iff the number of cookies that John ate exceeds 3. On the definition of POS in (211), 'John ate $\text{POS}_{\mathbf{s}}$-many cookies' is true iff the number of cookies that John ate exceeds **s**.

Now that POS is defined like a comparative, and now that its alternatives are also 'comparative'-like sentences, where the standard **s** is replaced with a degree name *d*, we expect POS to mirror the behavior of comparatives when they associate with *only*. An important finding from F&H is that comparatives can acceptably appear with *only* if they are embedded under universal operators, e.g. the modal *need*, but as we saw earlier, POS does not associate well with *only* across the predicate.

(212) ✓John only needs to ace more than 3_F courses to get into the program.

(213) #John only needs to ace [(very) many]$_F$ courses to get into the program.

(214) This program isn't so tough. In order the make the short list, you only need to have failed less than 3_F courses. (logical reading available: it is not necessary for you to have failed less than 2).

(215) #This program isn't so tough. In order to make the short list, you only need to have failed [very few]$_F$ courses. (logical reading unavailable: *it is not necessary for you have failed no courses).

Below I summarize the predictions of the two attempts sketched above. The checkmarks indicate where they succeed, and the $\times$s indicate where they do not.

			1: ALT($\textsc{pos}$) $= \emptyset$	2: Comparative $\textsc{pos}$
(195)	(i)	*[*only* [$\cdots$ $\textsc{pos}_{\mathrm{F}}$ $\cdots$]]	✓	✓
	(iia)	*[*only* [*need* [$\cdots$ $\textsc{pos}_{\mathrm{F}}$ $\cdots$]]]	✓	$\times$
	(iib)	*[*only* [*allowed* [$\cdots$ $\textsc{pos}_{\mathrm{F}}$ $\cdots$]]]	✓	✓
	(iii)	✓[*only* [*say/claim* [$\cdots$ $\textsc{pos}_{\mathrm{F}}$ $\cdots$]]]	$\times$	✓
	(iv)	✓[*only* [*if* [$\cdots$ $\textsc{pos}_{\mathrm{F}}$ $\cdots$] *then* $\cdots$]]	$\times$	✓

Chapter 5

At least/At most

In this chapter I consider two recent accounts of *at least*/*at most*. Much of the literature on modified numerals is dedicated to studying the ignorance inferences that the modifiers give rise to, see e.g. Krifka (1999), Geurts and Nouwen (2007), Nouwen (2010, 2012), Spector (2011), Mayr (2011), Krifka and Cohen (2011), and Schwarz et al. (2012). The ignorance inference of *at least* is shown in examples like (216), which is claimed to contrast with (217).

(216) #A triangle has at least 2 sides.

(217) A triangle has more than 2 sides.

A central concern of the works cited above is what makes (216) odd. Specifically, does *at least* force an ignorance interpretation of the sentences it appears in? And if so, why?

I restrict my attention to two proposals, Büring (2008) and Schwarz (2011), which as I will show are similar for our purposes. The answers provided by both accounts (regarding the ignorance inference) extend straightforwardly as answers to our question: why *at least*/*at most* cannot associate with *only*. We will see, however, that without further amendments, the accounts predict acceptability between *only* and *at least*/*at most* when association takes place across a universal operator. This predicts, as I will show, is incorrect. Ultimately I consider the possibility that, like POS and *very*, these modified numerals are mirative.

I take [*at least n*] (and [*at most n*]) to denote generalized quantifiers over degrees that, at LF, raise to a clausal position where they can be interpreted, and leave a trace of type *d* in their base position. The lexical entries for the two expressions are shown in (218,219). Note that I remain noncommittal here about the internal composition of the phrases: how *at least* composes semantically with the numeral to generate (218) is a question that I leave unaddressed (but see e.g. Krifka 1999, Geurts and Nouwen 2007, Krifka and Cohen 2011).

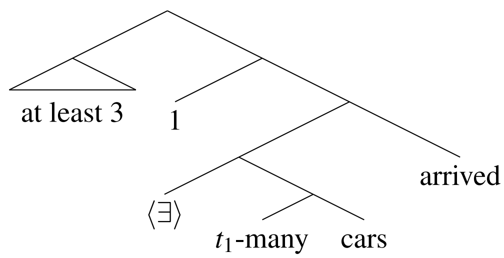
(218) $[\![at\ least\ n]\!] = \lambda D\,.\,\max(D) \geq n$

(219) $[\![at\ most\ n]\!] = \lambda D\,.\,\max(D) \leq n$

An example LF is shown in (220).

(220) At least 3 cars arrived.

(221)



(222) $[\![(221)]\!] = [\![\text{at least } n]\!](\lambda d.[\![1\ [\langle\exists\rangle\ t_1\text{-many cars] arrived}]\!]^{d/1})$
$= 1 \text{ iff } \max(\lambda d.[\![1\ [\langle\exists\rangle\ t_1\text{-many cars] arrived}]\!]^{d/1}) \geq n$
$= 1 \text{ iff } \max(\{d : d\text{-many cars arrived}\}) \geq n$

To Büring and Schwarz, the ignorance inference is pragmatic: Büring treats *at least* as a disjunction, and since disjunctions generally license ignorance inferences regarding their disjuncts, e.g. (223), *at least* likewise licenses ignorance if it is taken to look like (224) at LF.

(223) John spoke to Mary or Sue.
$\vDash$ speaker does not know if John spoke to Mary.
$\vDash$ speaker does not know if John spoke to Sue.

(224) [A triangle has at least 2 sides] = [A triangle has (Exh 2) or (more than 2) sides][1]

[1] An important remark is in order regarding (224): the exact reading of the first disjunct (the numeral 2) is derived by prefixing it with the exhaustification operator, which generates the 'exactly' reading of the numeral by negating its stronger alternatives. Without Exh, the disjunction would stand in violation of what is now commonly referred to as Hurford's constraint, which blocks a disjunction if one of its disjuncts entails the other (Hurford 1974), e.g.

(1) #John lives in France or in Paris.

If the grammar is assumed to freely generate both (224) and its unexhaustified variant, the former parse is preferred given Hurford's constraint, as neither of its disjuncts is entailed by the other.

$\vDash$ speaker does not know if a triangle has (Exh 2) sides,

$\vDash$ speaker does not know if a triangle has (more than 2) sides.

The ignorance inference of *at least*-sentence is discussed in the appendix to this chapter. For now let us recall from the discussion of symmetry in Chapter 1 that, if a sentence S has a pair of alternatives S_1 and S_2 whose disjunction is equivalent to S, then neither S_1 nor S_2 is innocently-excludable given S. The reason is that negating both alternatives amounts to negating S itself, so any set $M \in \mathbf{MC}(S)(A)$ that includes one of S_1, S_2 cannot include the other.

(225) For any sentence S, set of alternatives A, and $S_1, S_2 \in A$,
if $[\![S_1]\!] \vee [\![S_2]\!] = S$, then $S_1, S_2 \notin \mathrm{IE}(S)(A)$

If [*at least n*] is just a conventionalized way of pronouncing what (at LF) is a disjunction, then we expect neither disjunct—[Exh n], [*more than n*]—to be excluded by *only*, because neither disjunct is IE. More generally, we find that there cannot be *any* elements in $\mathrm{IE}(S)(A)$, if $S = [\textit{at least } n \ \cdots]$, and if A includes S's disjuncts, in addition to the result of replacing n with other numerals. I prove this in (226). For convenience I replace (Exh n) with ($=n$), and (*more than n*) with ($>n$).

(226) Let S be of the form $(=n \vee >n)$ and $A = \{>d : d \in \mathbf{D}_d\} \cup \{=d : d \in \mathbf{D}_d\}$. Then $\mathrm{IE}(S)(A) = \emptyset$:

Proof: Assume for reductio that it isn't. Then there must be a degree d' such that either $[>d'] \in \mathrm{IE}(S)(A)$, or $[=d'] \in \mathrm{IE}(S)(A)$.

(a): If $[>d'] \in \mathrm{IE}(S)(A)$, then $[>d']$ is in $\bigcap \mathbf{MC}(S)(A)$. Since for any d'', $[=d'']$ is an alternative to $[>n]$, then there is a set $M \in \mathbf{MC}(S)(A)$ that contains $[=d'']$ for every d'' s.t. $n < d'' \leq d'$. The conjunction $\bigwedge\{\neg\phi : \phi \in M\}$ is $\bigwedge\{\neq d'' : n < d'' \leq d'\}$, i.e. nowhere between n and d', and not d'. In conjunction with the utterance $[>n]$, this entails $[>d']$. So $[>d']$ cannot be added to M, because it would otherwise be negated along with every other member of M, resulting in $>d'$ and $\not> d'$. Therefore $>d'$ does not belong to every $M \in \mathbf{MC}(S)(A)$, so $\mathrm{IE}(S)(A)$ cannot contain alternatives of the form $[>d']$.

(b): if $(=d') \in \mathrm{IE}(S)(A)$, then every set $M \in \mathbf{MC}(S)(A)$ includes $[=d']$. Let some M contain $\{=d'' : n < d'' \,\&\, d'' \neq d'\}$. The conjunction $\bigwedge\{\neg\phi : \phi \in M\}$ is consistent with the utterance $[>n]$, and with it, the set entails $[=d']$. Therefore $[=d']$ cannot be in M because it would otherwise be negated, and would make the set inconsistent. Therefore $=d'$ does not belong to every $M \in \text{MCSE}(S)(A)$, so $\text{IE}(S)(A)$ cannot contain alternatives of the form $[=d']$.

As a result of (226), *only* makes no assertoric contribution when it associates with [*at least n*], and because of the vacuity constraint on *only* (see Chapter 1, Section 1.2.3), the utterance is correctly blocked.

Note that (226) can be extended to the case of [*at most n*] if, like [*at least n*], it is assumed to be a disjunction of $\text{Exh}(n)$ and *less than n*. The result in (226) may be repeated for [*at most*] by reversing the direction of the inequality sign where appropriate.

For our purposes, Schwarz's analysis is very similar to Büring's. It differs in that, instead of postulating a disjunctive LF for modified numerals, Schwarz proposes that *at least* has the modifier *exactly* as formal alternative, where *exactly* is treated also as a numeral modifier, with the semantics in (227).

(227) $[\![\textit{exactly } n]\!] = \lambda D\,.\,\max(D) = n$

It is fairly simple to show that, on Schwarz's account, the set $\text{IE}(S)(A)$ will also be empty for a sentence $S = [\textit{at least } n]$, and a set $A = \{[\textit{at least } d] : d \in \mathbf{D}_d\} \cup \{[\textit{exactly } d] : d \in \mathbf{D}_d\}$. The proof, which relies on (226), is shown below. Here *at least* is abbreviated as $\geq$.

(228) Let S be of the form $(\geq n)$ and $A = \{\geq d : d \in \mathbf{D}_d\} \cup \{=d : d \in \mathbf{D}_d\}$. Then $\text{IE}(S)(A) = \emptyset$:

Proof: Assume for reductio that it isn't. Then there must be a degree d' such that either $[\geq d'] \in \text{IE}(S)(A)$, or $[=d'] \in \text{IE}(S)(A)$.

(a): If $[\geq d'] \in \text{IE}(S)(A)$, then $[\ngeq d']$ is consistent with every set in $\bigcap \text{MCSE}(S)(A)$. Since $\ngeq d$ entails $\ngtr d$, the proposition $\ngtr d$ must be consistent with every set in $\bigcap \text{MCSE}(S)(A)$. But step (226a) shows that this is false.

(b): if $(=d') \in \text{IE}(S)(A)$, then every set $M \in \text{MCSE}(S)(A)$ includes $[=d']$. Let some M contain $\{=d'' : d < d'' \,\&\, d'' \neq d'\}$. The conjunction $\bigwedge\{\neg\phi : \phi \in M\}$ is consistent with the utterance $[>d]$, and with it, the set entails $[=d']$. Therefore $[=d']$ cannot be in M because it would otherwise be negated, and would make the set inconsistent. Therefore $=d'$ does not belong to every $M \in \text{MCSE}(S)(A)$, so $\text{IE}(S)(A)$ cannot contain alternatives of the form $[=d']$.

5.1 Obviation under universal modals?

In this section I show cases where association is predicted to be possible between *only* and [*at least n*]. The discussion will take us back to the *o*PN generalization, where I proposed that whenever *only* can take a positive associate (using the term 'positive' loosely), the negative variant is acceptable on its logical reading. We will see that the "symmetry" which was earlier shown to block association with *only*, is disabled under a universal operator. This, in turn, predicts that both *at least* and *at most* (under a universal modal) can appear as associates to the particle, because in both cases the set of IE-alternatives will be non-empty. But while the association is, as a consequence, predicted to have a logical reading, we will see that judgements do not confirm the prediction.

To see how symmetry is obviated under universal operators, consider (229). This is a schematized version of example (10), from Section 1.2.1: it is consistent to hold simultaneously that what is required is at least *one* of p and q, but that p is not required, and that q is not required. In short, $\Box(p \vee q)$ is consistent with the negations $\neg\Box p$ and $\neg\Box q$.

(229) a. $S = \Box(p \vee q)$

b. $A = \{\Box p, \Box q\}$

c. $\mathbf{MC}(S)(A) = \{A\}$

d. $\mathrm{IE}(S)(A) = \bigcap \mathbf{MC}(S)(A) = A$

Let us then consider cases where S is of the form $\Box$[*at least n*]. This embedding obviates the symmetry that, without the embedding $\Box$, leads to an empty IE. Here, just like in (229), it is consistent to hold all of the following simultaneously:

(230) $$\Box(\geq n) \;\&\; \forall d(d > n \rightarrow \neg\Box(\geq d)) \;\&$$
$$\forall d(d \in \mathbf{D}_d \rightarrow \neg\Box(=d))$$

The meaning in (230) states that in every world w the relevant amount is n or more, but not every world is one where the amount is greater than n, (which means that exactly n is allowed), and not every world is one where it is exactly d (for any given d). Because this is a consistent proposition, the set $\mathrm{IE}(S)(A)$, for our modalized S and its alternatives A, will not be empty.

(231) $S = \Box(\textit{at least } n)$,

$$A = \{\Box(\textit{at least } d) : d \in \mathbf{D}_d\} \cup \{\Box(\textit{exactly } d) : d \in \mathbf{D}_d\}$$

(232) $\mathrm{IE}(S)(A) = \{\Box(\textit{at least } d) : d > n\} \cup \{\Box(\textit{exactly } d) : d \in \mathbf{D}_d\}$[2]

Consider now example (233).

(233) ??You only need to exercise for [at least 2 hours]$_{\mathrm{F}}$ a week if you take this pill.

??$\models$ You do not need to exercise for exactly 2 hours

??$\models$ You do not need to exercise for at least 3 hours

Because the introduction of $\Box$ obviates symmetry, as shown in (230), applying Büring's and Schwarz's accounts to *only* predicts acceptability for (233). The sentence should mean that the only requirement (if you take the pill) is for you to exercise for two or more hours a week; it not required of you to exercise for *exactly* two hours, and it is not required of you to exercise for more. But judgements indicate that the sentence is odd, and though to some speakers there is a subtle contrast between (233) and the unembedded variant in (234), the contrast does not seem robust enough to support my attempted application of Büring/Schwarz to *only*.

(234) #John only exercised for [at least 2 hours]$_{\mathrm{F}}$ last week.

Let us set aside the theoretical analysis for a moment. What do we expect given this judgement, and given the *o*PN generalization? Because association with *at least* is odd even across a universal modal, we expect the parallel association with *at most* to strongly prefer a non-logical reading. This is indeed what we find.

(235) ??On the relaxed regulations you are only required to bring [at most \$10,000]$_{\mathrm{F}}$.

??$\models$ You are not required to bring at most \$9,000.

(236) Mary got a lot of points from Round 1, and John didn't. Now it's Round 2, and if John wants to advance to Round 3 he'll have to do really well, because he doesn't have much to stand on from R1. To advance, he needs to get at most 3 incorrect answers (?). Mary is more relaxed. She only needs to get at most 7 incorrect answers. (???)

Neither (235/236) is acceptable. This is not expected under Büring/Schwarz, but as an empirical finding the pattern is in line with the *o*PN. What, then, could be behind the oddness of the examples?

[2] Note that it is consistent to add the negation $\neg\Box(=d)$ even for d's that are smaller than n. This follows from the prejacent, whose meaning entails the stronger negation $\neg\Diamond(=d)$ for these smaller degrees.

5.2 *at least/at most* as mirative items

In Chapter 4 I took von Stechow's (2006) paraphrases of temporal modifiers to diagnose what Zeevat called "mirativity". We saw that sentences like (237) are odd.

(237) Only n_{F} students registered last semester. #And again this semester many students registered.

The judgement indicated that some component of *only* is incompatible with the inferences licensed by *many*, and this was narrowed down to the silent positive morpheme that is thought to accompany it.

The oddness of (238) suggests that *at most* is in a similar way incompatible *many*, and *at least* is likewise incompatible with *few*:

(238) At most n students registered last semester. #And again this semester many students registered.

(239) At least n students registered last semester. #And again this semester few students registered.

It seems then that *at least* licenses the inference that the quantity it selects is high, in the sense that it includes what counts as ordinary or expected. If we take this to indicate that $[\![\textit{at least } \phi]\!]$ entails $[\![\text{POS } \phi]\!]$, $[\![\textit{at most } \phi]\!]$ entails $[\![\text{POS-}\textit{few } \phi]\!]$, then we derive another source of incompatibility between *only* and the modifiers. But unlike symmetry, this conflict is not obviated by modals like *need* and *allow*, as we saw in Chapter 4.

(240) #You only need to have six_{F} children in order to qualify for this tax break.

(241) *You only need to have aced $[\text{(very) many}]_{\mathrm{F}}$ courses to get into the program.

(242) ??You only need to exercise for $[\text{at least } 2]_{\mathrm{F}}$ hours a week if you take this pill.

We did see, however, that mirativity is plugged under e.g. the verb *say*. If mirativity blocks *at least* and *at most* from associating with *only*, embedding them under these verbs should make their logical readings available. (243) shows that this is indeed the case.

(243) Mary only said that $[\text{at least } n]_{\mathrm{F}}$ students failed.
✓She didn't say that more failed.

(244) Mary only said that $[\text{at most } n]_{\mathrm{F}}$ students failed.
✓She didn't say that fewer failed.

I leave it open whether the mirativity of these modified numerals, if indeed present, is derivable from some other property that the constructions are thought to have.

Further issue: *n or more/less* and obviation under universal modals

In the beginning of the chapter I considered Büring's treatment of (*at leastn*)/(*at most n*) as disjunctions, of the form (*n or more*)/(*n or less*)). I showed that, if we assume that the individual disjuncts are alternatives to the disjunction (the LF representation of *at least*), then the set of IE-alternatives is empty. Because of this, association with *only* is predicted to be vacuous, hence ungrammatical. We saw also that when the disjunction is embedded under $\Box$, the set IE is not empty. This gave us the incorrect prediction that sentences like (245) are grammatical:

(245) ??You only need to exercise for [at least 2 hours]$_{\mathrm{F}}$ a week if you take this pill.

I want to point to another case where the same prediction is made, namely that where *only* takes an explicit disjunction as its focus associate.

(246) [**?**]You only need to exercise for [two or more hours]$_{\mathrm{F}}$ a week if you take this pill.

Though the status of (246) is unclear, I want to mention two questions that the example gives rise to. The first is, if the sentence is unacceptable, what makes it unacceptable? The move to claim *at least* and *at most* as mirative items is unlikely to be of help here. For one thing, it is difficult to locate the source of the mirative inference in a disjunction like this.[3] Moreover, it seems that the judgements cited earlier, which were claimed to support the mirative analysis of *at least/at most*, are less convincing in the case of *n or more*:

(247) [3 or more] students registered last semester. (?)And again this semester few students registered.

It does not seem that (247) is as bad as its analog in (248).

(248) At least 3 students registered last semester. #And again this semester few students registered.

The second question is what (246) means for the *o*PN: if the sentence is acceptable, then we expect to detect the logical reading from its antonym [*n or less*] in association with *only*. That is, we expect (249) to be as acceptable as (246).

(249) [**?**]On the relaxed regulations you are only required to bring [\$10,000 or less]$_{\mathrm{F}}$.

[3] As we will see in the next chapter, *only* can associate with comparatives across universal modals, unlike POS and *at least*. So we cannot say that *more* carries a mirative inference. Otherwise we expect it to be unassociable with *only* across verbs like *need*.

?$\models$ You are not required to bring \$9,000 or less.

I do not find any of these judgements reliable, and much more work is needed before the properties of these constructions are investigated in enough detail. This section is merely an illustration of how the data, once gathered properly, might fit into the picture that is outlined in this project.

5.3 Appendix: Ignorance inferences of *at least/at most* on Sauerland (2004)

Sauerland (2004) presents a two-step process for deriving implicatures. In the first step, the utterance S is taken to be believed by the speaker $B(S)$, and for each non-weaker alternative S' it is taken that the speaker does not believe S', $\neg B(S')$.[4] Sauerland calls these the Primary Implicatures.

(250) Given S, and set of alternatives A,
$\mathrm{PI}(S)(A) = \{\neg B(S') : S' \in A \text{ and } \neg B(S') \wedge B(S) \not\models \bot\}$

The second step, sometimes called the 'epistemic' step, promotes primary implicatures $\neg B(S')$ to secondary implicatures $B(\neg S')$. The promotion does not take place if it gives rise to inconsistency with $B(S)$ and the primary implicatures.

(251) Given S, and set of alternatives A, $\mathrm{SI}(S)(A) = \{B(\neg S') : S' \in A \text{ and } B(\neg S') \wedge B(S) \wedge \bigwedge \mathrm{PI}(S)(A) \not\models \bot\}$

Take a disjunctive $S = p \vee q$ as an example:

(252)
$$\begin{aligned}\mathrm{PI}(p \vee q)(\{p,q\}) &= \{\neg B(S') : S' \in \{p,q\} \text{ and } \neg B(S') \wedge B(p \vee q) \not\models \bot\}\\ &= \{\neg B(p), \neg B(q)\}\end{aligned}$$

Now consider the possibility of promoting the PI $\neg B(p)$ to an SI $B(\neg p)$: the promotion must be consistent with $B(p \vee q)$, and with the PIs $\neg B(p)$ and $\neg B(q)$. But a rational speaker who believes $\neg p$ and $p \vee q$ must believe q. Therefore from utterance $B(p \vee q)$ and the attempted promotion $B(\neg p)$ we infer $B(q)$. This contradicts the PI $\neg B(q)$, and so the promotion is blocked (the same holds of $B(\neg q)$).

[4] Note that the negation of the belief operator B does not translate exactly to negating the English predicate *believe*, which is a neg-raising predicate.

(253) $\text{SI}(p \vee q)(\{p,q\}) = \emptyset$

With an empty set of Secondary Implicatures, the pragmatic inferences regarding the individual disjuncts are left as Primary Implicatures, which then means that the speaker believes the disjunction $(p \vee q)$, but does not believe p, and does not believe q, i.e. allows for the possibility that $\neg p$, and allows for the possibility that $\neg q$. This is the ignorance inference.

In the case of *at least* things are more complicated. If we suppose that [*at least n*] has two kinds of alternatives, [*at least d*] and [*exactly d*], for all degrees d, then we will derive a contradictory set of SIs from Sauerland's algorithm. Take (254) as an example.

(254) John ate at least 3 cookies.

(255) $A = \{\geq d : d \in \mathbf{D}_d\} \cup \{=d : d \in \mathbf{D}_d\}$

The set of Primary Implicatures for (254), relative to the assumed alternatives in (255), is shown in (256).

(256) $\text{PI}(254)(255) = \{\neg B(\geq d) : d \in \mathbf{D}_d\} \cup \{\neg B(=d) : d \in \mathbf{D}_d\}$

We will now see that the definition of SIs in (251) will admit too many elements: it is consistent to add $B(\neg(=4))$, for example, to the PIs above and the inference about the utterance $B(\geq 3)$. The *individual* promotion of this implicature does not give rise to any contradictions, because it is rational to believe that John ate a number of cookies from 3 and up, but did not eat exactly 4. This would mean that John ate 3, or 5 or 6 etc. Therefore $B(\neg(=4)) \in \text{SI}(254)(255)$. But by the same logic the promotion is predicted for every alternative $(=d)$, because each promotion is by itself consistent with (believing) the utterance and the set of Primary implicatures. Together, however, the SIs will contradict utterance:

(257) $\text{SI}(254)(255) = \{B(\neq d) : d \geq 3\}$
$= \{B(\neq 3)), B(\neq 4)), B(\neq 5)), B(\neq 6)), \cdots\}$

The problem is bypassed if the SI condition is changed to admit only IE alternatives, of which there are none in the current case (see proof in (228)).

(258) $\text{SI}(S)(A) = \{B(\neg S') : S' \in \text{IE}(S)(A)\}$

Chapter 6

Comparatives, and the Universal Density of Measurement

The *o*PN Generalization claims that no logical reading can be found for *only*+NQ unless its positive counterpart is grammatical. In this chapter I focus on comparative constructions. I will base my discussion mainly on the empirical and theoretical findings of Fox and Hackl (2006), F&H hereafter, who argue that the ungrammaticality of *only*+comparative constructions follows if the domain of degrees is assumed to be formally dense.

The chapter is divided in two parts. In Part 1 I lay out my assumptions about the semantics of the comparative. I use Heim's (2006) analysis of comparison as set-inclusion, an analysis that finds theoretical appeal in its uniform treatment of positive/negative antonyms. I argue, however, that additional assumptions are needed in order to extend the account to *than*-clauses that contain numerals (what set does 3 refer to in 'John has more than 3 children'?) I use findings from Hackl (2000) in addressing the issue, and cite independent reasons to create a set denotation for numerals in *than*-clauses.

In Part 2 I move to (un)associability with *only*. The majority of the discussion is based on F&H. I present their findings first, and offer an informal description of their analysis (Section 6.3). In Section 6.4, I turn back to the *o*PN, and show that what we expect, based on F&H's empirical findings, is indeed what we find: all the cases where *only* cannot associate with a positive comparative will have a strictly non-logical reading when the positive is replaced with its negative antonym. In contrast, cases where the positive comparative *can* be associated with *only* results in an ambiguous sentence (logical/non-logical) when the negative replaces the positive. In Section 6.5 I discuss the consequences of F&H's account on the I(nnocent)E(xclusion)-treatment of *only*, which was motivated in Chapter 1. The discussion of IE continues in Section 6.6, where F&H's account of modalized prejacents is presented. In Section 6.8 I discuss the possibility of placing focus on the entire comparative phrase. The chapter is summarized in Section 6.9.

Part 1

6.1 The comparative as set inclusion

In this section I outline my assumptions about the semantics of the comparative. Because the positive/negative distinction is central to the topic of this study, I will adopt an analysis of the comparative that is uniformly applicable to both. I base my analysis on Heim (2006), who treats the comparative as denoting a relation between two sets of degrees, and requires that the first be a proper subset of the second.

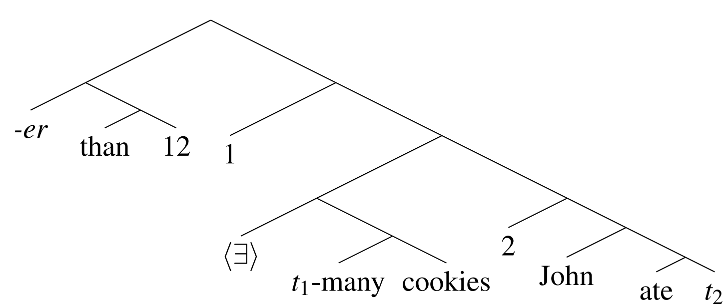
The comparative form applies to positive and negative antonyms alike, e.g. *more*/*fewer*, *taller*/*shorter*, etc. If we maintain the decompositional theory of antonymy (Chapter 3), then we want a analysis of the comparative that is uniformly applicable to both *many* and *few*, where *more* is composed of the morpheme *-er* and *many* (Bresnan 1973), and *fewer* is composed of *-er* and *few*. But since *few* is itself decomposed into ANT and *many*, we want *-er* to be semantically compatible with *many*, in the case of *more*, and with ANT, for *fewer*. *-er* must therefore be compatible with predicates of individuals, type $\langle e,t \rangle$, and with quantifiers over degrees, type $\langle dt,t \rangle$.

One way of achieving this is to treat *-er* (together with its *than*-clause) as denoting a quantifier over degrees, e.g. (259). As an object of type $\langle dt,t \rangle$, the comparative is interpreted at a clausal level at LF, as in (261,262), repeated from Chapter 1.

(259) $[\![\text{-}er]\!] = \lambda d.\lambda D.\max(D) > d$ (to be revised)

(260) John ate more than 12 cookies.

(261)



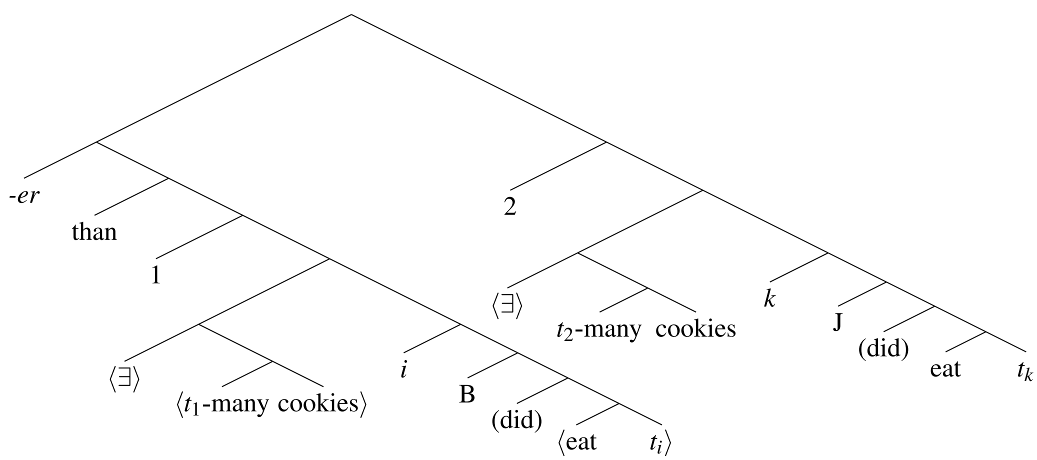
(262) $[\![(261)]\!] = [\![\textit{-er}]\!](12)(\lambda d.[\![[\langle\exists\rangle\ t_1\text{-many cookies}]\ [2\ [\text{J ate } t_2]]]\!]^{d/1})$
$= 1 \text{ iff } \max(\lambda d.[\![[\langle\exists\rangle\ t_1\text{-many cookies}]\ [2\ [\text{J ate } t_2]]]\!]^{d/1}) > 12$
$= 1 \text{ iff } \max(\lambda d.\exists x(|x| \geq d\ \&\ [\![\text{cookies}]\!](x) = 1\ \&\ [\![\text{J ate } t_2]\!]^{x/2} = 1)) > 12$

(260) is true iff the greatest degree d for which the following holds: that John ate a d-sized collection of cookies, is 12. The comparative can also take more complex *than*-clauses, e.g. (263).

(263) John ate more cookies than Bill (did).

Unlike in (260) where the degree 12 is explicitly mentioned, (263) compares degrees that are not lexically denoted by the material in the sentence. Conceptually, the degrees that the sentence compares are those d's that verify [John ate d-many cookies], and those that verify [Bill ate d-many cookies]. In (264) below, just like in (261), the matrix clause denotes the former set of degrees. To get the latter set (Bill's set) from the *than*-clause, we assume that it has the (unpronounced) clausal structure shown in (264). The unpronounced material, marked below with angled brackets, is elided via a process that's come to be known as Comparative Deletion (Bresnan 1973). The ellipsis is conditional on semantic identity with an antecedent phrase, in this case the VP in the matrix clause.[1]

(264)



In order to derive the correct meaning for (263/264), the entry for $[\![\textit{-er}]\!]$ must be rewritten as a relation between two sets of degrees, as in (265).

[1]The identity condition on ellipsis is a complex issue that will be set aside in what follows. In the context of comparatives see e.g. Lechner (2004).

(265) $[\![\textit{-er}]\!] = \lambda D \lambda D' . \max(D) < \max(D')$

(266) $[\![(264)]\!] = [\![\textit{-er}]\!](\lambda d [\![[\langle\exists\rangle\ t_1\text{-many Cs}]\ [i\ [\text{B ate } t_i]]]\!]^{d/1})(\lambda d [\![[\langle\exists\rangle\ t_2\text{-many Cs}]\ [k\ [\text{J ate } t_k]]]\!]^{d/2})$

$= 1 \text{ iff } \max(\lambda d . [\![[\langle\exists\rangle\ t_1\text{-many Cs}]\ [i\ [\text{B ate } t_i]]]\!]^{d/1}) > \max(\lambda d . [\![[\langle\exists\rangle\ t_2\text{-many Cs}]\ [k\ [\text{B ate } t_k]]]\!]^{d/2})$

The set of degrees denoted by the *than*-clause are those d for which some cookie-collection of size d was eaten by Bill. The matrix set of degrees is the same, but for John. If Bill ate 3, then the *than*-set will be $\{3,2,1\}$, and if John ate 5, the matrix set is $\{5,4,3,2,1\}$.

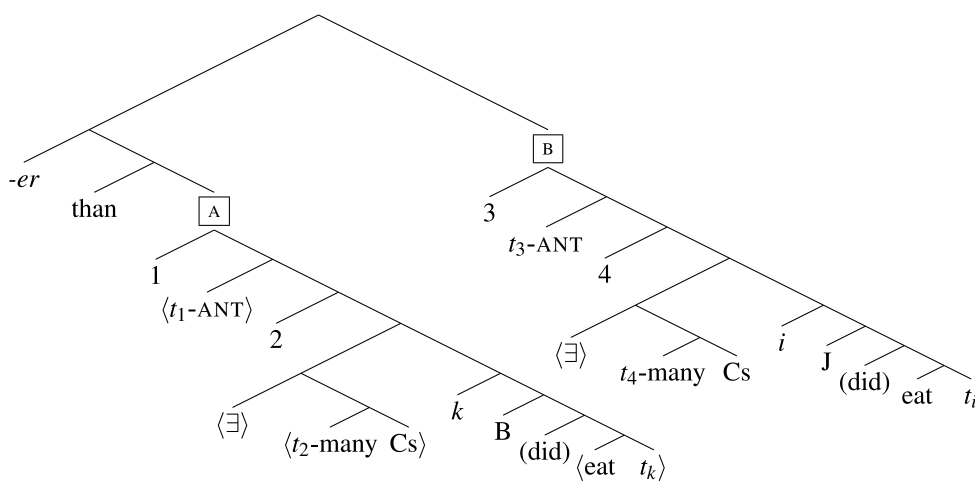
The entry for [*-er*] in (265) takes the maxima of its arguments, 3 and 5, and returns True iff the former falls below the latter. In our current example, where Bill ate 3 and John ate 5, the entry makes correct predictions. If we change the scenario and assume that Bill and John ate the same number of cookies (or that Bill ate 5 and John 3), the matrix set's maximum will not exceed that of the *than*-set, and [*-er*] will correctly return false.

Let us now turn to the negative case. We will see that here the definition in (265) derives incorrect truth conditions. Recall that ANT denotes a degree negation operator, so a negative comparative like (267) must feature ANT in both its *than*-clause and its matrix clause, if it is to comply with the identity that licenses ellipsis.

(267) John ate fewer/less cookies than Bill (did).

(268) $[\![\text{ANT}]\!] = \lambda d . \lambda D . D(d){=}0$

(269)



The sets denoted by the *than*-clause and the matrix clause consist of degrees d, d' for which Bill/John (respectively) did *not* eat d/d'-many cookies. If Bill ate 5 cookies, then he did not eat 6, nor 7, 8, etc. And if John ate 3 cookies, then he did not eat 4, 5, etc.:

(270) If Bill ate 5 cookies, and John ate 3, then

a. $[\![\boxed{\text{A}}]\!] = \{6, 7, \cdots\}$

b. $[\![\boxed{\text{B}}]\!] = \{4, 5, \cdots\}$

Neither one of these sets has a maximal element, so the current entry for the comparative will not do. Instead we want to compare the *minimal* elements of $[\![\boxed{\text{A}}]\!]$ and $[\![\boxed{\text{B}}]\!]$: in the current scenario it is true that John (who ate 3) ate fewer cookies than Bill (who ate 5). The comparative should therefore require that min($[\![\boxed{\text{A}}]\!]$), 6, *exceed* min($[\![\boxed{\text{B}}]\!]$), 4.

(271) $[\![\textit{-er}]\!] = \lambda D \lambda D' . \min(D) > \min(D')$

If we now compare (265) to (271), we find two differences: where the former refers to max, the latter refers to min, and where the former requires $>$, the latter requires $<$. The differences are both overcome by redefining $[\![\textit{-er}]\!]$ as in (272).

(272) $[\![\textit{-er}]\!] = \lambda D \lambda D' . D \subset D'$ (official entry, from Heim 2006)

(272) requires proper inclusion between its two arguments. If (272) holds of two sets of degrees D and D', then there must be a degree in D' that is not in D. In the case of *more*, *-er* takes two sets that share a lower bound, e.g. $[\lambda d\,.$ Bill ate d-many cookies$]$ and $[\lambda d\,.$ John ate d-many cookies$]$, so the only way for the first to be a proper subset of the second is for the second to have a *greater upper* bound. This is equivalent to the requirement that the second have a greater maximum than the first. In the case of *fewer*, *-er* compares two sets that extend infinitely upward. Proper inclusion is possible in this case only if the *lower* bound of the first set is *above* the lower bound of the second, which is the same as the requirement in (271).[2]

The main topic of the current chapter is the (in)compatibility between *only* and comparative focus associates, specifically comparatives where *than*-clauses contain bare numerals, e.g. (273), repeated.

(273) John ate more than three cookies.

But if (273) is to be interpreted using the semantic entry in (272), there needs to be two sets of degrees, one denoted by the matrix clause [λd.John ate d-many cookies], and the other denoted by the *than*-clause, the numeral 3. What set does a numeral in a *than*-clause denote? I turn to this question in the next section.

[2]In Chapter 4 we saw a similarly advantageous unification in the semantics of the positive morpheme.

Set inclusion and numerals in *than*-clauses

Hackl's (2000) account of the comparative calls for additional material in *than*-clauses, in light of the unacceptability of (274).

(274) a. #More than one person gathered.
b. #John introduced more than one person to each other.

Intuitively, (274a,b) is true iff the number of (gatherers)/(people introduced by John) is greater than 1. Our lexical knowledge tells us that the verb *gather*, and the predicate *be introduced to each other*, cannot combine with atomic/singular individuals. But the understood subject in (274) is not atomic, because the comparative requires that at least two people gather/be introduced to each other. So there is no obvious reason why the sentences should sound as odd as they do.[3]

To Hackl, this suggests that the internal structure of *than*-clauses contains more material than what appears at the surface. The intuition is that (274a,b) are actually paired with LFs that are like (275a,b).

(275) a. More people gathered than (would have been the case if) one person gathered.
b. John introduced more people to each other than (would have been the case if) John introduced one person to each other.

As I will show, this facilitates the move from the traditional account of the comparative, where two degrees are compared, to Heim's subset condition, where the comparative operates on two sets of degrees.

Consider (273) again.

(273) John ate more than three cookies.

In attempting to fit (273) into Heim's subset semantics, we might at first stab create a singleton consisting of the degree named by the numeral. [-*er*] will now require that this singleton be contained in $\{d$: John ate d-many cookies$\}$. If John ate 2 cookies, the sentence is predicted to be false—as desired—since $\{3\}$ is not contained in $\{1,2\}$. If John

[3]An attempt at explaining the oddness of (274) might appeal to the triviality of the predicted truth conditions: collective predicates like *gather* do not apply to atomic entities. So it is *always* true that more than one person gathered. But then we (incorrectly) expect to see the same oddness in (1).

(1) (At least) two people gathered/John introduced (at least) two people to each other.

ate 4, the sentence is predicted to be true—again as desired—since $\{3\}$ is contained in $\{1,2,3,4\}$. But the problem comes in the scenario where John ate 3 cookies: the singleton $\{3\}$ *is* properly contained in $\{1,2,3\}$, so the sentence is predicted be true, but it isn't; if John ate 3 cookies, then it is false that he ate more than 3.

The source of this problem is easy to see: the subset condition behaves well when we compare intervals that share one of their bounds. In example (263), repeated,

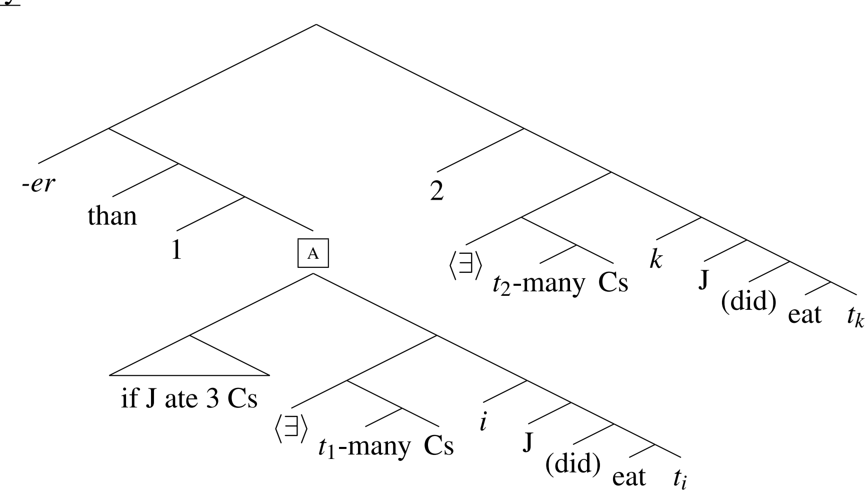
(263) John ate more cookies than Bill (did)

we compared John's numbers of eaten cookies to Bill's. The two sets have a common lower bound, and differ only in how far they extend in the scale of degrees. If John and Bill ate 3 cookies each, the sets are identical. This blocks proper subsethood, and in consequence the sentence is (correctly) predicted to be false. In the case of (273) we want to derive the same result, but the singleton approach chops away the bottom segment of one of [-*er*]'s arguments, and as a result it allows subsethood to hold for the wrong reasons. If we want to maintain Heim's proposal, we will want to abandon the singleton idea, and somehow retrieve the lower-bounded set $\{1,2,3\}$ from $\{3\}$; if we do, [-*er*] will produce correct results for our current example: because the two sets are identical, proper subsethood cannot hold between them, and the sentence will correctly come out false.

The remainder of this section is about deriving the desired set denotation for *than*-clause numerals. I show that Hackl's observations provide a key guideline in unifying Heim's subset-semantics with bare-numeral *than*-clauses.

My first step is to take the paraphrases in (275) more literally, and accordingly assign the following LF to (273).

(276) First try:



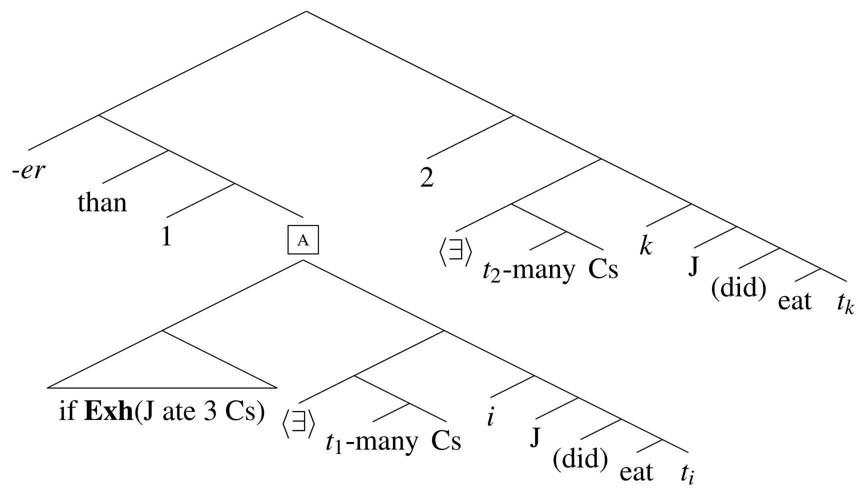
Let us construct the meaning of the *than*-clause in (276): it is a property that holds of a degree d iff it satisfies the denotation of $\boxed{\text{A}}$, i.e. the proposition in (277).

(277) If John ate (at least) 3 cookies, John ate d-many cookies

If John ate (at least) 3 cookies, then the consequent will hold if $d = 1$, because in that case John ate (at least) 1 cookie. Likewise for 2 and 3, but not for 4, because not all worlds where John ate at least 3 cookies are worlds where he ate 4.

Another, slightly different, interpretation of Hackl is one where an exhaustification operator is added to the embedded antecedent, as in (278). I will ultimately argue in favor of this analysis.

(278) <u>Final</u>:



The revision in (278) adds the operator Exh to the unpronounced *if*-clause. Exh is like *only*, but it asserts its prejacent together with the negation of its IE-alternatives (see (279)). This gives the antecedent in (278) the meaning that John ate 3 cookies and not more.

(279) $[\![\text{Exh}_A(S)]\!] = 1$ iff $[\![S]\!] = 1 \,\&\, \forall S'(S' \in \text{IE}(S)(A) \rightarrow [\![S']\!] = 0)$

(280) $[\![\text{Exh(John ate 3 cookies)}]\!] = [\![\text{John ate 3 cookies}]\!] \,\&\, \neg[\![\text{John ate 4/5/etc cookies}]\!]$

The exhaustified revision will have important consequences later, when we look at negative comparatives. Currently there is no difference between (276) and (278): a degree is admitted into the *than*-set iff it verifies

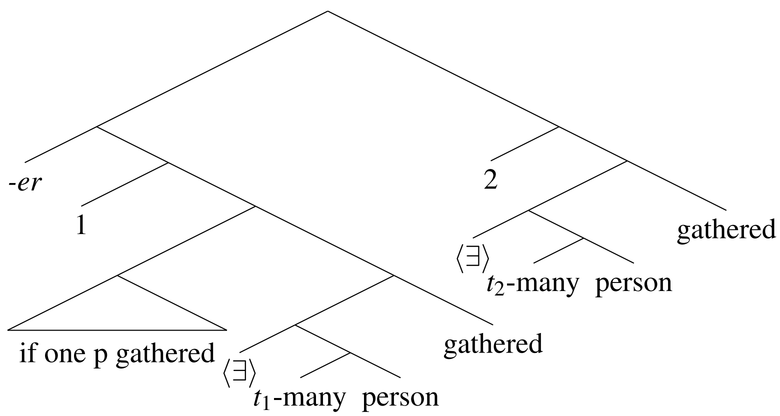
(281) If John Exh(ate 3 cookies), John ate d-many cookies.

In worlds where Exh(John ate 3 cookies), i.e. where John ate 3 cookies and not more, the number of cookies he ate is greater or equal to 1, greater or equal to 2, and greater or equal to 3. The *than*-set then denotes $\{1,2,3\}$, exactly the same as the result obtained without the use of Exh.

Before I show how this makes a difference in the case of NQs, let me quickly comment on Hackl's original problematic example. On the current proposal the sentence in (274a) will have the LF in (282).

(274a) #More than one person gathered

(282)

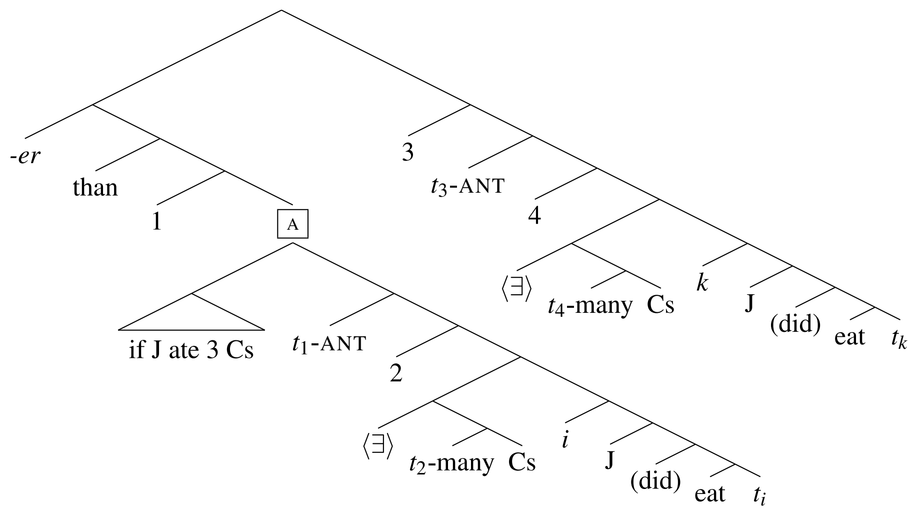


The conditional in (282) contains an antecedent where the subject conflicts with the demands of the collective predicate. The conditional is therefore undefined, as is the entire sentence in consequence. We get the same result if we choose (278) instead.

Let us now turn to negative comparatives. We will see that maintaining this analysis will favor the second of our two interpretations of Hackl, (278). To see the problem with the first, consider (284).

(283) John ate less/fewer than 3 cookies.

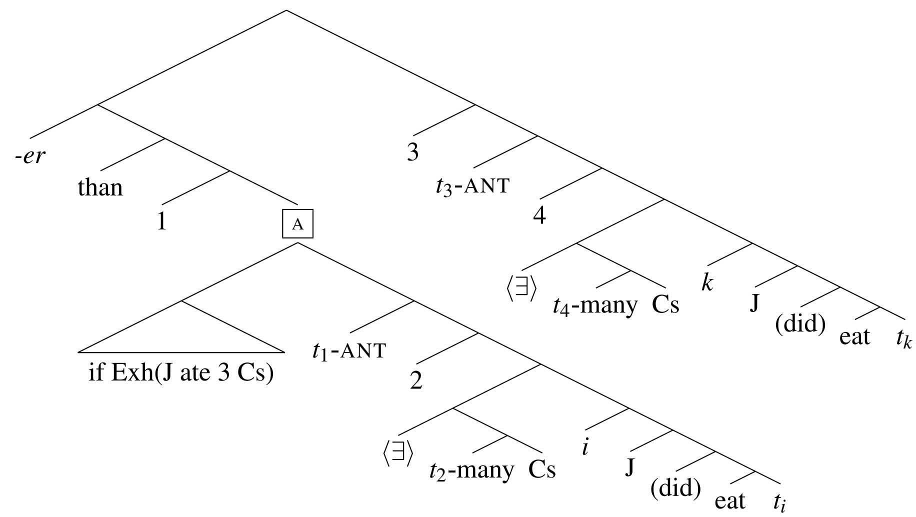
(284)



The conditional in node $\boxed{\text{A}}$ denotes the membership condition of the *than*-set, and the set is required by the comparative to be properly included in the matrix set. We will now see that $[\![\boxed{\text{A}}]\!]$ does not hold of any degrees at all, which makes the *than*-set in (284) empty. $[\![\boxed{\text{A}}]\!]$ holds of a degree d on the following condition: if John ate (at least) 3 cookies, then John did not eat d-many cookies. What degrees does the conditional hold of? Obviously none of 1, 2, 3, because if John ate (at least) 3 cookies, then he *did* eat at least 3, 2, and 1. But it turns out that the conditional does not hold of higher degrees either. Take 4, for example: the conditional holds of 4 provided that, if John ate (at least 3) cookies, then John did not eat 4 cookies. But this conditional is false. To see why, we only need to find one possible world (from those that the conditional ranges over) where John ate at least 3 cookies, but where he ate 4. Any world where he ate 4 or more cookies will do. The same holds of all higher degrees, and so $[\![\boxed{\text{A}}]\!]$ does not hold of any degree. The comparative is therefore predicted to hold trivially, since the empty set is properly included in every non-trivial set.

If we adopt (278) on the other hand, we get (285) for the antonym.

(285)



In (285), $[\![\boxed{\text{A}}]\!]$ holds of a degree d iff, if John ate 3 cookies and *did not eat more*, then John did not eat d-many cookies. This holds of 4, 5, 6, etc. The comparative is satisfied iff this set of degrees is properly included in the matrix set, i.e. iff the numbers of cookies that John *failed* to eat are $\{3,4,5,\cdots\}$, $\{2,3,4,\cdots\}$, etc. This holds iff John ate fewer than 4 cookies, as desired.

6.2 Summary of Part 1

In the previous section I stated my assumptions about the comparative, and appealed to Heim's subset-based semantics in order to treat positive and negative antonyms uniformly. In the course of motivating the analysis, I pointed to a problem concerning *than*-clauses that contain numerals, and proposed a way of assigning set denotations to them. I argued that the proposal finds independent support in a problem that was noted and discussed in Hackl (2000).

In the remainder of the chapter I will considerably simplify my treatment of the comparative. Now that a consistent analysis has been developed, where the comparative need not be customized depending on the polarity of its arguments, we may for brevity revert to a simpler (albeit less uniform) view of the morpheme, where the relations $<,>$ are referenced directly. The entries for [*more than*] and [*less than*] are shown in (286).

(286) a. $[\![\text{more than } n]\!] = \lambda D\,.\max(D) > [\![n]\!]$
b. $[\![\text{less than } n]\!] = \lambda D\,.\max(D) < [\![n]\!]$

E.g.

a. $[\![\text{John ate more than 3 cookies}]\!] = 1$ iff $\max\{d : \text{John ate } d\text{-many cookies}\} > 3$

b. $[\![\text{John ate less than 3 cookies}]\!] = 1$ iff $\max\{d : \text{John ate } d\text{-many cookies}\} < 3$

I now want to show that the move to the semantics in (286) is safe, that is, that whatever is discovered about the properties of (286), specifically in the context of *only*, the findings extend directly to the more detailed analysis that was developed earlier. To show this, we must be sure that the truth conditions in (286) and those derived in the previous section are equivalent.

Let us examine the positive case first. [*more than n*] was argued to be short for a structure where the comparative relates the set $\{1, \cdots, n\}$ (the *than*-set), to another, e.g. the degrees d such that John ate d-many cookies, $\{1, \cdots, d\}$. The comparative requires that the first set be a proper subset of the second, which is possible iff $d > n$, i.e. iff the maximal element of $\{d : \text{John ate } d\text{-many cookies}\}$ is greater than n. This satisfies the truth conditions of in (286a).

The negative case is more easily illustrated with an example. Take the sentence 'John ate less than 3 cookies' again. The comparative in this case takes two sets, each containing the antonymizer ANT. ANT effectively turns a set of degrees D into its complement $\overline{D}$. The *than*-clause, through ANT, denotes the complement of the set $\{3, 2, 1\}$. The matrix clause denotes the complement to the set $\{d : \text{John ate } d\text{-many cookies}\}$. The comparative requires that the *than*-set be properly included in the matrix set. This is true iff the *complement* of the *than*-set properly *includes* the complement of the matrix, i.e. iff $\{3, 2, 1\}$ includes $\{d : \text{John ate } d\text{-many cookies}\}$. This can only hold if $\max\{d : \text{John ate} \cdots\} < 3$, which is the condition in (286b).

Part 2

6.3 The Universal Density of Measurement – Fox and Hackl (2006)

Let us assume the following interpretation of the comparative:

(287) $[\![\text{John has more than 3 children}]\!] = 1$ iff $\max\{d : \text{John has } d\text{-many children}\} > 3$

Fox and Hackl (F&H) note that, unlike bare numerals, e.g. (288), comparatives, e.g. (289), do not give rise to scalar implicatures, and are not associable with *only*.

(288) a. John has three children.
$\Rightarrow$ John does not have four children
b. ✓John has only three$_{\mathrm{F}}$ children.

(289) a. John has more than three children.
$\not\Rightarrow$ John does not have more than four children
b. *John has only more than three$_{\mathrm{F}}$ children.

This pattern is also found when the relevant scale of degrees is intuitively dense, e.g. (290,291).

(290) a. John weighs 160 lbs.
$\Rightarrow$ John does not weigh 160+ε lbs.
b. ✓John only weighs 160$_{\mathrm{F}}$ lbs.

(291) a. John weighs more than 160 lbs.
$\not\Rightarrow$ John does not weigh more than 160+ε lbs.
b. *John only weighs more than 160$_{\mathrm{F}}$ lbs.

A dense scale is an ordering, e.g. $>$, where for any two elements d and d': if $d' > d$, then there is a third element d'' such that $d' > d'' > d$. Degrees of weight are intuitively ordered along a continuum, i.e. on a dense scale, so if John weighs more than 160 lbs. then whatever his exact weight is, 160+ε, there will be some other degree that stands between 160 and 160+ε, e.g. 160+$\frac{\varepsilon}{2}$. It is this property that blocks the unattested implicature in (291a): on the assumption that degree names are alternatives of one another, we get

A.	John weighs more than 160 lbs.	(Assumed example)
B.	There is a degree $d > 160$ such that John's (maximal) weight is d.	(Meaning of A)
C.	For all degrees $d' > 160$, $\neg$(John weighs more than d')	(Implicature of A)
D.	There is a degree d'' such that $d > d'' > 160$	(Density)
E.	$\neg$(John weighs more than d'')	(C,D)
F.	$\neg$(John's maximal/"exact" weight is d)	(D,E)
G.	$\perp$	(B,F)

The contradiction in step G depends crucially on the density assumption, for if there were no degrees of weight between 160 and John's actual weight (which exceeds 160), step D above would no longer be valid, and the contradiction would not arise. Take for example a sentence where the relevant degree is ordered on an intuitively *discrete* scale, e.g. numbers of children as in example (289).

	A.	John has more than 3 children.	(Assumed example)
	B.	There is a number $n > 3$ such that John has n-many children.	(Meaning of A)
	C.	For all numbers $m > 3$, $\neg$(John has more than m-many children)	(Implicature of A)
✖	D.	There is a number n' such that $n > n' > 3$	(✖Density)
✖	E.	$\neg$(John has more than n'-many children)	(C,✖D)
✖	F.	$\neg$(John has n-many children)	(✖D,✖E)
✖	G.	$\perp$	(B,✖F)

With discrete scales, step D is not valid; if John has 4 children, then it is true that he has more than 3, but there is no number (of children) between 3 and 4, contra the claim in D. Therefore, it is consistent to say that no number n greater than 3 is such that John has more than n children. The sentence in A is therefore predicted to mean that John has more than 3 children, but not more than 4, i.e. that John has exactly 4 children.[4]

This predicted meaning is incorrect, however. The examples in (289) show that even when the relevant degree is ordered on a discrete scale, no implicatures are detected, and association with *only* is ungrammtical.[5] So if we are to extend the density explanation to these cases, we would have to commit to a view where e.g. numbers of children are ordered on a dense scale as well, e.g. where there can be $3\frac{1}{2}$ children, and $3\frac{3}{4}$ children, and so on. F&H's claim is effectively just that: all degree scales, even those that are intuitively thought to be discrete, are in fact dense. Comparative sentences that explicitly refer to degrees are evaluated in a cognitive component that is blind to granularity. F&H call this component the Deductive System (DS). At DS, all degree scales are dense, and sentences containing degree terms must pass DS without giving rise to inconsistency. Information regarding granularity, whether contextual or lexical, enters the semantics at a later stage.[6]

[4]Without *only*, the 'exactly'-reading comes from scalar implicatures; with *only* it comes from the particle's assertoric component.

[5]One may explain the ungrammaticality by resorting to a version of Grice's maxim of manner: if the speaker intends that John has exactly 4 children, the speaker would have said 'John has 4 children' instead of 'John has more than 3'. But as F&H point out, this would also block implicatures in (1), even with full knowledge that Bill has, say, 2 children.

(1) John has 3 more children than Bill.

If the speaker intended to say that John has (exactly) 5 children, then he would have said so according to the maxim of manner. So (1) cannot mean that John has (exactly) 5 children. But this is clearly incorrect.

[6]F&H's analysis is not limited to the comparative. Their account is also intended to explain similar behavior in definite descriptions and negative islands in wh-questions. I refer the reader to their paper for details.

Association across modals – Empirical findings

An important observation in F&H is that implicatures are not always absent from comparative constructions, nor is association with *only* always ungrammatical:

(292) a. In order to qualify for the tournament, John needs to weigh more than 160 lbs.
$\Rightarrow$ John does not need to weigh more than 160+ε lbs.

b. ✓John only needs to weigh more than 160_{F} lbs.

Though association with *only* is possible across universal modals, as (292) shows, the same is not true of existential modals:

(293) a. According to the doctor, you are allowed to drink more than 1 oz. of wine a day.
$\not\Rightarrow$ You are not allowed to drink more than 1+ε oz. of wine

b. *You are only allowed to drink more than 1_{F} oz. of wine.

As we will see shortly, the density assumption predicts both (292) and (293). Let us now see what these empirical facts (summarized schematically in (294)) mean for the *o*PN.

(294) a. $*only(\mu > d_{\mathrm{F}})$

b. $*only\Diamond(\mu > d_{\mathrm{F}})$

c. $\checkmark only\Box(\mu > d_{\mathrm{F}})$

6.4 Consequences for the *Only* P/N Generalization

Recall that association between *only* and NQs is claimed by the *o*PN to be unambiguously non-logical, provided that association with the positive antonym is ungrammatical. When association with the positive is possible, we predict (absent other intervening factors), that association with the negative be interpretable logically, i.e. that it give rise to a reading where logically stronger alternatives are negated. Given F&H's observations, we predict that association with a negative comparative, e.g. *less/fewer*, be (i) strictly non-logical in unembedded cases, and also in cases where association crosses an existential modal, because of (294a,b) and (ii) optionally logical when association crosses a universal modal, because of (294c). These predictions are confirmed below.

(295) To qualify for the next round of this trivia contest, you have to answer at least half of the questions correctly. There are 60 questions: if you answer 30 or more, you advance; if you do not, you fail.

a. #John only got less than 35_F, so he passes/doesn't fail.

b. John only missed less than 25_F, so he passes/doesn't fail.

Let us first imagine what the status of (295a) would be if numbers of questions were not ordered on a dense scale. The prejacent of *only* says that John got less than 35 questions, and the exclusive component will negate every stronger alternative. If alternatives are ordered on a discrete scale, the sentence will assert that John did not get less than 34 questions, which together with the prejacent will mean that John answered exactly 34 questions, i.e. enough for him to advance. The sentence should therefore be acceptable. Compare this to how we intuitively understand (295b): John missed less than 25 questions, but did not miss *more* questions than that, i.e. he did not miss 26 or 27 etc. This is the non-logical reading: if the excluded alternatives are [*less than n*], for *n*s that are *greater* than 25, then exclusion is not operating logically, since these alternatives are *weaker* than the prejacent. What is noteworthy is that this (non-logical) interpretation of (295b) is roughly equivalent to the logical reading that (295a) would have if it were acceptable. (295a) is true iff John got 34 questions, i.e. missed 26, which is nearly the same as him missing (less than) 25. If we assume a discrete scale, there is no obvious reason why one sentence should be more acceptable than the other.

But if we follow F&H and assume density, we predict the unacceptability of (295a).

A.	John only got less than 35_F Qs	(Assumed example)
B.	John got less than 35 Qs	(A's prejacent)
C.	There is a degree $d < 35$ such that John got maximally d-many Qs.	(Meaning of B)
D.	For all degrees $d' < 35$, $\neg$(John got less than d'-many Qs)	(Exclusion by *only* in A)
E.	There is a degree d'' such that $d < d'' < 35$	(Density)
F.	$\neg$(John got less than d'')	(D,E)
G.	$\neg$(John got maximally d-many Qs)	(E,F)
H.	$\perp$	(C,G)

Likewise we expect that association across an existential modal be unacceptable on a logical reading:

(296) I heard that this (10-course) program was too easy on its students, and they can get away with lots of Fs. #But it turns that you're only allowed to pass less than 9_F courses.

Let us again imagine that numbers of courses are ordered on a discrete scale. On this assumption we predict the sentence to mean that (a) passing less than 9 courses is allowed, but (b) passing less than 8 courses is not. In other words, the sentence should describe a plausible scenario where students must pass at least 8 courses (out of ten) in order to graduate. Nevertheless, the sentence is ungrammatical. If we assume instead that the scale is formally dense, the unacceptability of (296) falls out.

Before I go through F&H's analysis of modals, let me first present the final case, where *only* can associate with the comparative, namely across a universal modal. We predict that here the logical reading of *only* be available, as is indeed the case:

(297) The entry regulations were relaxed in recent years, so now travelers are only required to carry less than $15,000$_{\mathrm{F}}$.

(298) (*John is a proud host who likes to provide more drinks than his guests do individually. But since he does not drink much, he provides little beer, and to satisfy his pride we have to bring even less. No one has a good time at John's parties*).

This time John feels like drinking a lot. His roommate told us how much he is planning to buy, and this could end up being enjoyable: we only need to bring less than 4_{F} liters.

Both (297,298) can be understood on their logical readings: (297) is set up so that in the past the entry regulations were more stringent. When association between *only* and a negative comparative appears later in the sentence, it is expected to describe more relaxed laws by negating *stronger* alternatives to the comparative prejacent. And because the comparative is negative, its stronger alternatives are those where the numeral is replaced with a smaller numeral. The intended meaning of the sentence is therefore that travelers are not required to carry e.g. less than $14.9k, $14.8k, etc. Example (298) behaves similarly: the guests want to bring more beer to the party, but they are held back by John's insistence to bring more. Now that John is up for a serious night of drinking, the pressure on the guests is less severe, and they only need to be bring less than 4_{F} liters of beer, not less than 3, 2, etc.

Note that association across a universal modal (in other contexts) can also have the non-logical reading, e.g. (299).

(299) The train travels at 250-300 km/h, so I will only need to spend less an hour$_{\mathrm{F}}$ in transit from the city to the airport.

The availability of the non-logical reading in (299) does not falsify the *o*PN, however; the claim of the *o*PN is that *un*associability with the positive entails non-logical readings for the negative. Cases where the positive is associable with *only* trivially satisfy this conditional.

6.5 Consequences for *only* and Innocent Exclusion

Recall from Chapter 1 that, given a prejacent S and a set of alternatives A, *only* excludes the members of $\text{IE}(S)(A)$, the innocently-excludable alternatives to S. Recall also that the set $\text{IE}(S)(A)$ is constructed by intersecting all the Maximal Consistent Sets of Excludables: $\text{IE}(S)(A) = \bigcap \mathbf{MC}(S)(A)$ (see Section 1.2.2).

In the case of comparatives, the relevant prejacent S is of the form $[\mu > n_{\text{F}}]$, where μ is a measure, e.g. John's weight, or the number of John's children etc., and the set of alternatives A is $\{\mu > d : d \in \mathbf{D}_d\}$. I will now show (informally) that in this case there can be no $M \in \mathbf{MC}(S)(A)$. The discussion is inspired by Gajewski's (in press) far more rigorous proof, to which I refer interested readers. An important observation made by Gajewski is that, without any $M \in \mathbf{MC}(S)(A)$, the set $\text{IE}(S)(A)$ consists of *all* sentences, which makes the contribution of *only* a contradiction.

Let us first see why there cannot be any $M \in \mathbf{MC}(S)(A)$, given e.g.

(300) $S = \mu > 3$,

$A = \{[\mu > d] : d \in \mathbf{D}_d$ where $\mathbf{D}_d$ is densely-ordered$\}$,

Assume for reductio that some M is a non-empty $\text{CSE}(S)(A)$, where M contains $S' = [\mu > 4]$. So far M is not maximal, because we can consistently negate $S'' = [\mu > 3.5]$. S'' must therefore be added to M in order for the set to be maximal. But adding S'' is still not enough, because we can add $[\mu > 3.25]$, and $[\mu > 3.125]$, and so on. From this, we conclude that M cannot be maximal if it has a member S' that is weaker than every other member, but stronger than S. In other words, if there is an $S' = [\mu > d']$ where d' is smaller than every d'' in other elements of M, but where d' is greater than 3, then M is not maximal. The reason is density: if M has a weakest element $S' = [\mu > d']$ which is stronger than S, then there is a degree between d' and 3 that can be consistently admitted to M. Therefore, every S' that is stronger than S must be in M in order for M to be maximal. But this is contradictory; from S we infer that some degree $n > 3$ is such that $\mu = n$. By density there is a degree n' between 3 and n. $[\mu > n']$ is stronger than S, so $[\mu > n'] \in M$. Therefore $(\mu > n')$ is false. But then $(\mu = n)$ is false, which contradicts the (arbitrarily-)assumed inference from the prejacent.[7]

The set IE contains an alternative S' if, for all sets M, if $M \in \mathbf{MC}(S)(A)$, then S' is in M. Since no set satisfies the antecedent of this conditional, every sentence S' verifies

[7] Gajewski's strategy is to build a non-empty set D_M out of every d' that appears in the members in M. He then shows that D_M cannot have a greatest lower bound, which contradicts a property of any non-empty subset of the real/rational numbers (Dedekind on the reals – see Gajewski's footnote 2 for a comment on the rationals).

the condition. Therefore, IE is the set of all sentences (the entire language L), and so the negation of every one of its elements (by *only*) is the negation of all propositions, i.e. the falsum.

We may then cite this contradiction as the reason behind *only*'s incompatibility with the comparative. Alternatively, we may strengthen the definition of IE so that its members are those S' that belong in *some* $M \in \mathbf{MC}$, and to all $M \in \mathbf{MC}$. This revision converts IE in the current case from all of L to the empty set; because there are no $M \in \mathbf{MC}$, there are no sentences S' that belong to *some* $M \in \mathbf{MC}(S)(A)$. The result is that *only* is in this case vacuous, which, as was argued in Chapter 1 (Section 1.2.3), blocks it from associating with the prejacent.

In the next section we will see that the proof extends straightforwardly to cases where an existential modal intervenes between *only* and the comparative, but not where a universal modal does. This correctly predicts unassociability in the former case, and not for the latter.

6.6 Association across modals – F&H's theoretical account

We may now resume the discussion that began in Section 6.3. The empirical findings are repeated below.

(294) a. $*only(\mu > d_{\mathrm{F}})$

b. $*only\Diamond(\mu > d_{\mathrm{F}})$

c. $\checkmark\, only\Box(\mu > d_{\mathrm{F}})$

Let us first see the (b) case. Assume that a prejacent $S = [\Diamond(\mu > d)]$ is true. Then there is a degree $d' > d$ such that $\Diamond(\mu \geq d')$. By density, there is a degree d'' between d and d'. The alternative $[\Diamond(\mu > d'')]$ is stronger than the prejacent, and is therefore negated by *only* (assuming for now the version of *only* that negates non-weaker alternatives in Section 1.2). But if $[\Diamond(\mu > d'')]$ is false, then $\Diamond(\mu = d')$ must be false, contrary to the initial assumption. I show below that, just as in the unembedded case, this makes it impossible for there to be $M \in \mathbf{MC}$ given a prejacent like S.

The (c) case incurs no contradiction, and is therefore not blocked if density is assumed. If $S = [\Box(\mu > d)]$ is true, then in every world w there is a (potentially different) degree $d' > d$ such that $\mu = d'$ in w. Let us take a model where a dense set of worlds w_i corresponds to each $d_i > d$. Let $\mu = d_i$ in w_i. Then the prejacent $\Box(\mu > d)$ is true, and no degree $d' > d$ is such that $\Box(\mu > d')$.

Given these findings, it can be shown that the prejacent in (294b) does not have any $M \in$ **MC** given its alternatives. A proof just like the one in the previous section should demonstrate this: let $S = [\Diamond(\mu > 3)]$, and assume (for reductio) some M which is a CSE$(S)(A)$, where M contains $S' = [\Diamond(\mu > 4)]$. M is not maximal, even if we add $[\Diamond(\mu > 3.5)]$, $[\Diamond(\mu > 3.25)]$, and $[\Diamond(\mu > 3.125)]$, etc. More abstractly, if there is an $S' = [\Diamond(\mu > d')]$ where $d' > 3$ and $d' < d''$ for all $[\Diamond(\mu > d'')]$ in M, then M is not maximal. The reason is density: if M has a weakest element $S' = [\Diamond(\mu > d')]$ which is stronger than S, then there is a degree between d' and 3 that can be consistently admitted to M. Therefore, every S' that is stronger than S must be in M in order for M to be maximal. But this is contradictory; from S we infer that some degree $n > 3$ is such that $\Diamond(\mu = n)$. By density there is a degree n' between 3 and n. $\Diamond(\mu > n')$ is stronger than S, so $[\Diamond(\mu > n')] \in M$. Therefore $\Diamond(\mu > n')$ is false. But then $\Diamond(\mu = n)$ is false, which contradicts the (arbitrarily-)assumed inference from the prejacent.

In the case of universal modals the second half of the proof does not go through. If $S = \Box(\mu > d)$, then a set $M \in \mathbf{MC}(S)(A)$ must contain every stronger alternative $\Box(\mu > d')$, otherwise M is not maximal (because of density). This then makes every element of M false, which (as was shown above) is not contradictory.

6.7 Summary

The findings so far are the following: given

(a) a dense ordering on the scale of degrees,

(b) a comparative sentence $S = \mu > d$, and

(c) a set of alternatives $A = \{\mu > d' : d' \in \mathbf{D}_d\}$,

then, the set $\mathbf{MC}(S)(A)$ is empty, (301i), as is the set $\mathbf{MC}(\Diamond S)(A_\Diamond)$, where the comparative sentence S is embedded under an existential modal, (301ii). When S is embedded under a universal modal, the set $\mathbf{MC}(\Box S)(A_\Box)$ is not empty, and the resulting set of IE-alternatives contains all those alternatives in which the numeral in S is replaced with a higher numeral, (301iii).[8]

(301) (i) $\mathbf{MC}(S)(A) = \emptyset$, therefore $\mathrm{IE}(S)(A) = \bigcap \emptyset = L$

(ii) $\mathbf{MC}(\Diamond S)(A_\Diamond) = \emptyset$, therefore $\mathrm{IE}(\Diamond S)(A_\Diamond) = \bigcap \emptyset = L$

(iii) $\mathbf{MC}(\Box S)(A_\Box) = \{\{\Box(\mu > d') : d' > d\}\}$,
therefore $\mathrm{IE}(\Box S)(A_\Box) = \{\Box(\mu > d') : d' > d\}$.

[8] $A_\Diamond = \{\Diamond S' : S' \in A\}$, and $A_\Box = \{\Box S' : S' \in A\}$.

On the IE-based definition of *only* developed in Chapter 1, (301) lead to (302).

(302)
$$\begin{aligned}
[\![only_A\ S]\!] &= 1 \text{ iff } \forall S'(S' \in \mathrm{IE}(S)(A) \to [\![S']\!] = 0)\\
&= 1 \text{ iff } \forall S'(S' \in L \to [\![S']\!] = 0);\\
[\![only_{A_\Diamond} \Diamond S]\!] &= 1 \text{ iff } \forall S'(S' \in \mathrm{IE}(\Diamond S)(A_\Diamond) \to [\![S']\!] = 0)\\
&= 1 \text{ iff } \forall S'(S' \in L \to [\![S']\!] = 0);\\
[\![only_{A_\Box} \Box S]\!] &= 1 \text{ iff } \forall S'(S' \in \mathrm{IE}(\Box S)(A_\Box) \to [\![S']\!] = 0)\\
&= 1 \text{ iff } \neg\Box(\mu > d') \text{ for all } d' > d.
\end{aligned}$$

Alternatively, if IE is made less susceptible to triviality, as in the modification in (303),

(303)
$$\mathrm{IE}(S)(A) = \{S' : \exists M \in \mathbf{MC}(S)(A) \,\&\, \forall M(M \in \mathbf{MC}(S)(A) \to S' \in M)\}$$

then the following changes apply to (302).

(302′)
$$\begin{aligned}
[\![only_A\ S]\!] &= 1 \text{ iff } \forall S'(S' \in \emptyset \to [\![S']\!] = 0)\\
&= 1;\\
[\![only_{A_\Diamond} \Diamond S]\!] &= 1 \text{ iff } \forall S'(S' \in \emptyset \to [\![S']\!] = 0)\\
&= 1;
\end{aligned}$$

Therefore, assumptions (a-c) above result in contradictory/trivial truth conditions for *only*, given S and A, or $\Diamond S$ and $A_\Diamond$. The conditions are contradictory if IE is the intersection of **MC**, and trivially true if IE is redefined as in (303).

6.8 Further Issue: the UDM with broad focus

F&H restrict their attention to cases where focus is marked on the numeral inside the comparative clause. In this section I want to discuss the effect of expanding focus to include the comparative in addition to the numeral.

By focus-marking [more than n], instead of the numeral alone, the prejacent of *only* may have more alternatives than it would with narrower focus, because other degree modifiers may qualify as formal alternatives to *more than*. But if we can show that the comparative does not have formal alternatives, or that other alternatives do not escape the density-based contradictions, we predict that both narrow and wide focus be equivalent.

I will consider two candidate alternatives to '*more than*': '*at least*', and '*exactly*'.[9] There is a third possibility which is different, but it is different only in form: following Katzir (2007), we may admit subconstituents of '*more than n*' into the set of alternatives. This will make the numeral itself an alternative of the focus phrase that contains it, giving us *n* for [*more than n*]. But note that, if numerals have 'at least'-like semantics, this will not be different from taking '*at least*' to be a formal alternative to '*more than*'. And if instead we take numerals to have 'exactly' readings, this will be the same as allowing '*exactly*' to be a formal alternative to '*more than*'. I will therefore discuss only the former two possibilities. My conclusion will be that allowing '*at least n*'/'*n*' into ALT('*more than*') does not obviate density violations.[10] The case of '*exactly*' is more complicated: admitting it as a formal alternative potentially generates unwanted inferences, but it does so in specific LF configurations that seem unavailable from independent considerations. The discussion draws on Fox (2007b), and findings from Alxatib (In progress).

at least* as an alternative to *more than

For any given numeral *n*, '*at least n*' ($\geq n$) is logically weaker than '*more than n*' ($>n$). The former can be thought of as a disjunction of two propositions: $(\geq n) \leftrightarrow (= n \vee > n)$. So negating the weaker proposition will be stronger than negating the stronger one: $\neg(\geq n)$ is stronger than $\neg(>n)$. This means that, if '*at least n*' (or the structurally simpler numeral *n* on its 'at least' reading) is excluded by *only*, the exclusion will entail $\neg(>n)$.

So let us assume that '*at least*' is a formal alternative of '*more than*'. Let us assume also that there is a degree name *m* such that '*only* [*more than n*]$_{\mathrm{F}}$' excludes '*at least m*'. Since only IE alternatives are excluded, '*at least m*' must be in the set IE(*more than n*)(A), where, by assumption,

(304) $A = \{$'*at least d*' $: d \in \mathbf{D}_d\} \cup \{$'*more than d*' $: d \in \mathbf{D}_d\}$

This means that every $M \in \mathbf{MC}$('*more than n*')(A) contains the sentence '*at least m*'. But since the negation of '*at least m*' entails the negation of '*more than m*', every $M \in$

[9] In principle, replacing *more than* with (*at least*)/*exactly* is possible by replacing the constituent [-*er* [*than* $\cdots$]], which is of type $\langle dt,t\rangle$, with [*at least n*] or [*exactly n*], both of which also of type $\langle dt,t\rangle$. If the original comparative sentence contained e.g. [λd.John ate *d*-many cookies] in the matrix clause, then replacing the [-*er*-*than*-clause] constituent with [*at least n*] will generate the sentence [John ate at least *n*-many cookies], likewise for *exactly*.

[10] In Chapter 7 I will make use of this alternative to derive what I have called the non-logical reading of [*only* ... *less than*$_{\mathrm{F}}$].

$\mathbf{MC}$(‘*more than n*’)(A) must also include ‘*more than m*’. So ‘*more than m*’ is also innocently excludable, and is therefore predicted to be part of the meaning of ‘*only* [*more than n*]$_{\text{F}}$’. But in Section 6.5 we saw that this is not possible: there cannot be any $M \in$ $\mathbf{MC}$(‘*more than n*’)(A) that contains ‘*more than m*’. Therefore the same $M \in \mathbf{MC}$ cannot contain alternatives of the form ‘*at least m*’. The conclusion then is

(305) If $A = \{$‘*at least d*’ $: d \in \mathbf{D}_d\} \cup \{$‘*more than d*’ $: d \in \mathbf{D}_d\}$, and
$A' = \{$‘*more than d*’ $: d \in \mathbf{D}_d\}$,
then $\text{IE}(\textit{more than } n)(A) = \text{IE}(\textit{more than } n)(A') = \mathbf{D}_{\langle s,t\rangle}$

exactly* as an alternative to *more than

If *exactly* is assumed to be a formal alternative to *more than*, then for a given $S = [>d]$ we have the set of alternatives A in (306) (from now on I will use the shorter $=d$, $\geq d$, and $>d$).

(306) $A = \{>d' : d' \in \mathbf{D}_d\} \cup \{=d' : d' \in \mathbf{D}_d\}$

I will now show that the set $\text{IE}(S)(A)$ is empty. The strategy is as follows: if we take a degree d' above d, and negate $>d'$ (i.e. $\not> d'$), then this is consistent with the prejacent $>d$, but it forces the relevant measure to fall between d and d': greater than d, but not greater than d'. Now suppose that we take every degree d'' between d and d', and negate the ‘*exactly d''*’ alternative. This set of exclusions is also consistent with the prejacent $>d$, but it forces the reading that the relevant measure is *above d'*. In this latter set, we cannot add $>d'$, because then there would be a contradiction: the measure would have to be above d, nowhere between d and d', and nowhere above d', which is impossible. The strategy is repeated also for the ‘*exactly d'*’ case:

> Assume for reductio that $\text{IE}(S)(A)$ is not empty. Then there must be a degree d' such that either $[>d'] \in \text{IE}(S)(A)$, or $[=d'] \in \text{IE}(S)(A)$.
>
> (a): If $[>d'] \in \text{IE}$, then $[>d']$ is in $\bigcap \mathbf{MC}(S)(A)$. Since for any d'', $[=d'']$ is an alternative to $[>d]$, then there is a set $M \in \mathbf{MC}(S)(A)$ that contains $[=d'']$ for every d'' s.t. $d < d'' \leq d'$. The conjunction $\bigwedge\{\neg\phi : \phi \in M\}$ is $\bigwedge\{\neq d'' : d < d'' \leq d'\}$, i.e. nowhere between d and d', and not d'. In conjunction with the utterance $[>d]$, this entails $[>d']$. So $[>d']$ cannot be added to M, because it would otherwise be negated along with every other member of M, resulting in $>d'$ and $\not> d'$. Therefore $>d'$ does not belong to every $M \in \mathbf{MC}(S)(A)$, so $\text{IE}(S)(A)$ cannot contain alternatives of the form $[>d']$.

(b): if $(=d') \in$ IE, then every set $M \in \mathbf{MC}(S)(A)$ includes $[=d']$. Let some M contain $\{=d'' : d<d'' \,\&\, d''\neq d'\}$. The conjunction $\bigwedge\{\neg\phi : \phi \in M\}$ is consistent with the utterance $[>d]$, and with it, the set entails $[=d']$. Therefore $[=d']$ cannot be in M because it would otherwise be negated, and would make the set inconsistent. Therefore $=d'$ does not belong to every $M \in \mathbf{MC}(S)(A)$, so IE$(S)(A)$ cannot contain alternatives of the form $[=d']$.

With an empty IE$(S)(A)$, *only* is vacuous, and this violates the nonvacuity condition formulated in Section 1.2.3.

It is worth noting that admitting '*exactly*' as a formal alternative changes the cause behind the incompatibility of *only* and the comparative. We have seen that F&H's account relies on density, but as it turns out, the incompatibility can be derived if we drop the density assumption, and instead use the symmetry between '*exactly*' and '*more than*'. This can be seen by observing that the proof above applies just as well to discretely-ordered scales of degrees. Note, moreover, that like F&H's analysis, this proposal predicts obviation under universal modals, but not under existential ones.[11]

Fox (2007b) cites Benjamin Spector as having proposed a solution along these lines, but he notes that the proposal comes with an unwelcome empirical result. I discuss the problem in the following subsection, and raise questions about whether it applies to *only*.

Detour: Symmetry for comparatives predicts unattested Free Choice inferences. While it is true that symmetric alternatives to a sentence S cannot belong to IE$(S)(A)$, under existential modals they give rise to FC inferences. Consider the following example from Fox:

(307) You are allowed to smoke 6 or more cigarettes. John is luckier. He is allowed to smoke 7 or more. #More specifically, he is allowed to smoke 7 but not more than 7.

The final addition in (307) is felt to conflict with the sentence that immediately precedes it, specifically, the inference that John is allowed to smoke 7 cigarettes *and* he is allowed to smoke more than 7. This is the Free Choice inference (FC): $\Diamond(p \vee q) \vDash \Diamond p \wedge \Diamond q$.

[11] If $S = \Box(>d)$, and $A = \{\Box(>d') : d' \in \mathbf{D}_d\} \cup \{\Box(=d') : d' \in \mathbf{D}_d\}$, then it is consistent to maintain S, and simultaneously negate $\Box(>d')$ and $\Box(=d')$ for all $d' > d$. Let S be 'you need to have more than 3 children (to qualify for the tax break)', then the maximal number of children you have must be 4 or more. The exclusions $\neg\Box(>d')$ and $\neg\Box(=d')$ add, for e.g. 5, 6, that you do not need to have more than 5 children, nor more than 6 children, nor exactly 5, nor exactly 6, etc.

The case of existential modals can be shown by repeating the proof above, and prefixing $\Diamond$ to S and to each of its alternatives.

The final sentence in (307) denies FC, hence the oddness of the monologue. I will not show the details of how this inference is derived here. I merely point to the fact that the inference could not be derived without the availability of the disjuncts $(=7)$ and (>7) as alternatives, as they are assumed to be. (See Section 1.2.1, where I motivate the assumption that disjuncts qualify as alternatives to the disjunction containing them). This gives us $\Diamond(=7), \Diamond(>7) \in \text{ALT}(\Diamond(=7 \vee >7))$.

If we now take '*exactly*' to be a formal alternative to '*more than*', then we will also derive the FC inference detected in (307). $[\Diamond >n]$ will have $[\Diamond =m]$ and $[\Diamond >m]$ as alternatives for any $m > n$. So $[\Diamond =7]$ and $[\Diamond >7]$ will be alternatives to $[\Diamond >6]$. But because we are assuming a discrete scale, $[\Diamond >6]$ is equivalent to $[\Diamond(=7 \vee >7)]$. So we have recreated the same setup that led to the FC inference in (307): the same pair of alternatives is available, and the utterances, $[\Diamond(=7 \vee >7)]$ and $[\Diamond >6]$, are equivalent (see appendix for the FC derivation). If the inference is derived for the disjunction in (307), we expect to see it in a similar sentence where the disjunction is replaced with a comparative. So if we make this modification to (307), we predict the result to be equally odd.

(308) You are allowed to smoke more than 5 cigarettes. John is luckier. He is allowed to smoke more than 6. More specifically, he is allowed to smoke 7 but not more than 7.

We find, however, that (308) is acceptable. Its acceptability indicates that the final sentence does not deny any earlier inferences. Since it denies FC (that smoking 7 is allowed, but smoking more than 7 is not), this shows that FC does not arise in prior discourse, which in turn shows that embedding the comparative under the existential modal does not license the same FC that (307) does.

Fox cites the contrast between (307) and (308) as evidence against Spector's suggestion. His conclusion is that postulating symmetric alternatives for comparatives, though it correctly gives rise to empty implicatures, and correctly predicts obviation under universal modals, it incorrectly permits FC inferences and, as a result, predicts no difference between sentences like (307) and (308). Fox concludes that Spector's proposal (symmetry-density) falls short of accounting for the data.

Given this finding, what danger is there for association with *only*? On Fox (2007a), FC arises from stacking two exhaustification operators on top of one another (see Chapter **??**). With *only*, FC is derived by substituting the particle for one of the two occurrence of Exh. The two possibilities are

(309) a. $\text{Exh}(\textit{only}\Diamond(>n)_{\text{F}})$

b. $only(\text{Exh}\Diamond(>n)_{\text{F}})$

Both LFs potentially generate FC inferences, and this is undesirable. We therefore want to find reasons to block them. In (309a) *only* is vacuous, since symmetry will block any alternatives $\Diamond(>d)$/$\Diamond(=d)$ from belonging in $\text{IE}(\Diamond > d)(A)$. (309a) is therefore not a source of worry, since it is ruled out by the non-vacuity condition on *only*. LF (309b) is not vacuous, but in Alxatib (In progress) I argue that its meaning (in the context of FC permission sentences like (310)) is never detected.

(310) You are only allowed to have cake or ice cream.
$\neq$ [*only*(Exh(you are allowed to have cake or ice cream))]

If this indicates that something blocks Exh from appearing in the scope of *only*, e.g. because the environment is Strawson-DE (von Fintel 1999), then neither LF is available.[12] As a result, the FC worry is not applicable to *only*, because deriving it is only possible on LFs where either *only* is vacuous, (309a), or where Exh appears somewhere where we independently do not expect it to, (309b).

6.9 Summary

This chapter concerned comparative constructions. It was first shown that a uniform analysis of positive and negative comparative, using subsethood (Heim 2006) is compatible with *than*-clauses containing bare numerals. The result was translated to a more intuitive semantics in preparation for a discussion of F&H, whose analysis provides concrete predictions regarding the *o*PN Generalization: environments where *only* is incompatible with a positive comparative are ones where negative comparatives can only be interpreted non-logically. Environments where positives can associate with *only* (across a universal modal) allow for the logical reading in the negative case. I finally noted that changing the focus of *only* to include the comparative (not just the numeral) is predicted to derive the same results, because other degree modifiers do not interfere with the results from narrow focus.

[12] Appealing to Strawson-DE is likely insufficient, since Exh is generally thought to be blocked *where its contribution to meaning is vacuous*, not simply in all DE environments. The fact that Exh contributes an otherwise absent FC inference in this environment should be enough to license it. Something else must therefore ban Exh from appearing under *only*.

Appendix: FC inferences for comparatives under symmetry

Assume that $S =$ '*more than n*', and its set of alternatives A is:

(311) $A = \text{ALT}(\text{'}\textit{more than d}\text{'}) = \{\text{'}\textit{more than } d'\text{'} : d' \in \mathbf{D}\} \cup \{\text{'}\textit{exactly } d'\text{'} : d' \in \mathbf{D}\}$

Then (i) $\text{IE}(A)(\text{'}\textit{more than d}\text{'}) = \emptyset$, and (ii) $[\![\text{Exh}_{A'}(\text{Exh}_A \Diamond(\textit{more than n}))]\!] = \forall d \in \mathbf{D}(\Diamond(= d))$.

(i) $\text{IE}(A)(\text{'}\textit{more than d}\text{'}) = \emptyset$. Assume for reductio that IE is not empty. Then there must be a degree d' such that either (a) '*more than* d'', $>d'$ for short, is in IE, or (b) '*exactly* d'', $=d'$, is in IE. (a): If $>d' \in$ IE, then $>d'$ is in every maximally consistent set with $>d$. Since for any d'', $(=d'')$ is an alternative to $(>d)$, then there is a maximally consistent set M that contains $=d''$ for every d'' s.t. $d < d'' \leq d'$. The conjunction $\bigwedge\{\neg\phi : \phi \in M\}$ is $\bigwedge\{\neq d'' : d < d'' \leq d'\}$, i.e. nowhere between d and d', and not d'. In conjunction with the utterance $>d$, this means $>d'$. So $>d'$ cannot belong to M, because if it did it would be negated along with every other member of M, and this will mean $>d'$ and $\not> d'$. Therefore $>d'$ does not belong to every maximally consistent set with the utterance, so IE cannot contain alternatives of the form '*more than* d''.
(b): if $(=d') \in$ IE, then every maximally consistent set includes $=d'$. Let M contain $\{=d'' : d < d'' \,\&\, d'' \neq d'\}$. The conjunction $\bigwedge\{\neg\phi : \phi \in M\}$ is consistent with the utterance $>d$, and with the utterance it amounts to $=d'$. Therefore $=d'$ cannot be in M. Since there is a maximally consistent set that excludes $=d'$, $=d'$ cannot be in IE.

(ii) $[\![\text{Exh}_{A'}(\text{Exh}_A \Diamond(\textit{more than n}))]\!] = \forall d \in \mathbf{D}(\Diamond(=d))$. First let us determine the contents of the set $\text{IE}(A')(\text{Exh}\Diamond(>n))$, where $A' = \{\text{Exh}_A(S) : S \in A\}$. The set A', which is the source of alternatives for the upper Exh, contains expressions of two forms. The first is $\text{Exh}_A\Diamond(>d)$, and the second is $\text{Exh}_A\Diamond(=d)$. So at the matrix level, exhaustification will assert $\text{Exh}_A\Diamond(>d)$, and negate whatever can be innocently negated of $\text{Exh}_A\Diamond(>d)$ and $\text{Exh}_A\Diamond(=d)$. Recall that for any degree d, $\text{Exh}_A\Diamond(>d)$ is vacuous, so matrix Exh will assert $\Diamond(>n)$, and deny the IE alternatives of the form $\Diamond(>d)$. But since there are no such alternatives, we only get the assertion of $\Diamond(>n)$ so far. Next we have alternatives of the form $\text{Exh}_A\Diamond(=d)$. For each d these will individually amount to $\Diamond(=d) \wedge \neg\Diamond(=d')$ for every other d', and so the negation of $\text{Exh}_A\Diamond(=d)$ by matrix $\text{Exh}_{A'}$ will admit $\Diamond(=d) \rightarrow \Diamond(=d')$ so long as it is consistent with every other alternative. All these conditionals are consistent with the utterance $\Diamond(>n)$, for they will add that every other degree is permitted: $\Diamond(>n) \wedge \forall d \forall d'(\Diamond(=d) \rightarrow \Diamond(=d'))$. Together this gives us the FC meaning $\forall d \Diamond(= d)$.

Chapter 7

Generating the non-logical readings for *only*-NQ

In this chapter I address question (80) in Section 2.3, repeated here as (312).

Given a pair of antonyms $\langle P, N\rangle$,

(312) What allows the non-logical reading of [*only* $[_S \cdots [\cdots N \cdots]_{\mathrm{F}} \cdots]]$:

a. What assumptions make the non-logical reading available for sentences of the form [*only* $[_S \cdots [\cdots N \cdots]_{\mathrm{F}} \cdots]]$?

b. Why do these assumptions not predict a non-logical reading for [*only* $[_S \cdots [\cdots P \cdots]_{\mathrm{F}} \cdots]]$?

My proposal, in a nutshell, is that sentences like 'John only ate at most 3 cookies' contain an unpronounced existential quantifier in the prejacent. The meaning of these LFs is paraphrased as follows: there is a plurality of cookies that John ate, and [at most 3] holds of the number/size of that plurality. The exclusive component of *only* adds that there are no similar pluralities of sizes greater than [at most 3]. As I will ultimately show, adding an existential quantifier changes the monotonicity of the prejacent, and with this change, the factors that block the association without $\exists$ (density for comparatives, mirativity for POS and *at most*) will not produce the same results in the presence of $\exists$. This is a brief overview of the answer I propose for (312a). In the case of the positive antonyms, question (312b), I will argue that adding $\exists$ does not make a difference, because the problems that block association with *only* without $\exists$ will continue to be problems when the quantifier is added. This, I argue, is the reason why *only*-*N* can have a grammatical, albeit non-logical reading, while *only*-*P* cannot.

As we will see, the technical details behind this informal sketch are involved, and they require a number of assumptions. Chief amongst these is the proposal that, when MCSEs are constructed for a given prejacent, so that its innocently-excludable alternatives may be determined, there is no access to information regarding distributivity, a notion that I define shortly. I will show that dropping distributivity changes an important and unwanted logical property of structures where the NQ is outscoped by an existential quantifier.

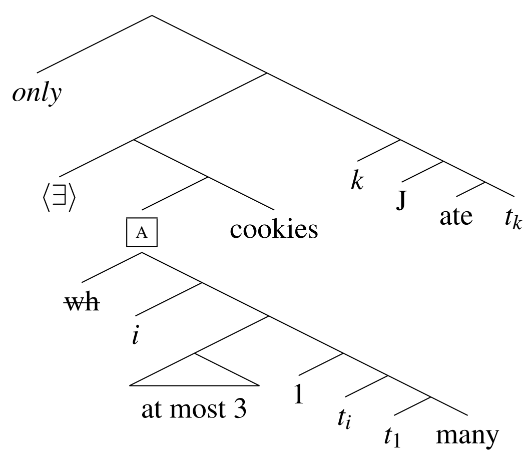
7.1 The non-logical reading under ∃

7.1.1 *at most/at least* – part 1

Let us begin with (313). In the LF in (314), existential closure applies above the degree quantifier [*at most* 3].

(313) John only ate at most 3 cookies.

(314)



Node $\boxed{\text{A}}$ denotes a property of individuals that holds of x iff its size is at most 3. The phrase is generated by raising a silent wh-like operator that then binds its individual-type trace. The denotation of $\boxed{\text{A}}$ composes by Predicate Modification with $[\![\text{cookies}]\!]$, to make the complex predicate $\lambda x.|x| \leq 3$ & $[\![\text{cookies}]\!](x) = 1$. The silent quantifier $\langle\exists\rangle$ takes this predicate as its first argument, and the property of being eaten by John as its second. The prejacent is therefore true iff there is a plurality of cookies that was eaten by John whose size is at most 3.

I now show that exclusion by *only* is too strong in (314), owing to the excessive weakness of the prejacent. First let us see what alternatives the prejacent has: in Chapter 5, I followed Schwarz (2011) and assumed that the modifier *exactly* is a formal alternative to *at most*. If we schematically represent the prejacent as $\exists x(|x| \leq 3 \,\&\, P(x))$, where P is the property of being cookies and being eaten by John, we get the set of alternatives in (315).

(315) $$\mathrm{ALT}(\exists x(|x| \leq 3 \,\&\, P(x))) = \{\exists x(|x| \leq d \,\&\, P(x)) : d \in \mathbf{D}_d\} \cup \{\exists x(|x| = d \,\&\, P(x)) : d \in \mathbf{D}_d\}$$

It turns out that none of the elements of the first set of alternatives are innocently excludable, because they are all equivalent to it. The reason is distributivity, as defined in (316).

(316) **Distributivity**:
$$\forall d \forall d'(\exists x(|x| = d \wedge P(x)) \wedge (d \geq d') \rightarrow \exists x(|x| = d' \wedge P(x))$$

(316) states that for any two degrees d, d' where $d \geq d'$, if there is an individual with property P and size d, then there is a (sub-)individual with property P and size d'. If we assume (316), we get the result that all propositions of the form $\exists x(|x| \leq d \& P(x))$ are equivalent:

(317) $\exists x(|x| \leq d) \& P(x)$ is true iff $\exists x(|x| \leq d') \& P(x)$ for any two degrees d, d', e.g. 3,4:

(i) $\boxed{\rightarrow}$ There is an x of cardinality of at most 3, and $P(x)$. Therefore, x's size is less than 4, which means that x is of cardinality of at most 4. Therefore there is an x of cardinality of at most 4, and $P(x)$.

(ii) $\boxed{\leftarrow}$ There is an x of cardinality of at most 4, and $P(x)$. Then either x's size is exactly 4, or at most 3. If exactly 4, then, by distributivity, there is a subindividual of x of size 3, i.e. at most 3. Therefore there is a x of cardinality of at most 3, and $P(x)$.

The result in (317) can be understood (slightly) differently as follows: $\exists x(|x| \leq 3 \;\&\; P(x))$ is equivalent to the disjunction $\exists x(|x| = 3 \;\&\; P(x)) \vee \exists x(|x| = 2 \;\&\; P(x)) \vee \exists x(|x| = 1 \;\&\; P(x))$.[1] By distributivity, the first of these three disjuncts asymmetrically entails the second, and the second asymmetrically entails the third. The disjunction is therefore equivalent to its weakest disjunct, which makes the expression $\exists x(|x| \leq 3 \;\&\; P(x))$ equivalent to $\exists x(|x| = 1 \;\&\; P(x))$. Crucially, the same result obtains when 3 is replaced with any other numeral, because the numeral adds stronger disjuncts to the expression, and the expression remains equivalent to the weakest.

It follows from this equivalence that none of the '*at most*'-alternatives in (315) can be excluded consistently with (313)'s prejacent; because they are each equivalent to it, negating any of them would contradict it.

[1] For clarity I assume granularity at the integer level. As far as I can see, the point remains the same at other levels.

(318) $\text{ALT}(\exists x(|x| \leq 3 \,\&\, P(x))) = \cancel{\{\exists x(|x| \leq d \,\&\, P(x)) \,:\, d \in \mathbf{D}_d\}} \cup$
$\{\exists x(|x| = d \,\&\, P(x)) \,:\, d \in \mathbf{D}_d\}$

This leaves the '*exactly*'-alternatives. Clearly, it is consistent to maintain our prejacent $(\exists x(|x| \leq 3) \& P(x))$ and add that no individual with property P is of cardinality 4, 5, 6, etc. This amounts to the claim that some x is P, e.g. cookies eaten by John, and x's size is 3, 2, or 1, and moreover, no x is P and is of size 4 or 5 or 6 etc. This is the meaning we want.

The problem begins when we consider other alternatives: can the alternative $\exists x(|x| = 3 \,\&\, P(x))$ be consistently excluded with the prejacent? In other words, is it consistent to say that there is a collection of cookies that John ate, whose size is 3, 2 or 1, *and* that there is no collection of such cookies of size 3? The answer is clearly yes, because this results in the consistent claim that there is a collection of size 2 or 1 (of cookies that John ate). We can go even further: x's size can be exactly 1 and the prejacent would still be held true. So the alternative $\exists x(|x| = 2 \,\&\, P(x))$, is consistently excludable. We may therefore construct an MCSE M, whose result is the negation of every proposition of the form $\exists x(|x| = n \,\&\, P(x))$ for ns that are greater than 1:

(319) $M = \{\exists x(|x| = n \,\&\, P(x)) \,:\, n > 1\}$

Since M is consistent, and since it contains everything that can be excluded given our prejacent, the set $\text{IE}(S)(A)$ is M: the *at most* alternatives cannot be negated, and all remaining excludables are in (319). The predicted meaning of (313) is therefore that John ate exactly 1 cookie, and, crucially, this result does not change if, instead of the numeral 3, the prejacent had the numeral 1,000: (313) would still mean that John ate exactly 1 cookie.

(320) $[\![(314)]\!] = 1$ iff John ate exactly 1 cookie.

Our goal is to somehow distinguish between, on the one hand, the alternatives $[\exists x(\cdots = 4)]$, $[\exists x(\cdots = 5)]$, etc, and on the other hand the alternatives $[\exists x(\cdots = 3)]$ and $[\exists x(\cdots = 2)]$. If these latter alternatives are somehow removed from the set M, the meaning of (313) will only negate that John had 4, 5, etc. cookies, and will not exclude the possibility that he had 3 or 2. This is exactly the meaning that we want. In the next section, I show that dropping distributivity affords us with just this restriction of M.

Constraining MCSEs by dropping distributivity

In this section I show that, if MCSEs are constructed blindly to distributivity, we correctly restrict the alternatives to our problematic prejacent. There are a number of ways of implementing blindness. One may postulate that the algorithm that generates CSEs views all individuals as atomic. This means that, as far as this module is concerned, every individual, whether intuitively singular or plural, is opaque. That is, for any given entity x there are no ys that satisfy $y \sqsubset x$. Let us call this the atomicity assumption. Another way is to say that the algorithm that builds CSEs has no access to the *lexical* properties of P, for any P that is predicated of the relevant individual. Call this the extra-lexical assumption.[2]

Either assumption defuses step (ii) in (317), repeated below:

(317ii) $\exists x(|x| \leq 4)$ & $P(x)$ entails $\exists x(|x| \leq 3) \& P(x)$

Justification: There is an x of cardinality of at most 4, and $P(x)$. Then either x's size is exactly 4, or at most 3. If exactly 4, then, by distributivity, there is a subindividual of x of size 3, i.e. at most 3. Therefore there is a x of cardinality of at most 3, and $P(x)$.

On the atomicity assumption the justification in (317ii) is not valid, because the existence of a smaller individual with property P is not a logical consequence of the existence of some larger x with property P (though it is consistent with it). On the extra-lexical assumption, step (ii) is licensed if the predicate P is distributive, but since the algorithm has no access to the properties of P, owing to its blindness to lexical information, (ii) can never hold. Note that on both assumptions, step (i) in (317)—repeated below—remains valid, for the simple reason that $\leq n+1$ follows *logically* from $\leq n$, for any n.[3]

(317i) $\exists x(|x| \leq 3)$ & $P(x)$ entails $\exists x(|x| \leq 4) \& P(x)$

Now alternatives of form $\exists x((\leq n)(x) \& P(x)$ are totally-ordered, as in (321).

(321) $\exists x((\leq \mathbf{n})(x) \& P(x) \Vdash \exists x((\leq \mathbf{n{+}1})(x) \& P(x) \Vdash \exists x((\leq \mathbf{n{+}2})(x) \& P(x) \Vdash \cdots$

Another result we obtain is that '$\exists$-*exactly*'-alternatives are independent of one another: if there is an individual x of size exactly n and property P, there need not be a bigger individual with property P, nor does there need be a smaller individual with property P (either because individuals are atomic, or because the (unavailable) meaning of P is needed before the sub-individual inference can be drawn). The consequence, as I will now show, is that the symmetry that played a key role in blocking the logical reading of '*at most*/*at*

[2]The extra-lexical assumption is due to Gajewski (2002). An earlier version of it is attributed to Kai von Fintel by Fox and Hackl (2006).

[3]There may be an inconsistency in invoking the extra-lexical principle here, since the system that builds CSEs must at least have access to the meaning of *at most*, which on our assumptions is interpreted along with the adjective *many*. Therefore, the semantics of *many* must be accessible in order for the CSE-builder to understand the logical properties of *at most*.

least'-sentences (minus existential quantification – see Chapter 5), is now going to hold between some of A's elements as well.[4]

7.1.2 *at least/at most* – part 2, and other *P/N* pairs

To see how our current set-up works, let us continue with our example (313), but now with different notation: the proposition $\exists x((\leq n)(x)\&P(x))$ holds just in case some individual x is P and x's size is somewhere between 1 and n (again assuming integer-level granularity for illustration). But since the existential quantifier commutes with disjunction, the proposition is equivalent to: some x is P and x's size is 1, or some x is P and is of size 2, and so on. As long as there is an x whose size is in the interval $[1,n]$, and $P(x)$, the sentence will be true. In what follows I use interval notation to mark the size(s) that are needed in order for the relevant proposition to hold:

$[1,4]$ is shorthand for $\exists x((\leq 4)(x)\ \&\ P(x))$;

$[3]$ is shorthand for $\exists x((=3)(x)\ \&\ P(x))$;

$[n,\infty)$ is shorthand for $\exists x((\geq n)(x)\ \&\ P(x))$.

Our example prejacent is $\exists x((\leq 3)(x)\&P(x))$, which on our current notation is $[1,3]$. The set of alternatives A for $[1,3]$ is divided into the four subsets in (322).

(322) $$\begin{aligned}\mathrm{ALT}(S) = {} & \cancel{\{[1,3],[1,4],[1,5],\cdots\}}\cup \\ & \{[1,2]\}\cup \\ & \{[3],[2],[1]\}\cup \\ & \boxed{\{[4],[5],[6],\cdots\}}\end{aligned}$$

The first row in (322) shows a set of alternatives that logically follow from the assumed prejacent: if some individual (with some relevant property) is of a size in $[1,3]$, then its size is in $[1,4]$, $[1,5]$, etc. No MCSE can contain these alternatives, because adding any of them, e.g. $[1,4]$, forces the absurd meaning that the relevant measure is in $[1,3]$ but outside of $[1,4]$.

The next two rows contain symmetric alternatives. The prejacent is consistent with the negation of $[1,2]$: this means that the measure is in $[1,3]$ but not in $[1,2]$, which puts it in $[3]$, i.e. there is a collection of size exactly 3 (of cookies that John ate). Now we cannot add to this same CSE the alternative $[3]$, because this will contradict the meaning that results

[4] Recall, however, that symmetry was not enough to account for the full facts about *at least/at most*.

from the set. So any CSE that includes $[1,2]$ must not include [3]. The reverse also holds: The prejacent $[1,3]$ is consistent with the negation of [3]. The result means $[1,2]$, i.e. there is a collection size 1 or 2 (of cookies that John ate). This set cannot contain the alternative $[1,2]$, because negating $[1,2]$ is inconsistent with the conjunction of $[1,3]$ with the negation of [3]. Therefore neither $[1,2]$ nor [3] is innocently-excludable.

We may also construct a CSE $\{[1],[3]\}$, from which [2] must be absent, and another CSE $\{[2],[3]\}$, from which [1] is absent. Therefore, neither [1] nor [2] is IE.

This leaves the boxed subset, whose members can all be consistently negated with every other excludable alternative in A. The set $\mathrm{IE}(S)(A)$, constructed under the atomicity/extra-lexical assumption, is therefore the boxed set in (322). The exclusion in our example will therefore mean that there is an individual which is cookies that John ate, whose size is at most 3, and no individual of size 4 or 5 etc. satisfies the same property. Once distributivity is re-introduced in the semantics, i.e. after atomicity is dropped, the inference to eating individual cookies becomes available:

(323) $only(\exists x((\leq 3)_{\mathrm{F}}(x)\ \&\ P(x)))$ is defined only if some stronger alternative $S' \in \mathrm{IE}(S)(A)$ is such that $\mathbf{N}(\lambda w[\![S']\!]^w)$,

i.e. only if $\mathbf{N}(\lambda w[\![\exists x((=4)(x)\ \&\ P(x))]\!]^w)$, or
$\mathbf{N}(\lambda w[\![\exists x((=5)(x)\ \&\ P(x))]\!]^w)$, or
$\mathbf{N}(\lambda w[\![\exists x((=6)(x)\ \&\ P(x))]\!]^w)$, etc.

A crucial thing to note about (323) is that $\lambda w[\![\exists x((=4)(x)\ \&\ P(x))]\!]^w$, for example, is assumed to be logically-stronger than the prejacent. If *only*'s mirative component is evaluated *after* distributivity is re-introduced, then this is true: if a quartet of cookies was eaten by John, then it follows that there is a cookie (sub-)collection, of size 3 or less, that John ate. But now we ask what happens if a *collective* predicate is used instead? If P were, say, the property of forming a circle, then surely it doesn't follow from *exactly 4 people formed a circle* that there is a sub-plurality of people (of size at most 3) who formed a circle. In this case, *only*'s mirative component cannot be satisfied, because none of the prejacent's IE-alternatives are logically stronger, even after atomicity/extralexicality is dropped. I will get back to this issue in Section 7.2.

Let us turn to the case of *at least*. The formal alternatives to the prejacent $S = [3,\infty)$ are divided into the subsets in (324).

(324) $\mathrm{ALT}(S) = \cancel{\{[3,\infty),[2,\infty),[1,\infty)\}} \cup$
$\{[4,\infty),[5,\infty),\cdots\} \cup$

$$\{[3],[4],\cdots\}\cup$$
$$\boxed{\{[2],[1],\cdots\}}$$

As in the case of (322), the top row contains non-excludable alternatives, because they all follow logically from the prejacent: if a measure's size is at least 3, i.e. somewhere between 3 and infinity, then its size is at least 2 (between 2 and infinity). Turning now to the second and third sets, we see that none of their elements are innocently-excludable. Take $[5,\infty)$ from the second set, for example. The alternative cannot be excluded consistently with [3] and [4], because the result would require that the measure be in $[3,\infty)$—the prejacent—but not in [3], [4], or $[5,\infty)$. Therefore $[5,\infty)$ cannot belong to any CSE that contains [3] and [4], and is therefore not innocently excludable. The same can be shown for any member of the second set.

The third set contains no IE alternatives either. To see why, note first that negating all of its elements will contradict the prejacent. But as soon as one of its elements is removed, e.g. [7], the result will be a CSE: the resulting set negates every alternative 'exactly n', with the exception of 'exactly 7'. These negations are consistent with the prejacent, because the resulting proposition is that there is a plurality of size 3 or more (of cookies that John ate), but not of size 3, 4, 5, 6, 8, etc. Therefore, the plurality must be of size 7. As a result, the alternative [7] cannot be added to the set, and this disqualifies it from being IE. The same can repeated for any [n] from 3 and up.

This leaves the boxed set at the bottom, whose members are consistent with the prejacent, and with the negation of every other alternative in A. As a result, we derive the following definedness condition:

(325) $only(\exists x((\geq 3)_{\mathrm{F}}(x) \mathbin{\&} P(x))) = 1$ is defined only if some stronger alternative $S' \in \mathrm{IE}(S)(A)$ is such that $\mathbf{N}(\lambda w [\![S']\!]^w)$,

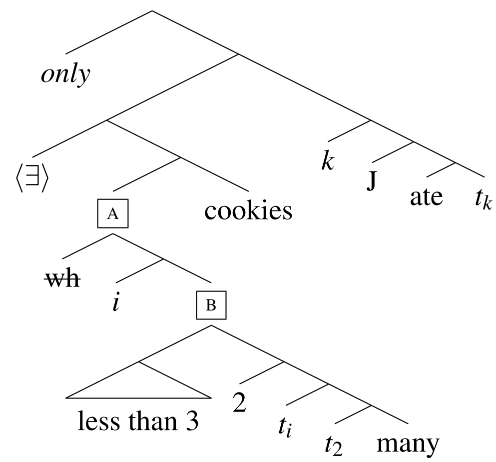
If we assume that (325) is checked after distributivity is activated, then it does not matter whether P is distributive or not. In the case of *at least* there cannot be any stronger IE-alternatives, let alone stronger ones that have the property $\mathbf{N}$ (recall the discussion in Chapter 5). As I said, the distributive/collective distinction will be discussed in Section 7.2.

The comparative

In this discussion, the only change to the LF in (313) is that the property of having a cardinality of at most 3, will now be replaced with the property of having cardinality less than 3. The *at most* phrase will therefore have to be replaced with a more complex negative comparative:

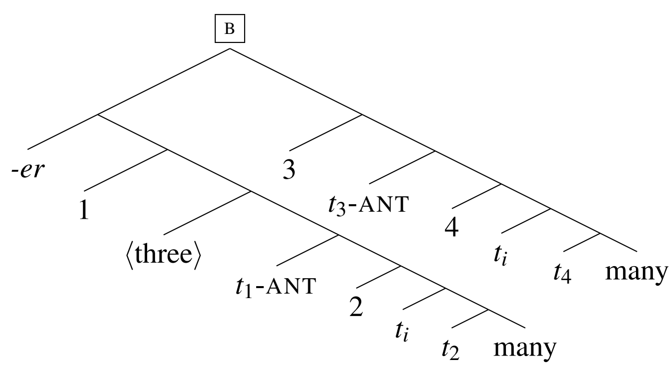
(326) John only ate less than 3 cookies.

(327)



Node $\boxed{\text{A}}$ is derived, as before, by abstracting over the trace of a raising wh-operator. The result is the property of measuring to less than 3 in size. Note: the comparative node $\boxed{\text{B}}$ is substantially abbreviated in (327). Its detailed structure (faithful to the structure assumed in Chapter 6) is shown in (328). Recall also the representation of unmodified numerals in *than*-clauses that was developed in Chapter 6, which is abbreviated in the LF as ⟨three⟩.[5]

(328)



[5] Note also that ⟨*three*⟩ contains the antecedent of a conditional, of the form 'if x's size were 3'. Here x is a bound variable, but what does it mean to talk about worlds where a *variable*'s size is 3? I believe that this is a technical issue that can be resolved if x varied over individual concepts instead of individuals (type $\langle s,e\rangle$ instead of type e). I thank Irene Heim for pointing this out.

We first note that distributivity gives us the same problem encountered earlier: the proposition $\exists x(<d)(x)$ & $P(x)$ is equivalent to $\exists x(<d')(x)$ & $P(x)$, for any d,d':

(329) $\exists x(|x|<d)\&P(x)$ is true iff $\exists x(|x|<d')\&P(x)$ for any two degrees d,d', e.g. 3,4:

(i) There is an x of cardinality less than 3, and $P(x)$. Therefore, x's size is less than 4, which means that x is of cardinality of less than 4. Therefore there is an x of cardinality of less than 4, and $P(x)$.

(ii) There is an x of cardinality of less than 4, and $P(x)$. Then either x's size is less than 3, or greater or equal to 3. If greater or equal to 3, then, by distributivity, there is a subindividual of x of size less than 3. Therefore there is a y of cardinality of less than 3, and $P(y)$.

(329) tells us that no alternative of the form $\exists x((|x|<n)$ & $P(x))$ is excludable given a prejacent of the same form, because they are all equivalent. What alternatives does that leave? In the case of *at most*, we had the *exactly*-alternatives. In the case of *less than* we may derive two kinds of alternatives: by removing the antonymizing component ANT, we may derive *more than*, and by simplifying the structure further we may derive the bare numeral, which gives us the *at least*-reading of n.[6] Negating a sentence $\exists x(|x|>n\&P(x))$ amounts to saying that every x which is P is of size at most n. Negating a sentence $\exists x(|x|\geq n\&P(x))$ amounts to saying that every x which is P is smaller than n.

Many of these negations are consistent with the prejacent, which we may now represent in interval notation as $(0,3)$—the parenthesis indicates that the measure is strictly below 3, i.e. less than 3. With this prejacent, it is consistent to negate '*at least* 2', for example, because this means that John ate less than 2 cookies, i.e. exactly 1 on integer-level granularity.

The meaning we *want* to derive should negate $[3,\infty)$, $[4,\infty)$, etc. but not those containing lower numerals, e.g. $[2,\infty)$. That is, we only want the strengthening to go as far as $(0,3)$. With distributivity this is not possible. But without it, some of the *less-than*-alternative will be activated, specifically those with the lower numerals, and their activation will disqualify the unwanted alternatives from membership in IE. In keeping with F&H's density assumption, I abandon granularity from now on.

Let the prejacent S be $\exists x(|x|<3\&P(x))$, $(0,3)$ for short. The set of alternatives A is divided into the following subsets.

(330) $A = \cancel{\{(0,d) : d \geq 3\}} \cup$

[6]Recall these results from Chapter 3 that the positive results from simplifying the negative antonym.

$$\{(0,d) : d<3\}\cup$$
$$\{(d,\infty) : d<3\}\cup$$
$$\{[d,\infty) : d<3\}\cup$$
$$\boxed{\{(d,\infty) : d\geq 3\}}\cup$$
$$\boxed{\{[d,\infty) : d\geq 3\}}$$

The first subset contains logical consequences of the prejacent, so none of its members are consistently excludable with it. The second, third, and fourth subsets contain excludable alternatives, but no innocently excludable ones.

Take for example an element S' of the second set, $S' = (0,d')$, for some d' that is lower than 3. Is there a CSE that cannot include S'? Yes: take from the fourth set $S'' = [d',\infty)$. Negating S'' is consistent with the prejacent, because the result means that some cookie-collection smaller than 3 was eaten by John, and no such collection exists of size d' or more, where d' is smaller than 3. This means that John ate less than d'-many cookies, i.e. $(0,d')$. So a CSE that contains S'' cannot contain S', because the meaning it gives rise to entails S'.

The third set is also entirely non-IE: take any member $S' = (d',\infty)$. Now take a degree $d''>d'$, and let $S'' = (0,d'')$. The negation of S'' entails $[d'',\infty)$, which is inconsistent with S'. Therefore, any CSE containing S'' cannot contain S'. Therefore S' is not IE.

Finally the fourth set: let $S' = [d',\infty)$. A CSE that contains $(0,d')$ entails S', and so it cannot contain it. Therefore, no element from the fourth set is IE.

This leaves the alternatives in the boxed subsets. From the bottom set take the alternative $S' = [3,\infty)$. To find a CSE the entails S', the CSE must include its complement in its alternatives. S''s complement is $[0,3)$, but this alternative is not excludable to begin with, because it negates the prejacent. Therefore, no CSE can entail S', so every CSE contains it. S' is therefore innocently-excludable. The same holds of the elements of the top boxed-set, whose negations are entailed by the negations of the bottom set. This, then, gives us the following IE$(S)(A)$, given an algorithm that builds MCSEs blindly to distributivity:

(331) $\text{IE}(\exists(\textit{less than } 3))(A) = \{\exists(\textit{at least } d) : d\geq 3\}$

(more formally): $\text{IE}(\exists x(|x|<3\,\&\,P(x)))(A) = \{\exists x(|x|\geq d\,\&\,P(x)) : d\geq 3\}$

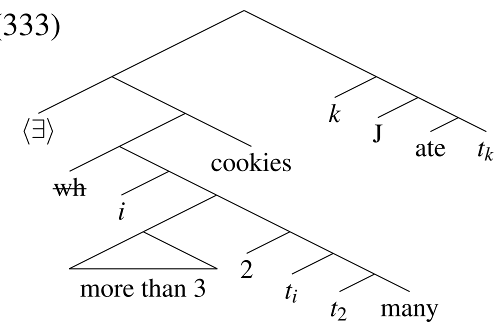
We therefore have the following definedness condition on a sentence like (326).

(332) $\textit{only}(\exists x((<3)_{\text{F}}(x)\ \&\ P(x)))$ is defined only if some stronger alternative $S'\in\text{IE}(S)(A)$ is such that $\mathbf{N}(\lambda w[\![S']\!]^w)$,

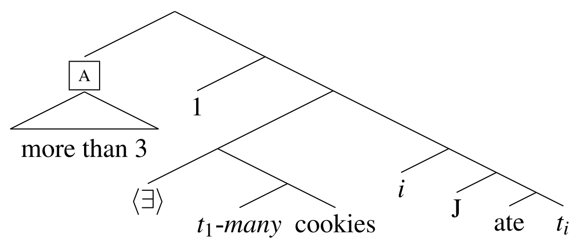
i.e. only if $\mathbf{N}(\lambda w[\![\exists x((\geq 3)(x)\ \&\ P(x))]\!]^w)$, or
$\mathbf{N}(\lambda w[\![\exists x((\geq 4)(x)\ \&\ P(x))]\!]^w)$, or
$\mathbf{N}(\lambda w[\![\exists x((\geq 5)(x)\ \&\ P(x))]\!]^w)$, etc.

In the case of *more than*, nothing changes with the addition of the existential quantifier. It is easy to see that, if there is a plurality of cookies that John ate, and whose size exceeds 3 ($\exists x(|x|>3\&P(x))$), then, and only then, the number of cookies that John ate exceeds 3, i.e. $[\![\text{more than }3]\!](\lambda d[\![\text{J ate }d\text{-many cookies}]\!])=1$. The two LFs in (333) and (334) are equivalent.

(333)



(334)



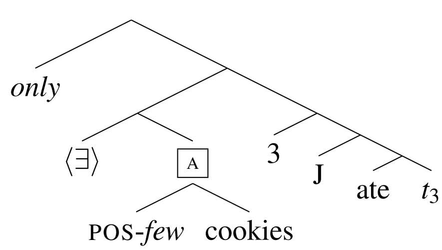
This means that the alternatives to (333) are each equivalent to their translation to (334), so whatever makes *only* incompatible with (334) should also make it incompatible with (333). In Chapter 6 we saw that if density is assumed, there cannot be any MCSEs given (334) and its alternatives, and because (333) is equivalent to (334), and its alternatives are *mutatis mutandis* equivalent to those of (334), there cannot be any MCSEs for (333) either.

POS/*very*

We finally turn to the case of POS. My example is (335), whose ∃-LF is shown in (336).

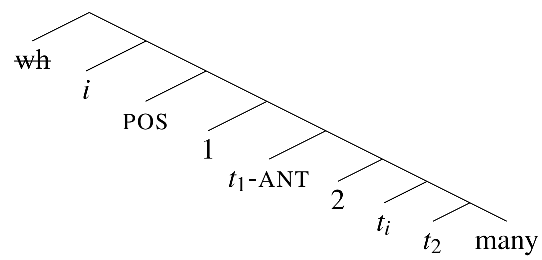
(335) John only ate (very/POS)-few cookies.

(336)



Note that the property of fewness above is derived from the more detailed structure of node $\boxed{\text{A}}$, shown below.

(337)



Having already covered the case of the comparative, we may abstract away from the compositional details in (336), by recalling that POS holds if its argument D properly contains the neutral (**N**) degrees.[7] In the case of *many*, POS holds iff the greatest neutral degree is below D's maximal element, and in the case of *few*, POS holds iff **N**'s *minimal* degree is *above* D's maximal element. Instead of referring to **N**, we may for our purposes refer to its maximal element, call it $\mathbf{s}_{many}$, and its minimal element, $\mathbf{s}_{few}$. The prejacent S in (336) holds if some plurality (of cookies etc.) is smaller in size than $\mathbf{s}_{few}$. Let us then abbreviate S as $(0, \mathbf{s}_{few})$. The alternatives to S are those where POS is replaced with a degree name (as assumed in Chapter 4), and where $[t_1\text{-ANT}]$ is structurally simplified into t_1, creating the positive meaning. The first replacement produces alternatives of the form $\exists x(P(x)\,\&\,[\![\text{few}]\!](d)(\lambda d'.|x|\geq d'))$, which are true iff some x has property P, and the degrees d' that are reached by x's size do not include d (this is the contribution of ANT). x must therefore be smaller than d, so in effect, the replacement gives us alternatives that are equivalent to $\exists x(P(x)\,\&\,|x|<d)$, for any d in the domain of degrees:

(338) $\{(0,d) : d \in \mathbf{D}_d\} \subseteq \text{ALT}(S)$

The other operation removes ANT. If POS is not replaced with a degree name, this gives us the [POS-*many*] alternative. If POS is replaced with a degree name d, we get the [d-*many*] alternative. This latter result, $\exists x(P(x)\,\&\,(\lambda d'.|x|\geq d')(d))$, holds iff some x is P and is at least d-big, i.e. iff $\exists x(P(x)\,\&\,|x|\geq d)$:

(339) $\{[d,\infty) : d \in \mathbf{D}_d\} \subseteq \text{ALT}(S)$[8]

[7] In Chapter 4, Section 4.2, this was redefined in terms of neutral propositions instead of degrees, and it was shown then that the change does not affect the truth-conditions.

[8] I will ignore the alternative [POS-*many*]; the alternative's truth conditions require that some x be P, and have a size that properly includes the neutral degrees **N**, i.e. have a size that exceeds $\mathbf{s}_{many}$ (**N**'s upper bound). [POS-*many*] is therefore equivalent to $[\mathbf{s}_{many}, \infty)$. I ignore the alternative because there is certainly an alternative S' in (339) such that $S' = [\mathbf{s}_{many}, \infty)$.

The sets in (338) and (339) can be divided in two halves each, as in (340). This is now reminiscent of the case of the comparative:

(340) $A = \cancel{\{[0,d) : d \geq \mathbf{s}_{few}\}} \cup$
$\{[0,d) : d < \mathbf{s}_{few}\} \cup$
$\{[d,\infty) : d < \mathbf{s}_{few}\} \cup$
$\boxed{\{[d,\infty) : d \geq \mathbf{s}_{few}\}}$

As in the earlier cases, the struck-out set contains no excludable alternatives. The negation of any element of the second set will conflict with the negation of some element of the third (and vice versa), in exactly the same way as the comparative, leaving only the alternatives in the boxed subset. The meaning after association with *only*, as desired, is that some group of less than $\mathbf{s}_{few}$ cookies were eaten by John, and no similar group of cookies are of size $\mathbf{s}_{few}$ or more.

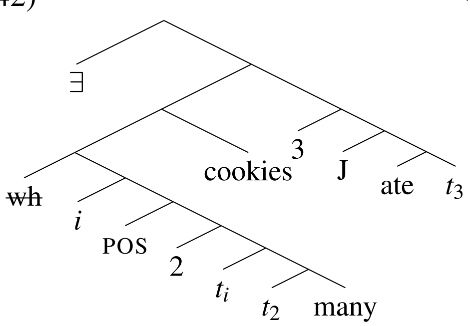
The definedness condition on a sentence like (335) is shown in (341).

(341) $only(\exists x((<\mathbf{s}_{few})_{\mathrm{F}}(x) \ \& \ P(x)))$ is defined only if some stronger alternative $S' \in \mathrm{IE}(S)(A)$ is such that $\mathbf{N}(\lambda w[\![S']\!]^w)$,

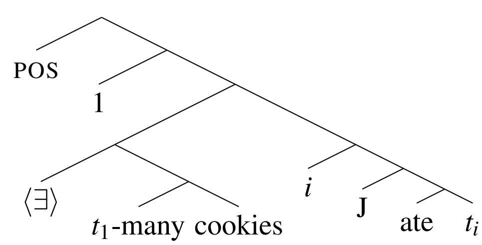
i.e. only if $\mathbf{N}(\lambda w[\![\exists x((\geq d)(x) \ \& \ P(x))]\!]^w)$, for some $d \geq \mathbf{s}_{few}$

The POS-*many* case remains ungrammatical under existential quantification: if some plurality of eaten cookies is bigger than $\mathbf{s}_{many}$ (the meaning of (342)), then the degrees d that satisfy [d-many cookies were eaten] must exceed $\mathbf{s}_{many}$ (the meaning of (343)). The reverse also holds. So whatever violation is committed by the second proposition (mirativity, as is argued in Chapter 4), the same violation is committed by the first.

(342)



(343)



7.1.3 Summary

This section began with LFs where NQs appeared in the scope of the silent existential quantifier $\langle\exists\rangle$. Initially a problem was diagnosed where the LFs were shown to have the same meaning, regardless of the numeral/degree expression contained in the NQ. This had an undesirable effect on the meaning of these LFs when they appear under *only*: because all alternatives that feature the NQ are equivalent, independently of the numeral, none of them is consistently excludable with respect to the prejacent. This opened the door for other excludable alternatives to be negated, and gave rise to an unattested meaning. However, the problem was shown to depend crucially on the distributivity assumption. If distributivity is dropped during the computation of CSEs, the set of exclusions is restricted, due to further competition between CSEs. The correct exclusions arise as a result. The IEs for each NQ are summarized below.

(344) **IE-sets for NQs under existential closure, and results for association with *only***

a. $\text{IE}(\exists(\textit{at most n})_{\text{F}})(A) = \{\exists(\textit{exactly } d) : d{>}n\}$
$[\![\textit{only}_A(\exists(\textit{at most n})_{\text{F}})]\!] = 1$ only if $\forall d((d{>}n) \rightarrow ([\![\exists(\textit{exactly } d)]\!] = 0))$

b. $\text{IE}(\exists(\textit{less than } n_{\text{F}}))(A) = \{\exists(d) : d{\geq}n\}$
$[\![\textit{only}_A(\exists(\textit{less than } n_{\text{F}}))]\!] = 1$ only if $\forall d((d{\geq}n) \rightarrow ([\![\exists(d)]\!] = 0))$

c. $\text{IE}(\exists(\text{POS-}\textit{few})_{\text{F}})(A) = \{\exists(d) : d{\geq}\mathbf{s}_{\textit{few}}\}$
$[\![\textit{only}_A(\exists(\text{POS-}\textit{few})_{\text{F}})]\!] = 1$ only if $\forall d((d{\geq}\mathbf{s}_{\textit{few}}) \rightarrow ([\![\exists(d)]\!] = 0)$

With the positive antonyms we saw that existential closure made no difference. The IE-sets for the positives are shown in (345).

(345) **IE-sets for PQs under existential closure, and results for association with *only***

a. $\text{IE}(\exists(\textit{at least n})_{\text{F}})(A) = \{\exists(\textit{exactly } d) : d{<}n\}$
$[\![\textit{only}_A(\exists(\textit{at most n})_{\text{F}})]\!] = 1$ only if $\forall d((d{<}n) \rightarrow ([\![\exists(\textit{exactly } d)]\!] = 0))$

b. $\text{IE}(\exists(\textit{more than } n_{\text{F}}))(A) = \emptyset$
$[\![\textit{only}_A(\exists(\textit{more than } n_{\text{F}}))]\!]$ is undefined.

c. $[\![\textit{only}_A(\exists(\text{POS-}\textit{many})_{\text{F}})]\!] = [\![\textit{only}_A(\text{POS-}\textit{many})_{\text{F}}]\!]$, which defined only if $\mathbf{N}(\text{POS})$, i.e. undefined.

7.2 Collective/Distributive?

The collective/distributive distinction emerges naturally in the theme of this chapter: how does the distinction affect the logical relationship between alternatives, and what effect does this have on the semantics of *only*?

Let me mention at the outset that the answer to this question, whatever it is, is orthogonal to the method of constraining alternatives that I presented. One of the main goals in this chapter was to identify a property that distinguishes two kinds of alternatives: given an LF where $\langle\exists\rangle$ outscopes an NQ like [$\leq$3], we wanted to separate alternatives where 3 is replaced with a lower numeral, from those where it is replaced with a higher one. It was shown that the former are logical consequences of the LF's denotation, while the latter follow only under distributivity. Dropping the assumption produced an asymmetry between the two families of sentences, in a way that was useful in building CSEs.

Naturally, information about distributivity/collectivity re-emerges when the truth conditions of the LF are computed. Now, what effect does this difference have on the semantics of *only*? The answer is that it is the same effect even if the CSEs are not constrained in the way I argued for.

Take, for example, a predicate like *form a circle*. The predicate is truly collective in that, whenever it holds of a group of a certain size, one infers neither that a smaller group formed a circle, nor that a bigger group did. Logically, both alternatives bear the same relation to the original proposition: they are logically-independent of it. So, as far as our predictions go, associating *only* with e.g. [$\langle\exists\rangle$ *at most*] should be unacceptable, as in e.g. (346).

(346) They only formed a circle of [at most 5]$_{\mathrm{F}}$ people. (should be undefined)

In (347) I repeat the mirative requirement of *only* for sentences like (346).

(347) *only*$(\exists x((\leq 5)_{\mathrm{F}}(x)\ \&\ [\![\text{circle}]\!](x)))$ is defined only if some stronger alternative $S' \in$ $\mathrm{IE}(S)(A)$ is such that $\mathbf{N}(\lambda w[\![S']\!]^{w})$

The reason why (347) is not met is that none of the IE-alternatives in (348) entail the prejacent. Therefore, (346) is predicted to be undefined.

(348) $[\exists x((=6)(x)\ \&\ [\![\text{circle}]\!](x))]$,
$[\exists x((=7)(x)\ \&\ [\![\text{circle}]\!](x))]$,
$[\exists x((=8)(x)\ \&\ [\![\text{circle}]\!](x))]$.

The prediction is not borne out, however, as there is no detectable difference between (349), which contains a collective predicate, and (350), where the predicate is distributive.

(349) They only formed a circle of [at most 5]$_F$ people. (should be undefined)

(350) John only ate [at most 5]$_F$ cookies. (should be defined)

The point I want to make is that this is a more general problem that arises in the absence of NQs. Consider (351), where *only*'s prejacent contains a bare numeral and a collective predicate.

(351) Only 5$_F$ people formed a circle.

Here, *only*'s mirative component requires that some *stronger* alternative be expected, but again the requirement cannot be satisfied, because replacing the numeral in the prejacent produces logically-independent alternatives. So there cannot be any *stronger* **N**-alternatives. Yet the sentence is perfectly acceptable, which means that, if the mirative condition is right, then something must coerce collective predicates to behave just like distributive ones. In other words, the mirative component, as it is currently defined, must be satisfied in some other way in cases like (349/351).

Possibly, this is a case where *only* operates on a truly non-logical scale, where alternatives are ordered not by logical strength, but according to the value of the numeral that they contain. If we make this move in correcting our predictions about (351), that is, if we invoke the scale of numbers (which in the case of collective predicates induces a non-logical ordering), then by the same move we correct our predictions for (346).

These remarks are tentative and require further work. The point I intended to convey is that, while collective predicates present a challenge to our predictions, the challenge arises in more general cases, e.g. bare numerals. I claimed that if the bare numeral problem can be handled, e.g. by invoking the non-logical reading of *only*, then the solution is applicable to the case of NQs as well. I leave it to future work to determine whether my claim is indeed true.

7.3 Remarks on the distribution of $\exists$

The LFs used in this chapter are needed in order for *only*-NQ associations to have their 'non-logical' reading. We see now that the term 'nonlogical'—at least in the case of distributive predicates—is something of a misnomer, since the reading turns out to be logical, but logical relative to a prejacent where existential closure outscopes the NQ ($\exists >$ NQ). I will refer to these LFs as the $\exists$-LFs from now on.

Structures where ∃ appears under the NQ (NQ $>$ ∃) were shown in previous chapters to be incompatible with *only*, leaving the ∃-LFs as the only interpretable variants. The positive antonyms were shown to be problematic on both parses, hence their unacceptability.

But if ∃-LFs are made available by the grammar, what blocks them from appearing in the absence of *only*? This question is often posed in treatments of NQs, and is contemporarily referred to as van Benthem's problem (after van Benthem 1986).[9] The meaning that the LFs give rise to is quite weak, since their truth-conditions are met even if some very large plurality satisfies the relevant property. For example, without *only*, the ∃-parse of [*John ate at most three cookies*] will be true even if John ate 15 cookies, because if he did, there is a subgroup of cookies that he ate whose size is 3 or less. So it may be this weakness that blocks the ∃-parses from being considered, particularly against the more informative parses where the NQ outscopes the existential quantifier.

What if the ∃-parse is embedded in the antecedent of a conditional, or the restrictor of a universal quantifier? Then the meaning of the conditional (or the universal statement) will be strong. For example, the meaning of a sentence like [*everyone who had* ∃ *at most 3 cookies was given a prize*] will incorrectly entail that everyone who had 40 cookies was given a prize. So if the grammar is assumed to prefer strong statements, then the ∃-parses will be allowed, incorrectly, as their placement inside a DE environment strengthens the sentence as a whole. To prevent this, the prohibition against uninformativity must be formulated as a local condition rather than a global one. This would then rule out the embedded restrictor in the current example, even if its weak meaning ends up strengthening the entire sentence.

What about association with *only*? Does the weakness of the ∃-prejacent not locally disallow it in this case as well? Here there are a number of answers that can be given. If ∃-parses compete with the other parses, and if preference is given to the informative *grammatical* alternative, then under *only* we allow ∃-prejacents, because otherwise association with *only* would be ungrammatical. We may add also that with focus marking, the locality condition may not be applicable. The presence of focus marking may block the condition, because focus requires evaluation of alternatives, and this is not done until a focus-sensitive operator like *only* is added into the parse. In other words, an ∃-parse containing focus marking cannot be checked for informativity, until the focus marking is used to determine the meaning of the parse.

[9] See e.g. Diesing (1992), Krifka (1999), Herburger (2000), Hackl (2000), de Swart (2001), Landman (2004), Solt (2009, 2013).

7.4 NQs as Negative Polarity Items – Beck (2012)

Recently, Beck (2012) proposed a decompositional analysis of NQs that is intended, in small part, to capture what I have called the nonlogical reading of *only*-NQ association. In this section I briefly sketch Beck's analysis, and compare its predictions to mine.

The technical details behind Beck's approach are complicated, but I believe the section heading is an accurate summary of her approach. Beck's concern is by no means restricted to capturing the nonlogical reading of *only*-NQs. Her proposal is an attempt at accounting for what is now called the Heim-Kennedy generalization. In Chapter 1 it was noted that degree operators behave as if they could take syntactic scope above modals, e.g. (352), repeated and simplified from (108) in Chapter 1.

(352) You will need to make few compromises.

a. You need to not make many compromises. ($\Box >$ *few*)

(*More accurately*): in every w the maximal degree d such that you make d-many compromises is lower than $\mathbf{s}_{few}$.

b. You do not need to make many compromises. (*few* $> \Box$)

(*More accurately*): the maximal d such that in every w you make d-many compromises is lower than $\mathbf{s}_{few}$.

However, if the universal modal in (352) is replaced with a quantifier over individuals, as in (353), the ambiguity does not arise.

(353) Everyone made few compromises.

a. Everyone is such that they did not make many compromises. ($\forall >$ *few*)

(*More accurately*): for every x, the maximal degree d such that x made d-many compromises is lower than $\mathbf{s}_{few}$.

b. *Not everyone is such that they made many compromises. (**few* $> \forall$)

(*More accurately*): the maximal degree d such that every x made d-many compromises is lower than $\mathbf{s}_{few}$.

If there were an LF where ANT outscoped the quantifier in (353), the sentence would be predicted to have the meaning paraphrased in (353b). Let us unpack the meaning carefully: the degrees that satisfy *every x made d-many compromises* are the degrees that the stingiest compromiser made in context. If the universal quantifier ranges over three individuals, who made 10, 4, and 3 compromises respectively, then they all made 1-*many*, 2-*many*, and

3-*many*. They did *not* all make 4-*many*, 5-*many*, etc. because the third compromiser—the stingy one—stopped at 3. Therefore, [λd.everyone made d-many compromises] is the set of degrees d that the stingiest compromiser made, irrespective of the higher degrees contributed by the more generous individuals. (353) is true, on this LF, iff *someone* made few compromises, which does not match intuition.

The unavailability of (353b) shows that the LF is blocked by the grammar. To my knowledge, little has been discovered about why this difference should exist between modals and quantifiers, but Beck observes that the modal/quantifier distinction figures also in intervention effects, e.g. in the licensing of Negative Polarity Items (NPIs).

Linebarger (1987) notes that, contra the Ladusaw-Fauconnier generalization (Ladusaw 1979, Fauconnier 1979), not all DE environments license NPIs. While the NPI is licensed in (354a), where a universal modal can appear between the NPI and the licensing negation, the NPI is not licensed under an intervening quantifier (354b).

(354) a. You don't need to talk to anyone.
b. *Not everyone talked to anyone.

The reasoning in Beck is as follows: since ANT cannot cross a quantifier, but can cross a modal, and since NPIs behave in the same way, the two phenomena may be subject to the same constraints. Beck takes the nonlogical reading of *only*-NQ as evidence that NQs are semantically *positive*. That is, *d-few* is semantically equivalent to *d-many*. The negative meaning of *d-few* arises because its distribution is restricted: it is licensed provided that a higher antonymizer is present in the structure. The feature that allows this licensing relation is the same as the feature that relates an NPI to a higher DE operator. I will now offer a very brief and informal introduction to Krifka's (1995b) account of NPIs, so that we may later assess Beck's attempt at applying the same behavior to NQs. Readers familiar with Krifka's proposal may skip this subsection.

<u>An informal introduction to NPIs</u>. To Krifka (1995b), an NPI like *any cookie* has the same ordinary semantics as the quantifier *some cookie*, but in addition, the NPI projects alternatives, call them *some cookie*$^+$, whose individual meaning is strictly stronger that *some cookie*, but whose joint disjunction is equivalent to it. This is just like the relation that disjuncts have to the disjunction that contains them: they are each strictly stronger, but disjunctively they are equivalent to the disjunction. Let us take the sentences in (355) as our examples:

(355) a. *John ate any cookie.

b. John didn't eat any cookie.

As the meaning of (355a) is put together, the semantics computes the ordinary meaning in the familiar way, but in addition, each alternative meaning that the NPI projects (the different *some*$^+$) is computed in parallel. These alternative computations are together grouped into a set of denotations, which we may call the *alternative* semantics of the sentence. With the ordinary semantics computed in one thread, and the alternative semantics computed in another, the composition finally reaches an exhaustivity operator for evaluation. The operator asserts the ordinary meaning, which in this case is *John ate some cookie*, and negates all of the stronger alternative meanings *John ate some cookie*$^+$. Since each alternative is (by assumption) stronger than the ordinary meaning *John ate some cookie*, the alternatives are negated. But because the alternative meanings are disjunctively equivalent to the ordinary meaning, their negation is equivalent to its negation. This results in a contradiction. This is why (355a) is ungrammatical.[10]

In (355b), where the NPI appears under negation, this changes. The composition proceeds in the same way, and with negation we get the ordinary meaning *not(John ate some cookie)*, and the alternative meanings *not (John ate some cookie*$^+$*)*. Now the exhaustivity operator asserts the ordinary meaning, *not(John ate some cookie)*, and negates all stronger alternatives. But none of the alternatives are stronger: *not some*$^+$ is weaker than *not some*, because *some*$^+$ is stronger than *some*. Therefore no contradiction arises, and the NPI is licensed under negation.

□ End of introduction to NPIs

Beck attempts to extend Krifka's analysis to NQs as follows: *d-few* has *d-many* as its ordinary meaning, but like NPIs it projects stronger alternative meanings, and the alternatives cannot be jointly negated with the ordinary meaning. The positive meaning of *d-few* (as *d-many*) is therefore only licensed in DE environments, e.g. under a c-commanding antonymizer. For this reason, *few* does not show its positive meaning in ordinary sentences like (356); the sentence would be contradictory (in the same way that (355a) is contradictory) if no antonymizer is there to license the NQ. The antonymizer reverses monotonicty in the same way as negation in (355b), and therefore prevents the contradiction. By reversing the monotonicity, ANT also gives *few* its negative meaning.

(356) John ate few cookies = ANT[John ate many cookies]

[10] A parallel case to this was used in Chapter 1 to motivate the IE modification to *only*: for any disjunction $p \vee q$ the disjuncts p and q are individually stronger, but their negation contradicts the disjunction, which is why IE was needed for *only*.

If, then, the licensing of NPIs by a higher negation is blocked by an intervening quantifier, we predict the same for NQs. This, on Beck analysis, is why NQs cannot split scope across quantifiers, but can do so across modals.[11] The account is impressive in how it ties the different phenomena together, but unfortunately it falls short of explaining the nonlogical reading under *only*: we have repeatedly seen cases where *few* and other NQs appear as *only*'s focus associate. So in order to maintain Beck's analogy to NPIs, we must ensure that NPIs are licensed as associates of *only*. (357) shows that they are not.

(357) *John only saw any student.

7.5 The status of *only*'s prejacent

In the opening chapter I offered a brief discussion of the presuppositional component of *only*. In presenting my analysis of the particle, I adopted Horn's (1969) proposal that *only* presupposes its prejacent, but I also pointed to alternative accounts (e.g. Atlas 1993, Horn 1996, Ippolito 2008).

For the sake of discussion, let us continue to assume Horn's idea. In this chapter I showed parses of NQs (∃-LFs) that do not conflict with the properties of *only*. An important feature of these parses is that they contain an existential quantifier above the NQ. When an ∃-prejacent is taken as *only*'s argument, we predict (following Horn) that the resulting sentence *presuppose* existence. But in Chapter 2 it was argued that, under *only*, the predicted existence inference of NQs does not arise, at least not in *only*'s *assertive* component. Now, with ∃-LFs, we expect to see existence as a presupposition. Is this problematic?

A theoretical answer to this question requires a thorough investigation of *only*'s presuppositional component, an issue that I do not deal with in this study. However, the answer need not depend on theoretical assumptions. What is needed is a comparison between the judgements from Chapter 2, and judgements with analogous sentences where existence is explicitly mentioned in the prejacent. Let's compare, for example, *few* with [*a few*]: while the former does not entail existence, the latter does:

(358) a. I had few visitors these past weeks. In fact I had none.

b. I had a few visitors these past weeks. #In fact I had none.

[11] This is simply an empirical generalization: whatever property quantifiers have that interferes in NPI licensing, the same property blocks LFs where ANT is interpreted above quantifiers. See Beck (2006), and the extensive literature on NPI licensing.

The oddness of (358b) indicates that the existence inference forms part of the *semantic* meaning of [*a few*]. So now we may run the following experiment: if we detect *uncancellable* existence when [*a few*] is taken as *only*'s prejacent, then we have no reason to expect our ∃-LFs to behave differently, because there too, existence forms part of the prejacent's semantic value. If, on the other hand, existence turns out to be just as cancellable for [*a few*] as it is for [*few*], then we have a consistent (though incomplete) picture. It turns out that, under *only*, the existence inference of [*a few*] is cancellable. In (359,360) I compare the item to *few*, based on examples from Chapter 2.

(359) a. I bet you that only [a few]$_{\mathrm{F}}$ students will show up.

b. I bet you that only [(very) few]$_{\mathrm{F}}$ students will show up.

(If no students show up, is the bet won? lost? suspended? <u>Same answer in both cases</u>)

(360) a. Everyone who submits few assignments will be in bad shape. #?Everyone who only submits [a few]$_{\mathrm{F}}$ assignments will be in slightly better shape.

b. Everyone who submits few assignments will be in bad shape. #?Everyone who only submits few$_{\mathrm{F}}$ assignments will be in slightly better shape.

(359,360) show that the existence inference is equally weak for [*a few*] and *few* under *only*. The inference to *only*'s prejacent must therefore be weaker than a presupposition proper.

So now that we've made this conclusion, that the existence inference for [*a few*] is rather weak under *only*, why was the problem of NQs claimed to be a problem in the first place? The answer is that, with NQs, existence was predicted to follow from the *assertoric* component of *only*, not from its presuppositional component. The move from ordinary LFs (NQ>$\langle\exists\rangle$), to ∃-LFs, ($\langle\exists\rangle$>NQ), effectively moves the existence inference from *only*'s assertive component to its presuppositional component. From examples (359,360) we learn that the move is safe, because even when [*a few*] is embedded under *only*, its existence inference is cancellable.

Chapter 8

An important shortcoming and concluding remarks

The subject of this dissertation is how *only* interacts with *few*, [*less than*], and [*at most*]. I claimed that, on a simple analysis of *only* where stronger alternatives to the prejacent are negated, the predicted interaction does not match intuitions. I observed that *only* does not associate acceptably with the positive counterparts to each of these constructions, and capitalized on this unacceptability in explaining why the initially predicted meanings for NQs are unavailable.

The attested meanings of *only*-NQ association were derived from what I called ∃-LFs, parses of NQ-constructions where the NQ is outscoped by an existential quantifier. But since the meaning of ∃-LFs is not attested in the absence of *only* (van Benthem's problem), using the LFs comes with the task of constraining their distribution. I speculated that ∃-LFs (when without focus) are blocked because of their weak truth conditions, particularly when they compete with the more informative parses where the NQ outscopes $\langle\exists\rangle$.

8.1 An important shortcoming: the SI of *few*?

As a final note, I want to turn the discussion back to *few* and its existence implicature. In Chapter 2 I argued that while existence seems to be an implicature of *few*, it is not inferred from sentences where *few* is associated with *only*. From the implicature, I moved first to the conclusion that *no* is an alternative to *few*, and in Chapter 3 I repeated the conclusion in terms of degrees: POS-*few* gives rise to the existence implicature if a stronger alternative d-*few* is negated, where d is lower than the standard of fewness. The same negation is predicted to be part of the assertoric contribution of *only*, though as I later argued, the construction is blocked because *only* is incompatible with POS.

I want to point out an important weakness in my overall proposal. The attested reading of *only*-(POS-*few*) requires the parse *only*-($\exists$-POS-*few*). The IE alternatives to this prejacent have the form [$\exists$-*d*-*many*], where *d* results from replacing POS, and *many* from removing the ANT node from the prejacent. But if *d*-*many* is always an alternative to *d*-*few*, then there cannot be *any* IE-alternative of the form *d*-*few*, because its negation is equivalent to the alternative *d*-*many*. Therefore, any CSE that contains *d*-*many* must not contain *d*-*few*, and any CSE that contains *d*-*few* must not contain *d*-*many*; the two alternatives cannot be simultaneously negated.

What follows from this? If *d*-*many* and *d*-*few* are both alternatives to POS/*very*-*few*, then [POS/*very*-*few*] will not have any IE-alternatives at all. Therefore, [POS/*very*-*few*] is predicted to be unassociable with *only*, simply because *only* would have no IE-alternatives to operate on. Perhaps, then, there is no need to resort to mirativity in order to explain the *only*-*few* incompatibility.

But this story runs into two related problems. First, if this is really what lies behind the *only*-*few* incompatibility, then in the scope of a universal operator we predict association to become acceptable. This is just like the case of *at most*: *d*-*few* and *d*-*many* cannot be negated together, just like [*at most d*] and [*exactly d*] cannot be negated together. Under a predicate like *need*, however, we find that the negations are simultaneously consistent: $\neg\Box(d\text{-}many)$ is consistent with $\neg\Box(d\text{-}few)$: the resulting proposition says that less than *d* (*d*-few) is not needed, and *d* or more (*d*-many) is not needed either. We therefore expect to see different judgements under *need*, but as we saw in Chapter 4, *need* does not make the logical-reading of *few* available.

The other problem has to do with SIs. If the symmetry of *many* and *few* blocks both kinds of alternatives from being IE—given a prejacent of the form [POS-*few* ...]—then *few* cannot have any scalar implicatures either.[1] But judgements show quite clearly that *few*, in the absence of *only*, does give rise to an existence inference, and the inference behaves just like a scalar implicature.

So, on the one hand, *many* is needed as an alternative to *few*, in order to generate the desired "non-logical" meaning from $\exists$-LFs with POS-*few*. On the other hand, admitting *many* as an alternative to *few* takes away the existence inference of *few* when it appears above $\langle\exists\rangle$, because the construction has no IE-alternatives that can be negated.

I do not have a solution to this problem. We might bite the bullet and say that *no* is an

[1] This assumes that, just like *only*, SIs are computed by negating IE alternatives, an assumption that adopted throughout this study.

alternative to *few*, but given the very different shapes taken by the two expressions at LF, it is far from obvious how one of them can be interpretably replaced with the other. The point remains, however, that if the existence SI is derived by negating an alternative to *few*, and if SI computation and focus computation are sensitive to the same set of alternatives, then the existence inference should be part of the semantic meaning of *only*-*few*, and it isn't.

Bibliography

ABELS, K. AND L. MARTÍ. 2010. A unified approach to split scope. *Natural Language Semantics* 18:435–470.

AIKHENVALD, A. Y. 2012. The essence of mirativity. *Linguistic Typology* 16:435–485.

ALXATIB, S. In progress. FC inferences under *only*. Work in preparation for NELS 44.

ATLAS, J. D. 1993. The importance of being *only*: Testing the neo-Gricean versus neo-entailment paradigms. *Journal of Semantics* 10:301–318.

BALE, A. AND J. COON. 2012. Classifiers are for numerals, not for nouns: Evidence from Mi'gmaq and Chol. To appear in the proceedings of NELS 43.

BARWISE, J. AND R. COOPER. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.

BEAVER, D. AND B. CLARK. 2008. *Sense and Sensitivity*. Wiley Blackwell.

BEAVER, D. AND H. ZEEVAT. 2007. Accommodation. In *The Oxford Handbook of Linguistic Interfaces*, ed. G. Ramchand and C. Reiss. New York: Oxford University Press.

BECH, G. 1955. *Studien über das deustche Verbum infinitum*. Copenhagen: Munksgaard.

BECK, S. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.

BECK, S. 2012. DegP scope revisited. *Natural Language Semantics* 20:227–272.

BENNETT, J. 2003. *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press.

VAN BENTHEM, J. 1986. *Essays in Logical Semantics*. Dordrecht: Reidel.

BIERWISCH, M. 1989. The semantics of gradation. In *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation*, ed. M. Bierwisch and E. Lang. Springer-Verlag.

BONOMI, A. AND P. CASALEGNO. 1993. *Only*: association with focus in event semantics. *Natural Language Semantics* 2:1–45.

BREAKSTONE, M. Y. 2012. Inherent evaluativity. In *SuB 16*, ed. A. A. Guevara, A. Chernilovskaya, and R. Nouwen, Vol. 1. Cambridge, MA: MITWPL.

BRESNAN, J. 1973. Syntax of the comparative clause in english. *Linguistic Inquiry* 4:275–

343.

BÜRING, D. 2007. Cross-polar nomalies. In *SALT XVII*, ed. T. Friedman and M. Gibson. Ithaca, NY: Cornell.

BÜRING, D. 2008. The least *at least* can do. In *WCCFL 26*, ed. C. B. Chang and H. J. Haynie. Somerville, MA: Cascadilla Press.

CHOMSKY, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

CRESSWELL, M. 1976. The semantics of degree. In *Montague Grammar*, ed. B. Partee. Academic Press.

DAVIS, W. 2013. Implicature. In *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. The Metaphysics Research Lab, Stanford University, spring 2013 edition.

DE SWART, H. 2000. Scope ambiguities with negative quantifiers. In *References and Anaphoric Relations*, ed. K. von Heusinger and U. Egli. Dordrecht: Kluwer Academic Publishers.

DE SWART, H. 2001. Weak readings of indefinites: Type shifting and closure. *The Linguistic Review* 18:69–96.

DELANCEY, S. 1997. Mirativity: The grammatical marking of unexpected information. *Linguistic Typology* 1:33–52.

DIESING, M. 1992. *Indefinites*. Cambridge, MA: MIT Press.

DOZON, A. 1879. *Manuel de la langue chkipe ou albanaise*. Paris: Leroux.

DUCROT, O. 1973. *La preuve et le dire*. Paris: Mame.

FARA, D. G. 2000. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics* 28:45–81. Originally published under the name "Delia Graff".

FAUCONNIER, G. 1979. Implication reversal in natural language. In *Formal Semantics and Pragmatics for Natural Language*, ed. F. Guenther and S. J. Schmidt. Dordrecht: Reidel.

FERNANDO, T. AND H. KAMP. 1996. Expecting many. In *SALT VI*, ed. T. Galloway and J. Spence. Ithaca, NY: Cornell University.

VON FINTEL, K. 1997. Bare plurals, bare conditionals, and *only*. *Journal of Semantics* 14:1–56.

VON FINTEL, K. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16:97–148.

FOX, D. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

FOX, D. 2003. Implicature calculation, *only*, and lumping: another look at the puzzle of disjunction. Handout of talk presented at Yale University.

FOX, D. 2007a. Free choice and the theory of scalar implicatures. In *Presupposition and*

Implicature in Compositional Semantics, ed. U. Sauerland and P. Stateva. Houndmills: Palgrave Macmillan.

FOX, D. 2007b. Too many alternatives: density, symmetry, and other predicaments. In *SALT XVII*, ed. T. Friedman and M. Gibson. Ithaca, NY: Cornell.

FOX, D. AND M. HACKL. 2006. The universal density of measurement. *Linguistics and Philosophy* 29:537–586.

FOX, D. AND R. KATZIR. 2011. On the characterization of alternatives. *Natural Language Semantics* 19:87–107.

FRIEDMAN, V. 1986. Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian. In *Evidentiality: The Linguistic Coding of Epistemology*, Vol. 20 of *Advances in Discourse Processes*. Norwood, NJ: Ablex.

FRIEDMAN, V. 2003. Evidentiality in the Balkans with special attention to Macedonian and Albanian. In *Studies in evidentiality*, ed. A. Aikhenvald and R. M. W. Dixon. Benjamins.

GAJEWSKI, J. 2002. On analyticity in natural language. Unpublished manuscript.

GAJEWSKI, J. in press. Innocent exclusion is not contradiction free. Squib submitted to *Linguistic Inquiry*.

GEURTS, B. AND R. NOUWEN. 2007. At least et al.: the semantics of scalar modifiers. *Language* 83:533–559.

GRICE, H. P. 1975. Logic and conversation. In *Syntax and semantics*, ed. P. Cole and J. Morgan, Vol. 3. San Diego, CA: Academic Press.

GROENENDIJK, J. AND M. STOKHOF. 1984. Studies on the Semantics of Questions and the Pragmatics of Answers. Doctoral Dissertation, University of Amsterdam.

HACKL, M. 2000. Comparative quantifiers. Doctoral Dissertation, MIT.

HEIM, I. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertation, UMass, Amherst.

HEIM, I. 2006. *Little*. In *SALT XVI*, ed. C. Tancredi, M. Kanazawa, I. Imani, and K. Kusumoto. Ithaca, NY: Cornell University.

HEIM, I. 2008. Decomposing antonyms? In *SuB 12*, ed. A. Grønn. Oslo: Department of Literature, Area Studies and European Languages, University of Oslo.

HERBURGER, E. 2000. *What Counts: Focus and Quantification*. Cambridge, MA: MIT Press.

HOEKSEMA, J. 1983. Plurality and conjunction. In *Studies in Model-theoretic Semantics*, ed. A. G. B. ter Meulen. Dordrecht: Foris.

HOEPELMAN, J. AND C. ROHRER. 1981. Remarks on *noch* and *schon* in german. In *Tense and Aspect*, ed. P. J. Tedeschi and A. Zaenen, Vol. 14 of *Syntax and Semantics*. New York: Academic Press.

HORN, L. R. 1969. A presuppositional analysis of *only* and *even*. In *CLS 5*, ed. R. I. Binnick, A. Davidson, G. M. Green, and J. L. Morgan. University of Chicago Department of Linguistics.

HORN, L. R. 1972. On the semantic properties of logical operators in English. Indiana University Linguistics Club.

HORN, L. R. 1989. *A Natural History of Negation*. Chicago: Chicago University Press.

HORN, L. R. 1996. Exclusive company: *only* and the dynamics of vertical inference. *Journal of Semantics* 13:1–40.

HURFORD, J. R. 1974. Exclusive or inclusive disjunction. *Foundations of Language* 11:409–411.

IPPOLITO, M. 2006. Remarks on *only*. In *SALT XVI*, ed. C. Tancredi, M. Kanazawa, I. Imani, and K. Kusumoto. Ithaca, NY: Cornell University.

IPPOLITO, M. 2008. On the meaning of *only*. *Journal of Semantics* 25:45–91.

JACOBS, J. 1983. *Fokus und Skalen: Zur Syntax und Semantik von Gradpartikeln im Deutschen*. Tübingen: Niemeyer.

JACOBSEN, W. 1964. A grammar of the Washo language. Doctoral Dissertation, University of California, Berkeley.

KAMP, H. 1981. A theory of truth and semantic representation. In *Formal methods in the study of language*, ed. J. Groenendijk, T. Janssen, and M. Stokhof. Mathematical Centre Tracts 135, Amsterdam.

KAMP, J. A. W. 1975. Two theories about adjectives. In *Formal Semantics of Natural Language*, ed. E. L. Keenan. Cambridge: Cambridge University Press.

KATZIR, R. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30:669–690.

KEENAN, E. L. AND J. STAVI. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9:253–326.

KENNEDY, C. AND L. MCNALLY. 2005a. Scale structure, degree modification, and the semantics of gradable adjectives. *Language* 81:345–381.

KENNEDY, C. AND L. MCNALLY. 2005b. The syntax and semantics of multiple degree modification in English. In *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*, ed. S. Müller. Stanford: CSLI Publications.

KLEIN, E. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4:1–45.

KLINEDINST, N. 2005. Scales and *Only*. Master's thesis, UCLA.

KRATZER, A. 1991. The representation of focus. In *Semantik*, ed. A. von Stechow and D. Wunderlich. Berlin: Walter de Gruyter.

KRIFKA, M. 1995a. Common nouns: a contrastive analysis of English and Chinese. In *The Generic Book*, ed. G. Carlson and F. J. Pelletier. Chicago: Chicago University Press.

KRIFKA, M. 1995b. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25:209–257.

KRIFKA, M. 1999. At least some determiners aren't determiners. In *The Semantics/Pragmatics Interface from Different Points of View*, ed. K. Turner, Vol. 1 of *Current Research in the Semantics/Pragmatics Interface*. Kidlington, Oxford: Elsevier Science Ltd.

KRIFKA, M. AND A. COHEN. 2011. Superlative quantifiers as modifiers of meta-speech acts. In *Formal Semantics and Pragmatics. Discourse, Context and Models*, ed. B. H. Partee, M. Glanzberg, and J. Šķilters, Vol. 6 of *The Baltic International Yearbook of Cognition, Logic and Communication*. Manhattan, KS: New Prairie Press.

KROCH, A. 1972. Lexical and inferred meanings for some time adverbs. In *Quarterly progress report of the Research Laboratory of Electronics*, 104. MIT.

LADUSAW, W. 1979. Polarity sensitivity as inherent scope relations. Doctoral Dissertation, University of Texas at Austin.

LANDMAN, F. 2004. *Indefinites and the Type of Sets*. London: Blackwell.

LECHNER, W. 2004. *Ellipsis in Comparatives*. Berlin: Mouton de Gruyter.

LEWIS, D. K. 1975. Adverbs of quantification. In *Formal Semantics of Natural Language*, ed. E. Keenan. Cambridge: Cambridge University Press.

LINEBARGER, M. 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10:325–387.

LÖBNER, S. 1989. German *Schon - Erst - Noch*: an integrated analysis. *Linguistics and Philosophy* 12:167–212.

MATSUMOTO, Y. 1995. The conversational condition on Horn Scales. *Linguistics and Philosophy* 18:21–60.

MAYR, C. 2011. Implicatures of modified numerals. Unpublished manuscript.

MCCAWLEY, J. 1981. *Everything that Linguists Have Always Wanted to Know about Logic but Were Ashamed to Ask*. Chicago: University of Chicago Press.

NOUWEN, R. 2010. Two kinds of modified numerals. *Semantics and Pragmatics* 3:1–41.

NOUWEN, R. 2012. Modified numerals: the epistemic effect. Unpublished manuscript.

PARTEE, B. 1989. Many quantifiers. In *Proceedings of the 5th Eastern States Conference on Linguistics*, ed. J. Powers and K. de Jong. Columbus: Ohio State University. Reprinted in Partee (2004).

PARTEE, B. 2004. *Compositionality in Formal Semantics: Selected Papers by Barbara H. Partee*. Oxford: Blackwell.

PENKA, D. 2011. *Negative Indefinites*. New York: OUP.

REINHART, T. 1997. Quantifier scope: how labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335–397.

RETT, J. 2008. Degree modification in natural language. Doctoral Dissertation, Rutgers.

RIESTER, A. 2006. Only scalar. In *Proceedings of the Eleventh ESSLLI student session*, ed. J. Huitink and S. Katrenko.

VAN ROOIJ, R. AND K. SCHULZ. 2007. *Only*: meaning and implicatures. In *Questions in Dynamic Semantics*, ed. M. Aloni, A. Butler, and P. Dekker. Oxford: Elsevier.

ROOTH, M. 1985. Association with focus. Doctoral Dissertation, UMass Amherst.

ROOTH, M. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.

ROOTH, M. AND B. PARTEE. 1982. Conjunction, type ambiguity, and wide scope *or*. In *WCCFL 1*, ed. D. Flickenger, M. Macken, and N. Wiegand. Linguistics Department, Stanford University.

RULLMAN, H. 1995. Maximality in the semantics of wh-constructions. Doctoral Dissertation, UMass Amherst.

SAUERLAND, U. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27:367–391.

SCHLESINGER, S. C. AND S. KINZER. 1999. *Bitter Fruit: The Story of the American Coup in Guatemala*. Cambridge, MA: The David Rockefeller Center for Latin American Studies, Harvard University, expanded edition.

SCHWARZ, B. 2011. Remarks on class-B numeral modifiers. Handout of talk delivered at Göttingen University.

SCHWARZ, B., B. BUCCOLA AND M. HAMILTON. 2012. Two types of class B numeral modifiers: A reply to Nouwen 2010. *Semantics and Pragmatics* 5:1–25.

SOLT, S. 2009. The Semantics of Adjectives of Quantity. Doctoral Dissertation, CUNY.

SOLT, S. 2013. Q-adjectives and the semantics of quantity. Unpublished manuscript.

SPECTOR, B. 2011. *At least* at last? Handout of talk delivered at the ENS, Paris.

VON STECHOW, A. 2001. Temporally opaque arguments in verbs of creation. In *Semantic Interfaces: Reference, Anaphora, Aspect*, ed. C. Cechetto, G. Chierchia, and M. T. Guasti. Stanford: CLSI publications.

VON STECHOW, A. 2006. Times as degrees: *früh(er)* ‘early(er)’, *spät(er)* ‘late(r)’, and phase adverbs. Revised and published as von Stechow (2009).

VON STECHOW, A. 2009. The temporal degree adjectives *früh(er)/spät(er)* ‘early(er)’/‘late(r)’ and the semantics of the positive. In *Quantification, definiteness, and nominalization*, ed. A. Giannakidou and M. Rathert. Oxford: Oxford University Press.

STEUBE, A. 1980. *Temporale Bedeutung im Deutschen, Studia Grammatica XX*. Berlin: Akademie-Verlag.

SWANSON, E. 2010. Structurally defined alternatives and lexicalizations of XOR. *Linguistics and Philosophy* 33:31–36.

WILHELM, A. 2008. Bare nouns and number in Dëne Sųłiné. *Natural Language Semantics* 16:39–68.

WINTER, Y. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20:399–467.

ZEEVAT, H. 2008. "Only" as a mirative particle. In *Focus at the Syntax-Semantics Interface*, ed. A. Riester and E. Onea. Working Papers of the SFB 732, Vol. 3, University of Stuttgart.