The Logic of Subtractives

or,

Barely anyone tried almost as hard as me

by

Christopher Baron

B.A., Linguistics and B.A., Philosophy
University of Massachusetts, Amherst (2015)

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2022

Author . . . .
Department of Linguistics & Philosophy
September 1, 2022

Certified by . . . .
Roger Schwarzschild
Professor Emeritus of Linguistics
Thesis Supervisor

Accepted by . . . .
Daniel Fox
Anshen-Chomsky Professor of Language & Thought
Department Head

The Logic of Subtractives

or,

Barely anyone tried almost as hard as me

by

Christopher Baron

Submitted to the Department of Linguistics & Philosophy
on September 1, 2022, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This dissertation is about the meaning and distribution of the modifiers *almost* and *barely*. We advocate for an analysis in which they are, across their uses, modifiers of quantifiers, encoding set subtraction; *barely*, but not *almost*, additionally contributes negation. Both modifiers remove elements from the arguments of quantifiers that they modify, and require exhaustification for their licensing as modifiers. We start with subtractive modified quantificational determiners, like *almost every* and *barely any*. We then push this analysis further, showing it very naturally extends to degree constructions like comparatives and equatives, and captures the facts better other options. This extension also provides an argument for the idea that all natural language scales are dense. Numeral constructions like *almost one hundred* and *barely one hundred* appear to complicate the idea that *almost* and *barely* are in complementary distribution, but we argue this shows us there's more than meets the eye in such constructions. We offer a theory of numeral constructions that captures the overlapping distribution. Finally, we suggest that subtractives are evidence in favor of a view of exhaustification in which the contribution of the latter is presuppositional, rather than truth conditional.

Thesis Supervisor: Roger Schwarzschild
Title: Professor Emeritus of Linguistics

Acknowledgments

I started this task months ago, and yet here I am, at the very last minute, unsure of what to say. It's very difficult to find adequate words, and I'm convinced to a reasonably high degree that it's an impossible endeavor. I don't want to write too little, and risk leaving too much unsaid, and I don't want to write too much, either, and end up with a laundry list that's no fun to read, though I usually end up saying altogether too much regardless. I'll try my best.

Let's start with my committee: Kai von Fintel, Martin Hackl, and Roger Schwarzschild.

When I asked Kai von Fintel why I should come to MIT, he said, without a beat skipped, "Because we're the coolest." He followed it up with more arguments, of course, but his enthusiasm, his cool, did a good deal of work to convince me. Kai shepherded me through my first year as a semantics teacher and co-investigator into Aktionsart and tenselessness, topics which I'm sorry to have left behind. In this dissertation, the influence of both his work and his criticism should obvious throughout; he is incredibly sharp and gets right to the point. I'm very glad and grateful he joined me on this journey.

Martin Hackl probably remembers, but maybe doesn't, the offhanded comment he made about the unsatisfactory analyses of *almost* that exist during a meeting about yet another topic I left behind. That comment is why I took this topic up. Meetings with Martin pushed me towards clarity and precision in argumentation—I'm prone to hand waving and rarely would he let me slide on that. His incisive questions about the big picture made this a better, more considered piece of work, too, though I'm not sure I ever fully addressed some of his harder questions. I hope I have!

What's there to say about Roger Schwarzschild? So much, honestly, that I'm not sure I can write it all. First, this document would be an absolute mess, if it even existed at all, without Roger. He reads everything, and I mean everything, so quickly, too. He was reading this thesis and giving me comments today, like, filing day. His commentary was, a non-zero amount of the time, longer than the prose I sent, and so consistently made arguments and logic better. Arguing with Roger, learning from Roger in long, long, long meetings was stressful and exhausting and so much fun. His way of thinking has profoundly shaped my way of thinking, even when we don't think at all alike about something. Even when I walked into or out of a meeting miserable and drained, for linguistic reasons or other reasons, I had a good time with Roger. This was an incredibly fruitful partnership for me. I'm not sure what else to say, so I'll leave it with this: How dare you?

The Linguistics & Philosophy department at MIT, as well as the Linguistics Departments at UMD and UMass, introduced me to so many other fantastic people. I learned about linguistics, and much more, from such sharp scientists, it's scary. I'm at the point where it's all a blur—who I had for what and when—so I'm sorry if you've been left off this list. At MIT, Sabine Iatridou, Donca Steriade, Adam Albright, Norvin Richards, Danny Fox, Irene Heim, Suzanne Flynn, and more, and the professionals who run this department from Headquarters past and present deserve all the recognition, too—Aggie Fernandes, Sheelah Ward, Matt Sikorski, Chris Naylor, Jen Purdy, Chrissy Graham, and especially, especially Mary Grenham; at UMD, my Baggett mentors Valentine Hacquard and Alexander Williams and Rachel Dudley; at UMass, my undergrad thesis advisor Seth Cable, and so many more, too. I learned a lot from you all.

MIT has been my home for six years, and it was made so much better by the people here. Yadav Gowda is one of my very best friends; we drove to Illinois and back twice and still like each other. Thanks for those trips (shout out to them), for sharing tweets and sometimes letting

my walls of texts just sit there as they should, for your humor and insight, for your general handiness and know-how, for yourself. Justin Colley got me so much more excited about cooking for and with others, since we began during the early days of our pandemic bubble. His culinary adventurousness, his wit and ribbing, and his love of the recipes of Yotam Ottolenghi made for many lovely evenings, and I hope very dearly we do it again soon. Our boy's got many more recipes we've yet to try. Anton Kukhto took a lot of grief from the rest of us, but he's a real trooper. He's so very thoughtful and a really great gift giver (I regret that I'm not nearly as good) and filled with jokes so abominable, so poorly constructed that I get deep chuckles still thinking about them. Thanks to you three for making me a movie-watching person after a long time of not watching them, for really no good reason; it's grand there's a whole world of them. Kelly Gaus is one of my parallel cohortmates from Philosophy, and we not only share pics of our pets and gripes about everything, we share a birthday, and that's always been a delight. She is so empathetic and intelligent and fun to be around. May the People's Republik live on in our memories! Mitya Privoznov was a wonderful roommate for four happy years; he's a fabulous cook, a caring friend to my cat, and truly one of my very favorite people in this entire world. There are many things big and small that I will cherish about our time living together.

I've had the same office since second year, and its residents past and present have made the time spent in a space with wild, unseasonable temperature fluctuations a real joy. As chronologically as I can get it: Ishani Guha, Snejana Iovtcheva, Newell Lewey, Erin Olson, Agnes Bi, Boer Fu, Maša Močnik, Filipe Hisao Kobayashi, Adèle Mortier, Ido Benbaji, Yash Sinha. It's been a delight. Ling-16—first year officemates, classmates, and friends throughout—also deserve special recognition. Neil Banerjee, Keny Chatain, Yadav Gowda, Cora Lesure, Newell Lewey, Elise Newman, Frank Staniszewski, and Danfeng Wu (also a defense exact-date-and-time buddy). What a journey we undertook.

There are so many other people, past and present, that deserve special words, so I'm sorry, it too will be a list, and likely a woefully incomplete one. At MIT: Michelle Yuan, Athulya Aravind, Carrie Spadine, Verena Hehl, Agnes Bi, Dóra Kata Takács, Filipe Hisao Kobayashi, Sherry Yong Chen, Maša Močnik, Tanya Bondarenko, Patrick Elliott, Chris Yang, Vincent Rouillard, Enrico Flor, John Dennis, Devon Denny, Keely New, Giovanni Roversi, Annauk Olin, Kat Hintikka, Mallory Webber, Tracy Kelley, Kenyon Branan, Sophie Moracchini, Ömer Demirok, Juliet Stanton, Hadas Kotek, and more; at UMD, Paulina Lyskawa, Rodrigo Ranero, Ted Levin, Rachel Dudley, Kasia Hitczenko, and more; more still from UMass and beyond.

There's always outside-of-linguistics folks that get us through linguistics. Here's some of mine. Stephanie Nuñez is one of the dearest people in the world to me; thanks for Crybabies, dancing, burn-filled beach days, and for your openness and depth and for flopping around with me for years. Susannah Gleason is also one of those very dearest; thanks to you and all the Gleasons and Sullivans, especially Mike and Mary Ann, for letting me be one of you for all these Thanksgivings (I look forward to more), thanks to you and Mark and Sarah for lending me your couches in Portland over the years, thanks to you especially for your constant friendship and profoundly caring spirit. Anthony Fischer and Barry Grace were my first new friends in Boston, and a lovely byproduct of their enduring and wonderful friendship is the great privilege and immense pleasure to know so many other folks. Erin McCarthy, Wes Mason (honk honk), KC Kerby Patel, Feanil Patel, Ivy Lee, Blake Sessions, Nicole Bolliger, Sarina Canelake (Percy, Rupert, and the Gremlin, too), David Laibstain, and especially the true friend that is Connor Hogan (MPH, RD, CPT, LDN; you gotta get more certifications so I can add them to my phone). Thanks for

Dinner parties, Cookbook Fantastic Club, Feb Club, Mahoney's trips, and everything else. It all brought me so much joy. I'll see some of you soon, and others not soon enough.

Thanks especially to Dr. Jill Colman, for everything; thanks to the Halflings, it's been real fun getting to know you; to friends back home and beyond; to my brothers Patrick, Ryan and Trevor. Thanks especially especially to my parents, Patti and Sam, who have been constant supporters; their love and guidance got me here.

Final thanks to Fern, my little kitten (sixteen years young) who's been with me since college. She made it, I made it!

Sorry, that wasn't final thanks. Thanks to you who might read this; I'm sorry for the mistakes and omissions and errors. They're all mine. Okay, now read!

Contents

Chapter 1

A very brief introduction

This chapter is really more of a prologue. In it, we'll lay out the basic goals of the dissertation and its outline, and we'll end with some notational conventions we'll make use of throughout dissertation. From there, we'll get started with our work.

1.1 The plot

This dissertation is about *almost* and *barely*. We'll start in Chapter 2 with a bit of historical and theoretical context for our investigation of *almost* and *barely*. Then, we'll take a close look at uses of these operators in which they seem to modify determiners.

(1) a. Almost every vampire in Spike's gang attacked Sunnydale High School.
  b. Barely any vampires in Angel's gang attacked Sunnydale High School.

(1a) conveys that the vast majority of vampires in Spikes gang attacked, but not all of them did. (1b) conveys that a few vampires in Angel's gang attacked the school, but not very many at all. We want to know how these interpretations arise. We also want to know why (2a) and (2b) are odd; what's going wrong with these cases?

(2) a. # Almost any vampires in Spike's gang attacked Sunnydale High School.
  b. # Barely every vampire in Angel's gang attacked Sunnydale High School.

We want to explain why such combinations are illicit. Chapter 2 provides such a theory, starting from and building on work by von Fintel (1993), Gajewski (2013), and Crnič (2018) on exceptive *but*, which is similar to *almost* in meaning and distribution.

(3) a. Every vampire in Spike's gang but Drusilla attacked Sunnydale High School.
  b. # Any vampires in Spike's gang but Drusilla attacked Sunnydale High School.

We take seriously the idea that *almost* and *barely* are, at their core, modifiers of quantificational determiners. We take the operators' core semantic contribution to be set subtraction—they subtract elements from the arguments of the quantifiers they modify. From now on, we'll call these operators *subtractives.*

Chapter 3 and Chapter 4 are, in some ways, the central contribution of the dissertation. Chapter 3 expands the coverage of this theory to degree constructions. In the context of *almost*, degree constructions like comparatives, equatives, and the positive have received little attention; in the context of *barely*, they have received none.

(4) a. Willow is {almost/#barely} as old as Buffy is.

b. Willow is {barely/#almost} older than Buffy is.

(5) a. Willow's room is {almost/#barely} clean.

b. Willow's room is {barely/#almost} dirty.

We show the theory developed and refined in Chapter 2 naturally accounts for these data, and provides support for some very interesting proposals in the semantics of degrees. Chapter 4 probes numeral constructions, where we see that complementarity in distribution between *almost* and *barely* disappears.

(6) {Almost/barely} one hundred vampires attacked Sunnydale High School.

We'll argue on the basis of the theory developed in Chapters 2 and 3 that there is structurally more than meets the eye in (6). We'll argue that subtractives give us evidence for a more complicated compositional structure in numeral constructions than is typically assumed; in particular, subtractives show us that there are multiple quantificational elements at play in the structure and meaning of numerals.

Chapter 5 returns to quantificational determiners, and complicates matters by considering structures in which subtractive-modified quantifiers are embedded under other quantificational elements. What we ultimately are doing is exploring the source and nature of the *polar* inference that subtractives give rise to—the inference in (1a) that not every vampire in Spike's gang attacked the high school, and the inference in (1b) that some vampires in Angel's gang did, in fact, attack the school. As will be discussed sooner than that, in Chapter 2, this inference is rather tricksy. We'll pin it down a bit more in Chapter 5.

The dissertation then concludes in Chapter 6. There will be a lot of territory that we do not cover here. The final chapter briefly discusses some of those important areas, and sketches how the theory put forth in the dissertation might be extended.

1.2 Notational conventions and shorthand

Let's get some notational conventions straight. Throughout the dissertation, I use italics in prose to indicate particular expressions of the object language, e.g. *vampire*. Titles of books, paintings, and movies, etc., when occurring in a larger expression of the object language, will be offset by normal font in prose, e.g. *Willow loves the spellbook* The Writings of Dramius *by Dramius*, but otherwise will be in italics consistently. Proper names of all kinds are taken to be rigid designators, and denote entities. These are given in Latin letters as well, typically abbreviated where it is clear and expeditious to do so.

(7) $[\![\text{Willow}]\!]^{g,c} = \text{W(illow)}$

(8) $[\![\textit{The Writings of Dramius}]\!]^{g,c} = \text{WoD}$

Variables ranging over expressions of the object language will be given in Greek letters, e.g. "For any expression α (of the object language)..." There are a few variables we will reserve for particular kinds of expressions. We will use χ as a variable for what we'll call 'exceptions.' We'll discuss these more momentarily. We will also restrict φ and ψ for exclusively referring to expressions denoting $\langle s,t\rangle$ functions, and Δ for quantifiers, i.e. for any type σ, expressions of type $\langle\langle\sigma,t\rangle,\langle\sigma t,t\rangle\rangle$. This includes expressions like *every*, also known as 'quantificational determiners,' and expressions like *-er*, what we'll call 'degree quantifiers.'

Greek letters will be used in two other ways. First, when discussing semantic types, we will use three Greek letters: σ, η, and τ. Second, when we discuss measure functions in degree constructions, we will use μ as a variable over functions of type $\langle e,d\rangle$, functions often called 'measure functions.' Hopefully, in both cases, it will be very clear that we are not referring to arbitrary object language expressions in these cases; in any case, we will not use these particular Greek letters as variables for object language expressions.

Metalanguage variables are given in Latin letters. For example, for any arbitrary expression α whose denotation is a function of type $\langle\sigma, t\rangle$, $[\![\alpha]\!] = \left[\lambda A_\sigma.\ A\ ...\ \right]$. We will also use these when we wish to talk about arbitrary expressions of the metalanguage, e.g. "For any function A..." Now, there are a few Latin letters that will be used for specific purposes in line with more general considerations. For example, unless otherwise marked, we will use A and B to refer to functions of type $\langle e,t\rangle$, and D, D′, D*, etc., for functions of type $\langle d,t\rangle$. There is one exception to this: we will use $\mathscr{D}$ as variable for the denotations of object language quantifiers Δ.

Functional constants in the metalanguage are given in small caps. I assume that common nouns, expressions like *vampire*, denote relations between worlds and individuals, and worlds are written as subscripts of other expressions in the metalanguage. So:

(9) $[\![\text{vampire}]\!] = \lambda w \lambda x.\ \textsc{vampire}_w(x)$

In many situations, it is useful and more intuitive for us to manipulate sets. Sets are not identical to functions, of course, but there is a one-to-one mapping between sets and the characteristic functions of those sets.

(10) Let $\mathbb{A}$ be a set; then $\textsc{char}_{\mathbb{A}}$, the *characteristic function of* $\mathbb{A}$, is that function A such that, for any $x \in \mathbb{A}$, $f(x) = 1$, and for any $x \notin \mathbb{A}$, $A(x) = 0$

(11) Let A be a function with range $\{0, 1\}$; then $\textsc{char}_{A}$, the *set characterized by* A, is $\{x \mid A(x) = 1\}$

We will often use the first letter of object language expressions to indicate the sets characterized by their denotation. For example, the object language expression *vampire* is of type $\langle s,et \rangle$, and we can use the following shorthand to talk about the set of vampires in a given world w.

(12) $\mathbb{V}_w := \{x \mid [\![\text{vampire}]\!](w)(x) = 1\}$

We will generalize this correspondence notationally across types in the following way.

(13) **RULES FOR SET NOTATION**

For any g, c, type σ, object language expression α, and function A:

a. If $[\![\alpha]\!]^{g,c}$ is of type $\langle s,\sigma t \rangle$, then for any world w, $\mathbb{A}_w$ is the set characterized by $[\![\alpha]\!]^{g,c}(w)$.

b. If $[\![\alpha]\!]^{g,c}$ is of type $\langle \sigma,t \rangle$, then $\mathbb{A}$ is the set characterized by $[\![\alpha]\!]^{g,c}$.

c. If A is of type $\langle \sigma,t \rangle$, then $\mathbb{A}$ is the set characterized by the value of A.

Here's a denotation that makes use of these notational choices. In the dissertation, we will make use of set-theoretic meanings for quantificational determiners like *every*.

(14) $[\![\text{every}]\!]^{g,c} = \lambda A_{\langle e,t \rangle} \lambda B_{\langle e,t \rangle}.\ \{x \mid A(x) = 1\} \subseteq \{y \mid B(x) = 1\}$ $\langle et,ett \rangle$

Given our shorthand, we can rewrite this as follows, since $\mathbb{A}$ is the set characterized by the value A.

(15) $[\![\text{every}]\!]^{g,c} = \lambda A_{\langle e,t \rangle} \lambda B_{\langle e,t \rangle}.\ \mathbb{A} \subseteq \mathbb{B}$

Strictly speaking, quantifiers like *every* denote relations between functions, and *every* denotes a relation between two $\langle e,t \rangle$ functions; this notation makes it seem like they denote relations between sets. We'll be cheating a bit, too, when we're trafficking in arbitrary quantifier denotations. That is, we'll use $\mathscr{D}$ as the metalanguage variable for the denotations of quantifiers, i.e. relations between functions, but we'll also use it in the metalanguage as a relation between sets. For example, *not* in *not every vampire is evil*, taking it to be a modifier of the quantificational determiner, could have the following denotation:

(16) $[\![\text{not}]\!]^{g,c} = \lambda \mathscr{D}_{\langle et,ett \rangle} \lambda A \lambda B.\ \mathscr{D}(\mathbb{A})(\mathbb{B}) = 0$

The two uses of $\mathscr{D}$ aren't the same, but the two ways are related, just as set notation $\mathbb{A}$ is related to A.

When we begin our analysis of *almost* and *barely* in earnest, we will see talk about 'exceptions,' as mentioned above, and we'll utilize χ to talk about them. Sometimes these expressions χ will bear indices. These are taken to be pronouns. So:

(17) For any n, $[\![\chi_n]\!]^{g,c} = g(n)$

If χ does not bear an index, then it is *not* a pronoun, but rather denotes an $\langle e,t \rangle$ function, and as such, we will utilize our set notation, reserving $\mathbb{X}$ for 'exception sets.' So, $[\![\chi]\!]^{g,c} = \mathbb{X}$.

When we turn to degree constructions, additional notational conventions will be utilized. First, we assume that adjectives encode measure functions which map entities in a world to a degree, e.g:

(18) $[\![\text{old}]\!] = \lambda w \lambda d \lambda x.\ \textsc{age}_w(x) \geq d$

When it is clear from the discussion which adjective and measure function are under discussion, we will use the following shorthand: for an arbitrary world *w*, $\textsc{age}_w$(Buffy) will be written $\mathcal{B}_w$, $\textsc{age}_w$(Willow) will be written $\mathcal{W}_w$, and so on as necessary.

Finally, we will use '#' in two ways in the dissertation. First, as a sort of neutral marker of sentential oddity, if we do not wish to commit ourselves to claiming something is syntactically ill-formed and ungrammatical, and second, as a truth value in a trivalent system wherein definedness conditions on some expression are not met, i.e. neither truth nor falsity. Other such conventions, as they arise, will be established.

Chapter 2

Quantificational determiners

2.1 Subtractives, an introduction

2.1.1 Polarity and proximity

Our analytical quarry are the expressions *almost* and *barely* and their close kin. Let's first consider the contribution *almost* makes.

(19) a. Buffy killed almost every vampire in the warehouse.

b. Buffy almost always beats her enemies.

c. Buffy almost killed the vampire last night.

(19a) conveys that Buffy didn't kill every vampire, but she came close to killing them all. Similarly, (19b) conveys that while there are some occasions where Buffy fails to beat her enemies, and hence that she doesn't always beat them, far more often than not she does beat them. (19c) is true where she failed to kill the vampire, but she came close to doing so—say, she tripped right before she plunged the wooden stake in its heart and it escaped as she scrambled up.

These intuitive characterizations illustrate the core contribution of *almost*: from *almost* φ we infer that φ does not hold, and further that φ is somehow 'close' to obtaining. We'll call these two inferences the POLAR and the PROXIMAL, following Sevi (1998) and Horn (2002).

(20) *almost* φ

a. $\neg\varphi$ — POLAR inference

b. $\textsc{close}(\varphi)$ — PROXIMAL inference

Some of the literature on *almost* investigates in parallel the modifier *barely*: it gives rise to inferences we can characterize quite like we characterized those *almost* gives rise to.

(21) a. Xander killed barely any vampires in the warehouse.

b. Xander barely ever beats his enemies.

c. Xander barely survived the vampire attack last night.

(21a) conveys that while Xander did kill at least one vampire, the total number is quite close to zero–maybe just one or two. (21b) conveys that Xander occasionally beats his enemies, but such events are few and far between. (21c) conveys that Xander did indeed survive the attack, but it was a close call. Had there been one misstep, he would have died. We can characterize, then, *barely*'s contribution quite like we do *almost*'s.

(22) *barely* φ

a. φ — POLAR inference

b. CLOSE(¬φ) — PROXIMAL inference

Before we can decide on lexical entries, there are a few important issues to consider. First, we need to figure out how these inferences are derived. Whether or not the polar and the proximal inferences have the same semantic and pragmatic status has been hotly debated since Sadock (1981). More specifically, there are disputes over whether the polar inference is truth conditional and contributed by *almost* and *barely*, or else is something like a presupposition or an implicature. The facts concerning this question are complicated and seemingly contradictory. Second, we need to consider what notion of closeness underwrites the proximal inference. It is sometimes taken to be a modal notion (Sadock 1981; Rapp & von Stechow 1999, Morzycki 2001, Nouwen 2006) and other times taken to be a scalar one (Hitzeman 1992; Rotstein & Winter 2004; Penka 2006). Let's consider each question in turn.

2.1.1.1 The polar inference

The status of the polar inference, as mentioned above, has been debated for a long while and for many reasons. First, it seems to be non-cancellable.

(23) a. # Buffy killed almost every vampire. In fact, she killed every one.

b. # Xander killed barely any vampires. In fact, he killed none.

(24) a. # Buffy almost killed the vampire last night. In fact, she did kill the vampire.

b. # Xander barely survived the vampire attack last night. In fact, he was killed in the scuffle.

These are all quite odd; the continuations in each contradict the polar inference. This is perfectly expected if the inference is a truth conditional contribution of *almost* and *barely*. Assuming that the proximal inference is also truth conditional, we might posit denotations for the operators as in (25a) and (25b).

(25) a. $[\![\text{almost}]\!]^{g,c} = \lambda p_{\langle s,t\rangle}\lambda w.\ \neg p(w)\ \&\ \textsc{close}(p(w))$

b. $[\![\text{barely}]\!]^{g,c} = \lambda p_{\langle s,t\rangle}\lambda w.\ p(w)\ \&\ \textsc{close}(\neg p(w))$

These entries in (25) take *almost* and *barely* to be clausal modifiers, operating semantically at the level of the proposition; this is a widespread view in the semantics of *almost*. These entries also, as is clear, encode both the polar and the proximal inferences as truth conditional. Such an analysis can be called CONJUNCTIVE (Horn's 2011 terminology) or INTEGRATED (Crnič's 2018 terminology)—it proposes the two inferences to have the same status.

A conjunctive analysis faces a few problems, though. *Barely*, but not *almost*, licenses NPIs.

(26) a. Buffy almost {never/#ever} loses a fight.

b. Buffy barely {ever/#never} loses a fight.

(27) a. # Buffy almost {budged/slept a wink/spoke to anyone}.

b. Buffy barely {budged/slept a wink/spoke to anyone}.

These contrasts are entirely unexpected given the conjunctive entries above. They both contain a positive and a negative conjunct, and while a precise characterization of CLOSE is needed to more accurately defend this, they are both non-monotonic on p as it stands. However, if we were to sever the polar inference from the truth conditions, and we gave a more precise definition of CLOSE, we could expect it.

There are other environments in which the polar inference seems to be inaccessible or backgrounded, and only the proximal inference accessible.

(28) a. Good news! My printer is almost functional.

b. ? Good news! My printer is barely functional.

(29) a. ? Bad news! My printer is almost functional.

b. Bad news! My printer is barely functional. Horn (2002:57)

The good news in (28a) must be that the printer is approaching functionality, not that it isn't working. (28b) is rather odd, generally, but in a context where the printer's inevitable decline means we get to go home from work early, it's acceptable. Regardless, the good news cannot be that my printer *is* working. Similar intuitions arise for (29a) and (29b). *Good news* and *bad news* only seem to have access to the proximal inference, not the polar. Now, these data should not be taken to suggest that *almost* and *good news* correlate, and *barely* and *bad news*correlate.

(30) a. Good news! My in-laws can barely stand me (so they plan to come around less).

b. Bad news! My in-laws are almost here (so we need to get out of here lickety split).

The good news in (30a) is that they are close to completely detesting me; the bad news in (30b) is their spatial proximity. Similar facts can be shown for other similar modifiers, e.g. *fortunately*, Nouwen 2006; *luckily*, *amazingly*, *depressingly*, and so on; we'll call all these 'speaker-oriented' adverbs.

In other constructions, too, this inference is invisible to operators—Ziegeler (2000) shows for *almost* that it is only the proximal inference, not the polar inference, that can support causal justification. We can show the same for *barely*.

(31) Buffy almost won the battle...

a. because of her skills and tenacity.

b. # because of her missteps and bad fortune.

(32) Willow barely understood the spellbook...

a. # because of her talent and research.

b. because the text was terribly damaged.

It appears that the proximal inference is all that *because* has access to. (31a) explains why Buffy came close to winning, but (31b) cannot be used as a continuation to explain why she ultimately did not win. Similarly, (32a) cannot be used as a continuation to explain why she did in fact understand a little of the spellbook, but (32b) can justify why it was so difficult. This asymmetry between the two inferences is as unexpected as NPI licensing is on a conjunctive analysis. These kinds of facts push some towards treating the polar inference as a presupposition (Ducrot 1973; Anscombre & Ducrot 1983), an implicature, either conversational (Sadock 1981; Ziegeler 2000) or conventional (Jayez & Tovena 2008), or else somewhere in between (Atlas 1984, 1997).

Zooming out, what we learned here is that any analysis of *almost* and *barely* has to account for a few facts: (i) that *barely*, but not *almost*, licenses NPIs; (ii) that the polar inference, while non-cancellable, is inaccessible to many operators. Throughout the dissertation, we'll pursue a particular division of labor between operators in *almost* and *barely* sentences that allows us to account for these facts. The licensing of NPIs and the non-cancellability of the polar inference will be discussed in this chapter, but the inaccessibility, the "invisibility" of the polar inference to certain operators, will be discussed in Chapter 5. Now, we'll turn to the proximal inference.

2.1.1.2 The proximal inference

The truth conditional status of the proximal inference has never been in doubt, as far as I can tell. That said, the precise characterization of what notion underwrites a proposition being CLOSE is up for discussion. There are roughly two main camps in the literature, largely focusing on *almost*. Abstractly, we can characterize these two as 'modal' closeness and 'scalar' closeness. With various instantiations debating details, modal closeness analyses take the essential contribution of *almost* to be intensional similarity: *almost* modifies a proposition and says that proposition is true at some world closely related to the world of evaluation. This kind of a view originates with Sadock (1981); Rapp & von Stechow (1999), Morzycki (2001), Nouwen (2006), among others, adopt such an analysis as well.

(33) $[\![\text{almost}_{\text{modal}}]\!]^{g,c} = \lambda p_{\langle s,t\rangle}\lambda w.\ \neg p(w)\ \&\ \exists w'\big[\textsc{close}(w)(w')\ \&\ p(w')\big]$ MODAL CLOSENESS

This *almost* takes as argument a proposition p and a world w, and asserts that p does not hold in w, but does in some world w′ accessible to w via the CLOSE relation. There are various ways one could spell out this intensional similarity. Nouwen (2006), for example, does so in terms of the number of (relevant) propositions that are shared between worlds. Two worlds w and w′ are identical for a given purpose if all relevant propositions are true in those worlds. If they differ only in terms of one relevant proposition, they are 1-removed; if they differ in terms of two propositions, they are 2-removed, and so on. Whether or not a world is close enough to verify the truth depends on the context: sometimes, a world that is 3-removed will be an appropriate verifier for *almost*, and other times it might not be.

An alternative takes *almost* to convey scalar, rather than modal, closeness. Penka (2006), following work by Hitzeman (1992) (see also Amaral 2007), takes *almost* φ to express not that φ is true at some world close to, but distinct from, the world of evaluation, but rather that some scalar alternative to φ is true in the world of evaluation. The modal entry for *almost* above doesn't make explicit the inferences about the truth of some proposition in the world of evaluation. That is, while (34) conveys that it's false in the world of evaluation that one hundred people lost their lives, it also conveys that the amount of people who did in fact perish in the world of evaluation is close to one hundred. There being a 'close' world where a hundred people perished doesn't guarantee that close to one hundred did.

(34) Almost one hundred people died of the disease. Penka (2006:279)

This suggests that *almost* φ asserts the truth of some proposition ψ which is a *weaker* scalar alternative to φ. Setting aside how the alternatives are generated for the moment, this leads to the following semantics for scalar *almost*.

(35) $[\![\text{almost}_{\text{scalar}}]\!]^{g,c} = \lambda p_{\langle s,t\rangle}\lambda w.\ \neg p(w)\ \&\ \exists q\big[q \in \textbf{ALT}(p)\ \&\ q(w)\big]$ SCALAR CLOSENESS

This entry takes a proposition p and a world w and returns truth *iff* p is not true in w, and if there is some proposition q, in the set of alternatives to p, and q is true in w. Now, as Penka notes, this kind of an entry merely needs to require that the alternatives are ones which are 'close'—the alternatives to p aren't required to be stronger or weaker than p. So, considering (35), the relevant alternatives are the propositions expressed by *ninety-eight people died of the disease, ninety-nine people died of the disease, one hundred and one people died of the disease,* and so on so long as they are 'close.' If the alternatives entail p, then they cannot be the propositions that verify *almost,* i.e. the close q proposition. This is because *almost* asserts the falsity of p at the world of evaluation in this semantics, and so any stronger alternative cannot be true at that same world; the only possible verifying alternatives are those that are weaker.

These two camps do not exhaust the logical space of possible characterizations of closeness. In fact, we'll ultimately turn to something a little different later on. That said, this closeness is indeed something that is intuitively central to the meaning of *almost* and *barely*, and it must be a part of any adequate analysis thereof. It is also, generally, a source of vagueness: what counts as close in one context very well might not in another, and certainly the scale on which we're approximating matters. *Almost one hundred people died of the disease* could be true because ninety-five people died, and that's close enough to one hundred. *Almost ten people died* feels funny to say generally, even if nine people died. In the latter case, the raw number of exceptions is smaller, but it's percentage wise greater than the former case. Closeness, however it's encoded, must be flexible enough to capture for these distinctions. Using the term 'exceptions' is a good transition.

2.1.2 Exceptives and subtraction

Many of the examples above featured *almost* and *barely* appearing to modify a quantificational determiner, e.g. an expression like *every*, *no*, or *any*. Here are a few more.

(36) a. Almost every member of the Scooby Gang has lived at 1630 Revello.
     b. Xander has had romantic feelings for almost every friend he has.

(37) a. Barely any members of Cordelia's former clique speak to her anymore.
     b. Willow left barely any spellbooks untouched in her thirst for knowledge.

They are not the only modifiers of quantificational determiners, however, and in fact, *almost* in particular shares selectional restrictions with so-called exceptive *but* (Hoeksema 1987; von Fintel 1993; Crnič 2018).

(38) a. Every spellbook but *The Brekenkrieg Grimoire* is worth reading according to Willow.
     b. No book but *The Brekenkrieg Grimoire* is worth reading according to Willow.
     c. # Some book but *The Brekenkrieg Grimoire* is worth reading according to Willow.
     d. Any$_{\text{FCI}}$ book but *The Brekenkrieg Grimoire* is worth reading according to Willow.

(39) a. Almost every book is worth reading according to Willow.

b. Almost no book is worth reading according to Willow.

c. # Almost some book is worth reading according to Willow.

d. Almost any_{FCI} book is worth reading according to Willow.

What we see here is that exceptive *but* is compatible with universal *every* and negative existential *no*, but not simple existential *some* (Horn 1989; von Fintel 1993, a.o.); furthermore, both are compatible with free-choice *any*, arguably a universal quantifier (Eisner 1994; Dayal 1998, a.o.). That *almost* is compatible with universals and negative existentials, but not simple existentials, is also well-known; compatibility with *almost* has in fact been used as a diagnostic for universality (Carlson 1981; Zanuttini 1991, a.o.), though the validity has been debated (Partee 1986; Błaszczak 2000; Penka 2006, a.o.). Regardless, intuitively the two modifiers make similar contributions. As von Fintel (1993:141) says, we can understand both exceptive *but* and *almost* in these constructions as "marking the existence of exceptions to a generalization." *But* overtly marks the exception(s) in question, as we see in (38), but *almost* leaves it open just who or what the exceptions are. Let's call this idea **VON FINTEL'S CONJECTURE**.

(40) ***VON FINTEL'S (1993) CONJECTURE***[1]

Almost and *but* mark the existence of an exception to a quantificational generalization.

Barely completes the paradigm in a way. We have already seen that, unlike *almost*, *barely* licenses NPI quantificational expressions like *any* and *ever*. What's more is that *barely*'s distribution with respect to the quantificational determiners above seems complementary to that of *almost* and *but*.

(41) a. # Barely every book is worth reading according to Giles.

b. # Barely no book is worth reading according to Willow.

c. # Barely some book is worth reading according to Willow.

d. Barely any_{NPI} book is worth reading according to Willow.

Barely really only seems to be compatible with $\textit{any}_{\text{NPI}}$; the other examples are quite odd. Taken all together, the parallels and complementarity are striking. Let's add another conjecture.

[1]von Fintel himself attributes, in his 1991 UMass generals paper, this idea to a term paper by Paul Portner ("Non-partitive Determiner Combinations"):

> "One likely possibility is that [adverbs like *almost* or *nearly*] specify that the sentence may have only a small number of exceptions. That is, *Almost all 100 people were killed* should mean something like *All 100 people, except a few, were killed*".

Perhaps, then, we should call this the Portner-von Fintel conjecture.

(42) ***OUR CONJECTURE***

Barely marks the existence of an exception to a quantificational generalization

What comes now is a compositional implementation of this idea, but first, there are even more parallels to note.

The inferences that quantificational sentences containing *but* give rise to also parallel those containing *almost*. Consider first (43).

(43) Every spellbook but *The Brekenkrieg Grimoire* is worth reading according to Willow.

What (43) conveys is that, in Willow's view, every spellbook that isn't *The Brekenkrieg Grimoire* (BG) is worth reading, and that BG is not worth reading. In other words, BG is the only spellbook not worth reading. Now consider (44).

(44) No book but *The Brekenkrieg Grimoire* is worth reading according to Willow.

This conveys just the opposite of (43): in Willow's view, BG is the only spellbook worth reading, no other spellbook is. Each of these sentences carries a negative polar inference: from (43) we infer than in Willow's views it's false that every spellbook is worth reading, and from (44), we infer that there is some spellbook worth reading.

We saw environments where the polar inferences of *almost* and *barely* were invisible to certain adverbials; *but* shows similar properties.

(45) Good news! Every student of mine but Xander passed the test.

(46) Every student of mine but Xander passed test…

a. because they studied hard enough.

b. # because he didn't study hard enough.

In (45), the good news is that the vast majority of my students passed the test, not that Xander didn't. In the same vein, we can't use the continuation (46b) as justification for why Xander didn't pass; we can, however, use (46a) to justify why everyone else did pass. The polar inference in *but* sentences is invisible in some sense to the *because*-clauses.

There are even environments where the polar inferences of all three expressions can disappear. Nouwen (2006) shows that, in the consequents of conditionals, the polar inference of *almost* is "backgrounded," in a non-technical sense, or else wholly absent; we can show this, too, for *but* and *barely*.

(47) a. If Willow wants to get an A on the test, she has to answer every question but the last one.

b. If Willow wants to get an A on the test, she has to answer almost every question.

c. If Willow wants to fail the test, she has to answer barely any questions.

(47a) does not convey, at least not necessarily, that getting an A on the test requires *not* answering the last question, nor does (47b) convey that there are a small number of questions she must not answer in order to get an A. Of course, it's entirely possible that those are the requirements for passing, but that requires a special context. The negative polar inference doesn't seem obligatory here. Similarly, the positive polar inference typically conveyed by *barely* is absent in (47c); Willow needn't answer any questions if she wants to fail, she merely can answer a very small number at most. The consequents of conditionals aren't the only environments that can make these inferences disappear; antecedents of conditionals, too (Kilbourn-Ceron 2016 on *almost*) and under modals in ellipsis contexts (Crnič 2018 on both *but* and *almost*[2]).

There is something of the proximal inference in *but*, too, though this hasn't, to my knowledge, been discussed in the literature on exceptives prior. Assume that there are ten students in my class taking an exam—Anya, Buffy, Cordelia, Dawn, Ethan, Faith, Giles, Harmony, Kendra, and Larry. Next, assume that Faith and Larry are the only students who failed the test.

(48) Every student but Faith and Larry passed.

This is perfectly acceptable. Two of my students failed, but all eight others passed. Now imagine that Anya, Buffy, and Cordelia are the only students who passed the exam.

(49) # Every student but Dawn, Ethan, Faith, Giles, Harmony, Kendra, and Larry passed.

This is extremely odd. While we've not yet given an explicit semantics for *but*, let's simply assume for a moment that its contribution is to say its complement is the totality of exceptions to the quantificational claim it modifies. That is, when we set aside Dawn, Ethan, Faith, Giles, Harmony, Kendra, and Larry, all remaining students passed, and those students did not. This context is one that should verify (49), but the sentence is quite odd. It seems like *but* in these cases requires that the actual number of passing students be close to the total number of students taking the test. This, if we're characterizing it correctly, is a kind of proximal inference, just like *almost* and *barely* give rise to.

Taken together, the parallels between *almost*, *barely*, and exceptive *but* are striking. *Almost* and *but* appear to have quite similar distributions with respect to quantificational determiners, and *barely*'s appears to be complementary. They all seem to give rise to both polar and proximal

[2] It's possible to derive the weaker inference for (47b) (and the others) by having the DP take wide scope with respect to the modal *has to*. This weakens his argument. However, Crnič (2018) pushes this further by showing that the inference disappears in environments where you can't rely on scope; we'll discuss this in more detail in §2.3.

inferences, the former of which can disappear depending on the syntactic and semantic environment. Our analysis of *almost* and *barely*, then, should be closely related to that of *but*. Unlike the entries for *almost* and *barely* above, though, *but* is typically taken to be either a determiner modifier (von Fintel 1993, Crnič 2018) or else to simply denote a subtraction operation (Gajewski 2008, 2013).

A quick note on the distribution of *almost*, *barely* and *but*. The distribution isn't exactly parallel when we look beyond the quantificational determiners discussed. For example, *almost* and *barely* can modify numerals, whereas *but* cannot.

(50) a. $\{$Almost/barely$\}$ one hundred students attended graduation.
b. # One hundred students but Willow attended graduation.

This is unexpected on two counts. First, *almost* and *barely* were complementary in the data discussed above, but here they are both perfectly acceptable. Second, if *almost* and *but* are really exactly the same, it's not clear why we don't get *but* as a modifier here. Set aside these questions for now. Numeral constructions are the topic of Chapter 4, and there, we'll argue that there's good reason to think these contrasts are perfectly expected. So, for the remainder of the chapter, we'll focus on quantificational determiners *every*, *some*, and *no* throughout the remainder of the chapter.

2.1.3 Summary and roadmap

As we said above, most analyses of *almost* and *barely* take as their starting point the idea that these are clausal modifiers, that they operate semantically at the level of the proposition. Most analyses of exceptive *but* take it to operate as a determiner modifier or encode subtraction. The analytical starting point for our analysis of *almost* and *barely* is the striking parallels between *almost* and *barely* on the one hand and *but* on the other, taking very seriously the idea that *almost* and *barely* are in fact modifiers of quantificational determiners. We'll propose that they all share a common semantic core, encoding the proximal inference as a matter of their truth conditional import, and the polar inference derived as a kind of scalar implicature. From there, we'll turn to the core contribution of this dissertation: subtractives in degree constructions. We'll show how this analysis extends to other occurrences of *almost* and *barely*, and what we learn from those extensions.

§2 gives the basic proposal for *almost* and *barely*. Building on an influential analysis of exceptive *but* due originally to von Fintel (1993), we propose treating *almost* and *barely* as modifiers of quantificational determiners, and take them to encode set subtraction, removing elements from the restrictor of the quantifier. This, with the addition of a constraint on just how much can be removed from the restrictor (novel for exceptive *but*) comprises the proximal inference encoded by all three operators. We take the polar inference to be derived as a kind of scalar implicature, derived through exhaustification (Gajewski 2008, 2013; Spector 2014; Hirsch 2016; Crnič 2018). We explain how the distribution of the modifiers naturally falls out of our analysis as a result of the ingredients above, alongside facts like NPI licensing and the disappearance of the polar inference. §3 shows how facts like NPI licensing and the disappearance of the polar inference work on our analysis. §4 compares our analysis to others on the market. §5 concludes.

A note on synonyms and nomenclature. *Almost* is (almost) synonymous with expressions like *nearly*, *practically*, *just about*; *barely* is close to *only just*, *hardly*. Quirk et al. (1985) (1985:597) call *almost* an APPROXIMATOR, expressions which "serve to express an approximation to the force of the verb, while indicating that the verb concerned expresses more than is relevant." Other such expressions include the synonyms listed above, but outside of the verbal domain, it likely includes expressions like *approximately* and *roughly*. They call *barely* a MINIMIZER, what they term a "negative maximizer," alongside expressions like *little* and *a bit*. Both classes are subsumed by the class of DOWNTONERS, which have a "general lowering effect." These characterizations, of course, readily apply to verbal uses of the modifiers, but all the same they make sense enough for uses of the expressions as quantifier modifiers. Exceptive *but* is often called, well, an EXCEPTIVE, i.e. an expression marking an exception to some quantificational claim; *other than*, *except for*, and other such expressions also are in this class. Work that assumes *almost* is quite like *but* in semantic contribution might call the modifier an exceptive as well; we'll use something slightly different. We'll refer to this expression, alongside *barely* and *but*, as SUBTRACTIVES, since their core semantic contribution is set subtraction. In the end, then, the theory developed and defended here might indicate that some of the categories above are inappropriate for our quarry. We also should discuss, at least a bit, the differences between *almost* and *barely* and their kin, which, though subtle, are nonetheless important. We'll worry about these finer details later.

2.2 Subtractive *but*, *almost* and *barely*

Developing an analysis of *almost* and *barely* as subtractives operators, modeled after *but*, requires a semantics for the latter. This is where we'll start.

2.2.1 The meaning of *but* sentences

The semantics of *but* sentences is our starting point. (51a) conveys that *The Brekenkrieg Grimoire* is the only spellbook not worth reading, and (51b) that it is the only one worth reading. In both cases, it is the only exception to the unmodified quantificational claims, i.e. their versions without ‘*but BG*’.

(51) a. Every spellbook but *The Brekenkrieg Grimoire* is worth reading.

b. No spellbook but *The Brekenkrieg Grimoire* is worth reading.

There are two key components to *but* sentences. First, *but* removes an EXCEPTION, the entity or entities denoted by its complement (here, BG), from a set (here, the set of spellbooks). Second, the exception must be the smallest non-empty set removable such that the resulting modified quantificational claim is true. That is, $\{\mathrm{BG}\}$ is the smallest set $\mathbb{X}$ you can subtract from the set of spellbooks and claim that {*every*/*no*} *(remaining) spellbook is worth reading* is true. Schematically, for any quantifier Δ with a denotation $\mathscr{D}$, and expressions α, β, and χ whose denotations characterize the sets $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{X}$, respectively:

(52) $[\![\Delta\ \alpha \text{ but } \chi\ \beta]\!]^{g,c} = 1$ *iff* $\underbrace{\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})}_{\textbf{Subtraction}}\ \&\ \underbrace{\forall\mathbb{X}'\left[\mathbb{X} \nsubseteq \mathbb{X}' \rightarrow \neg\mathscr{D}(\mathbb{A}-\mathbb{X}')(\mathbb{B})\right]}_{\textbf{Uniqueness/Leastness}}$

These are the truth conditions from von Fintel (1993). The first conjunct in (52) is (domain) subtraction: when we remove $\mathbb{X}$ from $\mathbb{A}$, the retrictor of $\mathscr{D}$, we get truth. The second conjunct is *Uniqueness* (von Fintel 1993) or *Leastness* (Gajewski 2008)—for every set $\mathbb{X}'$ that is not a superset of $\mathbb{X}'$, subtracting $\mathbb{X}'$ from $\mathbb{A}$ makes the modified quantification false, or equivalently, every $\mathbb{X}'$ that makes the modified quantification true must be a superset of $\mathbb{X}$. That is, $\mathbb{X}$ must be the smallest, non-empty exception that makes the quantification true. Now, non-emptiness doesn't actually follow; if $\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B}) = 1$ because $\mathbb{X} = \emptyset$ and $\mathscr{D}(\mathbb{A})(\mathbb{B}) = 1$, then Uniqueness is trivially satisfied, since every $\mathbb{X}'$ is a superset of $\emptyset$. We'll say more on this later, but take it as an assumption for the moment. We define set subtraction as follows.

(53) ***SET SUBTRACTION*** ‘$-$’

a. For arbitrary sets $\mathbb{A}$, $\mathbb{B}$: $\mathbb{A}-\mathbb{B} := \{x \mid x \in \mathbb{A}\ \&\ x \notin \mathbb{B}\}$.
($\mathbb{A}$ is the *minuend* and $\mathbb{B}$ the *subtrahend*; the set $\mathbb{C}$ such that $\mathbb{A}-\mathbb{B} = \mathbb{C}$ is the *difference* or *remainder*.)

b. For any arbitrary expressions α, β, of type $\langle e,t\rangle$, let $\mathbb{A}$ and $\mathbb{B}$ be the sets characterized by $[\![\alpha]\!]^{g,c}$ and $[\![\beta]\!]^{g,c}$, and let ‘$\alpha - \beta$’ be a wellformed expression of type $\langle e,t\rangle$ such that $[\![\alpha - \beta]\!]^{g,c}$ is a function characterizing the set $\mathbb{A}-\mathbb{B}$

Now, to see how this all works applied to (51a), we need some entries. We'll utilize a set-based semantics for quantificational determiners, and we'll take NPs like *spellbook* to be world-dependent expressions.

(54) $[\![\text{every}]\!]^{g,c} = \lambda A_{\langle e,t\rangle}\lambda B_{\langle e,t\rangle}.\ \mathbb{A} \subseteq \mathbb{B}$

(55) $[\![\text{spellbook}]\!]^{g,c} = \lambda w \lambda x.\ \textsc{spellbook}_w(x)$

Set characterized by $[\![\text{spellbook}]\!]^{g,c}(w)$: $\mathbb{S}_w$

(56) $[\![\text{be worth reading}]\!]^{g,c} = \lambda w \lambda x.\ \textsc{worth-reading}_w(x)$

Set characterized by $[\![\text{be worth reading}]\!]^{g,c}(w)$: $\mathbb{WR}_w$

Putting these pieces together with the truth conditions above:

(57) $[\![(51a)]\!]^{g,c}(w) = 1$ *iff* $\underbrace{\mathbb{S}_w - \{BG\} \subseteq \mathbb{WR}_w}_{\textbf{Subtraction}}\ \&\ \underbrace{\forall \mathbb{X}' \left[\{BG\} \not\subseteq \mathbb{X}' \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \mathbb{WR}_w \right]}_{\textbf{Uniqueness}}$

Suppose that *The Brekenkrieg Grimoire* is a spellbook. The first conjunct, Subtraction, requires that when $\{BG\}$ is removed from the set of spellbooks, the remainder is a subset of all things worth reading. Hence, every non-BG spellbook must be worth reading. The second conjunct, Uniqueness, requires that every set $\mathbb{X}'$ that does not contain BG makes the modified quantificational claim false. Consider that for every non-superset $\mathbb{X}'$ of $\{BG\}$, there is something in $\{BG\}$ that is not in $\mathbb{X}'$. Since $\{BG\}$ is a singleton, none of these $\mathbb{X}'$ is a supertset of $\{BG\}$, and since $\{BG\} \subseteq \mathbb{S}_w$, for all such $\mathbb{X}'$, the definition of subtraction guarantees that $BG \in \mathbb{S}_w - \mathbb{X}'$. Note that wnothing requires any such $\mathbb{X}'$ to be a subset of $\mathbb{S}_w$, or even overlap with it—it simply must not contain BG. Now, for each such $\mathbb{X}'$, Uniqueness requires that there is some $y \in \mathbb{S}_w - \mathbb{X}'$ that is not in $\mathbb{WR}_w$. Since Subtraction requires that every non-BG spellbook be worth reading, the culprit $y \in \mathbb{S}_w - \mathbb{X}'$ not in $\mathbb{WR}_w$ must be BG. This, of course, requires that BG is a spellbook not worth reading, that it is the only such spellbook, and hence that not every spellbook is worth reading. This last bit is especially relevant when considering the similarities between *but* and *almost*: that (51a) requires that not every spellbook be worth reading is an instantiation of the polar inference. Here, we derived it through Uniqueness. This can be made even clearer: an $\mathbb{X}'$ such that $\{BG\} \not\subseteq \mathbb{X}'$ is $\emptyset$. Uniqueness thus requires that $\mathbb{S}_w - \emptyset \not\subseteq \mathbb{WR}_w$, i.e. $\mathbb{S}_w \not\subseteq \mathbb{WR}_w$. These truth conditions thus assert the polar inference, albeit indirectly.

The case of *but* with *no* is not all that different from *every.*

(58) $[\![\text{no}]\!]^{g,c} = \lambda A_{\langle e,t \rangle} \lambda B_{\langle e,t \rangle}.\ \mathbb{A} \cap \mathbb{B} = \emptyset$

(51b) No spellbook but *The Brekenkrieg Grimoire* is worth reading.

a. $[\![(51b)]\!]^{g,c}(w) = 1$ *iff*

$\mathbb{S}_w - \{BG\} \cap \mathbb{WR}_w = \emptyset\ \&\ \forall \mathbb{X}' \left[\{BG\} \not\subseteq \mathbb{X}' \rightarrow \mathbb{S}_w - \mathbb{X}' \cap \mathbb{WR}_w \neq \emptyset \right]$

For the first conjunct to be true, every non-BG spellbook must be such that it's not worth reading. For the second conjunct to be true, every set $\mathbb{X}'$ that doesn't contain BG must leave in $\mathbb{S}_w - \mathbb{X}'$ some spellbook that is worth reading. This can only be BG. Therefore, there is some spellbook that is worth reading, and BG is the only such spellbook; everything else is trash. As before, Uniqueness entails the polar inference.

The discussions above assumed that *The Brekenkrieg Grimoire* was a spellbook. If it weren't, intuitively (51a) and (51b) would be quite odd. Consider an example that presupposes less niche fictional lore.

(59) # Every novel but the *Mona Lisa* by da Vinci is worth reading.

The *Mona Lisa*, of course, is a painting, not a novel, so *but*'s subtraction is in a sense vacuous here. The analysis predicts (59) is false if in w, as in our world, the painting is a painting.

(60) $[\![(59)]\!]^{g,c}(w) = 1$ *iff* $\mathbb{N}_w - \{ML\} \subseteq \mathbb{WR}_w \;\&\; \forall \mathbb{X}' \big[\{ML\} \not\subseteq \mathbb{X}' \rightarrow \mathbb{N}_w - \mathbb{X}' \not\subseteq \mathbb{WR}_w\big]$

Since the *Mona Lisa* isn't a novel, the remainder of $\mathbb{N}_w - \{ML\}$ is just $\mathbb{N}_w$, so the first conjunct requires that every novel be worth reading. However, then there can be no $\mathbb{X}'$ that satisfies Uniqueness. For $\mathbb{N}_w - \mathbb{X}' \not\subseteq \mathbb{WR}_w$ to be true, there must be some element in $\mathbb{N}_w - \mathbb{X}'$ not in $\mathbb{WR}_w$, but by Subtraction $\mathbb{N}_w \subseteq \mathbb{WR}_w$. Set subtraction cannot add elements to $\mathbb{N}_w$! $\mathbb{N}_w$ is therefore an $\mathbb{X}'$ that always falsifies Uniqueness. Of course, falsity alone doesn't guarantee the oddity of the sentence. Let's add a presupposition, then, to *but* sentences.

(61) $[\![\Delta\ \alpha \text{ but } \chi\ \beta]\!]^{g,c}$ is defined only if $\mathbb{X} \subseteq \mathbb{A}$; where defined:
$[\![\Delta\ \alpha \text{ but } \chi\ \beta]\!]^{g,c} = 1$ *iff* $\underbrace{\mathscr{D}(\mathbb{A} - \mathbb{X})(\mathbb{B})}_{\textbf{Subtraction}} \;\&\; \underbrace{\forall \mathbb{X}' \big[\mathbb{X} \not\subseteq \mathbb{X}' \rightarrow \neg \mathscr{D}(\mathbb{A} - \mathbb{X}')(\mathbb{B})\big]}_{\textbf{Uniqueness}}$

This presupposition thus predicts (59) is infelicitous, allowing us to rule it and similar cases out on these grounds.

Now, let's return to an odd case we noted in §2.2. It seems that the exception introduced by *but* can't be too big. Assume that there are ten students in my class taking an exam—Anya, Buffy, Cordelia, Dawn, Ethan, Faith, Giles, Harmony, Kendra, and Larry. Anya, Buffy, and Cordelia are the only students who passed the exam.

(62) # Every student but Dawn, Ethan, Faith, Giles, Harmony, Kendra, and Larry passed.

This is odd, but nothing in the semantics predicts this.

(63) $[\![(62)]\!]^{g,c}(w)$ is defined only if $\{D, E, F, G, H, K, L\} \subseteq \mathbb{S}_w$; where defined:
$[\![(62)]\!]^{g,c}(w) = 1$ *iff*
$\mathbb{S}_w - \{D, E, F, G, H, K, L\} \subseteq \mathbb{A}_w \;\&\; \forall \mathbb{X}' \big[\{D, E, F, G, H, K, L\} \not\subseteq X' \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \mathbb{A}_w\big]$

The first conjunct requires that Anya, Buffy, and Cordelia all passed the test, true in the context; the second requires that $\{D, E, F, G, H, K, L\}$ be the smallest exception. This is also true in the context. Every non-superset of the exception lacks at least one element of the exception, and since that element is a student, it is in the remainder of $\mathbb{S}_w - \mathbb{X}'$. By assumption, those students didn't pass, and hence for all such $\mathbb{X}'$, $\mathbb{S}_w - \mathbb{X}' \not\subseteq \mathbb{A}_w$. What accounts for the oddity of (62), then?

Intuitively, we use *but* when a quantificational claim is false, but not false by much, and so we save it by naming the exception(s). The idea is that (62) is odd because the exception takes us too far away from the bare universal claim—there are just too many exceptions. The situation we're in isn't close enough to a situation in which the bare universal claim is true. This characterization is quite like the 'closeness' of the proximal inference encoded by *almost*—*almost every student passed* only seems true if the failers are few in number. For *but*, the proposal is to add a constraint on the size of the exception. It must count as 'small' in the context of utterance.

(64) $[\![\Delta\ \alpha \text{ but } \chi\ \beta]\!]^{g,c}$ is defined only if $\mathbb{X} \subseteq \mathbb{A}$; where defined:
$$[\![\Delta\ \alpha \text{ but } \chi\ \beta]\!]^{g,c} = 1 \textit{ iff } \underbrace{\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})}_{\textbf{Subtraction}} \ \& \ \underbrace{\textsc{small}_c(\mathbb{X})}_{\textbf{Size constraint}} \ \& \ \underbrace{\forall \mathbb{X}' \big[\mathbb{X} \not\subseteq \mathbb{X}' \rightarrow \neg\mathscr{D}(\mathbb{A}-\mathbb{X}')(\mathbb{B})\big]}_{\textbf{Uniqueness}}$$

Is this really part of the truth conditions? Given how the intuitions about (62) were characterized, it perhaps should be something more like a presupposition or an implicature. Either alternative is ultimately compatible with the final analysis, as far as I can tell. We'll not worry too much about that for the moment, and simply leave it as truth conditional.

How should we define the function SMALL? When is an exception too big? That depends. First, the size of an exception $\mathbb{X}$, relative to the size of the set $\mathbb{A}$ from which it is subtracted, matters. In the context of count nouns like *students in my class* or *spellbooks*, the size of the sets intuitively corresponds to its cardinality. The size of $\{\text{D, E, F, G, H, K, L}\}$ is thus the cardinality of the set—here, seven. In the context described, the set $\mathbb{S}_w$ has ten members, and so $\{\text{D, E, F, G, H, K, L}\}$ is 70% of $\mathbb{S}_w$. One way SMALL could operate, then, is by requiring that exceptions $\mathbb{X}$ be no more than, say, 20% of the set $\mathbb{A}$ from which they're subtracted. Context matters, too, though. We might allow (65a) to be acceptable if I have ten students, but (65b) sounds a little worse in a context where I have one hundred students, even though the ratio of exception to subtrahend is the same.

(65) a. Every student of mine but two passed the test.
b. ? Every student of mine but twenty passed the test.

It seems like stipulating smallness in terms of a cut-and-dry percentage won't quite work—the context plays a huge role here, and the boundary of smallness is fuzzy. That said, it seems intuitive that to count as small, the size of an exception $\mathbb{X}$ must be less than half the size of the set from which it is subtracted. If five of my ten students passed, I cannot say *Every student of mine but five passed the test*, because the exception is at least half the size of the restrictor. Here's a definition:

(66) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, μ_c be a measure function, determined by the context, mapping sets to numerical values representing their size, and n_c be a contextually determined numerical threshold for size.

$\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\mu_c(\mathbb{X}) \leq n_c$

(67) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

For any context c, an utterance of φ in c is felicitous only if $n_c < \frac{1}{2}(\mu_c(\mathbb{A}))$

The measure function μ is left unspecified here. In the kinds of cases we're discussing in this chapter, i.e. quantificational determiners taking as argument functions with atomic entities in their domain, we can construe this as a cardinality function; we posit a constraint on the maximal size of the exception relative to the argument from which it is subtracted. Applied to perhaps the most minimal case, this means that if the cardinality of some set $\mathbb{A}$ is less than 3, then in no context is there a small, non-empty $\mathbb{X}$ that can be subtracted from $\mathbb{A}$. All substantive non-vacuous subtraction from such a $\mathbb{A}$ would subtract, well, too much.

One final case: *but* with an existential quantifier, *some*. This always sounds odd.

(68) # Some spellbook but *The Brekenkrieg Grimoire* is worth reading.

a. $[\![(68)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$; where defined: $[\![(68)]\!]^{g,c}(w) = 1$ *iff* $\mathbb{S}_w - \{BG\} \cap \mathbb{WR}_w \neq \emptyset$ & $\textsc{small}_c(\{BG\})$ & $\forall \mathbb{X}' [\{BG\} \not\subseteq \mathbb{X}' \rightarrow \mathbb{S}_w - \mathbb{X}' \cap \mathbb{WR}_w = \emptyset]$

First, assume that the subset presupposition is met. Subtraction requires that at least one of the non-BG spellbooks is worth reading; the size constraint requires that $\{BG\}$ be considered small in c. Assume both are true: given the formulation of the size constraint, if $\{BG\}$ counts as small, it must contain less than half the spellbooks. For a singleton set to count as small, there must therefore be at least three spellbooks in w. Uniqueness requires that subtraction of all non-supersets of $\{BG\}$ from $\mathbb{S}_w$ yield an empty intersection with $\mathbb{WR}_w$. Whether or not BG is worth reading, we'll get falsity here. Suppose BG is worth reading: there will be a non-superset that doesn't contain BG, but which leaves BG in the set of spellbooks, and so the second conjunct will be false. Now suppose BG isn't worth reading, then there are at least one of the two remaining spellbooks must be worth reading. If both are worth reading, either singleton will do to make the second conjunct false. If only one is worth reading, then a singleton set that contains a non-BG spellbook not worth reading will make the second conjunct false. The conclusion, therefore, is that (68) can never be true.

Let's take stock for a moment. Here are the truth conditions of *but*-sentences that we assume.

(69) $[\![\Delta\ \alpha \text{ but } \chi\ \beta]\!]^{g,c}$ is defined only if $\mathbb{X} \subseteq \mathbb{A}$; where defined:

$[\![\Delta\ \alpha \text{ but } \chi\ \beta]\!]^{g,c} = 1$ *iff* $\underbrace{\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})}_{\textbf{Subtraction}}$ & $\underbrace{\textsc{small}_c(\mathbb{X})}_{\textbf{Size constraint}}$ & $\underbrace{\forall \mathbb{X}' [\mathbb{X} \not\subseteq \mathbb{X}' \rightarrow \neg\mathscr{D}(\mathbb{A}-\mathbb{X}')(\mathbb{B})]}_{\textbf{Uniqueness}}$

At this point, we've mostly recapped the meaning of *but* sentences, largely following the tradition started by von Fintel (1993) and taken up by Gajewski (2008, 2016), Crnič (2018), and others. *But*, in this view, encodes first set subtraction: it removes a set from the restrictor of the quantificational determiner that it modifies. We've added two additional constraints. First, a presupposition that the exception is a subset of the set from which it is subtracted, i.e. the restrictor of the modified determiner, and second, a requirement that an exception introduced by *but* be small in the context of utterance. The size constraint, in tandem with set subtraction, looks quite a bit like the proximal inference we see in *almost* sentences. We also saw that *but* requires the exception to be the smallest such exception, and that this ultimately entails that the unmodified quantificational claim is false. This is quite like the polar inference in *almost* sentences. In the next section, we'll extend a particular version of this analysis to *almost* and *barely* alongside *but*.

2.2.2 Set subtraction and exhaustification

2.2.2.1 Composition

The core of our analysis of subtractives comes from Gajewski (2008, 2013) and Crnič (2018). We are assuming a DISTRIBUTED analysis of subtractives, in Crnič's terminology, in proposing that the meanings of subtractive sentences are derived from the contributions of two distinct operators. We'll assume that the proximal inferences encoded in *but* sentences—both set subtraction and the size constraint—are contributed by *but* directly. We split off the calculation of the polar inference, deriving it through the application of a separate operator responsible for generating scalar implicatures.

(70) $[\![\text{but}]\!]^{g,c} = \lambda X \lambda A \lambda \mathscr{D}_{\langle et,ett\rangle} \lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}.\ \mathscr{D}(\mathbb{A} - \mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$ $\qquad \langle et,\langle et,\langle\langle et,ett\rangle,\langle et,t\rangle\rangle$

This high-type semantics ultimately allows *but* to modify the quantificational determiner by subtracting from its restrictor the exception $\mathbb{X}$; it then feeds the determiner its normal nuclear scope. The size constraint is novel, and requires that $\mathbb{X}$ count as small in the context of utterance. We'll simply assume that expressions like *The Brekenkrieg Grimoire* denote sets; plausibly they are lifted by a type-shifting operator. It's not really important for us here to decide this matter.

Almost has the same truth conditional import that *but* does: it removes from the restrictor of a quantificational determiner an exception $\mathbb{X}$ (Crnič 2018) and requires that $\mathbb{X}$ be small. This comprises, of course, the proximal inference *almost* sentences give rise to[3]. Our novel proposal for *barely* is that it has the exact same function as *almost* and *but*, modifying a quantificational

[3]McKenzie & Newkirk (2019) also propose a similar kind of size constraint as a part of the derivation of *almost*'s proximal inference.

determiner via set subtraction, but additionally contributes negation. Reflecting the fact that *almost* and *barely* linearly precede the quantificational determiners they modify, they take their arguments in a slightly different order from *but*.

(71) a. $[\![\text{almost}]\!]^{g,c} = \lambda X_{\langle e,t\rangle}\lambda\mathscr{D}_{\langle et,ett\rangle}\lambda A_{\langle e,t\rangle}\lambda B_{\langle e,t\rangle}\colon \mathbb{X} \subseteq \mathbb{A}.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

b. $[\![\text{barely}]\!]^{g,c} = \lambda X_{\langle e,t\rangle}\lambda\mathscr{D}_{\langle et,ett\rangle}\lambda A_{\langle e,t\rangle}\lambda B_{\langle e,t\rangle}\colon \mathbb{X} \subseteq \mathbb{A}.\ \neg\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle et,\langle et,\langle\langle et,ett\rangle,\langle et,t\rangle\rangle$

Unlike *but*, *almost* and *barely* have no overt exception, so what saturates this argument, what *is* the exception? We'll take it to be a covert pronoun of type $\langle e,t\rangle$, receiving its value from the assignment function g. We'll represent this pronoun as χ_n, where *n* is an index that the assignment function maps to a set of individuals. It is crucial that the exception be specific; we'll discuss why in a little while, but grant for now that the precise identity of the exception needn't be known to the utterer of some true *almost* sentence.

These contributions constitute the proximal inference in *almost* and *barely* sentences on this analysis. For some determiner Δ, and arbitrary expressions α, β, and χ, $[\![\textit{almost}\ \chi\ \Delta\ \alpha\ \beta]\!]^{g,c}$ conveys that all that separates us from the truth of $[\![\Delta\ \alpha\ \beta]\!]^{g,c}$ is that small exception $\mathbb{X}$. Similarly, $[\![\textit{barely}\ \chi\ \Delta\ \alpha\ \beta]\!]^{g,c}$ conveys that that small $\mathbb{X}$ is all that separates us from the falsity of $[\![\Delta\ \alpha\ \beta]\!]^{g,c}$. These entries give us license to state some truth conditional equivalences between the subtractives. Let α, β, and χ, be arbitrary expressions whose denotations characterize the sets $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{X}$, respectively; let g be such that $g(i) = \mathbb{X}$.

(72) $[\![[\text{every}\ [\alpha\ [\text{but}\ \chi]]]]\!]^{g,c} = [\![[[[\text{almost}\ \chi_i]\ \text{every}]\ \alpha]]\!]^{g,c} =$
$\lambda B\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ (\mathbb{A}-\mathbb{X}) \subseteq \mathbb{B}\ \&\ \textsc{small}_c(\mathbb{X})$

(73) $[\![[\text{some}\ [\alpha\ [\text{but}\ \chi]]]]\!]^{g,c} = [\![[[[\text{almost}\ \chi_i]\ \text{some}]\ \alpha]]\!]^{g,c} = [\![[[[\text{barely}\ \chi_i]\ \text{no}]\ \alpha]]\!]^{g,c} =$
$\lambda B\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ (\mathbb{A}-\mathbb{X}) \cap \mathbb{B} \neq \emptyset\ \&\ \textsc{small}_c(\mathbb{X})$

(74) $[\![[\text{no}\ [\alpha\ [\text{but}\ \chi]]]]\!]^{g,c} = [\![[[[\text{almost}\ \chi_i]\ \text{no}]\ \alpha]]\!]^{g,c} = [\![[[[\text{barely}\ \chi_i]\ \text{any}_{\text{NPI}}]\ \alpha]]\!]^{g,c} =$
$\lambda B\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ (\mathbb{A}-\mathbb{X}) \cap \mathbb{B} = \emptyset\ \&\ \textsc{small}_c(\mathbb{X})$

(75) $[\![[[[[\text{almost}\ \chi_i]\ \text{some}]\ \alpha]\ \beta]]\!]^{g,c}$ and $[\![[[[[\text{barely}\ \chi_i]\ \text{every}]\ \alpha]\ [\text{not}\ \beta]]]\!]^{g,c}$ are defined only if $\mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset$; where defined,
$[\![[[[[\text{almost}\ \chi_i]\ \text{some}]\ \alpha]\ \beta]]\!]^{g,c} = [\![[[[[\text{barely}\ \chi_i]\ \text{every}]\ \alpha]\ [\text{not}\ \beta]]]\!]^{g,c}$
$= 1$ *iff* $(\mathbb{A}-\mathbb{X}) \cap \mathbb{B} \neq \emptyset\ \&\ \textsc{small}_c(\mathbb{X})$

(76) $[\![[[[[\text{almost}\ \chi_i]\ \text{no}]\ \alpha]\ \beta]]\!]^{g,c}$ and $[\![[[[[\text{almost}\ \chi_i]\ \text{every}]\ \alpha]\ [\text{not}\ \beta]]]\!]^{g,c}$ are defined only if $\mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset$; where defined,
$[\![[[[[\text{almost}\ \chi_i]\ \text{no}]\ \alpha]\ \beta]]\!]^{g,c} = [\![[[[[\text{almost}\ \chi_i]\ \text{every}]\ \alpha]\ [\text{not}\ \beta]]]\!]^{g,c}$
$= 1$ *iff* $(\mathbb{A}-\mathbb{X}) \subseteq \overline{\mathbb{B}}\ \&\ \textsc{small}_c(\mathbb{X})$

These equivalences allow us to streamline the discussion by focussing primarily on *almost* sentences of this form. That is, conclusions drawn on the basis of *almost* χ *every* α will extend to *every* α *but* χ, and so on following the equivalences above. Note, though, that it should be clear why *barely*, but not *almost* or *but*, licenses NPI determiners: its negation makes it a Downward Entailing operator, and NPIs like *any* and *ever* are licensed in the scope of such operators (Ladusaw 1979); neither *almost* nor *but* themselves are Downward Entailing, and hence, can't license NPIs in and of themselves.

The polar inference for all three subtractives will be derived separately from the proximal inference now. The exhaustivity operator EXH, commonly utilized in the derivation of scalar implicatures (Fox 2007, i.a.), is posited as the operator deriving these inferences. A necessary condition for the licensing of all three subtractives is their being in the scope of this operator (Gajewski 2008, 2013[4]; Spector 2014; Crnič 2018). The meaning is syncategorematically given below.

(77) For any φ, and any admissible g, c:
$$[\![\textbf{EXH } \varphi]\!]^{g,c}(w) = 1 \textit{ iff } [\![\varphi]\!]^{g,c}(w) = 1 \;\&\; \forall\psi\left[\psi \in \textbf{ALT}(\varphi) \;\&\; [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\right]$$

(78) ***STRAWSON ENTAILMENT*** '$\Rightarrow_{ST}$' — von Fintel (1999)

For any φ and ψ of type $\langle s,t \rangle$, and any context c and assignment function g:
$$[\![\varphi]\!]^{g,c} \Rightarrow_{ST} [\![\psi]\!]^{g,c} \textit{ iff } \forall w\left[[\![\varphi]\!]^{g,c}(w) = 1 \;\&\; [\![\psi]\!]^{g,c}(w) \neq \# \rightarrow [\![\psi]\!]^{g,c}(w) = 1\right]$$

The sister of EXH, φ, is called the *prejacent* of EXH. φ is a logical form, and implicit in the definition is that φ denotes a proposition, a function from worlds to truth values. $[\![\textbf{EXH } \varphi]\!]^{g,c}(w)$ requires that φ be true at the world of evaluation w, and that all ψ in the set of alternatives to φ (i.e., $\textbf{ALT}(\varphi)$) that are not Strawson entailed by φ are not true in the world of evaluation[5]. We'll furthermore adopt the following constraint on the distribution of EXH, which blocks its presence in Downward Entailing environments (Spector 2014; see also Chierchia et al. 2012; Fox & Spector 2009, 2018).

(79) ***ECONOMY CONSTRAINT ON THE DISTRIBUTION OF*** **EXH**

Let φ be an arbitrary expression containing an occurrence of the EXH operator. EXH is not licensed in φ if eliminating this occurrence leads to an expression ψ such that for any admissible g, c, $[\![\psi]\!]^{g,c} \Rightarrow [\![\varphi]\!]^{g,c}$.

---

[4] Gajewski's (2008) operator is closely related to EXH, but crucially not equivalent to it. It is LEAST:

(1) $[\![\text{LEAST}]\!]^{g,c} = \lambda X \lambda \mathscr{Q}_{et,t}.\ \mathscr{Q}(\mathbb{X}) \;\&\; \forall \mathbb{X}'\left[\mathscr{Q}(\mathbb{X}') \rightarrow \mathbb{X} \subseteq \mathbb{X}'\right]$

It will become clear that in most situations, this operator yields the same results as EXH. Gajewski (2013) revisits his earlier work and tries to utilize EXH in lieu of LEAST, but runs into some issues. We'll discuss this in more detail in §2.3.3.

[5] Strawson entailment is necessary for our purposes because of the presuppositions encoded by subtractives.

This condition bars occurrences of **exh** where it doesn't add anything to the meaning of the sentence it modifies. If, for example, all alternatives to some φ are entailed by $[\![\varphi]\!]^{g,c}$, then $[\![\varphi]\!]^{g,c}$ will entail $[\![\textbf{EXH}\ \varphi]\!]^{g,c}$, and **exh** will not be licensed.

The connection between **exh** and subtractives will be made tighter. We'll take subtractives to be licensed *only* if they are within the scope of an **exh** operator.

(80) **exh** ***and the licensing of subtractive operators***
Subtractive operators are only licensed in the scope of an **exh** operator.

This, obviously, blocks the occurrence of a subtractive should they occur in a LF *sans* **exh**. Such a constraint has been posited by Gajewski (2008, 2013), Hirsch (2016), and Crnič (2018) for *but*, and Spector (2014) for *almost*[6], taken up by Crnič as well[7]. One might reasonably ask why some operators should only be licensed in the scope of an **exh** operator. Gajewski (2008) ties it to focus, with *but* focus-marking its exception and introducing focus alternatives; **exh** then operates on those alternatives. Perhaps the presence of focus marking forces the presence of **exh**. This is an important question, but one we'll set aside for now, and simply assume the condition above applies.

One might reasonably balk at the very high types of subtractives, so a brief note on this subject is necessary. For von Fintel, a low meaning is not possible, since Uniqueness is directly contributed by *but*, and this requires access to the determiner and both its arguments. However, once we go the **exh** route, we have wiggle room. Gajewski (2008, 2013) takes the wiggle, proposing a lower meaning for *but*:

(81) $[\![\text{but}]\!]^{g,c} = \lambda X \lambda A.\ \mathbb{A} - \mathbb{X}$

This semantics makes crystal clear what subtractives are: modifiers encoding set subtraction, and only set subtraction. This is appealing. There are, however, a few reasons we should consider the higher-type route.

[6]Spector's main quarry is the contrast between French disjunctions *soit...soit* and *ou*; the former obligatorily triggers exhaustivity inferences and is anti-licensed in globally Downward Entailing environments, whereas the latter does not necessarily trigger such inferences, and is anti-licensed in locally Downward Entailing environments. This, he argues, boils down to *soit...soit* necessarily being in the scope of **exh**, itself not licensed in Downward Entailing environments. He then extends the analysis to *almost*. We'll see in Chapter 5 that there are wrinkles to *almost*'s being anti-licensed in globally Downward Entailing environments, and we'll return to a discussion of this licensing condition there.

[7]Some care is needed, ultimately, in claiming this is something *all* subtractives require. Crnič (2018) gives the example of *other than* as a subtractive that does not require exhaustification.

(i) Every witch other than Willow passed the test.

(ii) Some witch other than Willow passed the test.

(i) doesn't seem to require that Willow didn't pass the test, though it's certainly a viable interpretation. (ii), notably, is perfectly acceptable, though of course *but* in place of *other than* is not. This can be tied to the *absence* of **exh** in (ii). Why there should be such a distinction, and what drives it, is left for future investigation.

The first is linear order. *But*, at the level of the surface syntax, sits in between a predicate and an exception. The low type semantics reflects that, but since *almost* and *barely* precede the determiners they modify, the low-type semantics is less plausible. If we take the surface syntax to directly reflect LF, the lower types cannot compose properly. That's a big if, though. There are plenty of places where we don't take surface syntax to directly reflect to LF—Quantifier Raising for interpretability, for example, involves a disconnect between surface syntax and LF.

The second reason is the size constraint, which as is makes it difficult to square with the lower types for all three subtractives depending on its status.

(82) $[\![\text{but}]\!]^{g,c} = \lambda X \lambda A.\ \mathbb{A} - \mathbb{X}\ \&\ \textsc{small}_c(\mathbb{X})$ INCOHERENT!

The output of (82) is a set; the size constraint, given a set, returns a truth value, so if it is part of the truth conditional import of *but*, composition won't work. To save this, we'd have to treat the size constraint as a presupposition, as in (83).

(83) $[\![\text{but}]\!]^{g,c} = \lambda X \lambda A{:}\ \textsc{small}_c(\mathbb{X}).\ \mathbb{A} - \mathbb{X}$

This would allow appropriate composition for all three subtractives, so whether this is viable depends on what we think the appropriate status is of the size constraint.

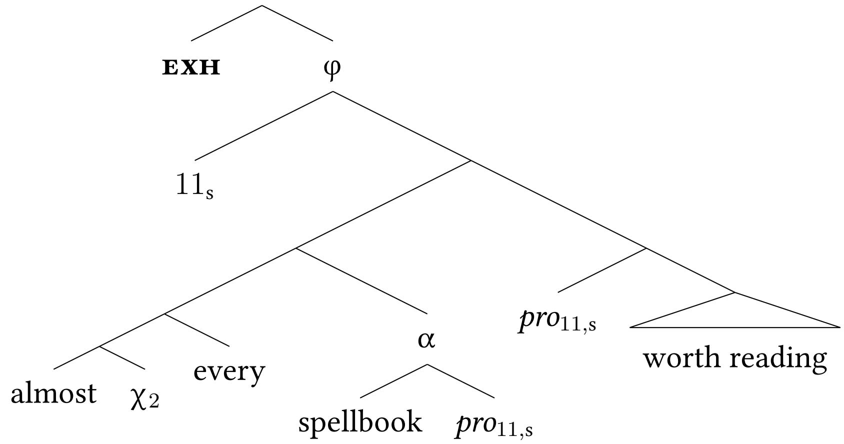
The final reason is perhaps the most compelling, although it only applies to *barely*. It is very difficult to see how the negation contributed by *barely* on our analysis would be encoded if *barely* has a low type. This requires access to the determiner, at least at first blush, and hence our higher type meanings would be necessary. It's possible that there is some way to make this work; perhaps low-type *barely* cooccurs not only with **EXH**, but with a negation operator that does modify the determiner itself. Should such a low-type analysis be made to work, for *barely* and the others, it should be compared against the theory developed here. At this point, though, we'll assume the high-type semantics for all operators.

2.2.2.2 Alternatives and entailment

Let's put some pieces together, starting with *almost*; note that the label for **EXH**'s sister is φ. Worlds are taken to be pronouns at LF, written as *pro* with an index and a type, and they are bound by higher operators (Percus 2000).

(84) Almost every spellbook is worth reading.

a.



b. $[\![(84a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$; where defined:
$[\![(84a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\mathbb{S}_w - g(2) \subseteq \mathbb{WR}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall\psi\left[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\right]$

(84bi) are the truth conditions of $[\![\varphi]\!]^{g,c}(w)$. When g(2) is removed from the set of spellbooks in w, everything in remainder must be worth reading in w; in the context c, g(2) must be considered small. To evaluate (84bii), we need to know what the alternatives to φ are, and which ones are Strawson entailed.

What are the alternatives for sentences containing our subtractive operators? We'll take them to be structural alternatives in the sense of Katzir 2007—that is, they are constructed by replacing the exception χ_n with other expressions χ' of the same type. Crnič (2018) proposes this for *but*, and we propose extending this to both *almost* and *barely*[8,9].

(85) ***SUBTRACTIVE LFs***

A LF φ is a *subtractive LF* just in case φ is a LF containing exactly one subtractive operator Σ introducing an exception χ, and no other alternative-triggering expression, and φ is of type ⟨s,t⟩.

(86) ***SUBTRACTIVE ALTERNATIVES*** (first version)

[8]Following Spector (2014), Crnič proposes treating the set of alternatives to '*almost* φ' to be merely φ. We'll see in §2.3.4 why this cannot be so for *barely*, and in the following chapter why it cannot be so for *almost*.

[9]There certainly can be multiple subtractives in an LF, and other alternative trigger expressions, yet we restrict ourselves to those with only one subtractive and no other alternative triggering expression. The reason for this is simplicity. First, multiple alternative-triggering expressions raises questions about whether you have just a single **EXH** operator, or one for every alternative-triggering expression. Second, regardless of whether we deal with one **EXH** or more, the logic of multiple alternative-triggering expressions is far more complicated and intricate than just a single one, and a discussing such complexities is beyond our scope here. Restricting ourselves to *subtractive LFs* as defined is expository, but it also allows us to make statements about what the alternatives are to a particular subtractive LF. See Gotzner & Romoli (2022) and references therein for discussion of the alternatives of sentences with multiple alternative-triggering expressions.

For a subtractive LF φ, with an exception χ,

$\textbf{ALT}(\varphi) := \{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_e\}$

($\varphi[\chi/\chi']$ is identical to φ except χ is replaced by χ')

The idea is that when we say *Almost every spellbook is worth reading* or *Every spellbook but* The Brekenkrieg Grimoire *is worth reading*, the alternatives under consideration are derived by substituting different expressions χ' of the same type as the actual exception for that exception. It is important for us that the alternative χ's are not expressions whose extensions are dependent on the assignment function for their value. For *but*, they must be overt exceptions, i.e. expressions that pick out different sets of entities. For *almost* and *barely*, they must be covert expressions also picking out different sets of entities. Let χ_n be the 'actual' exception in a subtractive LF. If the alternative χ's were pronouns bearing distinct indices from χ_n, it's entirely plausible that the assignment function wouldn't have a value for those indices, and they would be undefined. We could instead, perhaps, define those alternative χ's in terms of different values for the index on χ_n; as far as I can tell, this would be equivalent to what we propose[10,11].

We might not want to let *all* expressions denoting sets of individuals into the fray; we consider requiring that the alternatives be *grammatical*. That is, since *almost* doesn't permit overt exceptions, so for *Almost every spellbook is worth reading*, if we restrict ourselves to grammat-

[10]This general idea of subtractive exception 'substitution' could also be treated in a perhaps more traditional way, in terms focus and a structured meaning analysis of focus alternatives (Rooth 1992, e.g.) Such an approach has been pursued before for *but* by Gajewski (2008), who takes *but* to introduce focus alternatives to its complement. With *almost*, it's clear that focus plays some role. (i), with focus on direct object, conveys that he sprinted throughout the majority of the race, and then didn't towards the very end; (ii), with focus on the verb, suggests the manner in which he ran the whole race was quite like a sprint, but wasn't, in fact.

i. Oz almost sprinted [THE WHOLE RACE$_F$].

ii. Oz almost [SPRINTED$_F$] the whole race.

These cases are outside the purview of this chapter, but they motivate more generally a better incorporation of focus and its semantics in a theory of *almost* (and *barely*). We'll set them aside for now, as well as a structured-meaning approach. As far as I can tell, the essential contributions of the theory could be couched in terms of such an analysis without much change to the rest of the theory.

[11]Some facets of this proposal require clarification. First, Katzir takes alternatives ψ to some φ to be formed, in part, by substituting expressions in φ for other expressions of the same 'category' from the lexicon of the language of φ. So, we would need to define 'categories' for replacement in such a way that allowed exception pronouns and covert expressions of type $\langle e,t \rangle$ to be in the same category, and thus that the latter can be substituted for the former. We also need to grant that there *are* such covert expressions χ as part of the lexicon. Now, these questions apply not only here, but for a structural alternatives theory when overt pronouns are the loci of alternative construction. Marty (2017) is the only discussion of which I'm aware that discusses this particular question; a fuller investigation must be left for future work.

Second, the alternatives to *but*-sentences in particular would be much more restricted than discussed here, since on Katzir's theory, alternatives ψ to some φ cannot be more structurally complex than φ. This means, more or less, that alternative exceptions can be no bigger, structurally speaking, than the actual exception. These issues are important, and a more complete theory invoking structural alternatives would need to address them, but we will make the simpler assumptions above throughout the rest of the dissertation.

ical alternatives, then all alternatives will be formed with a covert expression. Similarly, *Every spellbook but* The Brekenkrieg Grimoire *is worth reading* couldn't have as an alternative *Every spellbook but smokes is worth reading*, since it isn't a grammatical string. We'll adopt this assumption.

(87) ***SUBTRACTIVE ALTERNATIVES*** (final version)
For a subtractive LF φ, with an exception χ,
$\textsc{alt}(\varphi) := \{\varphi[\chi/\chi'] \mid \mathscr{X}' \subseteq D_e \;\&\; \varphi[\chi/\chi'] \text{ is grammatical}\}$

Now, since all we're varying across alternatives is the exception, the question of which alternatives some subtractive φ containing an exception χ Strawson entails reduces to whether, for some χ′, $[\![\varphi]\!]^{g,c}$ Strawson entails $[\![\varphi[\chi/\chi']]\!]^{g,c}$. We can thus rewrite the contribution of **EXH** a little bit to suit our purposes.

(88) For a subtractive LF φ,
$[\![\textbf{EXH}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1 \;\&\; \forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}(w) \neq 1\right]$

This will be how we write the contribution of **EXH** generally in what follows.

The determiners that our subtractives can modify are important to the determination of Strawson entailed alternatives. Quantificational determiners generally give rise to particular inference patterns as a matter of their semantics. For both *every* and *no*, if (89a) is true, then (89b) must also be true, but from (89a) we cannot conclude that (89c) is true.

(89) a. {Every/no} spellbook is worth reading.
b. {Every/no} long spellbook is worth reading.
c. {Every/no} book is worth reading.

These inferences arise because of the relationship between the three sets (the set of long spellbooks is a subset of the set of spellbooks, in turn a subset of the set of books), and because *every* denotes a function with the following property (Barwise & Cooper 1981; Keenan & Stavi 1986).

(90) A determiner meaning $\mathscr{D}$ is LEFT DOWNWARD ENTAILING *iff* for any $\mathbb{A}$, $\mathbb{A}'$, and $\mathbb{B}$:
If $\mathscr{D}(\mathbb{A})(\mathbb{B}) = 1 \;\&\; \mathbb{A}' \subseteq \mathbb{A}$, then $\mathscr{D}(\mathbb{A}')(\mathbb{B}) = 1$.

Neither determiner is Left Upward Entailing:

(91) A determiner meaning $\mathscr{D}$ is LEFT UPWARD ENTAILING *iff* for any $\mathbb{A}$, $\mathbb{A}'$, and $\mathbb{B}$:
If $\mathscr{D}(\mathbb{A})(\mathbb{B}) = 1 \;\&\; \mathbb{A} \subseteq \mathbb{A}'$, then $\mathscr{D}(\mathbb{A}')(\mathbb{B}) = 1$.

Some, on the other hand is Left Upward Entailing, but not Left Downward Entailing. If (92a) is true, then (92c) must also be true, but from (92a) we cannot conclude that (92b) is true.

(92) a. Some spellbook is worth reading.

b. Some long spellbook is worth reading.

c. Some book is worth reading.

Other operators can affect these monotonic inferences: negation, for example, itself is Downward Entailing, and inverts the derived inferences.

(93) a. It's not the case that $\{$every/no$\}$ spellbook is worth reading.

b. It's not the case that $\{$every/no$\}$ long spellbook is worth reading.

c. It's not the case that $\{$every/no$\}$ book is worth reading.

(94) a. It's not the case that $\{$some$\}$ spellbook is worth reading.

b. It's not the case that $\{$some$\}$ long spellbook is worth reading.

c. It's not the case that $\{$some$\}$ book is worth reading.

If (93a) is true, then (93c) is true, but not necessarily (93b); if (94a) is true, then (94b) is true, but not necessarily (94c). The monotonic inferences flip. Whereas the whole of (89a) is Downward Entailing with respect to *spellbook*, in (93a), it is Upward Entailing; we can state a more general, syntactic notion of entailment that allows us to capture the interplay of operators like *negation* and *every* on monotonic inferences in terms of substitution.

(95) ***CONJOINABLE TYPES***

a. t is a conjoinable type.

b. If τ is a conjoinable type, then for all types σ, $\langle\sigma,\tau\rangle$ is a conjoinable type.

(96) ***CROSS-CATEGORIAL ENTAILMENT*** '$\Rightarrow$'

a. For p, q of type t: $p \Rightarrow q$ *iff* $p = 1 \rightarrow q = 1$.

b. For f, g of conjoinable type $\langle\sigma,\tau\rangle$: $f \Rightarrow g$ *iff* for every x of type σ, $f(x) \Rightarrow g(x)$.

(97) ***CROSS-CATEGORIAL STRAWSON ENTAILMENT*** '$\Rightarrow_{ST}$'

a. For p, q of type t: $p \Rightarrow_{ST} q$ *iff* $[p = 1 \;\&\; q \neq \#] \rightarrow q = 1$.

b. For f, g of conjoinable type $\langle\sigma,\tau\rangle$: $f \Rightarrow_{ST} g$ *iff*
for every x of type σ, $f(x) \Rightarrow_{ST} g(x)$.

(98) ***SUBSTITUTION-BASED STATEMENT OF (STRAWSON) ENTAILMENT*** (after Gajewski 2011)
For a constituent φ and some subconstituent α:

a. φ is (Strawson) Downward Entailing with respect to α *iff*
$\forall\beta\{[\![\beta]\!]^{g,c} \Rightarrow_{(ST)} [\![\alpha]\!]^{g,c} \rightarrow [\![\varphi]\!]^{g,c} \Rightarrow_{(ST)} [\![\varphi[\alpha/\beta]]\!]^{g,c}\}$

b. φ is (Strawson) Upward Entailing with respect to α *iff*
$\forall\beta\{[\![\alpha]\!]^{g,c} \Rightarrow_{(ST)} [\![\beta]\!]^{g,c} \rightarrow [\![\varphi]\!]^{g,c} \Rightarrow_{(ST)} [\![\varphi[\alpha/\beta]]\!]^{g,c}\}$

Set subtraction is also Downward Entailing with respect to the subtrahend, the exception; a formal proof can be found in the appendix, but this is also intuitive.

(99) ***SUBTRACTION IS DOWNWARD ENTAILING WITH RESPECT TO THE SUBTRAHEND***
Let $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{A}$ be arbitrary sets; $\mathbb{X} \subseteq \mathbb{Y} \rightarrow \mathbb{A}-\mathbb{Y} \subseteq \mathbb{A}-\mathbb{X}$.

Taking these facts all together, let's consider the following pair.

(100) a. **EXH** $[_{\varphi}$ every spellbook but BG and *The Writings of Dramius* is worth reading$]$
b. # **EXH** $[_{\varphi}$ some spellbook but BG and *The Writings of Dramius* is worth reading$]$

Given how we've defined *but*, alternatives to φ in (100a) formed by replacing $\{$BG, WoD$\}$ with a small superset will be entailed. This is because the largest constituent φ within the scope of **EXH** ('the whole shebang') is Downward Entailing with respect to the restrictor of *every*, and subtraction is Downward Entailing on the exception—within the scope of **EXH**, φ is *Upward* Entailing with respect to the exception. Now, for (100b), where the whole shebang is Downward Entailing with respect to the restrictor of *some*, because that determiner is Upward Entailing with respect to its restrictor, alternatives formed with *subsets* of $\{$BG, WoD$\}$ will be entailed. φ, in this case, is *Downward* Entailing with respect to the exception. These facts are generalize, and we can state some relevant theorems below.

(101) ***SUBTRACTIVE ENTAILMENT THEOREMS FOR DETERMINERS***
Let Σ be a subtractive operator, Δ a quantificational determiner, and α and χ be arbitrary expressions of type $\langle e,t\rangle$. Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$. Let g be an assignment function in context c.

a. If φ is Upward Entailing (and not Downward Entailing) with respect to the constituent α, then for all χ' such that $\mathbb{X}' \subseteq \mathbb{X}$, then $[\![\varphi]\!]^{g,c} \Rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}$.
b. If φ is Downward Entailing (and not Upward Entailing) with respect to the constituent α, then for all χ' such that $\mathbb{X} \subseteq \mathbb{X}'$ & $\mathbb{X}'$ is small in c, then $[\![\varphi]\!]^{g,c} \Rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}$.

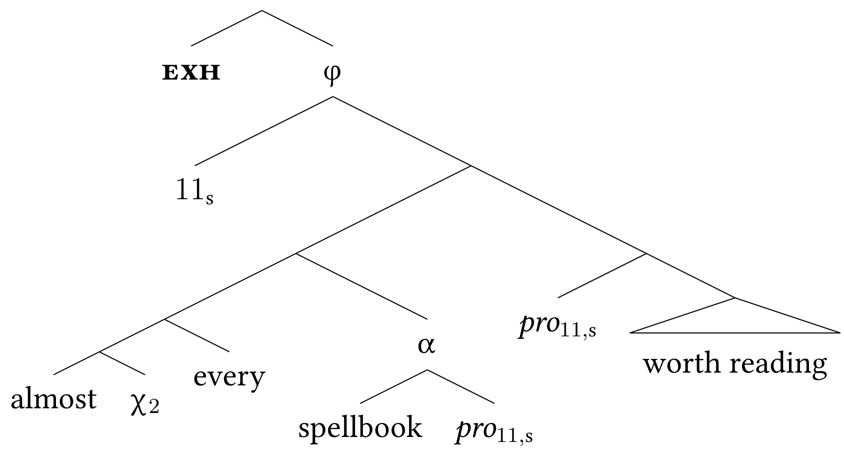
These theorems are proven in the appendix more thoroughly, and we will utilize them in the remainder of the dissertation. However, it's important make a few comments on these theorems. First, note that they are not exhaustive: for a given quantificational determiner, there very well could be additional entailments, so we'll need to consider those alternatives not already covered by the theorems when necessary. Second, the size constraint plays an important role where supersets are entailed. For our subtractive operators, which require their exception to be small, it's not just any old supersets that can be entailed—it's *small* supersets[12]. This will come up later. Finally, note that for any φ, ψ, if φ strictly entails ψ, φ certainly Strawson entails ψ. If φ doesn't Strawson entail ψ, then φ doesn't strictly entail φ, either. So, the entailment theorems above, defined in terms of strict entailment, imply Strawson entailment.

2.2.2.3 Some assembly required

Now let's put the pieces together.

(102) Almost every spellbook is worth reading.

a.



b. $\llbracket(102\text{a})\rrbracket^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$; where defined:

$\llbracket(102\text{a})\rrbracket^{g,c}(w) = 1$ *iff* (i) & (ii) = 1

i. $\mathbb{S}_w - g(2) \subseteq \mathbb{WR}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi'\big[\llbracket\varphi\rrbracket^{g,c} \not\Rightarrow_{\text{ST}} \llbracket\varphi[\chi_2/\chi']\rrbracket^{g,c} \rightarrow \llbracket\varphi[\chi/\chi']\rrbracket^{g,c}(w) \neq 1\big]$

(102bi) is true just in case g(2) is small in c and if every spellbook in w, less g(2), is worth reading in w. Now we need to know which alternatives are entailed. Since φ is Downward Entailing with respect to α, the restrictor of *every*, we know that alternatives that are formed with small supersets of g(2) will be entailed, and therefore Strawson entailed, given the theorems in (101). Are alternatives formed with any other χ' Strawson entailed?

Suppose that being a spellbook entails being worth reading, i.e. in all worlds w′, $\mathbb{S}_{w'} \subseteq \mathbb{WR}_{w'}$. In such a situation, $\llbracket\varphi\rrbracket^{g,c}$ entails $\llbracket\textbf{EXH}\ \varphi\rrbracket^{g,c}$. Here's why. $\llbracket\varphi\rrbracket^{g,c}(w)$, i.e. (102bi), is true just in case g(2) is small. (102bii) is true necessarily, since for all $\mathbb{X}'$, if $\mathbb{X}'$ is small, then the antecedent is false, and if it's not small, the consequent is true. $\llbracket\textbf{EXH}\ \varphi\rrbracket^{g,c}$ is thus entailed by $\llbracket\varphi\rrbracket^{g,c}(w)$, and so by the Economy Constraint, **EXH** is not licensed, and by *almost* is not licensed, either! So, *almost* cannot be licensed when modifying a necessarily true quantificational claim[13]. Let's make the following assumption:

[12]Of course, a subset of any small subset is itself small.

[13]Here's another, perhaps more intuitive case that reflects the kind of situation we're describing:

(103) ***TRUTH VALUES FOR UNLICENSED SUBTRACTIVES***

For all worlds w, if a subtractive operator present in some φ is not licensed in φ, then $[\![\varphi]\!]^{g,c}(w) = \#$.

Ideally, we'd have a compositional system that allows us to derive this result. At this point it's just stipulative, but it's meant to reflect that an LF that is ill-formed cannot receive a truth value.

Now let's assume, as is perhaps reasonable, that being a spellbook doesn't entail being worth reading. So, there is a world w′ where $\mathbb{S}_{w'} \not\subseteq \mathbb{WR}_{w'}$. Alternatives formed with χ' such that $g(2) \subseteq \mathbb{X}'$ and which are small in c are entailed, given our entailment theorem; any χ' such that $\mathbb{X}'$ isn't small in c isn't entailed, but it will not falsify (102bii). All $\mathbb{X}'$ that have an empty intersection with $\mathbb{S}_w$ also will not falsify (102bii).

What about small non-supersets $\mathbb{X}'$—are they entailed? No. To show this, given a context c, we need to find a world w′ such that $\mathbb{S}_{w'} - g(2) \subseteq \mathbb{WR}_{w'}$ & $\textsc{small}_c(g(2))$ is true, but for an arbitrary χ' such that $\mathbb{X} \not\subseteq \mathbb{X}'$ & $\textsc{small}_c(\mathbb{X}')$, $\mathbb{S}_{w'} - \mathbb{X}' \subseteq \mathbb{WR}_{w'}$ is not. A world w* where $\mathbb{X} = \mathbb{S}_{w*} \cap \overline{\mathbb{WR}_{w*}}$ is such a world. Since every $\mathbb{X}'$ such that $\mathbb{X} \not\subseteq \mathbb{X}'$ lacks an element in $\mathbb{X}$, call it *a*, $\mathbb{S}_{w*} - \mathbb{X}'$ contains *a*, but since $a \notin \mathbb{WR}_{w*}$, $\mathbb{S}_{w*} - \mathbb{X}' \not\subseteq \mathbb{WR}_{w*}$. Non-supersets are *not* entailed (as with other proofs, a fuller version can be found in the appendix). We can rewrite the truth conditions in (102b) as in (102b′).

(102b′) $[\![(102)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$; where defined:

$[\![(102)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i $\mathbb{S}_w - g(2) \subseteq \mathbb{WR}_w$ & $\textsc{small}_c(g(2))$

ii $\forall \mathbb{X}'\big[\big[g(2) \not\subseteq \mathbb{X}'$ & $\textsc{small}_c(\mathbb{X}')$ & $\mathbb{X}' \subseteq \mathbb{S}_w\big] \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \mathbb{WR}_w\big]$

These are just the truth conditions proposed in von Fintel (1993), with the additional contribution of the size constraint and the subset presupposition. (i) & (ii) thus both hold *iff* the exception, g(2), contains all and only the spellbooks not worth reading in w, and is small in the context of utterance.

First, assume everything in g(2) is a spellbook; (102b′) is undefined otherwise. (i), of course, requires that every other spellbook be worth reading, and that g(2) is a small-enough exception in the context. This alone doesn't require that everything in g(2) is not worth reading, but (ii) does. For each small $\mathbb{X}'$ that isn't a superset of g(2), $\mathbb{S}_w - \mathbb{X}'$ contains at least one element in g(2), and for (ii) to be true, all such $\mathbb{X}'$ must leave in $\mathbb{S}_w$ some spellbook not worth reading. Since (i) requires that every spellbook not in g(2) be worth reading, the unworthy spellbook must be in

(1) # Every novel but *Anna Karenina* is a book.

In all worlds, every novel is, of course, a book. Subtraction here contributes nothing; the logic described above thus applies. See the appendix for a fuller proof.

g(2). (i) & (ii) thus require that not every spellbook is worth reading (the POLAR inference) but that the small size of the exception means we're not far from the truth of *every spellbook is worth reading* (the PROXIMAL inference).

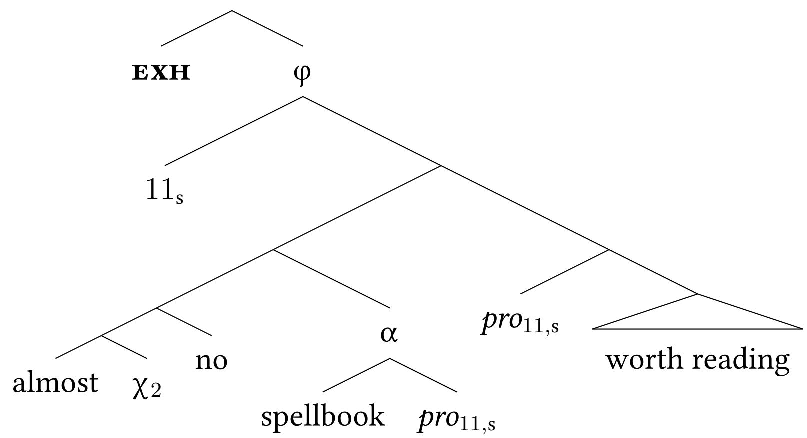
Note that we predict falsity if every spellbook is in fact worth reading in w, even though being a spellbook doesn't entail being worth reading. Assume that every spellbook is worth reading in w: (102b′i) is true, but then all small $\mathbb{X}'$ that don't contain everything in g(2) make (102b′ii) false—they would need to add elements to $\mathbb{S}_w$ to satisfy the consequent, but that's simply not in subtraction's power to do. So, if (102b′i) is true in w because every spellbook is worth reading, (102b′ii) is false, and therefore their conjunction is false.

Now let's try *no*.

(104) $[\![\text{no}]\!] = \left[\lambda A \lambda B.\ \mathbb{A} \cap \mathbb{B} = \emptyset\right] = \left[\lambda A \lambda B.\ \mathbb{A} \subseteq \overline{\mathbb{B}}\right]$

(105) Almost no spellbook is worth reading.

a.



b. $[\![(105\text{a})]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$; where defined:

$[\![(105\text{a})]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\mathbb{S}_w - g(2) \cap \mathbb{WR}_w = \emptyset$ & $\textsc{small}_c(g(2))$

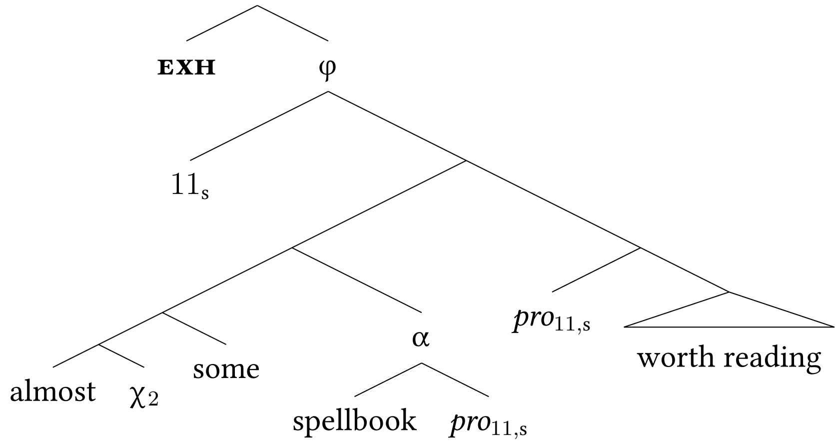
ii. $\forall \chi' \left[[\![\varphi]\!]^{g,c} \Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

For arbitrary α, β, and χ, $[\![\left[\left[\left[[\text{almost}\ \chi]\ \text{no}\right]\ \alpha\right]\ \beta\right]]\!]^{g,c} = [\![\left[\left[\left[[\text{almost}\ \chi]\ \text{every}\right]\ \alpha\right]\ [\text{not}\ \beta]\right]]\!]^{g,c}$. So, the logic for *almost every* applies here; we won't rehash it.

Exceptive *but* doesn't modify existential *some*, as we've seen. What does our analysis predict for *almost*?

(106) # Almost some spellbook is worth reading.

a.



b. $[\![(106a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$; where defined:

$[\![(106a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\mathbb{S}_w - g(2) \cap \mathbb{WR}_w \neq \emptyset$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \big]$

As we might expect, if being a spellbook entails being worth reading, (106bi) is true just in case g(2) is small, i.e. it contains less than half the spellbooks. Grant that this is the case. Every small $\mathbb{X}'$ is entailed, then, and any big $\mathbb{X}'$ satisfies (106bii). The truth of (106bi) guarantees the truth of (106bii) if being a spellbook entails being worth reading, and so **exh** is otiose, and isn't licensed. Therefore, *almost* isn't licensed here.

Setting aside such an entailment between spellbooks and worthwhile reading, then, (106bi) is true in worlds where some non-g(2) spellbook is worth reading, as long as g(2) is small in c. Sets $\mathbb{X}'$ that are too big aren't entailed, but that's just fine—they'll satisfy (106bii). φ is Upward Entailing with respect to α, since *some* is Upward Entailing on its restrictor, and given the entailment theorems, those χ' denoting subsets of g(2) will yield entailed alternatives. What about small, non-subsets of g(2)? These include proper supersets and wholly disjoint sets[14]. For such sets $\mathbb{X}'$, all we need to do is construct a world w′ where $\mathbb{S}_{w'} - g(2) \cap \mathbb{WR}_{w'} \neq \emptyset$ and $\mathbb{X}' \supseteq \mathbb{S}_{w'} \cap \mathbb{WR}_{w'}$ are both true in order to show that such $\mathbb{X}'$ are not entailed; see the appendix for proofs of their existence. Only subsets of g(2) are entailed, then. The truth conditions above are equivalent to the following:

(106b′) $[\![(106a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$; where defined:

$[\![(106a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\mathbb{S}_w - g(2) \cap \mathbb{WR}_w \neq \emptyset$ & $\textsc{small}_c(g(2))$

ii. $\forall \mathbb{X}' \big[\big[\mathbb{X}' \not\subseteq g(2)$ & $\textsc{small}_c(\mathbb{X}')$ & $\mathbb{X}' \subseteq \mathbb{S}_w \big] \rightarrow \mathbb{S}_w - \mathbb{X}' \cap \mathbb{WR}_w = \emptyset \big]$

[14]If the exception $\mathbb{X}$ is larger than a singleton, such sets would also include those that merely overlap $\mathbb{X}$.

These truth conditions can never be true, provided that g(2) is non-empty. We'll return to this non-emptiness assumption momentarily, but make it now.

Now, suppose that in w, every spellbook is worth reading. This is, of course, distinct from the case of spellbookhood entailing worthiness, but a reader's suspicion that this won't turn out well for the truth conditions would be justified. In such a situation, (106b′i) will be true just in case g(2) is small. This means that there must be at least three spellbooks, given the felicity conditions imposed by the size constraint. Since only subsets of g(2) are entailed, subtraction of any small $\mathbb{X}'$ that does not contain an element in g(2), will make (106b′ii) false. Call one of the other spellbooks **a**: $\mathbb{S}_w - \{a\}$ contains an element in g(2), by assumption a member of $\mathbb{WR}_w$. We'll predict falsity, then, if every spellbook is worth reading, in w.

Let's grant, then, that not every spellbook is worth reading in w. Given (106b′i), $\mathbb{S}_w$ must include one spellbook that is in $\mathbb{WR}$ and is not in g(2). Call it **a**. The size constraint means that $\mathbb{S}_w$ must include at least two spellbooks not in g(2); that means there is at least one other spellbook in $\mathbb{S}_w$ not in g(2), call it *b*. Let $\mathbb{X}' = \{b\}$: this is not a subset of g(2), and so satisfies the antecedent of (106b′ii), but that alternative is false, since $\mathbb{S}_w - \mathbb{X}'$ includes **a**, a spellbook worth reading.

It is important to note that the size constraint does crucial work here in restricting us to worlds where there are least three spellbooks. Suppose now that the size constraint didn't exist; we actually predict these truth conditions to be coherent if there are exactly two spellbooks, one of which is worth reading, as Gajewski (2013) noticed[15]. In such a world w*, where g(2) contains the unworthy read, (106b′i) is true, and there are two alternatives that are subsets of $\mathbb{S}_{w*}$: the set containing the worthy read, and the empty set. The latter is entailed, and the former is not, but it does not falsify (106b′ii), since it removes the only spellbook worth reading from $\mathbb{S}_{w*}$. So we'd predict truth without the size constraint, and hence that this should be acceptable. Note that this issue arises from the fact that **EXH** negates non-entailed alternatives. The monotonicity of the determiner crucially changes the entailed alternatives, and hence which alternatives **EXH** requires are not true—here, they are non-subsets. The Fintelian semantics is static: his *but*, and an analogous *almost*, always quantifies over non-supersets of the exception, so the empty set always causes falsity for existential quantificational determiners modified by subtractives, even in when the restrictor has a domain of two. It's a function of the move to **EXH** in the end, so it's good that we already have independent motivation for the size constraint's role.

Now we need to address that non-emptiness assumption. It turns out that it is crucial. Suppose that g(2) is empty. If there is exactly one spellbook, and it's worth reading, we can expect truth here, in spite of the small domain. The empty set has no cardinality, so plausibly its size is 0. That is small in any context given our definitions.

(107) ***THE SIZE CONSTRAINT***

[15] Gajewski's solution is different, and discussed in §2.2.3.3.

Let c be a context of utterance, μ_c be a measure function, determined by the context, mapping sets to numerical values representing their size, and n_c be a contextually determined numerical threshold for size.

$\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\mu_c(\mathbb{X}) \leq n_c$.

(108) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

For any context c, an utterance of φ in c is felicitous only if $n_c < \frac{1}{2}(\mu_c(\mathbb{A}))$

This means, though, that there is only one alternative $\mathbb{X}'$ that is a subset of $\mathbb{S}_w$ and which satisfies the antecedent in (106b′ii), and because it is not Strawson-entailed, its subtraction must remove all worthy reads from the set of spellbooks, which it does, of course. We predict truth, unfortunately[16]!

Here's how we'll rectify this. First note that some requirement like this is probably natural anyways—to utter *Almost everyone is here* is odd if you know that everyone is, in fact, here, and it can naturally elicit a response of *Who isn't here?*, which suggests that a hearer should infer the existence of exceptions from the original statement. If we were more careful, we would take time to investigate the status of this requirement further, but for now, we'll simply take it to be a presupposition on our subtractive operators. It would have been nice to derive this, instead of stipulate it, but here we are.

(109) $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B$: $\mathbb{X} \subseteq \mathbb{A}$ & $\underbrace{\mathbb{X} \neq \emptyset}_{\textsc{new!}}$. $\mathscr{D}(\mathbb{A} - \mathbb{X})(\mathbb{B})$ & $\textsc{small}_c(\mathbb{X})$

Barely and *but* will change accordingly. Now, this changes the logic of our analysis in two ways. First, obviously, we don't need to reason about empty exceptions anymore. That saves us some work! Second, we no longer can point to the empty set as the alternative exception that gets us the polar inference for *almost every* α β and *almost no* α β, for example. On either option, it's hard-wired in that such an exception yields falsity or undefinedness, and we don't derive that that's the case. Instead, then, the analysis generates the polar inference even more indirectly—the negation of all non-entailed alternatives, conjoined together, requires that the polar inference holds.

We gave *almost some* the best shot, but on this analysis, we cannot generate truth. Either *almost* is not licensed, in the event that being a spellbook entails being worth reading, or it yields

[16] This is, as far as I can tell, only an issue for *almost some/some...but*, and only where there is exactly one element in the restrictor *and* the exception is empty, but it's an irksome problem nonetheless, and the solutions will extend should it arise with other determiners.

falsity across the board. The latter occurs precisely because of the way the monotonicity of *some* affects entailments. With *every* and *no*, the analysis requires that the exception $\mathbb{X}$ must be the smallest exception because (small) supersets $\mathbb{X}'$ of that exception are entailed, and because **EXH** requires that all non-supersets, must yield falsity when swapped out for the exception. In particular, substituting subsets of $\mathbb{X}$—smaller sets—must make the subtractive sentence false. With *some*, we can find no such smallest exception. Entailments are reversed, so subsets of $\mathbb{X}$ are entailed, but then there are a good too many small, non-entailed $\mathbb{X}'$ that work just as well as $\mathbb{X}$, but which muck up **EXH**. For *almost some*, there is no smallest non-empty exception, and there couldn't be any.

Let's turn to *barely*. Its distribution is complementary to *almost* and *but* with respect to quantificational determiners. It doesn't modify *every* or *no*, and while it doesn't modify *some*, it does modify NPI *any*, an existential quantifier (Kadmon & Landman 1993, a.o.). This distribution, as we'll see, falls naturally from the idea that *barely* is a subtractive operator just like *almost* and *but*. The central difference is that *barely* further contributes negation (note the addition of the non-emptiness presupposition).

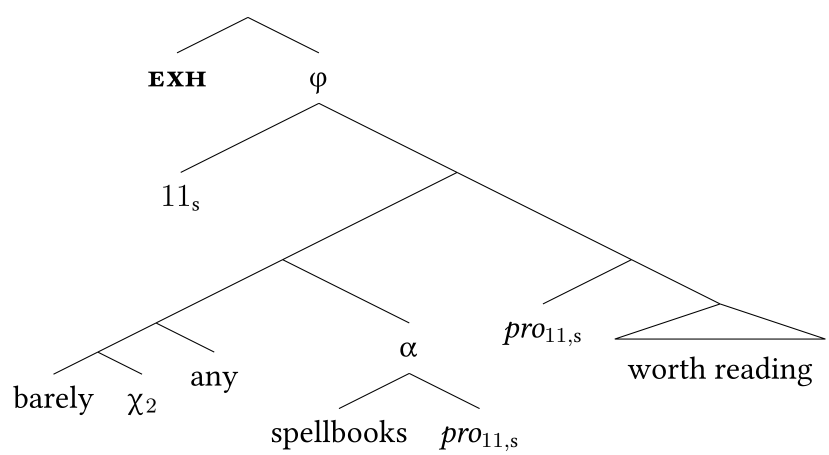
(110) $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B \text{: } \mathbb{X} \subseteq \mathbb{A} \;\&\; \underbrace{\mathbb{X} \neq \emptyset}_{\text{NEW!}} .\; \neg\mathscr{D}(\mathbb{A} - \mathbb{X})(\mathbb{B}) \;\&\; \text{SMALL}_c(\mathbb{X})$

Barely, on this view, modifies a quantificational claim by encoding that the small exception is all that prevents the quantification from being false. Intuitively, (112) is true just in case some spellbook, maybe a few, are worth reading, but if we set aside those, no spellbook is[17]. Putting the pieces together, we'll derive the truth conditions in (112b).

(111) $[\![\text{any}_{\text{NPI}}]\!] = \lambda A_{\langle e,t\rangle} \lambda B_{\langle e,t\rangle} .\; \mathbb{A} \cap \mathbb{B} \neq \emptyset$

(112) Barely any spellbooks are worth reading.

a.



[17]Plural morphology seems to be necessary for some speakers, but those that admit both singular and plural have not reported any intuitive difference; we'll ignore it for now, but will return to it when we discuss numerals in Chapter 4.

b. $[\![(112\text{a})]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$ & $g(2) \neq \emptyset$; where defined:
$[\![(112\text{a})]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

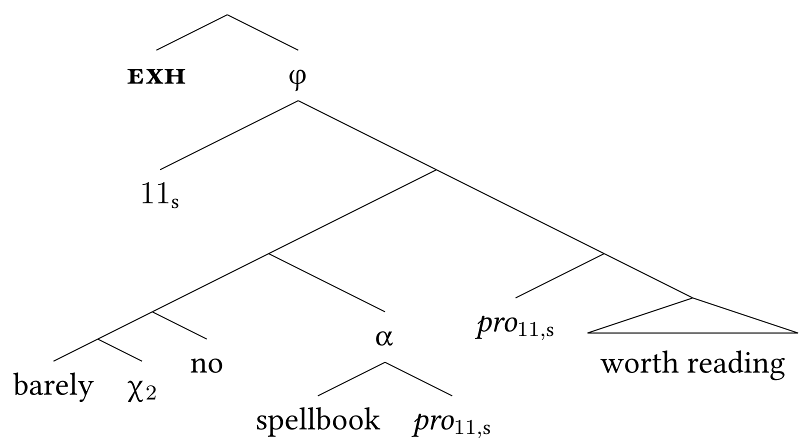
i. $\mathbb{S}_w - g(2) \cap \mathbb{WR}_w = \emptyset$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \big[[\![\varphi]\!]^{g,c} \Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \big]$

These truth conditions are equivalent to those we derived for *Almost no spellbook is worth reading*! Because *barely* contributes negation, when it modifies an existential quantifier, the result is equivalent to *almost no*. This equivalence makes sense intuitively. (105) and (112) might have different pragmatic effects, or contexts where one is more felicitous than the other, but more or less they seem to convey the same thing: precious few spellbooks are worth reading. The logic of the truth conditions follows that of *almost no*, which, as we discussed above, follow the logic of *almost every*. As the reader might reasonably thus expect, if *barely any* is semantically equivalent to *almost no*, then *barely no* is equivalent to *almost some*. This is borne out.

(113) # Barely no spellbook is worth reading.

a.



b. $[\![(113\text{a})]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$ & $g(2) \neq \emptyset$; where defined:
$[\![(113\text{a})]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\mathbb{S}_w - g(2) \cap \mathbb{WR}_w \neq \emptyset$ & $\textsc{small}_c(g(2))$

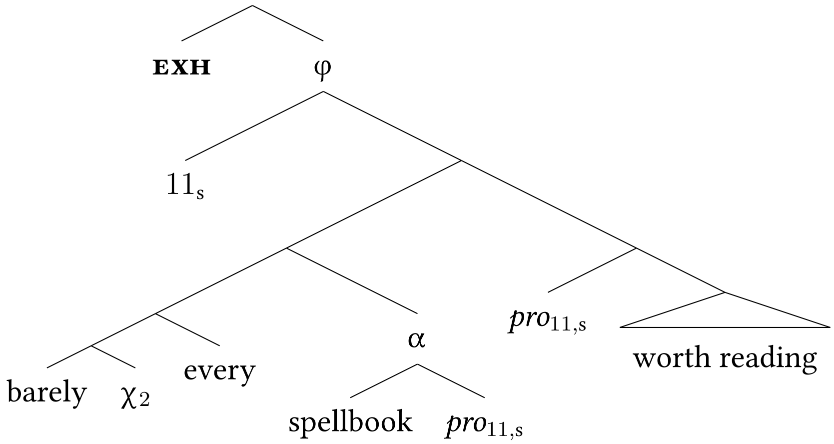
ii. $\forall \mathbb{X}' \big[\big[\mathbb{X}' \not\subseteq g(2)$ & $\textsc{small}_c(\mathbb{X}')$ & $\mathbb{X}' \subseteq \mathbb{S}_w \big] \rightarrow \mathbb{S}_w - \mathbb{X}' \cap \mathbb{WR}_w = \emptyset \big]$

Barely's negation reverses the monotonicity of *no*, and α is now in an Upward Entailing environment with respect to φ. As with *almost some*, then, these truth conditions are contradictory, and necessarily so, but we leave it to the reader to verify this.

This analysis has no problem predicting the across-the-board infelicity of *barely every*, as we should expect.

(114) # Barely every spellbook is worth reading.

a.



b. $[\![(114a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w \& \; g(2) \neq \emptyset$; where defined:
$[\![(114a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) = 1

i. $\mathbb{S}_w - g(2) \nsubseteq \mathbb{WR}_w \; \& \; \textsc{small}_c(g(2))$

ii. $\forall \mathbb{X}' \big[\big[\mathbb{X}' \nsubseteq g(2) \; \& \; \textsc{small}_c(\mathbb{X}') \; \& \; \mathbb{X}' \subseteq \mathbb{S}_w\big] \rightarrow \mathbb{S}_w - \mathbb{X}' \subseteq \mathbb{WR}_w\big]$

These truth conditions are equivalent to *Almost some spellbook is unworthy of reading*, where $[\![\text{unworthy of reading}]\!]^{g,c}(w) = \overline{\mathbb{WR}_w}$. As such, these truth conditions cannot yield truth.

A key thread running through this discussion is monotonicity. In a subtractive LF φ that is sister to **EXH**, the monotonicity of φ with respect to the restrictor α from which the exception is subtracted constrains which alternatives to φ are entailed, as we've seen. This tells us something essential about where we might expect subtractives to occur—that is, which determiners our subtractives can modify—as a function of entailed alternative exceptions. We can find a smallest, non-empty exception when (small) supersets of an exception are entailed, and this is the case when φ is Downward Entailing with respect to α. This obtains, as we've seen, when *almost* and *but* modify *every* and *no*, and *barely*, entailment-reversing operator that it is, modifies NPI *any*. When subsets of an exception are entailed, i.e. when φ is Upward Entailing with respect to α, we have no luck finding a smallest, non-empty exception. This is the case when *almost* and *but* modify *some*, and *barely* modifies *every* and *no*. The generalization seems to be the following:

(115) A subtractive operator Σ can modify a quantificational determiner Δ in an LF φ only if φ is Downward Entailing with respect to the constituent α such that α is the restrictor of Δ.

Now, this is a stronger generalization than some readers might find warranted—after all, we've ignored non-monotonic determiners, like *exactly one hundred*. Let's take a brief look.

(116) $[\![\text{exactly one hundred}]\!]^{g,c} = \lambda P \lambda Q.\, |\mathbb{P} \cap \mathbb{Q}| = 100$

(117) a. # Exactly one hundred vampires but Spike attacked the high school.

b. Almost exactly one hundred vampires attacked the high school.

These data are, at first glance, perplexing. First, we've claimed that for all intents and purposes, *almost* and *but* are equivalent, so why is there a contrast in their acceptability here? Second, since *exactly one hundred* is non-monotonic, why are we seeing *almost* compatible at all? We won't have a complete answer to this question in the dissertation, but we can give a partial answer. First, note that in a subtractive LF φ, if φ is non-monotonic with respect to the restrictor α of a determiner modified by a subtractive operator, then we can't conclude that any alternative exceptions are entailed prima facie. Now, if the determiner is *exactly one hundred*, it is easy to see that subtraction of a particular exception set $\mathbb{X}$ from the restrictor does not allow us to conclude that any alternative exceptions in particular are entailed. That's not inherently a problem, but rather, just a fact.

Now note that *exactly one hundred* is a complex determiner. There are two parts: *exactly* and the numeral. The entry above obscures that, and treats *exactly one hundred* as an unanalyzable whole. That's probably not right, even without thinking about subtractives. Furthermore, non-monotonicity seems to be a property of compositionally complex determiners—other cases include *some but not all*, *between two and four*, and *an even number of*—though I'm not sure I would defend this in a court of law at this juncture. However, the point I'd like to make is that there is more than enough on the surface to suggest that we need to be more careful when we think about non-monotonic "determiners." The compositional structure of *exactly one hundred* might lend itself to modification by *almost*, but not *but*. We'll discuss numerals in more detail in Chapter 4, and we'll have some cursory thoughts about *exactly* towards the end of the dissertation. The reader of this brief, woefully insufficient discussion of non-monotonicity will be left with a challenge: find a true, mono-morphemic determiner that is non-monotonic, and then see if *almost* or *but* can modify it. We're going to assume this is not possible, and go ahead with the generalization above.

2.2.3 Further details

2.2.3.1 How do we get unacceptability?

So far, we've not actually stated how we get the unacceptability, the # judgement, for *some* α *but* χ β, *almost* χ *some* α β, *barely* χ *no* α β, and *barely* χ *every* α β. As we've seen, they simply cannot be true; either the subtractive operators themselves aren't licensed, in which case we call them undefined, or else they generate falsity necessarily. Necessary falsity alone doesn't guarantee ungrammaticality, though, and Gajewski (2002), aiming to explain precisely why on von Fintel (1993)'s theory of exceptives that *some* α *but* χ β is ungrammatical, proposes that when we abstract away from world-dependent expressions, the LF of such a sentence is false under all possible replacements of α, β, and χ. It is what he calls 'L(ogically)-analytic,' and he proposes the following principle.

(118) ***L-ANALYTICITY*** (first version)

A LF constituent α of type $\langle s,t \rangle$ is L-Analytic *iff* for all worlds w', α's logical skeleton receives the denotation 1 under every variable assignment in w', or it receives $0/\#$.

The logical skeleton of an LF is derived through the following algorithm.

(119) a. Identify the maximal constituents containing no logical items.

b. Replace each such constituent with a distinct variable of the same type.

Finally, the following principle links L-analyticity and ungrammaticality.

(120) ***L-ANALYTICITY IMPLIES UNGRAMMATICALITY***

A sentence is ungrammatical if its Logical Form contains a L-analytic constituent.

This would allow us to more formally connect the fact that the relevant combinations of subtractive operators and quantificational determiners yield ungrammaticality, rather than just simply falsity across the board. However, the size constraint is problematic for an appeal to L-Analyticity. The contextual, world-dependent information we need to know in order to evaluate whether the size constraint is satisfied means that more than just *logical* facts are necessary to determine the infelicity of *Almost some spellbook is worth reading.* It's not necessarily true or necessarily false in virtue of its logical skeleton—it's concrete facts that make it infelicitous. So, while there are perhaps some cases where we can appeal to L-Analyticity, this is not one of them. We'll have to accept, for now, that *Almost some spellbook is worth reading* isn't ungrammatical, but rather, that it simply cannot be true on any assignment or in any context. Perhaps that's enough for now.

2.2.3.2 The covert exception pronoun

In our proposal for *almost* and *barely*, the exception is taken to be a covert pronoun, and in particular, a free pronoun. Free pronouns, like *she* below, require a context that makes a particular referent salient.

(121) She killed the vampire.

One might think, then, that the utterer of (122) must know precisely which students skipped out on the pep rally.

(122) Almost every student at Sunnydale High attended the pep rally.

While it's entirely possible that the speaker of (122) knows who is missing, it's obviously not a requirement. One can imagine looking out onto the bleachers where the students sit and noticing that some pockets aren't completely packed; this is enough to warrant uttering (122). The same goes for *barely*, in exactly the same context.

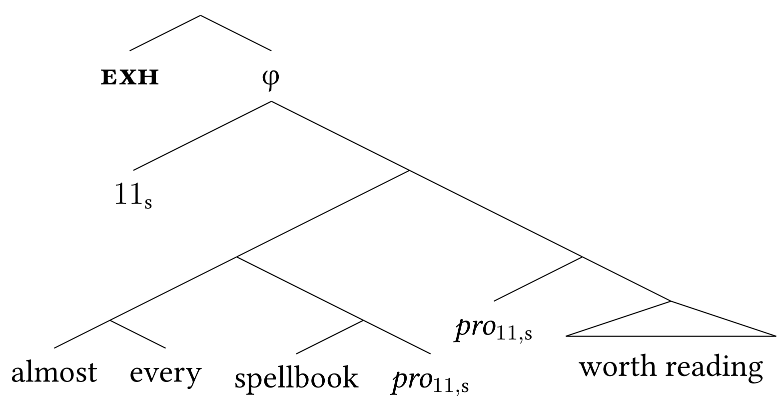
(123) Barely any students at Sunnydale High skipped the pep rally.

These considerations might push us to treat the exception as existentially quantified, rather than pronominal. Should this work, such a proposal would capture the intuition that one needn't know the identity of the exception in an *almost* or *barely* sentence to utter it, though of course one could. In what follows, we'll show that some initially plausible versions of such an alternative are untenable, and where they are, they're functionally equivalent to our proposal. We'll then suggest that, following work in the literature on tense, it's reasonable to assume that at least some kinds of pronouns don't require the precise identity of their extensions to be known in order to use them appropriately.

What does an existential alternative to our proposal look like? Well, *almost* itself could existentially quantify over exceptions, or else it could be closed within the scope of **EXH**. Let's start with the first option[18].

(124) $[\![\text{almost}]\!]^{g,c} = \lambda\mathscr{D}\lambda A\lambda B.\ \exists\mathbb{X}\big[\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})\big]$

(125) Almost every spellbook is worth reading.

a.


b. $[\![(125)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\exists\mathbb{X}\big[\mathbb{S}_w - \mathbb{X} \subseteq \mathbb{WR}_w\ \&\ \textsc{small}_c(\mathbb{X})\big]$

ii. $\forall\psi\big[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\big]$

(125bi) says there is some small set $\mathbb{X}$ exception such that, once it is subtracted from $\mathbb{S}_w$, every remaining spellbook is worth reading. A problem arises, though, with (125bii). It's not clear that

[18]Our presuppositions are suppressed here for simplicity's sake.

we can form alternatives in the way we have so far, since there is no constituent χ that can be replaced if the exception is existentially introduced by *almost* itself. Our alternatives are structural! We defined them in terms of substitution of expressions χ; for some subtractive φ containing χ:

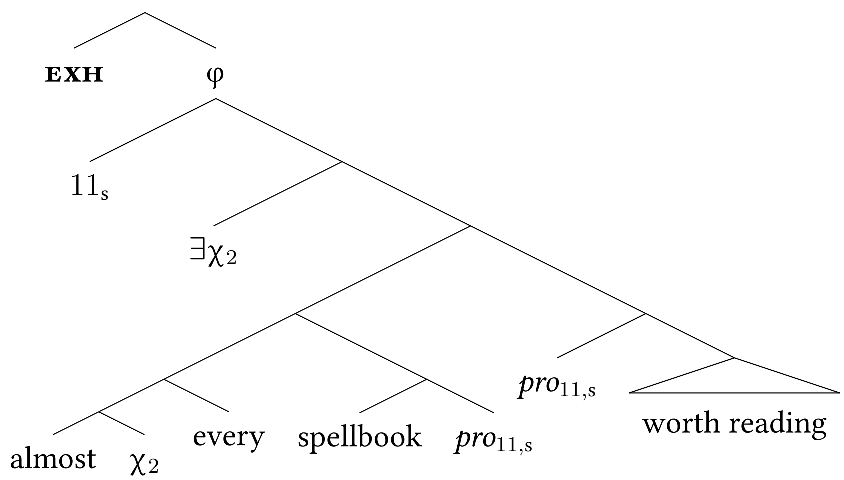
(126) $\textbf{ALT}(\varphi) = \{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_e \;\&\; \varphi[\chi/\chi'] \text{ is grammatical}\}$

($\varphi[\chi/\chi']$ is identical to φ except $\chi$is replaced by χ')

There simply is no constituent we can substitute. If we want to hold onto this formulation for alternatives—we'll see in the next section why we need to—we'll have to do something else.

So what happens if instead, we had existential closure higher, but within **EXH**? We'll revert to our earlier *almost*, and let's assume that the index on $\exists\chi_2$ acts as a binder for χ_2 below it.

(127)



(128) $[\![(127)]\!]^{g,c}(w) = 1$ *iff* (a) & (b) $= 1$

a. $\exists\mathbb{X}[\mathbb{S}_w - \mathbb{X} \subseteq \mathbb{WR}_w \;\&\; \textsc{small}_c(\mathbb{X})]$

b. $\forall\psi[\psi \in \textbf{ALT}(\varphi) \;\&\; [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \to [\![\psi]\!]^{g,c}(w) \neq 1]$

It is, I think, still rather difficult to reconcile (128) with our proposal for subtractive alternatives. The exception χ_2 is ultimately bound by $\exists\chi_2$, but it's still an expression at LF that can be substituted. If it is substituted all on its own, leaving $\exists\chi_2$ to quantify vacuously, then we can use our subtractive alternatives. Essentially:

(129) (128b) $= \forall\chi'[[\![\varphi]\!]^{g,c} \not\subseteq [\![\varphi[\chi_2/\chi']]\!]^{g,c} \to [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1]$

Now the task is to figure out which χ's, substituted for χ_2, yield entailed alternatives. The proposition denoted by φ—given in (130a)—doesn't give us access to anything about $\mathbb{X}$ beyond its existence. The vacuity of $\exists\chi_2$ when we move to the alternatives means that for any χ', the substitution of χ_2 for χ', the proposition denoted by $\varphi[\chi_2/\chi']$ is as in (130b).

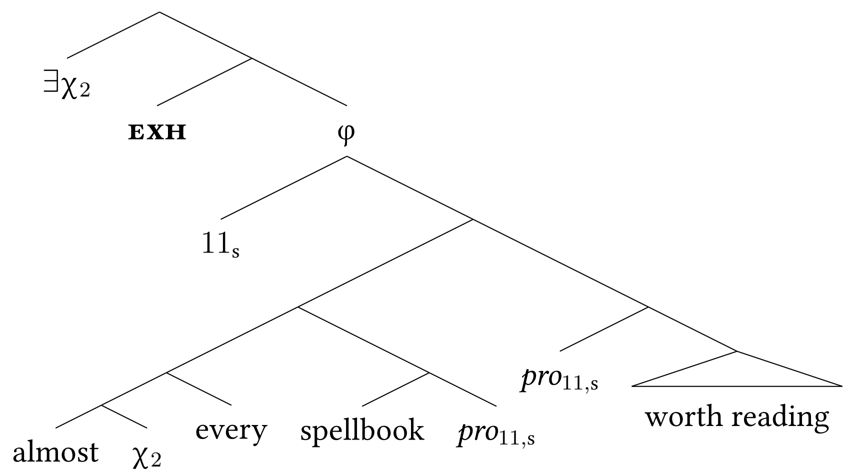
(130) a. $[\![\varphi]\!]^{g,c} = \{w \mid \exists\mathbb{X}[\mathbb{S}_w - \mathbb{X} \subseteq \mathbb{WR}_w \;\&\; \textsc{small}_c(\mathbb{X})]\}$

b. For any χ', $[\![\varphi[\chi_2/\chi']]\!]^{g,c} = \{w \mid \mathbb{S}_w - \mathbb{X}' \subseteq \mathbb{WR}_w \;\&\; \textsc{small}_c(\mathbb{X}')\}$

It seems that in fact for each χ' such that (130b) is true, (130a) is true, but not vice versa. It's actually not clear, though, that we can conclude anything from the existential quantification to the specific χ's—in fact, it's the reverse. Each alternative in (130b) entails (130a). It seems that (130a) doesn't entail any alternatives, and since no alternatives are entailed, we'll have to negate them all. We'll derive a contradiction, then: there must be some $\mathbb{X}$ such that its removal from $\mathbb{S}_w$ yields a subset of $\mathbb{WR}_w$ by (130bi), but no $\mathbb{X}'$ is such that its removal from $\mathbb{S}_w$ yields a subset of $\mathbb{WR}_w$ by (130b). This cannot work.

What does work is existential closure above **exh**; $\mathbb{X}$ in (132a) and (132b) is within the scope of $\exists\mathbb{X}$.

(131)



(132) $[\![(131)]\!]^{g,c}(w) = 1$ *iff* $\exists\mathbb{X}\big[(a) \;\&\; (b) = 1\big]$

a. $\mathbb{S}_w - \mathbb{X} \subseteq \mathbb{WR}_w \;\&\; \textsc{small}_c(\mathbb{X})$

b. $\forall\psi\big[\psi \in \textbf{alt}(\varphi) \;\&\; [\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\big]$

Here our subtractive alternatives work. Within the scope of **exh**, entailment is checked for some particular $\mathbb{X}$, and everything else proceeds as if we merely had the pronoun. So,

(133) $[\![(131)]\!]^{g,c}(w) = 1$ *iff* $\exists\mathbb{X}\big[(a) \;\&\; (b) = 1\big]$

a. $\mathbb{S}_w - \mathbb{X} \subseteq \mathbb{WR}_w \;\&\; \textsc{small}_c(\mathbb{X})$

b. $\forall\mathbb{X}'\big[\mathbb{X} \not\subseteq \mathbb{X}' \;\&\; \textsc{small}_c(\mathbb{X}') \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \mathbb{WR}_w\big]$

These truth conditions, of course, are just what we derived on the pronominal view of the exception, but we've just bound the pronoun outside the scope of **exh**. Not much else seems different.

At the end of the day, what's the difference between choosing between such high existential closure and assuming there's a pronoun, as we have? Really, it doesn't appear that there's much of any. There are parallels in other domains. Consider this:

(134) Buffy killed the vampire.

To utter (134), presumably there is some salient window in which we know she did the deed, but we needn't know precisely when any such an event occurred. Shockingly, there are both *quantificational* and *pronominal* theories of tense that allow us to capture these intuitions[19]. Both take the past tense, encoded by a mopheme PAST, to relate the time of the utterance to some past time at which a predicate held. Here are two entries for such a morpheme, quantificational and pronominal, with a temporally-sensitive meaning of the tenseless *kill the vampire* given as well (let's be vague about which preposition is most appropriate for reading the denotation; if you're really worried about it, assume it's 'at'). The semantic type of times is *i*.

(135) $[\![\textsc{past}_{R_c}]\!]^{g,c} = \lambda A_{\langle i,t\rangle}.\ \exists t' \left[t' < t_c \ \&\ t' \in R_c \ \&\ t' \in \mathbb{A}\right]$ $\langle$it,t$\rangle$

(136) $[\![\textsc{past}_i]\!]^{g,c}$ is defined only if $g(i) < t_c$; where defined:
$[\![\textsc{past}_i]\!]^{g,c} = g(i)$ i

(137) $[\![\text{Buffy kill the vampire}]\!]^{g,c} = \lambda t_i.\ \textsc{kill}(\text{vampire})(\text{Buffy})(t)$

Both proposals require that the time of Buffy's killing precede the time of the context for (134). The quantificational theory requires that some time t' that is before the time of the context (t_c) and within the contextually-determined restricted set of times (R_c) such that $t' \in \mathbb{A}$. The pronominal view restricts times via the presupposition, and given its sensitivity to the (contextually-determined!) assignment function, it is plausibly contextually restricted as well. Does either proposal ultimately require explicit knowledge of the precise time, down to the moment? It would seem not—all that's required on either view is that the speaker know there is some past time when the killing occurred, or some time window in which it did.

Even personal pronouns don't always require that we be able to identify with precision their extension in order to use them[20]. Consider an utterance of the following.

(138) Many vampires attacked Sunnydale High. Now they regret their decision.

We needn't be able to identify the attacking vampires in a lineup in order for the use of *they* to be felicitous. We can, incidentally, identify them with a definite description:

(139) Many vampires attacked Sunnydale High. Now the vampires that attacked the school regret their decision.

It seems reasonable, then, when considering exceptions in *almost* and *barely* sentences, to assume a pronominal view without it requiring utterers to know the exact identity of the pronoun. It

[19]Prior (1962) is an early proponent of the first view; Partee (1973) for the latter. The literature on the semantics of tense is dense, and I won't be able to do it any kind of justice here; see Ogihara & Kusumoto (2020) for an overview.

[20]Thanks to Roger Schwarzschild for this argument.

suffices that the pronoun *does* have an identity, and we can reason well enough based on that assumption. Perhaps this is, in tandem with the size constraint, is what gives us the vagueness of *almost* and *barely*—if we have one hundred students, *almost every student attended graduation* could be truthfully uttered if, looking out onto the masses, we see the vast majority of them. Some are missing, sure, but granting that there is some small group of them (four or five, maybe) missing, we can reason based on the assumption that there is a definite group missing. We just don't know precisely what the group is.

2.2.3.3 The utility of the size constraint

In this section, we will address a useful contribution the size constraint makes as part of the meaning of *but* and *almost* to rule out a potentially strange case. This case arises because of some particular assumptions made by two alternative versions of our proposal—Gajewski (2013) and Hirsch (2016)—and we submit that the most parsimonious way to solve this problem is assume the size constraint is active.

Recall our proposal that *but* imposes contextually-determined constraints on the size of the exception was motivated in part by the oddity of (140a).

(140) CONTEXT: Anya, Buffy, Cordelia, Dawn, Ethan, Faith, Giles, Harmony, Kendra, and Larry are the ten students in my class; Anya, Buffy, and Cordelia are the only students who passed the exam.

a. # Every student but Dawn, Ethan, Faith, Giles, Harmony, Kendra, and Larry passed.

We proposed that *but* encodes in addition to set subtraction the size constraint below.

(141) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, μ_c be a measure function, determined by the context, mapping sets to numerical values representing their size, and n_c be a contextually determined numerical threshold for size.

$\text{SMALL}_c(\mathbb{X}) = 1$ *iff* $\mu_c(\mathbb{X}) \leq n_c$.

(142) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

For any context c, an utterance of φ in c is felicitous only if $n_c < \frac{1}{2}(\mu_c(\mathbb{A}))$

The size constraint requires that the size of $\mathbb{X}$ be less than half the size of $\mathbb{A}$. When considering sets of atomic entities, the size is plausibly determined by cardinality; if the cardinality of $\mathbb{A}$ is 10, then in order to count as small, $\mathbb{X}$ must contain 4 or fewer atoms. This, of course, is the largest the set can be—the context, we presume, can require it to be even smaller.

Gajewski (2013) solves the problem of domains of two by restricting the subtractive alternatives over which **EXH** quantifies further than we have assumed. Gajewski takes *but* to have the following denotation, of a lower type than ours[21].

(143) $[\![\text{but}_G]\!]^{g,c} = \lambda X \lambda A\colon \mathbb{X} \subseteq \mathbb{A}.\ \mathbb{A} - \mathbb{X}$

This entry presupposes that the exception is a subset of the set from which it is subtracted, just like our entry, and the rest of the composition proceeds as ours did up until **EXH**. Gajewski restricts possible alternatives to those formed with subsets of the *exception.*

(144) ***SUBTRACTIVE ALTERNATIVES*** (Gajewski's version)
For a subtractive LF φ, with an exception χ, $\mathbf{ALT}(\varphi) = \{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq \mathbb{X}\}$
($\varphi[\chi/\chi']$ is identical to φ except χ is replaced by χ')

(145) # Some spellbook but_G BG is worth reading.

a. $[\![(145)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$; where defined:
$[\![(145)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$
   i. $\mathbb{S}_w - \{BG\} \cap \mathbb{WR}_w \neq \emptyset$
   ii. $\forall \psi [\psi \in \mathbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1]$

The contribution of **EXH** is different: given that the alternatives are restricted to subsets of the actual exception, it reflects that the presupposition encoded by *but* projects for each alternative. If we grant that this is a viable alternative, it's entailed, in which case there are no alternatives to negate. If we don't allow it as a viable alternative, then still, there are none to negate. Note that for any exception larger than a singleton, all alternatives will be entailed, simply as a matter of the monotonicity of *some* in φ. This holds, too, even when the restrictor has a cardinality of two, since all that matters is the exception and its subsets. So no matter what, when we've got *some*, there will be no non-entailed alternatives. So, since $[\![\varphi]\!]^{g,c}$ entails $[\![\mathbf{EXH}\ \varphi]\!]^{g,c}$ in this case, **EXH** is not licensed, and we predict infelicity. Gajewski's explanation, of course, relies essentially on the strong restriction to alternatives. Since we take the size constraint to be independently relevant and active, we don't find the additional restrictions to alternatives necessary at this junction. However, adopting it would indeed allow us to appeal to L-Analyticity in order to claim (145) is ungrammatical.

Hirsch (2016) also aims to solve the problem of domains of two via the following constraint on existential quantifiers.

[21]See the discussion in §2.2.2.1.

(146) ***ANTI-SINGLETON***

Existential quantification is infelicitous when the speaker and hearer can know that the restrictor of the existential is necessarily a singleton without knowing the extension of the restricting NP or the conversationally determined domain of quantification.

I have nothing terribly insightful to say about this principle. It does what it should: when the restrictor of *some* has a cardinality of two, then subtractives cannot modify it, because the resulting existential quantification would range over a singleton. Our size constraint is, as we've claimed, independently warranted, and takes care of the problem that is created by domains of two without needing any additional restrictions on the alternatives or constraints on the kinds of domains over which we can quantify, however, so we'll continue to assume that it is relevant and useful.

2.2.3.4 Subtractive alternatives are necessary for *barely*

In proposing a semantics for *almost* and *barely* modeled after that of *but*, we took the alternatives to be constructed just as they are for *but*, structurally, by replacing the exception with other expressions of the same type.

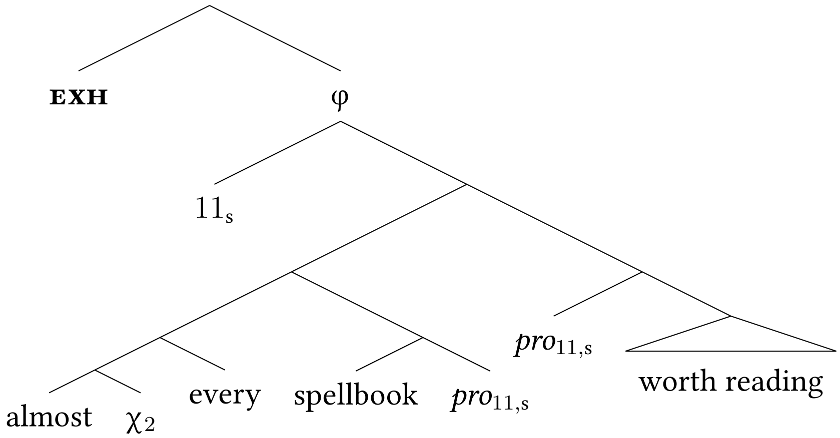
(147) ***SUBTRACTIVE ALTERNATIVES***

For a subtractive LF φ, with an exception χ,
$\textbf{ALT}(\varphi) := \left\{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_e \;\&\; \varphi[\chi/\chi'] \text{ is grammatical}\right\}$

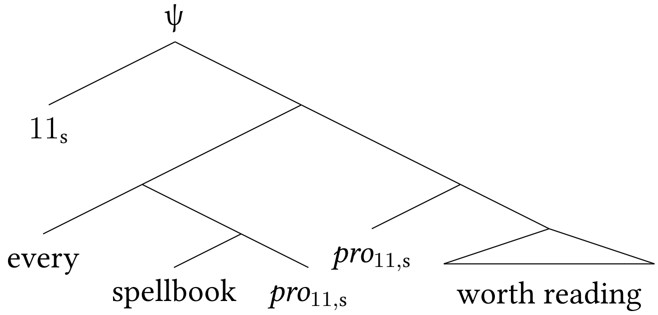
This is a crucial component of the analysis; this is precisely how we derive the polar inference for all three subtractive operators, and derive that the exception is the smallest (non-empty!) exception. This is absolutely necessary for *but*, as shown by von Fintel (1993), but it is novel to propose for *almost* and *barely*, and needs defending.

In their analyses of *almost*, both Spector (2014) and Crnič (2018) propose that the lone (relevant) alternative to some φ containing *almost* is formed by eliminating that subtractive modifier from the structure. Let's call this a *classical* view of the modifier's alternatives. It's essentially what the literature has generally assumed about *almost*, and it is intuitive in a sense—part of why we say *almost* φ is to convey that φ is false. We simply use **EXH** to derive that inference. In the context of our current proposal, we'll assume the classical view requires eliminating the exception pronoun as well. That is, the alternative to φ in (148) is just ψ in (149).

(148)



(149)



exh, of course, requires non-entailed alternatives to be false or undefined, and since $[\![\varphi]\!]^{g,c}$ doesn't entail $[\![\psi]\!]^{g,c}$, the latter must not be true. For *almost* sentences, this proposal works just fine.

(150) $[\![(148)]\!]^{g,c}(w) = 1$ *iff* $\underbrace{\mathbb{S}_w - g(2) \subseteq \mathbb{WR}_w \;\&\; \textsc{small}_c(g(2))}_{[\![\varphi]\!]^{g,c}(w)=1} \;\&\; \underbrace{\mathbb{S}_w \not\subseteq \mathbb{WR}_w}_{[\![\psi]\!]^{g,c}(w)\neq 1}$

These truth conditions seem adequate. The same result obtains for *almost no*:

(151) $[\![$**EXH** *Almost* χ_2 *no spellbook is worth reading*$]\!]^{g,c}(w) = 1$ *iff*
$\underbrace{\mathbb{S}_w - g(2) \cap \mathbb{WR}_w = \emptyset \;\&\; \textsc{small}_c(g(2))}_{[\![\varphi]\!]^{g,c}(w)=1} \;\&\; \underbrace{\mathbb{S}_w \cap \mathbb{WR}_w \neq \emptyset}_{[\![\psi]\!]^{g,c}(w)\neq 1}$

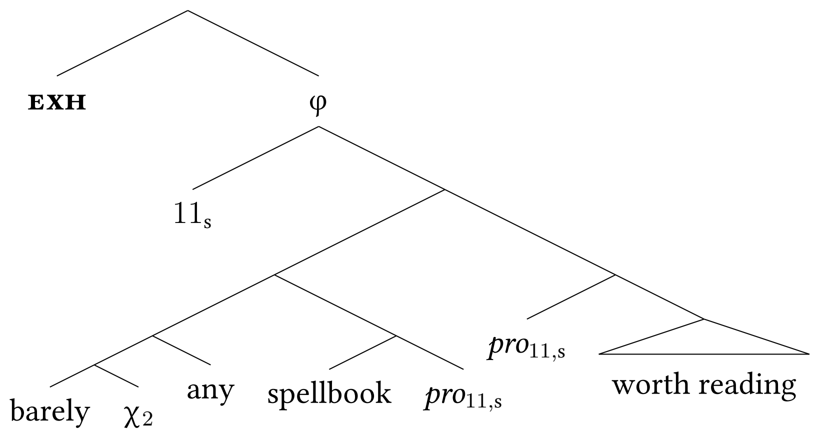
For *almost some spellbook is worth reading*, the alternative is just simply the the existential claim that some spellbook is worth reading, but that *is* entailed. By assumption there are no other alternatives to negate, so:

(152) $[\![$**EXH** *Almost* χ_2 *some spellbook is worth reading*$]\!]^{g,c}(w) = 1$ *iff*
$\underbrace{\mathbb{S}_w - g(2) \cap \mathbb{WR}_w \neq \emptyset \;\&\; \textsc{small}_c(g(2))}_{[\![\varphi]\!]^{g,c}(w)=1}$

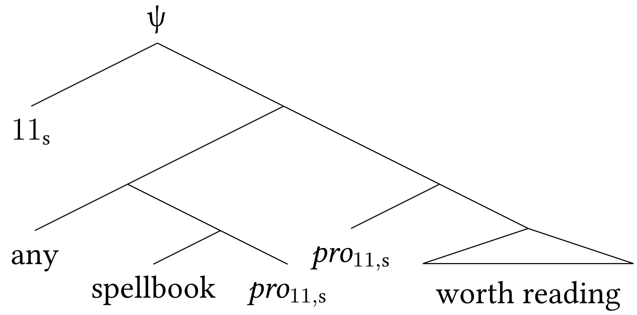
EXH is vacuous here, so it is not licensed, but then neither is *almost*. This predicts unacceptability.

The problem is *barely*. Part of our goal to analyze *almost* and *barely* is to treat them as closely connected in that both contribute set subtraction and imposing restrictions on the size of that subtrahend, but *barely* additionally contributes negation. The classical view of its alternatives is a little less intuitive, since *barely* φ has standardly been taken to convey φ, not its negation, but still, parity between the two operators seems plausible (and classical, too!). However, the classical view fails on two counts: it cannot derive the positive polar inference for *barely any*, and it doesn't predict unacceptability for *barely every* or *barely no*.

(153)



(154)



Any is an existential quantifier, and $[\![\varphi]\!]^{g,c}$ doesn't entail $[\![\psi]\!]^{g,c}$—the emptiness of the intersection of $\mathbb{S}_w - g(2)$ and $\mathbb{WR}_w$ doesn't mean there has to be some spellbook worth reading. EXH negates the existential alternative, then, deriving the following truth conditions.

(155) $[\![(153)]\!]^{g,c}(w) = 1$ *iff*
$$\underbrace{\mathbb{S}_w - g(2) \cap \mathbb{WR}_w = \emptyset \;\&\; \textsc{small}_c(g(2))}_{[\![\varphi]\!]^{g,c}(w)=1} \;\&\; \underbrace{\mathbb{S}_w \cap \mathbb{WR}_w = \emptyset}_{[\![\psi]\!]^{g,c}(w)\neq 1}$$

These truth conditions hold in w just in case g(2) is small in c and there are no spellbooks worth reading in w. This is perfectly coherent, but not the meaning of *barely any spellbooks are worth reading*. It would seem that this proposal doesn't work for *barely*.

One problem might arise with this argument: the lone alternative to *Barely any spellbooks are worth reading* is plausibly not grammatical, and therefore that the sentence doesn't have any alternatives.

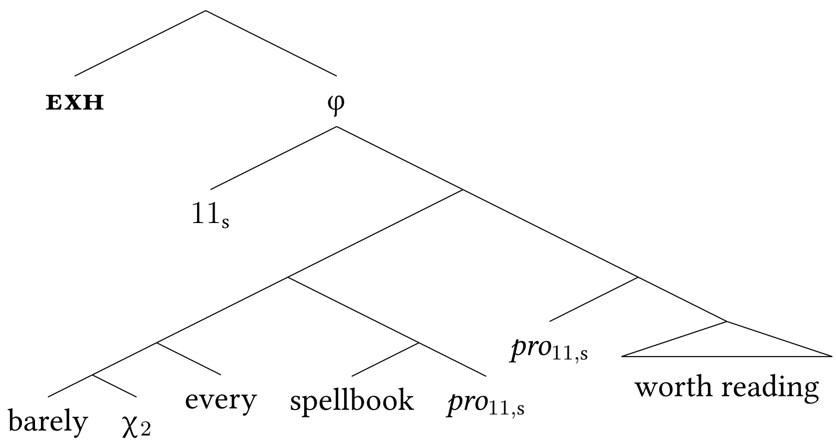
(156) ?? Any spellbooks are worth reading.

Granting that the *any* in *barely any* is an NPI, it's not clear what licenses *any* in (156) above. This can actually be understood as an argument *for* our proposal. If (156) is not a grammatical alternative, then there are in fact *no* alternatives to (153), and **exh** isn't licensed. We therefore would expect *Barely any spellbooks are worth reading* to be unacceptable, but it's perfectly good.

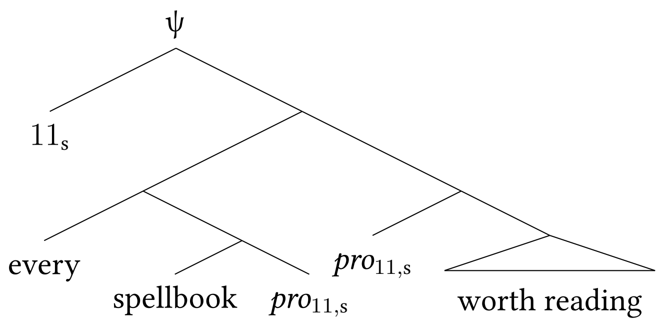
We can further show that the classical view doesn't make the right predictions for *barely every* and *barely no*.

(157) # Barely every spellbook is worth reading.

a.



b.

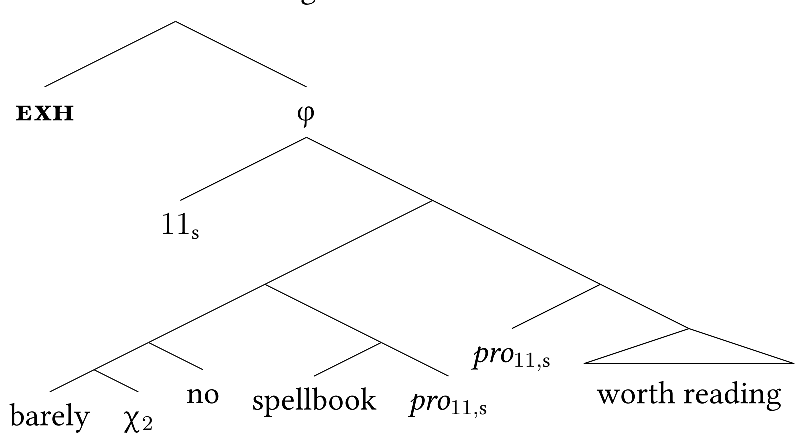


c. $[\![(157\text{a})]\!]^{g,c}(w) = 1$ *iff*

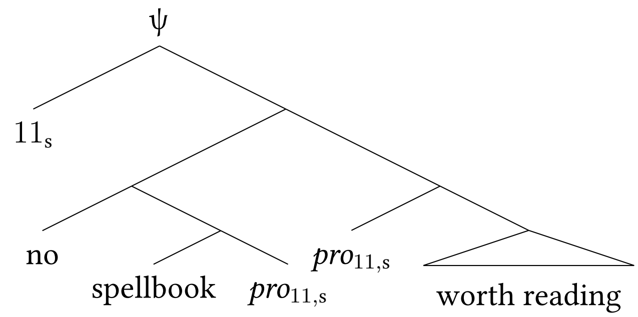
$\underbrace{\mathbb{S}_w - g(2) \not\subseteq \mathbb{WR}_w \ \& \ \textsc{small}_c(g(2))}_{[\![\varphi]\!]^{g,c}(w)=1} \ \& \ \underbrace{\mathbb{S}_w \not\subseteq \mathbb{WR}_w}_{[\![\psi]\!]^{g,c}(w)\neq 1}$

(158) # Barely no spellbook is worth reading.

a.



b.



c. $[\![(158a)]\!]^{g,c}(w) = 1$ *iff*

$$\underbrace{\mathbb{S}_w - g(2) \cap \mathbb{WR}_w \neq \emptyset \ \& \ \textsc{small}_c(g(2))}_{[\![\varphi]\!]^{g,c}(w)=1} \ \& \ \underbrace{\mathbb{S}_w \cap \mathbb{WR}_w \neq \emptyset}_{[\![\psi]\!]^{g,c}(w)\neq 1}$$

In neither case does $[\![\varphi]\!]^{g,c}$ entail their respective $[\![\psi]\!]^{g,c}$, and so they're negateable; both negations are coherent. (157) is true if not every spellbook is worth reading, and (158) is true if some spellbook is. These are not the right meanings, and it is hard to see how we'd get ourselves out of this. However, we saw that structuring our alternatives in terms of substitution of the exception led to the right results, and uniformly. Subtractive alternatives appear to be necessary[22]

The arguments above apply to *barely*; what about *almost*? At this point, we've seen no definitive evidence in favor of assuming the exceptive alternatives for the modifier. It just works, and quite well! While we will see in the next chapter that the classical view fails in a way that the exceptive view does not, right now it's simply parsimony that drives the extension. *Almost* and *barely* are taken to be identical in all respects save negation in the latter, so why shouldn't their alternatives be constructed in the same manner? If *but* needs these kinds of alternatives, too, then why shouldn't *almost*, which is, so far, identical in all important respects save the lack of an overt exception?

2.2.4 Summing up: the essentials of the analysis

Let's do some summarizing of what we've seen so far. To develop an analysis of the modifiers *almost* and *barely*, we started by looking at the exceptive modifier *but*, and its parallels with *almost*. They were striking: both are are capable of modifying universal *every* and negative existential *no*, but not positive existential *some* or NPI *any*. We adopt the proposal originating with von Fintel (1993) and further developed by Gajewski (2008) that *but* is a modifier of quantificational determiners, and following Crnič (2018), we take *almost* to be the same. Both operators take

[22] This discussion doesn't go into other ways we could define **exh**, or alternative alternatives. To the former, we could try and go the route of Fox (2007) and Bar-Lev (2018), i.a., and incorporate Innocent Exclusion and Inclusion. To the latter, we could adhere more strictly to Katzir's theory of structural alternatives, which would allow substitution of *almost* for *barely* and vice versa, different determiners for other determiners, etc. As far as we can tell, these routes are not promising, but a fuller discussion would take us too far afield, and frankly, might be a little tedious. We leave that for future work.

'exceptions' χ, *but* overtly and *almost* covertly, and remove them from the restrictor of the determiner they modify. The covert exceptions are taken to be pronominal in nature, receiving their value from the assignment function. We additionally propose a novel requirement: the exception χ must be small relative to the context of utterance. For *almost*, these contributions together—set subtraction and the size constraint—comprise what in the literature has been called the proximal inference,*almost* φ conveying, in part, that φ is 'close' to being true. For *but*, the observation that the exception χ must be small is, as far as I can tell, novel.

Furthermore, we extended this style of analysis to *barely*, which hasn't, as far as I'm aware, received an analysis as a modifier of determiners before. Its distribution with respect to the determiners listed above is complementary to *but* and *almost*—it is only compatible with NPI *any* of that set. *Barely* also gives rise to a proximal inference; *barely* φ conveys that φ was 'close' to being false. It is reasonable, then, to think of *almost* and *barely* as sort of mirror images of one another: *barely* is just the 'negative' variant of *almost.*

All three operators are taken to be licensed only within the scope of an operator generating implicatures, **EXH**, which quantifies over alternatives to its sister, requiring that they all be entailed or not true in the world of evaluation. We take the proposal for *but* from Crnič that the relevant alternatives are generated structurally, by replacing the exception with other expressions of the same type; we additionally propose that both *almost* and *barely* make use of these alternatives as well (for the latter, we've shown this is necessary, but the former not quite yet). **EXH** is ultimately responsible for the derivation of the polar inference for *almost* and *barely*, and where all three modifiers are acceptable, their exceptions χ must be the smallest, non-empty exceptions to the quantificational claims they modify.

Monotonicity, in tandem with **EXH**, tells us why this is the distribution we see: subtractives are only acceptable when the monotonicity of the φ in which they occur is Downward Entailing with respect to the constituent α from which they subtract their exception, i.e. the restrictor of the determiner they modify. When a subtractive in some LF φ subtracts an exception χ from an argument α in a Downward Entailing environment, the entailed alternatives φ are formed with (small) supersets of the exception χ. **EXH** then requires all other alternatives to be not true. This means χ must be the smallest, non-empty exception. When α is in an Upward Entailing environment in φ, alternatives formed with subsets of χ are entailed, and nothing else. This doesn't allow us to find a smallest, non-empty exception—the smallest exception is the empty set, and above that, if there is one true χ there is at least one more χ' that is also true, but not Strawson entailed.

Every and *no* are both Downward Entailing on their restrictors, and absent any other entailment-reversing operators, we predict that they should be compatible with *almost* and *but.* They ought not be compatible with *some* or NPI *any*, since these are Upward Entailing on their restrictors, and contradiction ensues. *Barely* is different precisely because it contributes negation—the modifier itself is an entailment-reversing operator. So, when it modifies NPI *any*, Upward Entailing on both its arguments, in fact, the net result is that its restrictor is in a downward entailing environment; *some* is presumably unacceptable because it is a PPI. When attempting to modify *every* or *no*, *barely*'s contribution makes their restrictors Upward Entailing environments, and therefore the exception it introduces cannot be the smallest, non-empty exception.

A few final facts are helpful to recall. We saw that when some quantificational claim is necessarily true, modification by a subtractive operator was not possible. That is, *almost every spellbook is a book* is infelicitous because being spellbook entails being a book. In such a situation, the subtraction contributed by *almost* doesn't do anything substantive, and the sentence with **EXH**, call it φ, is entailed by the sentence ψ identical to φ except that **EXH** is eliminated. In such a situation, **EXH** isn't licensed, and hence the subtractive isn't, either. We also saw that when a quantificational claim is true independent of subtraction in a particular world, **EXH** generates falsity. That is, if in w *every spellbook is worth reading* is true, then *almost every spellbook is worth reading* is false in w. Subtractives must do some real semantic work in order to be licensed and yield truth.

2.3 Why EXH?

The question isn't really "Why **EXH**?" but rather, why is something like it necessary? Our analysis is a DISTRIBUTED one, and we assumed it without argument. Why should we separate out the means by which we derive the polar inference for *almost* and *barely*, Uniqueness for *but* (for they really are one and the same), from the meanings of these modifiers? We began the analysis by coming to some conclusions about the meanings of *but* sentences, but we might wonder what goes wrong if we pack all that meaning into *but*, as originally proposed by von Fintel (1993), and likewise *almost* and *barely* (suppressing the subset and non-emptiness presuppositions for space).

(159) a. $[\![\text{but}]\!]^{g,c} = \lambda X \lambda A \lambda \mathscr{D} \lambda B.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})\ \&\ \forall \mathbb{X}' \left[\mathbb{X} \not\subseteq \mathbb{X}' \rightarrow \neg\mathscr{D}(\mathbb{A}-\mathbb{X}')(\mathbb{B})\right]$

b. $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})\ \&\ \forall \mathbb{X}' \left[\mathbb{X} \not\subseteq \mathbb{X}' \rightarrow \neg\mathscr{D}(\mathbb{A}-\mathbb{X}')(\mathbb{B})\right]$

c. $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B.\ \neg\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})\ \&\ \forall \mathbb{X}' \left[\mathbb{X} \not\subseteq \mathbb{X}' \rightarrow \mathscr{D}(\mathbb{A}-\mathbb{X}')(\mathbb{B})\right]$

These are INTEGRATED entries for the subtractives. It makes at least some intuitive sense to put the burden of all the complex meanings of subtractive sentences into the subtractive operators themselves. When thinking about *almost* and *barely*, it would capture, for example, the difficulty of cancelling the polar inference without much work:

(160) a. Every student but Willow failed the test; in fact, #every student did.

b. Almost every student failed the test; in fact, #every student did.

c. Barely any students failed the test; in fact, #no student did.

Of course, we saw that our analysis predicted this too—the continuations are contradicted by the entailments of the first sentences, so maybe an integrated analysis isn't justified quite yet. It turns out there is concrete reason we shouldn't even bother with such a view.

One of the starting points for Gajewski's (2008) revision of von Fintel's (1993) theory of exceptive sentences is the fact that *but* is licensed in certain environments that are unexpected on the latter's theory. Supposing NPI *any* is an existential quantifier, it must take narrow scope with respect to negation in (161).

(161) There aren't any girls but Cordelia in the auditorium.
(cf. *There aren't any girls in the auditorium.*)

As Gajewski points out, an integrated semantics for *but* as in (159a) will predict trivial truth conditions for this sentence: applying this *but* to NPI *any* existential determiner yields necessarily false truth conditions, and the negation licensing the NPI yields triviality from that.

(162) $[\![(161)]\!]^{g,c}(w) = 1$ *iff*
$$\neg\big[\mathbb{G}_w - \{C\} \cap \mathbb{A}_w \neq \emptyset \;\&\; \textsc{small}_c(\{C\}) \;\&\; \forall \mathbb{X}'\big[\{C\} \not\subseteq \mathbb{X}' \rightarrow \neg \mathbb{G}_w - \mathbb{X}' \cap \mathbb{A}_w \neq \emptyset\big]\big]$$
$$= \mathbb{G}_w - \{C\} \cap \mathbb{A}_w = \emptyset \vee \neg\textsc{small}_c(\{C\}) \vee \exists \mathbb{X}'\big[\{C\} \not\subseteq \mathbb{X}' \;\&\; \mathbb{G}_w - \mathbb{X}' \cap \mathbb{A}_w \neq \emptyset\big]$$

These truth conditions cannot be false, but (161) certainly can be; it's only true if Cordelia is in the auditorium, and no other girl is.

We can construct a (perhaps marginal) parallel example for *almost*.

(163) There aren't almost χ_2 any girls in the auditorium.

While some speakers don't like (163), others report this to be acceptable (see also Horn 2005, 2011[23]). If we straightforwardly apply the integrated semantics for *almost* above, we'll make the same predictions we did for (161), i.e. that (163) cannot be false.

(164) $[\![(163)]\!]^{g,c}(w) = 1$ *iff*
$$= \mathbb{G}_w - g(2) \cap \mathbb{A}_w = \emptyset \vee \neg\textsc{small}_c(g(2)) \vee \exists \mathbb{X}'\big[g(2) \not\subseteq \mathbb{X}' \;\&\; \mathbb{G}_w - \mathbb{X}' \cap \mathbb{A}_w \neq \emptyset\big]$$

This doesn't tell us anything for speakers that reject (163)—we'll have to leave this open for now.

There are more problems of this kind for an integrated approach to the meaning of subtractives.

[23]Horn's (2011) discussion presumes this kind of case to involve wide-scope *almost*, and thus on our proposal, wide-scope *any*, too. In this case, *any* would have to be its free-choice, universal variant, even though this is a *there*-insertion context. If this were plausible, we'd still need to explain why some speakers reject (163). In the end, this data point is helpful, but the argument for separating out Uniqueness from the meaning of subtractives doesn't hinge on this.

(165) a. If she wants to get an A on the test, Willow has to answer every question but the last one.

b. If she wants to get an A on the text, Willow has to answer almost χ_2 every question.

(165a) doesn't require that Willow skip (i.e. not answer) the last exam question in order to get an A; she merely can skip it and still pass. Nouwen (2006) points out the same fact for (165b). However, if *but* and *almost* have an integrated semantics, and are interpreted within the scope of the universal necessity modal in the consequent, we predict that the requirements are she skip at least the last question in the case of (165a), or at least one question in the case of (165b). Let's make this more concrete. We'll adopt the standard assumption that *if*-clauses serve as restrictors to modals in the consequent (Kratzer 1986; see von Fintel 2011 for an overview on conditionals), and we'll give the modal a very simple semantics, taking it to denote a universal quantifier over worlds where particular requirements are met. For any w, w′, let $\textsc{req}_w(w')$ be read as (and mean!) w′ is compatible with the laws and regulations in w.

(166) a. $[\![\text{have to}]\!]^{g,c} = \lambda w \lambda p_{s,t} \lambda q_{s,t}.\ \{w' \mid \textsc{req}_w(w') \ \&\ p(w') = 1\} \subseteq \{w' \mid q(w') = 1\}$

b. $[\![\text{Willow get an A on the exam}]\!]^{g,c} = \lambda w.\ \text{GAA}_w(\text{the exam})(\text{Willow})$

c. $\mathbb{G}_w = \{x \mid x \text{ is a question on the exam in } w\}$

d. $\mathbb{A}_w = \{x \mid \text{Willow answers } x \text{ in } w\}$

e. $\{L\} = \{\text{the last question on the exam}\}$

If *almost every question* and *every question but the last one* have an integrated semantics, and are within the scope of *has to*, we derive the following meaning for (165a).

(167) $[\![(165a)]\!]^{g,c}(w) = 1$ *iff* $\{w' \mid \textsc{req}(w)(w') \ \&\ \text{GAA}_{w'}(\text{the exam})(\text{Willow})\} \subseteq$
$\{w' \mid \mathbb{B}_{w'} - \{L\} \subseteq \mathbb{A}_{w'} \ \&\ \textsc{small}_c(\{L\}) \ \&\ \forall \mathbb{X}'[\{L\} \not\subseteq \mathbb{X}' \rightarrow \mathbb{B}_{w'} - \mathbb{X}' \subseteq \mathbb{A}_{w'}]\}$

These truth conditions hold in w if, for all w′ compatible with the rules and regulations in w and where Willow gets an A on the exam in w′, Willow answers every question that isn't the last one, and does *not* answer the last one. The requirements are stronger here because the subtractive is caught within the scope of the modal. The same holds for (165b), relevant changes made, of course. That these aren't the immediately intuitive truth conditions—though indeed, one can felicitously utter these two sentences if the requirements were a bit more odd—is a problem for the integrated semantics[24].

[24] It's plausible that we could scope just the quantifier above the modal; we'd derive the right meaning:

(i) $[\![(165a)]\!]^{g,c}(w) = 1$ *iff* $\textsc{small}_c(\{L\})$ &
$\mathbb{G}_w - \{L\} \subseteq \{y \mid \{w' \mid \textsc{req}(w)(w') \ \&\ \text{GAA}_{w'}(\text{the exam})(\text{Willow})\} \subseteq \{w' \mid y \in \mathbb{A}_{w'}\}\}$ &
$\forall \mathbb{X}'[\{L\} \not\subseteq \mathbb{X}' \rightarrow [\mathbb{G}_w - \mathbb{X}' \not\subseteq \{y \mid \{w' \mid \textsc{req}(w)(w') \ \&\ \text{GAA}_{w'}(\text{the exam})(\text{Willow})\} \subseteq \{w' \mid y \in \mathbb{A}_{w'}\}\}]]$

Crnič (2018) sharpens the problem by showing similar facts hold in ellipsis contexts, where crucially the wide-scope route described above isn't available.

(168) a. Willow answered every test question but the last. To get an A, she had to.

b. Willow answered almost every test question. To get an A, she had to.

For both examples, while requirements certainly could dictate that Willow skip at least one question to receive an A, intuitively these cases are compatible with her answering every question and still passing. The trouble is that we can't derive that with an integrated semantics for the subtractives.

It is well known that VP ellipsis is taken to be subject to (something along the lines of) the following condition (Rooth 1992; Fiengo & May 1994; Heim 1996; Fox 2000).

(169) **CONDITION ON VP ELLIPSIS:**

If a quantificational expression is interpreted in the antecedent VP, a semantically equivalent expression must be interpreted in a parallel position in the elided VP.

Crnič (2018:746)

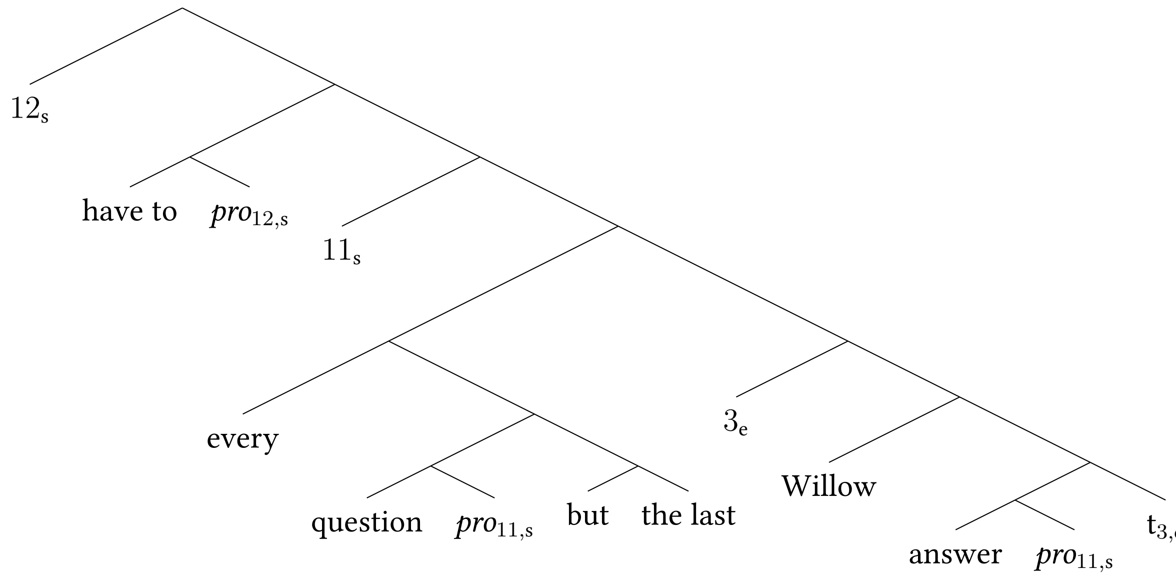
Now consider the following.

(170) Every girl in the class passed a test. Every boy did, too.

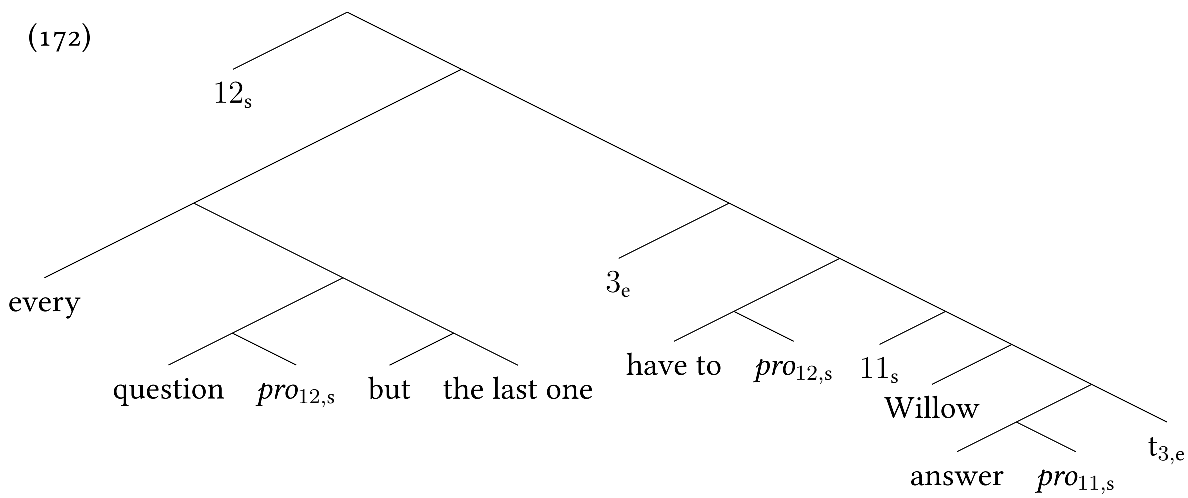
There are two admissible scopal configurations in the antecedent: $\forall>\exists$, and $\exists>\forall$. While both are possible in (170) under ellipsis, both in antecedent and in the ellipsis site, they must be identical. That is, if the antecedent conveys that there was a particular test passed by every girl in the class, the continuation could not convey that for every boy, that boy passed a distinct test. The Condition describes this restriction. When we go to the *almost* and *but* cases, the LFs of the antecedent clauses in (168) both lack a modal, and so in the ellipsis site, we couldn't scope the modified quantifiers above the modal—that is, it can only have the LF in (171), not (172).

This says Willow had to answer every question that wasn't the last one to get an A, and she didn't have to answer the last one to get an A. Good! Now, the pronoun saturating the world argument of *question* has changed so as not to be totally wild and free. Allowing *pro*s with indicies to denote worlds, and letting them be free, is a simplification and possibly an incorrect one, and here we potentially run into *de dicto* and *de re* issues that we are not prepared to deal with. We hope this won't do too much damage.

(171)



(172)

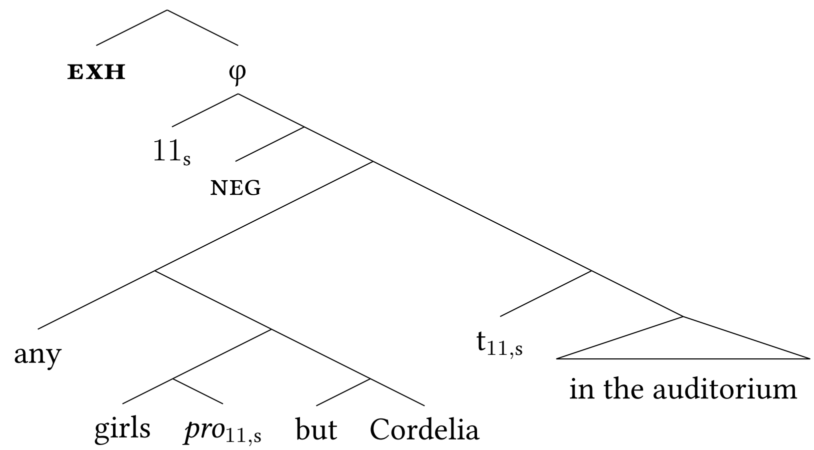


The truth conditions we'll derive for (171) are equivalent to those we did for (165a), less the contribution of the *if*-clause. The integrated approach faces serious problems, then!

Gajewksi's and Crnič's solution to this problem is precisely what we've been assuming all along: Uniqueness/the polar inference for the subtractives is derived separately, for us through EXH. For the Gajewski cases, EXH can scope above negation. In general, this makes sense—EXH is commonly taken to be banned from occurring in downward entailing environments anyways (Spector 2014, e.g., though see §4.1 and §4.2 for some further discussion). Crnič proposes that EXH doesn't originate within the VP at all, but rather, higher in the clausal structure of these cases; it can scope immediately above or below the modal. Where EXH scopes above negation in (161), the new LF of which is given below; φ is Downward Entailing with respect the restrictor of *any*, and our usual entailments hold. We assume *there* is vacuous, and omit it from the structure. we derive the truth conditions in (173b).

(173) There aren't any girls but Cordelia in the auditorium.

a.



b. $[\![(173a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

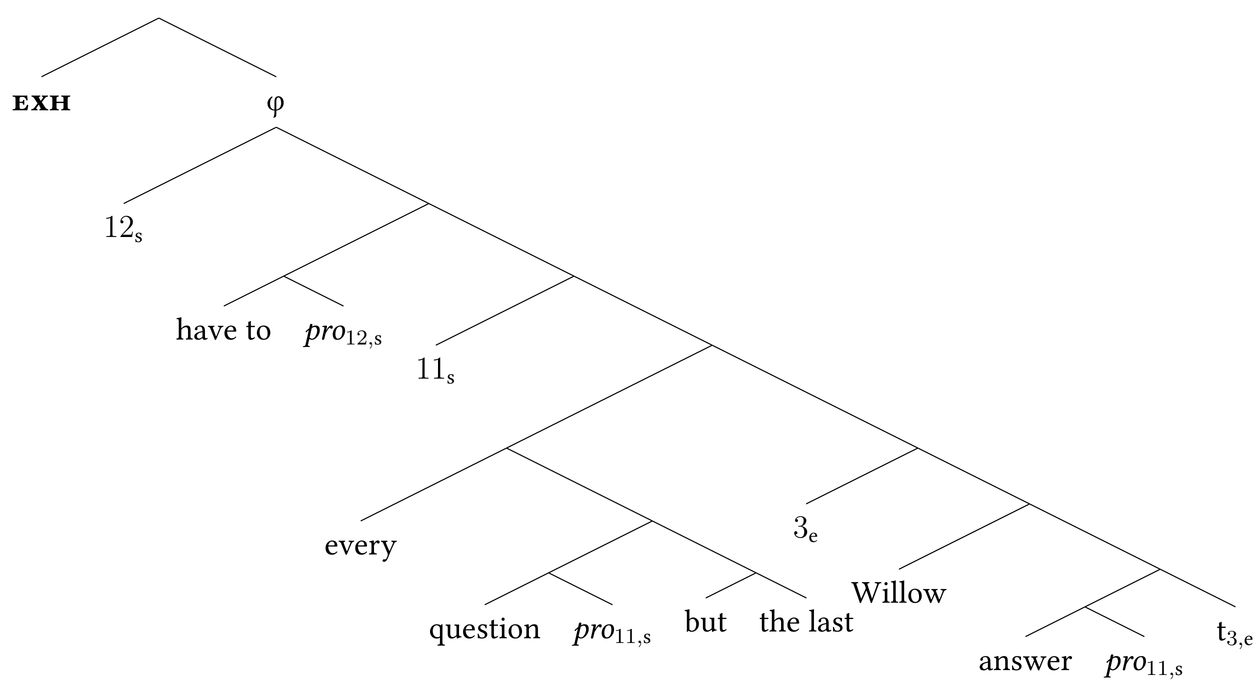
i. $\mathbb{G}_w - \{C\} \cap \mathbb{A}_w = \emptyset$ & $\textsc{small}_c(\{C\})$

ii. $\forall \mathbb{X}' [\{C\} \not\subseteq \mathbb{X}'$ & $\textsc{small}_c(\mathbb{X}') \rightarrow \mathbb{G}_w - \mathbb{X}' \cap \mathbb{A}_w \neq \emptyset]$

These truth conditions are just what we should expect, equivalent to *No girls but Cordelia are in the auditorium*. We also derive the right truth conditions for Crnič's cases; φ is again Downward Entailing with respect to the restrictor of *every*, so entailments follow as they should.

(174) (Willow answered every question on the exam but the last. To get an A,) she had to ~~answer every question but the last.~~

a.



b. $[\![(174a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\{w' \mid \textsc{req}(w)(w')$ & $\textsc{GAA}_{w'}(\text{the exam})(\text{Willow})\} \subseteq$
$\{w' \mid \mathbb{B}_{w'} - \{L\} \subseteq \mathbb{A}_{w'}$ & $\textsc{small}_c(\{L\})\}$

ii. $\forall \mathbb{X}' \big[\{L\} \not\subseteq \mathbb{X}' \text{ \& } \textsc{small}_c(\mathbb{X}') \rightarrow$
$\big\{w' \mid \textsc{req}(w)(w') \text{ \& } \mathrm{GAA}_{w'}(\text{the exam})(\text{Willow})\big\} \not\subseteq$
$\big\{w' \mid \mathbb{B}_{w'} - \mathbb{X}' \subseteq \mathbb{A}_{w'} \text{ \& } \textsc{small}_c(\mathbb{X}')\big\}\big]$

(175) $[\![$Willow have to answer almost χ_2 every question$]\!]^{g,c}(w) = 1$ *iff* (a) & (b) = 1

a. $\big\{w' \mid \textsc{req}(w)(w') \text{ \& } \mathrm{GAA}_{w'}(\text{the exam})(\text{Willow})\big\} \subseteq$
$\big\{w' \mid \mathbb{B}_{w'} - g(2) \subseteq \mathbb{A}_{w'} \text{ \& } \textsc{small}_c(g(2))\big\}$

b. $\forall \mathbb{X}' \big[g(2) \not\subseteq \mathbb{X}' \text{ \& } \textsc{small}_c(\mathbb{X}') \rightarrow$
$\big\{w' \mid \textsc{req}(w)(w') \text{ \& } \mathrm{GAA}_{w'}(\text{the exam})(\text{Willow})\big\} \not\subseteq$
$\big\{w' \mid \mathbb{B}_{w'} - \mathbb{X}' \subseteq \mathbb{A}_{w'} \text{ \& } \textsc{small}_c(\mathbb{X}')\big\}\big]$

No surprises here—these truth conditions capture that the requirement is Willow answer every question but the last, leaving open that she could answer it. The same goes for *almost*; simply substitute an exception pronoun for *the last one* and the logic will be the same. The distributed theory makes the right predictions for these cases precisely because its separability allows us to negate here the modal claim for each alternative, sufficiently loosening the interpretation from what it would be if **EXH** scoped below the modal.

Our main contribution now is that *barely* also requires this severance.

(176) Willow skipped barely any questions on the exam. To get an A, she had to.

This is naturally interpreted as saying Willow skipped few questions, and the requirements are such that she couldn't skip more than a few. The distributed theory derives this, whereas the integrated analysis doesn't.

(177) **INTEGRATED THEORY**

a. $[\![(176)]\!]^{g,c}(w) = 1$ *iff* $\big\{w' \mid \textsc{req}(w)(w') \text{ \& } \mathrm{GAA}_{w'}(\text{the exam})(\text{Willow})\big\} \subseteq$
$\big\{w' \mid \mathbb{B}_{w'} - \{L\} \cap \mathbb{S}_{w'} = \emptyset \text{ \& } \textsc{small}_c(\{L\}) \text{ \& } \forall \mathbb{X}' \big[\{L\} \not\subseteq \mathbb{X}' \rightarrow \mathbb{B}_{w'} - \mathbb{X}' \cap \mathbb{S}_{w'} \neq \emptyset\big\}\big]$

(178) **DISTRIBUTED THEORY**

a. $[\![(176)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) = 1

i. $\big\{w' \mid \textsc{req}(w)(w') \text{ \& } \mathrm{GAA}_{w'}(\text{the exam})(\text{Willow})\big\} \subseteq$
$\big\{w' \mid \mathbb{B}_{w'} - \{L\} \cap \mathbb{S}_{w'} = \emptyset \text{ \& } \textsc{small}_c(\{L\})\big\}$

ii. $\forall \mathbb{X}' \big[\{L\} \not\subseteq \mathbb{X}' \text{ \& } \textsc{small}_c(\mathbb{X}') \rightarrow$
$\big\{w' \mid \textsc{req}(w)(w') \text{ \& } \mathrm{GAA}_{w'}(\text{the exam})(\text{Willow})\big\} \not\subseteq$
$\big\{w' \mid \mathbb{B}_{w'} - \mathbb{X}' \cap \mathbb{S}_{w'} = \emptyset \text{ \& } \textsc{small}_c(\mathbb{X}')\big\}\big]$

It should be clear that the integrated truth conditions require Willow's skipping at least the last question in order get an A, whereas the distributed truth conditions do not; this indicates that *barely*, too, requires a distributed semantics.

This section essentially provided support for the particular division of labor we assumed in the §2. Except for the novel observation that *barely* is like *but* and *almost* in requiring this, everything else has been discussed previously (Gajewski 2008, 2013; Nouwen 2006; Crnič 2018). The obvious conclusion is that other analyses of *almost* and *barely*—regardless of their fine details—fail to capture the data discussed above. This applies to many analyses out there, e.g. Amaral (2007), Hitzeman (1992), Penka (2006), Rapp & von Stechow (1999), Sevi (1998), and on.

The Uniqueness inference that *but*-sentences carry is sensitive to scope. Since we've tied the derivation of the polar inference in *almost* and *barely* sentences to the exact same mechanism, it is sensitive to scope, too. These are obligatory inferences, derived by **EXH**, but their precise shape depends on the location of **EXH** with respect to subtractives and any other possible intervening operators. Now, when we think about what is in the literature called the polar inference for *almost* and *barely*, it should be clear that the name is misleading on our analysis. *almost* φ doesn't explicitly or directly convey $\neg\varphi$, and *barely* φ doesn't directly or explicitly convey φ. The scalar implicatures generated by negating subtractive alternatives ultimately entails that inference, but the 'classical' polar inference isn't a part of the meaning of *almost* sentences. When we move away from monoclausal declaratives, this is even more clear; the weakening or disappearance of the 'classical' polar inference naturally falls out of our theory—it 'disappears' in syntactic contexts where **EXH** has non-local scope with respect to *almost* and *barely*. That scalar implicature is still generated, but crucially it has a different shape in such contexts.

2.4 Closing words and moving forward

This chapter has developed an analysis of *almost* and *barely* that builds on an analysis of the subtractive modifier *but*. This analysis has two key components. First, all three operators are taken to be modifiers of quantificational determiners, removing elements from those determiners' restrictors, and requiring that what is removed be small relative to the context[25].

(179) $[\![\text{but}]\!]^{g,c} = \lambda X\lambda A\lambda\mathscr{D}\lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

(180) $[\![\text{almost}]\!]^{g,c} = \lambda X\lambda\mathscr{D}\lambda A\lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

(181) $[\![\text{barely}]\!]^{g,c} = \lambda X\lambda\mathscr{D}\lambda A\lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle et,\langle\langle et,ett\rangle,\langle et,\langle et,t\rangle\rangle$

(182) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, μ_c be a measure function, determined by the context, mapping sets to numerical values representing their size, and n_c be a contextually determined numerical threshold for size.

[25]The formulation of the size constraint now reflects the syntax not only of *but*, but also *almost* and *barely*.

$\text{SMALL}_c(\mathbb{X}) = 1$ *iff* $\mu_c(\mathbb{X}) \leq n_c$

(183) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

For any context c, an utterance of φ in c is felicitous only if $n_c < \frac{1}{2}(\mu_c(\mathbb{A}))$

Second, all three subtractive operators are required to be within the scope of **EXH**, which operates on the alternatives ψ to the LF φ containing those operators. The alternatives are derived structurally, by replacing the exception with other expressions of the same type. **EXH** requires all non-Strawson entailed alternatives be false or undefined.

(184) $[\![\textbf{EXH}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\psi\left[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\right]$

For expository purposes, we restricted our attention to 'subtractive' LFs and to 'subtractive' alternatives, as defined below, which allow us to more succinctly state the contribution of **EXH** given such LFs and alternatives.

(185) ***SUBTRACTIVE LFs***

A LF φ is a *subtractive LF* just in case φ is a LF containing exactly one subtractive operator Σ introducing an exception χ, and no other alternative-triggering expression, and φ is of type $\langle s,t\rangle$.

(186) **SUBTRACTIVE ALTERNATIVES**

For a subtractive LF φ, with an exception χ,
$\textbf{ALT}(\varphi) := \left\{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_d\ \&\ \varphi[\chi/\chi'] \text{ is grammatical}\right\}$

(187) For a subtractive LF φ,
$[\![\textbf{EXH}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}(w) \neq 1\right]$

We saw that monotonicity affects which alternatives are entailed.

(188) ***SUBTRACTIVE ENTAILMENT THEOREMS FOR DETERMINERS***

Let Σ be a subtractive operator, Δ a quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively. Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$. Let g be an assignment function in context c.

Let $\textbf{ALT}(\varphi) = \left\{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_e\ \&\ \varphi[\chi/\chi'] \text{ is grammatical}\right\}$.

a. If φ is Upward Entailing (and not Downward Entailing) with respect to the constituent α, then for all χ' such that $\mathbb{X}' \subseteq \mathbb{X}$, then $[\![\varphi]\!]^{g,c} \Rightarrow_{ST} [\![\varphi[\chi/\chi']]\!]^{g,c}$.

b. If φ is Downward Entailing (and not Upward Entailing) with respect to the constituent α, then for all χ' such that $\mathbb{X} \subseteq \mathbb{X}'$ & $\mathbb{X}'$ is small in c, then $[\![\varphi]\!]^{g,c} \Rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}$.

We saw that these facts come together to predict why *almost* and *but* can modify *every* and *no* but not *some*, and why *barely* can modify NPI *any* but neither *every* nor *no*. A subtractive can felicitously modify some determiner Δ in φ only if the exception it introduces is the smallest, non-empty exception to the quantificational claim that is made in φ. This is possible only if φ is Downward Entailing with respect to the restrictor of Δ. Where no entailment-reversing operator intervenes between EXH and Σ, this just boils down to the monotonicity of Δ for *almost* and *but*: *every* and *no* are both Downward Entailing on their restrictors. *Some* is not; there will never be a smallest exception in such a case. *Barely*, itself a downward entailing operator, reverses the monotonicity of NPI *any*, and as a result, its restrictor is in a Downward Entailing environment. For the other determiners discussed, its negative contribution creates problems.

This particular division of labor between EXH and subtractives has been argued in the literature previously; we showed that it was necessary for *barely*, too, and furthermore showed that *barely* requires subtractive alternatives as defined above. We don't yet have an argument for the necessity of such alternatives for *almost*—it is merely adequate. Such an argument will come in the next chapter. The split between subtractives and EXH allows us some flexibility in predicting why the classical polar inference typically assumed to be encoded by *almost* and *barely* seems to disappear in certain environments—the flexibility of the scope of EXH means that in certain environments, the inference is weakened, but the mechanism by which it appears is necessarily present. We'll discuss this more in Chapter 5. As we move forward, we'll take this analysis as a given. We'll show how it naturally extends to other domains, in particular degree constructions and numerals, and the consequences and payoff of such an analysis.

Let's end this chapter with some open questions that won't be addressed in the rest of the dissertation, at least not in any satisfactory way. First, as we've discussed, we've restricted ourselves to LFs with one subtractive and no other alternative-triggering operator. At this point, we have no way of dealing with stacking of subtractive operators, for instance, on a single determiner.

(189) a. Almost $\{$every/no$\}$ vampire but Spike attacked the school.

b. Barely any vampires but Spike attacked the school.

I'm not sure what my own judgments are of these cases, but if they are good, we need to figure out how to integrate them into our analysis. That means taking a deeper dive into questions about the nature of alternatives to such sentences. We'll have to leave this open for the future.

A second question is a bit deeper. We've claimed that *almost* is just *but* with a covert exception. If that's right, then does *barely* have a parallel operator, a negative version of *but*? I'm not sure. A more careful investigation into English, as well as other languages, might find such an operator.

I have no good reason to think one shouldn't exist, so if we cannot find one, then we have to explain why.

Chapter 3

Comparatives, equatives, and the positive construction

3.1 Introduction

In the previous chapter, we investigated the semantics of the *almost* and *barely* through the lens of the quantificational determiners *every*, *no*, *some*, and *any*. We proposed that *almost* and *barely* are *subtractive* operators that modify quantificational determiners by subtracting elements from the restrictor of those determiners. Those elements are supplied by a covert exception, and both modifiers require that exception to be small in the context of utterance, the PROXIMAL inference. *Almost* then says that were it not for the exception, the unmodified quantificational assertion would be true, and *barely* says that were it not for the exception, the unmodified quantificational assertion would be false, the POLAR inferences. Both operators must be within the scope of an **EXH** operator; **EXH**, of course, requires that the alternatives to the LF φ it takes as argument be either entailed or not true in the world of evaluation. Our attention is restricted for simplicity's sake to what we've called subtractive LFs—those LFs φ containing exactly one subtractive operator and exception X and no other alternative-triggering expressions—and the alternatives over which **EXH** quantifies are restricted to those ψ where the exception is replaced by an expression of the same type.

In this chapter, we'll bring this analysis to bear on another area in which subtractives surface: degree constructions. In particular, we'll look closely at the distribution of *almost* and *barely* with respect to comparatives, equatives, and the positive construction. We'll argue that the analysis developed in the previous chapter allows us to quite neatly account for the distribution of *almost* and *barely* in these constructions when we assume that degree morphemes are quantificational expressions. We'll learn quite a bit more from this investigation, too. We saw earlier that *barely* requires alternatives formed by replacing the exception with expressions of the same kind in order to derive the inferences it gives rise to intuitively, and for reasons of parsimony, we made the assumption that *almost* invoked such alternatives, too. We'll see that the distribution of *almost* with respect to comparative and equatives justifies that assumption. Subtractives also provide a novel argument for the idea that all measurement scales in natural language semantics must be dense, the *Universal Density of Measurement* hypothesis of Fox & Hackl (2006). *Barely* equatives show us that our meaning for *barely* hasn't been quite right, and that *almost* and *barely* aren't the "duals" purported sometimes in the literature. We'll see that we need to make some assumptions about the relationships between positive and negative antonyms in order to maintain maximality in the semantics of the comparative. Finally, we'll see what the semantics of the positive construction has to look like in light of the contribution of subtractives.

3.1.1 Parallels

Quantificational determiners are not the only expressions which are subject to modification by *almost* and *barely*—the modifiers are often found in degree constructions. *Almost* and *barely* surface in equatives and comparatives, but with apparently complementary distribution.

(190) a. Buffy is almost as old as Willow is.
b. # Buffy is barely as old as Willow is.

(191) a. # Willow is almost older than Buffy is.
b. Willow is barely older than Buffy is.

Almost appears to go with the equative, and not the comparative, whereas *barely* goes with the comparative and not the equative. Intuitively, (190a) *almost* is true where Buffy isn't as old as Willow is (the polar inference), but the difference between their ages is small (the proximal). Paraphrasing: were it not for the small difference between their ages, it would be true that Buffy is as old as Willow is. (191b) is also true where Willow is older than Buffy is (the polar), but only just so (the proximal). Again, paraphrasing: were it not for the small difference between their ages, it would be false that Willow is older than Buffy is.

We see similar complementarity in distribution with certain kinds of adjectives.

(192) a. Dawn's room is almost clean.
b. # Dawn's room is barely clean.

(193) a. # Buffy's room is almost dirty.
b. Buffy's room is barely dirty.

Almost can modify adjectives like *clean*, *safe*, and *impossible*, so-called TOTAL (Yoon 1996) or MAXIMUM STANDARD (Kennedy & McNally 2005) adjectives, but not their PARTIAL or MINIMUM STANDARD adjective counterparts. *Barely* is the opposite. (192a) is true if Dawn's room isn't clean, but the difference between Dawn's room and a clean room is minimal; (192b) if the difference between Buffy's room and a non-dirty room (i.e., a clean room), is likewise minimal, and her room is indeed dirty.

The *almost* facts—that it can modify the equative but not the comparative, and that it can modify total but not partial adjectives—are relatively underexplored, save by Rotstein & Winter (2004). The *barely* facts, as far as I can tell, are novel. They require explanation. However, intuitively, the contributions made by *almost* and *barely* in these constructions are quite like their contributions as when modifying quantificational determiners; our goal in this chapter is to show that the subtractive analysis predicts these facts. We're motivated to pursue this not only by the distributional complementarity and parallel interpretive contributions of the subtractives: we saw in the previous chapter that sometimes the polar inference was "absent" in certain constructions, like the consequents of conditionals, and this persists in degree constructions.

(194) a. If Spike wants to defeat Buffy, he has to be almost as strong as Buffy is.

b. Spike was almost as strong as Buffy was in their fight. To win, he really had to be.

Intuitively, (194a) doesn't require that Spike actually be weaker than Buffy if he wants to succeed, which is exactly the meaning one would derive if *almost* encoded both subtraction and exhaustivity itself. (194b) is the same. We also saw in the previous chapter that the polar inference is invisible to certain operators, like speaker-oriented adverbials and *because*; again, this is true in degree constructions as well.

(195) a. Unfortunately, Sunnydale is barely cooler than Los Angeles is.

b. Unfortunately, Sunnydale is almost as warm as Los Angeles is.

(196) a. Buffy left Sunnydale because it was barely inhabitable.

b. Buffy left Sunnydale because it was almost uninhabitable.

In (195a) and (195b), what's intuitively unfortunate is that the difference between the climate in Sunnydale and Los Angeles is negligible, not that Sunnydale is cooler than (or not as warm as) Los Angeles. The polar inferences these modifiers give rise to as a result of exhaustivity is outside the semantic scope of *unfortunately*. Similarly, a reason for Buffy's departure in (196a) and (196b) is that Sunnydale is nearing uninhabitability—she didn't leave because it was inhabitable. These operators only have access to the proximal inferences.

Our goal is to extend the analysis of subtractives as modifiers of quantificational determiners to the degree domain. We'll claim that in comparatives, equatives, and the positive construction, *almost* and *barely* modify expressions that denote relations between degree predicates, i.e. expressions of type $\langle$dt,dtt$\rangle$. Let's call these expressions *degree quantifiers*. The comparative and equative morphemes themselves are often analyzed as degree quantifiers, and we'll posit a novel analysis of the positive construction that works in just the same way. *Almost* and *barely*'s meanings are generalized just a bit to allow modification of degree quantifiers, as a first step in accounting for their cross-categorial distribution.

Pursuing an analysis of *almost* and *barely* that unifies the treatment of degree quantifiers with that of (nominal!) quantificational determiners in the previous chapter is already an improvement on piecemeal analyses. It is also notable that we pursue the analysis in the direction that we do—that is, analogizing degree constructions to nominal quantifiers. There are very few analyses in the literature on *almost* that deal in adjectival semantics generally; with the exception of Hitzeman (1992) and Rotstein & Winter (2004), what is out there has largely not investigated in detail the interaction of *almost* with positive adjectives. The syntax and semantics of the comparative and equative cases have also not been covered in any serious detail, though Rotstein & Winter (2004) discuss *almost* with comparatives and equatives briefly. Hitzeman (1992) and Rotstein & Winter (2004), which we will cover in more detail later on, suggest the opposite generalization strategy. What we will see is that our analysis requires comparatively little in order to generalize, whereas it's less clear that their proposal would generalize without major stipulations.

Before we get to explaining the data above, we need to introduce the semantics for degree constructions we'll use throughout the chapter.

3.1.2 Formal preliminaries

3.1.2.1 Degrees, scales and adjectives

Adjectives like *old*, *young*, *beautiful*, and *intelligent*—GRADABLE adjectives—intuitively describe properties that hold of entities to certain amounts. These amounts are called *degrees*, and we add to the our semantic ontology the type d, with the domain D_d. Degrees are points, and they represent measurement along some dimension—age, height, and beauty are all such dimensions. Dimensions of measurement are conceptual, and we can utilize measure functions as their formal correlates. Measure functions are relations between entities and degrees: AGE, HEIGHT, and *beauty* are measure functions. Adjectives are taken to denote relations between individuals and degrees on a *scale*.

(197) Let $\mathcal{S}$ be a set of degrees, μ be a measure function, and $\leq_\mu$ be an ordering on $\mathcal{S}$ relative to μ. $\mathcal{S}$ is a *scale* just in case $\forall$d,d$' \in \mathcal{S}$:

a. $d \leq_\mu d' \vee d' \leq_\mu d$ *total order*

b. $d <_{\mu} d' \rightarrow \exists d'' \big[d'' \in \mathcal{S}\ \&\ d <_{\mu} d'' <_{\mu} d' \big]$ *density*

For some measure function μ, we'll write '$\mathcal{S}_{\mu}$' for the scale associated with μ. For example, $\mathcal{S}_{\text{AGE}}$ is the age scale, the set of all possible ages; every age degree is totally ordered with respect to every other on that scale[1]. The assumption that a set of degrees is a scale only if it is dense, i.e. (197b), is not uncontroversial, but we follow Fox & Hackl (2006) in adopting this assumption, the *Universal Density of Measurement* hypothesis (UDM). We'll see in §3.2.2.1 that subtractives provide a novel argument for the UDM, but for now, take it for granted.

The scales associated with certain adjectives, the RELATIVE adjectives like *old* and *tall*, have been claimed to be 'open' in the following sense: they do not include any upper or lower bound (Kennedy & McNally 2005, i.a.). The intuition about upper bounds is that, in principle, there is no limit to how tall, how old, how intelligent, etc., an entity can be, i.e., there is no maximal height, age, or degree of intelligence. The measure functions HEIGHT, AGE, and *intelligence*, then, must be able to map entities in a world to any degree, no matter how high. The intuition about lower bounds concerns the fact that there are some kinds of entities for which it doesn't really make sense to say they have any height at all, even 0—ideas, rays of sunshine, and the calling of being a Vampire Slayer are not things that have height. We don't want HEIGHT in any world to map these objects to degrees. We can do this by stipulating that $\mathcal{S}_{\text{HEIGHT}}$ does not include its lower bound—the measure function is undefined for objects that lack a physical extent. We define the shape of these relative scales in the following way.

(198) **RELATIVE ADJECTIVE SCALES**
Let *A* be a relative adjective, and μ_A be the measure function *A* encodes.
$\mathcal{S}_{\mu_A}$ is the scale associated with *A*; $\mathcal{S}_{\mu_A} := \big\{ d \mid 0 <_{\mu_A} d \leq_{\mu_A} \infty \big\}$

Now we can spell out a meaning for adjectives.

We'll make the common assumption that gradable adjectives denote degree-individual relations; we'll add in world-dependence as well. Adjectives encode, in part, *measure functions*. Measure functions μ take a world w and an entity and return its unique measurement on a particular scale. The measure function AGE, in a world w, maps entities to the degree d on $\mathcal{S}_{\text{AGE}}$ corresponding to their age in w[2].

(199) a. $[\![\text{old}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{AGE}}.\ \text{AGE}_w(x) \geq d$ $\langle s,det \rangle$

b. $[\![\text{tall}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{HEIGHT}}.\ \text{HEIGHT}_w(x) \geq d$ $\langle s,det \rangle$

[1] We'll assume degrees are primitive objects in our ontology, but there are alternative views that construct them via, for example, equivalence classes of individuals (Cresswell 1976; Bale 2008, 2011; Schwarz 2010) or sets of possible individuals (Schwarzschild 2013).

[2] Minimally, I'm using the same notation for a dimension of measurement and a measure function, i.e. small caps. The distinction doesn't really matter here.

c. $[\![\text{beautiful}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{BEAUTY}}.\ \textsc{beauty}_w(x) \geq d$ $\langle s,det \rangle$

All adjectives carry definedness conditions restricting the degrees they take as argument to particular scales; in general we will take this to hold for all adjectives, but will suppress it nonetheless. The denotations above have an important property: they are *degree (downward) monotone.*

(200) A function A of type $\langle s,\langle d,et\rangle\rangle$ is *degree downward monotone* just in case
$\forall w \forall x \forall d \forall d' \left[A(w)(d)(x) = 1 \text{ \& } d' <_{\mu_A} d \rightarrow A(w)(d')(x) = 1\right]$ Heim (2000:2)

This property affects the shape of the set of degrees that an adjective *A* characterizes given an individual *x* and a world *w*, i.e. $\left\{d \mid [\![A]\!](w)(d)(x) = 1\right\}$—given the scale structure for $\mathcal{S}_{\text{AGE}}$ discussed above, for any *x*, *w*, $\left\{d \mid [\![\text{old}]\!](w)(d)(x) = 1\right\}$ is the set of degrees containing every degree on $\mathcal{S}_{\text{AGE}}$ from *x*'s age in *w* downwards. The same goes for *tall* and *beautiful.*

Now we can turn to an analysis of the comparative and equative.

3.1.2.2 Semantics and syntax for the comparative and equative

The comparative can be analyzed as expressing a strict ordering between degrees, so that (201a) is true just in case (201b) is true; the equative expresses a weaker ordering, so (202a) is true just in case (202b) is.

(201) a. Willow is older than Buffy is.

b. Willow's age $>$ Buffy's age

(202) a. Willow is as old as Buffy is.

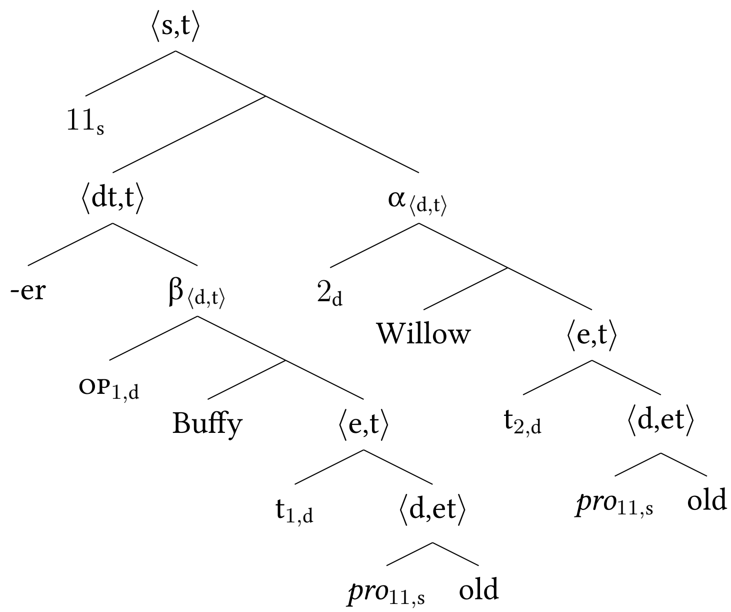
b. Willow's age $\geq$ Buffy's age

As we mentioned, we're going to assume that the comparative morpheme *-er/more* and the equative *as* denote functions that relate predicates of degrees, i.e. $\langle dt,dtt \rangle$ functions. Syntactically, we assume that degree quantifiers and their *subordinate* clauses originate in the degree argument position of the adjective A in the *matrix* clause, and move for interpretability (Heim 2000, et seq.). They leave behind an indexed trace of type d and drop a coindexed binder just below their landing site; this is just Quantifier Raising for generalized degree quantifiers, of course. The subordinate clause is essentially structurally parallel to the matrix clause: it contains a copy of the adjective, ultimately elided by Comparative Deletion (Bresnan 1973, 1975). Within the subordinate clause, an operator originating in the degree argument position of the adjective moves to the edge of the clause, leaving a trace of type d and binding it.

On this proposal, (201) has the structure below.

(203) Willow is older than Buffy is ~~old~~.

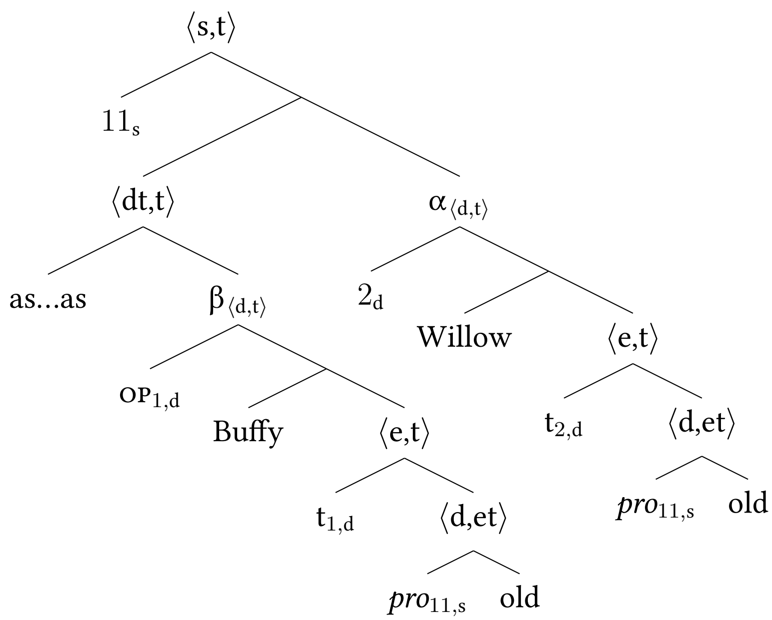
(204)



A parallel structure is assumed for equatives.

(205) Willow is as old as Buffy is ~~old~~.

(206)



Than is typically taken to be semantically vacuous in comparatives, and we'll simply omit it at LF. The world argument of the matrix adjective is bound by a coindexed binder at the top of the clause; we assume that the subordinate clause contains a pronoun coindexed with this binder, as before.

Old, as stated just above, denotes a degree downward monotone function, so for every degree on $\mathcal{S}_{\text{AGE}}$ below Buffy's age in a world w, the above function returns true. Such a set of degrees is an *interval*: a *convex* set of degrees[3].

(207) ***INTERVALS***

Let $\mathbb{D}$ be a set of degrees and $\mathcal{S}$ be a scale; $\mathbb{D}$ is an *interval* of $\mathcal{S}$ iff $\mathbb{D}$ is a convex subset of $\mathcal{S}$.

(208) ***CONVEXITY***

Let $\mathcal{S}$ be a scale totally ordered by $\leq$, and let $\mathbb{D} \subseteq \mathcal{S}$.
$\mathbb{D}$ is *convex* just in case $\forall \mathrm{d,d',d''}\big[\mathrm{d,d''} \in \mathbb{D}\ \&\ \mathrm{d'} \in \mathcal{S}\ \&\ \mathrm{d} \leq \mathrm{d'} \leq \mathrm{d''} \rightarrow \mathrm{d'} \in \mathbb{D}\big]$.

In general, intervals may or may not include their endpoints. An interval is lower-closed if it includes its greatest lower bound, lower-open if it does not; it is upper-closed if it includes its least upper bound, and upper-open if it does not. An interval that is lower- and upper-closed is (totally) closed, and an interval that is lower- and upper-open is (totally) open. We'll use the following notation for intervals.

(209) For some points m, n, such that $\mathrm{m} \leq \mathrm{n}$:

a. $\big\{\mathrm{x} \mid \mathrm{m} \leq \mathrm{x} \leq \mathrm{n}\big\} = \big[\mathrm{m,n}\big]$ (totally) closed interval
b. $\big\{\mathrm{x} \mid \mathrm{m} \leq \mathrm{x} < \mathrm{n}\big\} = \big[\mathrm{m,n}\big)$ lower-closed, upper-open interval
c. $\big\{\mathrm{x} \mid \mathrm{m} < \mathrm{x} \leq \mathrm{n}\big\} = \big(\mathrm{m,n}\big]$ lower-open, upper-closed interval
d. $\big\{\mathrm{x} \mid < \mathrm{x} < \mathrm{n}\big\} = \big(\mathrm{m,n}\big)$ (totally) open interval

Square brackets mark that a boundary is included in the interval, and curved brackets mark that that boundary is not included.

The intervals supplied to *-er* and *as...as* via their arguments are upper-closed, *initial* intervals.

(210) For any interval $\mathbb{D}$ and scale $\mathcal{S}$ such that $\mathbb{D} \subseteq \mathcal{S}$:
$\mathbb{D}$ is an *initial* interval of $\mathcal{S}$ *iff* $\exists \mathrm{d}\big[\mathrm{d} \in \mathbb{D}\ \&\ \forall \mathrm{d'}\big[\big[\mathrm{d'} \in \mathcal{S}\ \&\ \mathrm{d'} \leq \mathrm{d}\big] \leftrightarrow \mathrm{d'} \in \mathbb{D}\big]\big]$

Certain operators, like the comparative and equative morphemes, might want to pull out the endpoint of such an interval, so the function **MAX** is recruited (von Stechow 1984a; Rullmann (1995)[4]).

[3] Note that 'interval' is used sometimes to label both convex and non-convex sets of degrees, e.g. in Schwarzschild & Wilkinson (2002). We will restrict ourselves to *convex* intervals, in general dropping the modifier, unless otherwise indicated.

[4] Our **MAX** is stronger than the **MAX** operator that Rullmann proposes; his requires neither convexity nor initiality.

(1) $\textbf{MAX}_{\mathrm{Rullmann}}(\mathbb{D}) := \iota \mathrm{d}\big[\mathrm{d} \in \mathbb{D}\ \&\ \forall \mathrm{d'}\big[\mathrm{d'} \in \mathbb{D} \rightarrow \mathrm{d} \geq \mathrm{d'}\big]\big]$

Our version of **MAX** affects how the reasoning proceeds when analyzing subtractive degree constructions. Let $\mathbb{D}$ and $\mathbb{X}$ be arbitrary sets of degrees, let $\mathbb{D}' = \mathbb{D} - \mathbb{X}$, and let $\mathbb{D}'$ include its upper bound. If $\mathbb{D}'$ is not convex, or is not an initial interval of a scale, $\textbf{MAX}(\mathbb{D}')$ will be undefined with our **MAX**, but not Rullmann's. We make use of this stronger version of **MAX** for expository purposes. As we will see in §3.2.1.6, this is not necessary, but we want .

(211) For any set of degrees $\mathbb{D}$ on a scale $\mathcal{S}$:

$$\textbf{MAX}(\mathbb{D}) := \iota d\big[d \in \mathbb{D} \mathbin{\&} \forall d'\big[\big[d' \in \mathcal{S} \mathbin{\&} d' \leq d\big] \leftrightarrow d' \in \mathbb{D}\big]\big]$$

This operator picks out from an initial interval on a scale the maximal degree in that interval, should there be one; it is undefined otherwise. Two more notational notes. First, we'll use curly brackets when utilizing interval notation for the argument of **MAX**. Second, when it is clear in the context which adjective/measure function is under discussion, we will use the following shorthand: for an arbitrary world *w*, $\textsc{age}_w$(Buffy) will be written $\mathcal{B}_w$, $\textsc{age}_w$(Willow) will be written $\mathcal{W}_w$, and so on as necessary.

(212) $\textbf{MAX}\big\{\big(0, \mathcal{B}_w\big]\big\} = \mathcal{B}_w = \textsc{age}_w(\text{Buffy})$

Now we're in a position to define entries for the comparative and equative morphemes.

A reasonably standard analysis takes them, as we've said, to express relations between sets of degrees; more specifically, we'll posit that that relation holds between the maxima, as defined above, of two sets of degrees (von Stechow 1984a; Rullmann 1995[5]).

(213) $[\![\text{-er/more}]\!] = \lambda D_{\langle d,t\rangle}\lambda D'_{\langle d,t\rangle}.\ \textbf{MAX}(\mathbb{D}) < \textbf{MAX}(\mathbb{D}')$

(214) $[\![\text{as}]\!] = \lambda D_{\langle d,t\rangle}\lambda D'_{\langle d,t\rangle}.\ \textbf{MAX}(\mathbb{D}) \leq \textbf{MAX}(\mathbb{D}')$

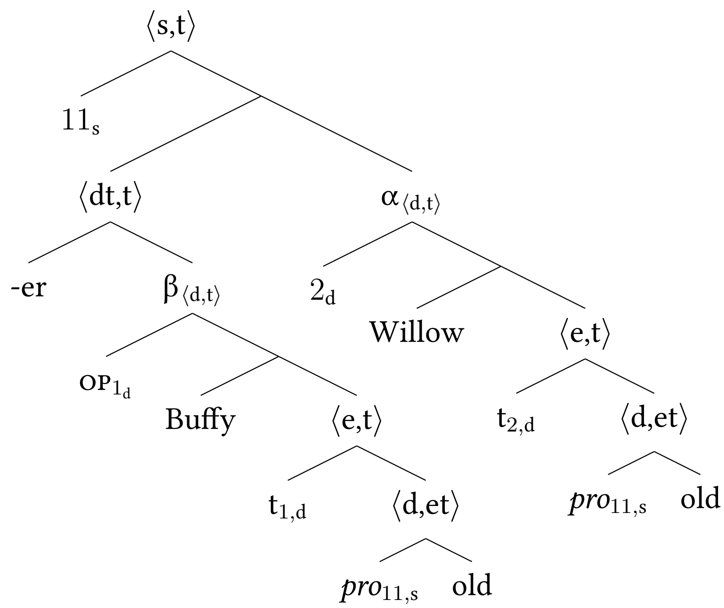
To illustrate:

(215) Willow is older than Buffy is.

[5]This semantics for the comparative and equative will create well-known difficulties for deriving the right truth conditions when quantifiers occur in the subordinate clauses. There is a host of recent literature on this matter; see Fleisher 2016 for an overview. We will be able to provide neither an adequate account of either subtractive-modified comparatives and equatives with quantifiers in subordinate clauses, nor an account of subtractive-modified quantifiers within the subordinate clauses of comparatives and equatives. These are topics left for future investigation.

We'll see in §3.3 that this maximality-based analysis requires us to think hard about how we analyze negative antonyms of relative adjectives anyways, and we'll discuss those issues and some changes there.

a.

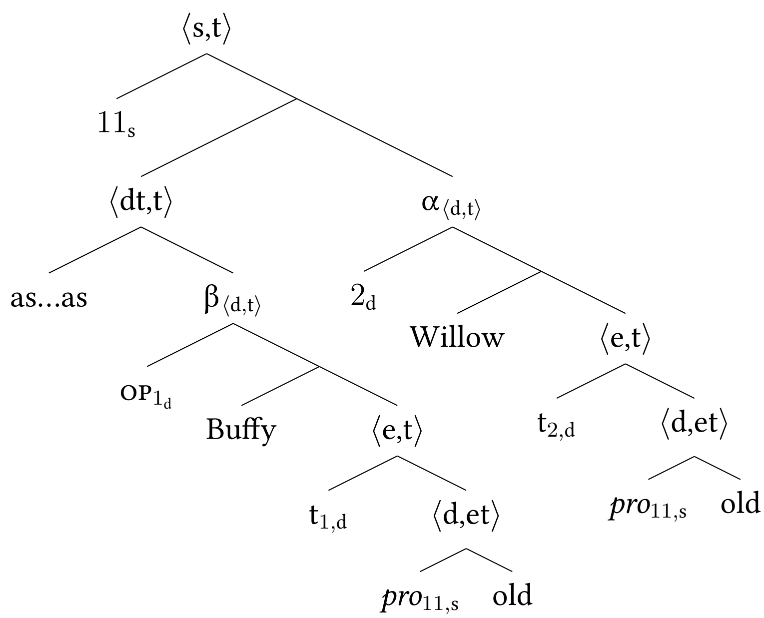


b. $[\![(215)\text{a}]\!](\text{w}) = 1$ *iff* $\textbf{MAX}\{(0, \mathcal{B}_{\text{w}}]\} < \textbf{MAX}\{(0, \mathcal{W}_{\text{w}}]\}$
i.e. $\textsc{age}_{\text{w}}(\text{Buffy}) < \textsc{age}_{\text{w}}(\text{Willow})$
SHORTHAND: $\mathcal{B}_{\text{w}} < \mathcal{W}_{\text{w}}$

Now, putting this together with our syntax and semantics for equatives[6]:

(216) Willow is as old as Buffy is.

a.



b. $[\![(216)\text{a}]\!](\text{w}) = 1$ *iff* $\textbf{MAX}\{(0, \mathcal{B}_{\text{w}}]\} \leq \textbf{MAX}\{(0, \mathcal{W}_{\text{w}}]\}$
i.e. $\textsc{age}_{\text{w}}(\text{Buffy}) \leq \textsc{age}_{\text{w}}(\text{Willow})$
SHORTHAND: $\mathcal{B}_{\text{w}} \leq \mathcal{W}_{\text{w}}$

[6]We've omitted a discussion of the positive construction in this section; this will be the subject of §3.3.

We predict truth just in case Willow is at least as old as Buffy is. These truth conditions allow for Willow to be older, in fact; the general intuition, however, is that equatives convey, well, equality between two measures. That is, (216) typically conveys that Willow and Buffy are exactly the same age. This has been argued to be a scalar implicature (e.g. Horn 1972; see Rett (2015:19-20) for discussion), plausibly generated through application of **EXH**, though this is not uncontroversial. We'll have nothing to say on this particular issue, but will assume the entry above.

Now we can turn to comparatives and equatives with *almost* and *barely*.

3.2 Comparatives and equatives

Our semantics for *almost* and *barely* isn't quite right to compose with the comparative morpheme and the equative morpheme just yet, but it's not hard to see that a minimal change is all that is necessary. Here are the entries from the previous chapter.

(217) $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A} - \mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

(218) $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A} - \mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle et,\langle et,\langle\langle et,ett\rangle,\langle et,t\rangle\rangle$

The minimal switch should be obvious: we'll swap individuals for degrees (and change the names for λ-bound variables).

(219) $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle dt,dtt\rangle} \lambda D_{\langle d,t\rangle} \lambda D'_{\langle d,t\rangle}\text{: } \mathbb{X} \subseteq \mathbb{D}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{D} - \mathbb{X})(\mathbb{D}')\ \&\ \textsc{small}_c(\mathbb{X})$

(220) $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle dt,dtt\rangle} \lambda D_{\langle d,t\rangle} \lambda D'_{\langle d,t\rangle}\text{: } \mathbb{X} \subseteq \mathbb{D}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{D} - \mathbb{X})(\mathbb{D}')\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle dt,\langle dt,\langle\langle dt,dtt\rangle,\langle dt,t\rangle\rangle$

These operators will compose just as they did with determiner quantifiers, at least type-wise. Terminological proposal: let's refer to cases where *almost* directly modifies the comparative morpheme ***almost* comparatives**, equatives directly modified by *almost* ***almost* equatives**, so on for *barely*. The exception pronoun now must denote a set of degrees, rather than a set of individuals. Of course, we'll continue to assume that subtractives are only licensed within the scope of an **EXH** operator, and we form alternatives structurally via substituting the exception for expressions of the same type. We'll continue to restrict our attention to LFs with exactly one subtractive operator, and no other alternative-triggering expressions.

(221) $[\![\textbf{EXH}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\psi\left[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\right]$

(222) ***SUBTRACTIVE LFs***

A LF φ is a *subtractive LF* just in case φ is a LF containing exactly one subtractive operator Σ introducing an exception χ, and no other alternative-triggering expression, and φ is of type ⟨s,t⟩.

(223) ***SUBTRACTIVE ALTERNATIVES***

For a subtractive LF φ, with an exception χ,

$\textsc{alt}(\varphi) := \left\{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_d \ \&\ \varphi[\chi/\chi'] \text{ is grammatical}\right\}$

(224) For a subtractive LF φ,

$[\![\textbf{EXH}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}(w) \neq 1\right]$

In §3.2.1 we'll see that this semantics for *almost* (and its concomitant parts) predicts that it can modify the equative but not the comparative. We'll also discuss how even in intuitively discrete—i.e. not dense—domains of measurement, the perhaps counter-intuitive assumption that *all* scales referenced in natural language semantics are dense is necessary (Fox & Hackl 2006). Then, we'll discuss why some apparently good *almost* comparatives are misleading, and shouldn't be analyzed as *almost* comparatives at all. §3.2.2 will take a close look at *barely* comparatives and equatives. We'll see that our semantics for *barely* needs a tweak: instead of subtracting from the restrictor of the modified quantifier, we need to subtract from its scope, and we'll discuss how this does and doesn't change the analysis of subtractives and why we didn't notice it before.

3.2.1 *Almost* in action

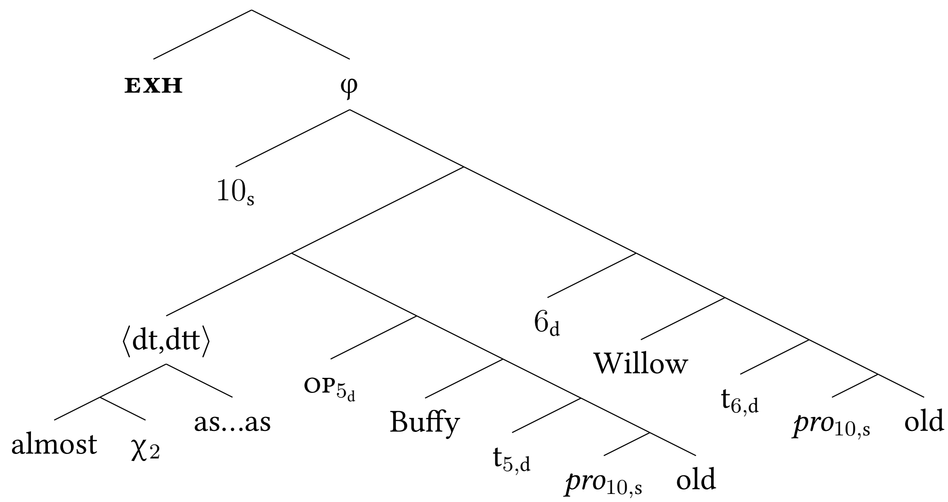
3.2.1.1 Equatives work

Let's put this to work with equatives. The semantics of the equative isn't exactly universal quantification, but it's not too far—for any two $\langle d,t\rangle$ functions D, D′, $[\![$as…as$]\!]$(D)(D′) yields truth just in case every degree d that makes D(d) true also makes D′(d) true. We've just added in maximality. It makes sense, prima facie, that *almost* should go with the equative, but we're in a new ontological domain, so we ought to be sure that the analysis does work.

Intuitively, (225) is true just in case Willow is younger than Buffy, but the difference between their ages is small. Another paraphrase: were Buffy a little less old, Willow and Buffy would be the same age. Given our assumptions, (225) has the LF in (225a), and we derive the truth conditions in (225b). (225bi) is the meaning of φ in w, and (225bii) requires that all non-entailed alternatives to φ be false or undefined in w.

(225) Willow is almost as old as Buffy is.

a.



b. $[\![(225a)]\!]^{g,c}(w)$ is defined only if $\textsc{max}\{(0,\mathcal{B}_w]-g(2)\}$ is defined, $\textsc{max}\{(0,\mathcal{W}_w]\}$ is defined, $g(2) \subseteq (0,\mathcal{B}_w]$, and $g(2) \neq \emptyset$; where defined:
$[\![(225a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) = 1

i. $\textsc{max}\{(0,\mathcal{B}_w]-g(2)\} \leq \mathcal{W}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

There are three main aspects of the meaning we need to work through to evaluate this properly: (i) the definedness conditions; (ii) the entailments between φ and its alternatives; (iii) the identity of g(2). In general, the logic proceeds as it did in the previous chapter, but let's go through these carefully in turn.

3.2.1.1.1 The definedness conditions of *almost* equatives The definedness conditions are contributed both by *almost* and *as…as*. First, *almost*'s subset presupposition requires that $g(2) \subseteq (0,\mathcal{B}_w]$; no degree in g(2) can exceed $\mathcal{B}_w$. It must also be non-empty. Let's grant that these hold. Second, given the semantics of the **max** operator proposed, for some $\mathbb{D}$, $\textsc{max}(\mathbb{D})$ is defined only if it is an initial interval of a scale that contains its upper bound. $\textsc{max}\{(0,\mathcal{W}_w]\}$ is defined, then.

For $\textsc{max}\{(0,\mathcal{B}_w]-g(2)\}$ to be defined, $(0,\mathcal{B}_w]-g(2)$ must be an initial interval of $\mathcal{S}_{\textsc{age}}$ that includes its upper bound. This means that g(2) must have the following properties, provided it's a non-empty subset of $(0,\mathcal{B}_w]$.

1. g(2) contains $\mathcal{B}_w$

2. g(2) does not include its lower bound

3. g(2) is convex

These are all necessary conditions. Let's go through each in turn.

Suppose $\left[1.\right]$ is false. $\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)$ contains $\mathcal{B}_{\text{w}}$, but then it is not convex, since g(2) removes degrees lower than $\mathcal{B}_{\text{w}}$. $\textbf{MAX}\left\{\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)\right\}$ is undefined. Now suppose $\left[2.\right]$ is false. Assume the lower bound of g(2) is *n*; its upper bound is $\mathcal{B}_{\text{w}}$; $\text{g}(2) = \left[\text{n},\mathcal{B}_{\text{w}}\right]$. $\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)$ is thus the totally open interval $\left(0,\text{n}\right)$, and $\textbf{MAX}\left\{\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)\right\}$ is undefined. Finally, suppose $\left[3.\right]$ is false. Since g(2) is not convex, $\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)$ is not convex, and so $\textbf{MAX}\left\{\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)\right\}$ is undefined.

The net result is that if g(2) is not empty, it *must* be a **final** subinterval of $\left(0,\mathcal{B}_{\text{w}}\right]$, and must not include its lower bound (i.e. it must be open at the bottom). This is a more general fact: if $\mathbb{D}^*$ is an initial interval of some scale $\mathcal{S}$ that includes its upper bound, and $\mathbb{X}$ is some non-empty set of degrees on $\mathcal{S}$, $\textbf{MAX}\left\{\mathbb{D}^*-\mathbb{X}\right\}$ will only be defined if $\mathbb{X}$ removes a final subinterval of $\mathbb{D}^*$, and $\mathbb{X}$ does not contain its lower bound. In the context of *almost* comparatives and equatives, subtraction must, and indeed can only, *lower* the threshold set by the subordinate clause. This is intuitively a good thing—(225) conveys that, were Buffy a bit younger, were we to lower her age, Willow would be as old as her.

3.2.1.1.2 Entailments

Here are the truth conditions of (225) again.

(225b) $[\![(225\text{a})]\!]^{\text{g,c}}(\text{w})$ is defined only if $\textbf{MAX}\left\{\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)\right\}$ is defined, $\textbf{MAX}\left\{\left(0,\mathcal{W}_{\text{w}}\right]\right\}$ is defined, $\text{g}(2) \subseteq \left(0,\mathcal{B}_{\text{w}}\right]$, and $\text{g}(2) \neq \emptyset$; where defined:

$[\![(225\text{a})]\!]^{\text{g,c}}(\text{w}) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}\left\{\left(0,\mathcal{B}_{\text{w}}\right]-\text{g}(2)\right\} \leq \mathcal{W}_{\text{w}}$ & $\textsc{small}_{\text{c}}(\text{g}(2))$

ii. $\forall\chi'\left[[\![\varphi]\!]^{\text{g,c}} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{\text{g,c}} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{\text{g,c}}(\text{w}) \neq 1\right]$

We know from the previous chapter that if *almost* modifies a necessarily true quantificational generalization φ, **EXH** is vacuous, and therefore neither **EXH** nor *almost* will be licensed. This pattern generalizes to degree constructions. If in all worlds w′, Buffy is at most as old as Willow is, i.e. $\forall\text{w}'\left[\mathcal{B}_{\text{w}'} \leq \mathcal{W}_{\text{w}'}\right]$, g(2) can pick out any small final subinterval of $\left(0,\mathcal{B}_{\text{w}}\right]$ that is lower-open and (225bi) will be true; all other sets will make it false or undefined. This means that the truth of (225bi) guarantees the truth of (225bii)—**EXH** adds nothing, and is not licensed. Neither is *almost*.

Let's assume, then, that it's not necessarily the case that Willow is at least as old as Buffy is. For a given value for g(2) that makes (225bi) true, which alternative χ′s are such that $[\![\varphi]\!]^{\text{g,c}}$, the intension of (225bi), entails $[\![\varphi[\chi_2/\chi']]\!]^{\text{g,c}}$? That depends, of course, on the entailment properties of *as...as*. Because of the definedness conditions imposed by the degree quantifier, many alternatives will be undefined; we need to find those alternatives whose presuppositions are satisfied in order to know whether they're entailed or not. $[\![\text{as...as}]\!]$ is Strawson Downward Entailing on the first, Strawson Upward Entailing on the second. Modulo the presuppositions rendering it Strawson entailing, this is just like *every*!

Let's show that this claim is right. Suppose first that for any world w, D_w and D'_w characterize initial intervals on a scale $\mathcal{S}$, and that in world w*, $[\![\text{as...as}]\!](D_{w*})(D'_{w*})$ is true. Let $\textbf{MAX}(\mathbb{D}_{w*}) = n$, and $\textbf{MAX}(\mathbb{D}'_{w*}) = m$. Now suppose that D''_{w*} characterizes an initial interval $\mathbb{D}''_{w*}$ of $\mathcal{S}$ such that $\mathbb{D}''_{w*}$ is a proper subset of $\mathbb{D}_{w*}$, and $\textbf{MAX}(\mathbb{D}''_{w*})$ is defined. It must be the case that $\textbf{MAX}(\mathbb{D}''_{w*}) <$ n, and so if $[\![\text{as...as}]\!](D_{w*})(D'_{w*})$ is true, then $[\![\text{as...as}]\!](D''_{w*})(D'_{w*})$ is true, too. If D''_{w*} characterizes a set $\mathbb{D}''_{w*}$ that is a proper superset of $\mathbb{D}_{w*}$, and it still characterizes an initial interval of $\mathcal{S}$, we cannot conclude that $[\![\text{as...as}]\!](D''_{w*})(D'_{w*})$ is true from the truth of $[\![\text{as...as}]\!](D_{w*})(D'_{w*})$. $\textbf{MAX}(\mathbb{D}''_{w*})$ $>$ m in this case, and it's entirely possible that it is greater than n, too. When both its arguments characterize initial intervals on a scale, $[\![\text{as...as}]\!]$ appears to be Strawson Downward Entailing on its first argument; it's easy to see that it will be Strawson Upward Entailing on its second.

An important note is warranted now. When considering whether an exception is small or not, because we're dealing with dense sets of degrees, the relevant measure function we use to determine the size of an exception cannot be cardinality. Any dense set of degrees has an infinite cardinality. That's fine, though; we never meant for cardinality to be the only measure used by the size constraint.

(226) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, μ_c be a measure function, determined by the context, mapping sets to numerical values representing their size, and n_c be a contextually determined numerical threshold for size.

$\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\mu_c(\mathbb{X}) \leq n_c$.

The contextually-determined measure function needs to be one with which we can check the relative size of an interval with respect to another. Perhaps it's something like LENGTH, measuring an interval, mapping it to a new scale, that allows a crisp comparison of size. We can do some revisions:

(227) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, and n_c be a contextually determined numerical threshold for size.

Where $\mathbb{X}$ is a dense set of degrees, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\textsc{length}(\mathbb{X}) \leq n_c$;

Otherwise, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $|\mathbb{X}| \leq n_c$

(228) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

An utterance of φ in c is felicitous only if, for any context c:

Where $\mathbb{A}$ is a dense set of degrees, $n_c < \frac{1}{2}(\text{LENGTH}(\mathbb{A}))$;

Otherwise, $n_c < \frac{1}{2}|\mathbb{A}|$

As far as I can tell, this should be sufficient for our purposes.

Now let's get back to the *almost* equative we're trying to analyze.

(225b) $[\![(225a)]\!]^{g,c}(w)$ is defined only if $\textbf{MAX}\{(0,\mathcal{B}_w]-g(2)\}$ is defined, $\textbf{MAX}\{(0,\mathcal{W}_w]\}$ is defined, $g(2) \subseteq (0,\mathcal{B}_w]$ and $g(2) \neq \emptyset$; where defined:
$[\![(225a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}\{(0,\mathcal{B}_w]-g(2)\} \leq \mathcal{W}_w$ & $\text{SMALL}_c(g(2))$
ii. $\forall \chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \right]$

Here's the updated state of affairs.

1. $g(2)$ must be a non-empty, lower-open, convex, final subinterval of $(0,\mathcal{B}_w]$ *definedness conditions*
2. $g(2)$ must be small *truth conditions*
3. $[\![\text{as...as}]\!]$ is Strawson Downward Entailing with respect to the subordinate clause; φ is, too *subordinate clause is an initial interval*

Given these facts, here's what we know about alternative χ's.

4. For all χ' that lack at least one property in [1.] or [2.], $[\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1$
5. For all χ' that have *all* properties in [1.] and [2.], and $g(2) \Rightarrow_{ST} \mathbb{X}'$, $[\![\varphi]\!]^{g,c} \subseteq [\![\varphi[\chi_2/\chi']]\!]^{g,c}$

[4.] should be obvious—the same properties that constrain appropriate values for g(2) constrain alternative χ's. [5.] follows from [3.]. Since φ is Downward Entailing with respect to the restrictor of *as...as*, small supersets of the exception will be entailed, provided they yield definedness. (Those supersets that aren't small are covered under [4.].) Now, are those $\mathbb{X}'$ that are non-supersets of g(2), but which have all the properties in [1.], entailed? We can exclude any non-subsets from the non-supersets we're investigating; none of them can be final subintervals of $(0,\mathcal{B}_w]$, and so none, when subtracted from will $(0,\mathcal{B}_w]$, yield an interval for which **MAX** is defined. All we need to know, then, is whether or not subsets of g(2) are entailed.

The answer is no. For any non-empty g(2) that has all the properties in [1.], there will always be a world w^* where for some subset $\mathbb{X}'$ of g(2) which also has all the properties in [1.], $\textbf{MAX}\{(0,\mathcal{B}_{w*}]-g(2)\} \leq \mathcal{W}_{w*}$ is true and $\textbf{MAX}\{(0,\mathcal{B}_{w*}]-g(2)\} \leq \mathcal{W}_{w*}$ is false. This is easy to see:

let w^* be such that $g(2) = \big(0{,}\mathcal{B}_{w*}\big] - \big(0{,}\mathcal{W}_{w*}\big]$; a subset $\mathbb{X}'$ of g(2) that yields definedness will have a lower bound greater than the lower bound of g(2), and so $\textsc{max}\big\{\big(0{,}\mathcal{B}_{w*}\big] - g(2)\big\} > \mathcal{W}_{w*}$. For a given g(2) that makes the pre-exhaustified meaning of an *almost* equative true, $\mathbb{X}'$s that pick out a proper subset of g(2) will not be entailed, and hence to know if the exhaustified meaning is true, we must verify that all such subsets aren't true. Everything else is already entailed or not true.

3.2.1.1.3 The exception is the difference Knowing all this, here's what must be the case in order for (225a) to be true. If non-empty, g(2) must be a small, final subinterval of $\big(0{,}\mathcal{B}_{w*}\big]$ subtracting enough from $\big(0{,}\mathcal{B}_{w*}\big]$ so that $\textsc{max}\big\{\big(0{,}\mathcal{B}_{w*}\big] - g(2)\big\} \leq \mathcal{W}_{w*}$; **EXH** requires that all non-entailed alternatives make $\textsc{max}\big\{\big(0{,}\mathcal{B}_{w*}\big] - \mathbb{X}'\big\} \leq \mathcal{W}_{w}{}^*$ false or undefined.

Suppose that $g(2) = \emptyset$. In such a context, if (225bi) is true, then (225bii) is true. All alternatives $\mathbb{X}'$ are supersets of $\emptyset$, so if they're small, they're entailed, and if they're not, they satisfy **EXH**. If (225bi) is false, then of course (225b) is false. So, $[\![\varphi]\!]^{g,c}$ entails $[\![\textbf{EXH}\ \varphi]\!]^{g,c}$—given the economy constraint on **EXH**, the latter is not licensed, and neither is *almost*. So, g(2) must be non-empty. Now, if $\mathcal{B}_{w*} \leq \mathcal{W}_{w*}$, every defined alternative is still true, but not every one is Strawson entailed. In particular, subsets of g(2) will not be entailed, and so **EXH** will be falsified by those subsets. Not only must g(2) be non-empty, then, but in the world of evaluation w^*, it must be that $\mathcal{B}_{w*} > \mathcal{W}_{w*}$.

In such a world w^*, g(2) must contain the difference between $\big(0{,}\mathcal{B}_{w*}\big]$ and $\big(0{,}\mathcal{W}_{w*}\big]$, i.e. $\big(\mathcal{W}_{w*}{,}\mathcal{B}_{w*}\big]$, and that difference must be small relative to the context. (225bi) will then be true. (225bii) requires that no subset of g(2) also make (225bi) true; this will only be the case if $g(2) = \big(\mathcal{W}_{w*}{,}\mathcal{B}_{w*}\big]$. Should g(2) subtract more than that, i.e. g(2) picks out a proper superset of $\big(\mathcal{W}_{w*}{,}\mathcal{B}_{w*}\big]$, (225bi) will true provided it's still small, but then there is a non-entailed subset that also makes (225bi) true, namely $\big(\mathcal{W}_{w*}{,}\mathcal{B}_{w*}\big]$. **EXH** will yield falsity, then. The exception must pick out the difference between Buffy and Willow's ages.

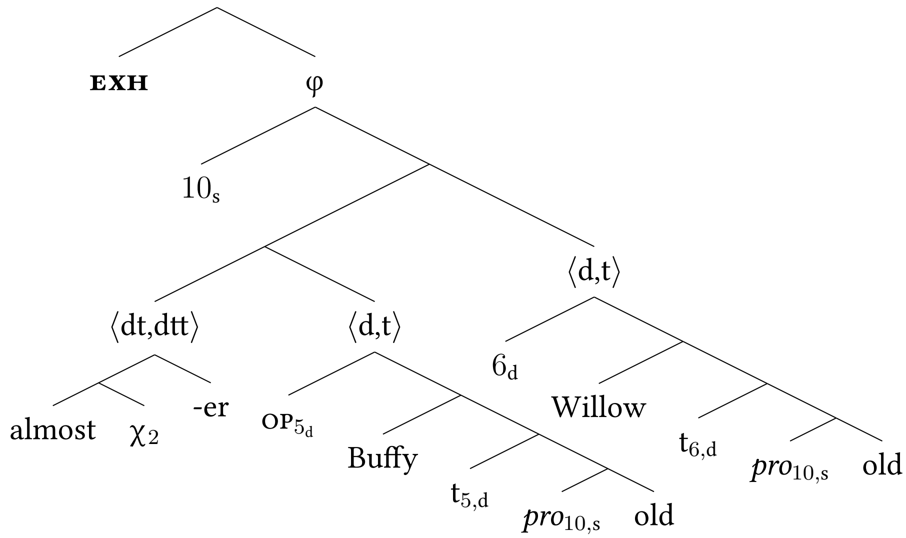
Let's zoom out for a moment. The analysis, or at least the discussion of it, runs a little on the technical side, but the key aspects should be clear. In an *almost* equative with the exception introduced by the subtractive must be a small, final subinterval of the interval provided by the subordinate clause. Anything else yields undefinedness. The exception must be equivalent to the difference between the intervals supplied to *as…as* by subordinate and the main clauses; by 'lowering' the subordinate clause interval, we make the modified equative true. The exception, that difference between two intervals, must be small, and it is all that stands between us and the truth of the equative. So, *almost* says, get rid of it! Were we to remove any less than that difference, it wouldn't be enough; any more, and it'd too much. The difference is just right. As we turn to *almost* comparatives, and then on to *barely*, these lessons are important. What matters in subtractive-modified degree quantification is the difference between two intervals.

3.2.1.2 Comparatives don't work!

The *almost* comparative in (229) is extremely odd, and our analysis predicts this.

(229) # Willow is almost older than Buffy is.

a.



b. $[\![(229a)]\!]^{g,c}(w)$ is defined only if $\textsc{max}\{(0,\mathcal{B}_w]-g(2)\}$ is defined, $\textsc{max}\{(0,\mathcal{W}_w]\}$ is defined, $g(2) \subseteq (0,\mathcal{B}_w]$, and $g(2) \neq \emptyset$; where defined:
$[\![(229a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textsc{max}\{(0,\mathcal{B}_w]-g(2)\} < \mathcal{W}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

The key difference between this semantics and that of the *almost* equative discussed in the previous section is, of course, that the comparative requires a strict ordering between two degrees. The exception must lower Buffy's age to a point strictly less than Willow's age. As it turns out, that's enough to make it impossible for (229b) to be true. Let's see why.

The logic discussed in the previous section applies here, so we'll go over it briefly. If it is necessarily true that Willow is older than Buffy is, then the truth of (229bi) entails the truth of (229bii), and therefore **EXH**, and by consequence *almost*, will not be licensed; we grant, then, that there is no such necessity. Like *as...as*, *-er* is Strawson Downward Entailing on its first argument and Strawson Upward Entailing on its second, and so it will be Strawson Upward Entailing on the exception. As before, g(2) must be a lower-open, final subinterval of $(0,\mathcal{B}_w]$ for $\textsc{max}\{(0,\mathcal{B}_w]-g(2)\}$ to be defined; for a given assignment for g(2) that makes (229bi) true, alternatives formed with small supersets of g(2), provided they are also lower-open, final subintervals of $(0,\mathcal{B}_w]$ for $\textsc{max}\{(0,\mathcal{B}_w]-g(2)\}$ will be entailed. There are only subsets left once we consider the definedness conditions, and they aren't entailed. If $g(2) = \emptyset$, then again, all alternatives are small and entailed, or not small and false. The truth of (229bi) thus guarantees the truth of (229bii), and **EXH** has nothing to add. It's vacuous, not licensed, and neither is *almost.* The exception must be non-empty.

What must be the case, then, is that Buffy is at least as old as Willow in the world of evaluation. g(2) must be small in the context, but still subtract a sufficiently large, lower-open, final subinterval from $\left(0,\mathcal{B}_{\mathrm{w}}\right]$ so that the biggest degree in the remainder is strictly less than $\mathcal{W}_{\mathrm{w}}$. Now, if $\mathcal{B}_{\mathrm{w}} = \mathcal{W}_{\mathrm{w}}$, then g(2) must remove $\mathcal{B}_{\mathrm{w}}$, but it can't remove *only* $\mathcal{B}_{\mathrm{w}}$. The remainder would be an open interval: $\left(0,\mathcal{B}_{\mathrm{w}}\right] - \left[\mathcal{B}_{\mathrm{w}},\mathcal{B}_{\mathrm{w}}\right] = \left(0,\mathcal{B}_{\mathrm{w}}\right)$. That won't do for **MAX**. g(2) must go further: for some small ε, $g(2) = \left(\mathcal{B}_{\mathrm{w}}-\varepsilon,\mathcal{B}_{\mathrm{w}}\right]$. The same goes if $\mathcal{B}_{\mathrm{w}} > \mathcal{W}_{\mathrm{w}}$; it must remove $\mathcal{W}_{\mathrm{w}}$ from $\left(0,\mathcal{B}_{\mathrm{w}}\right]$, and go a little farther, for the remainder to be strictly lower than $\mathcal{W}_{\mathrm{w}}$ and still defined.

That we have to go a bit farther than $\mathcal{W}_{\mathrm{w}}$ is the problem when we consider (non)-entailment. We discussed in the preliminaries the assumption that scales are *dense*:

(197) Let $\mathcal{S}$ be a set of degrees, μ be a dimension of measurement, and $\leq_\mu$ be an ordering on $\mathcal{S}$ relative to μ. $\mathcal{S}$ is a *scale* just in case $\forall d,d' \in \mathcal{S}$:

a. $d \leq_\mu d' \vee d' \leq_\mu d$ *total order*

b. $d <_\mu d' \rightarrow \exists d''\left[d'' \in \mathcal{S}\ \&\ d <_\mu d'' <_\mu d'\right]$ *density*

This is intuitive when we consider ages: age is dependent on time, and surely between any two points in time is another point in time, so between any two ages is another, distinct age. This is in fact the problem for *almost* comparatives. g(2) needs to remove enough from $\left(0,\mathcal{B}_{\mathrm{w}}\right]$ so that its maximal element is *just below* $\mathcal{W}_{\mathrm{w}}$. For some small ε, $g(2) = \left(\mathcal{W}_{\mathrm{w}}-\varepsilon,\mathcal{B}_{\mathrm{w}}\right]$. This has a proper subset that also removes enough from $\left(0,\mathcal{B}_{\mathrm{w}}\right]$ so that the maximal element is strictly less than $\mathcal{W}_{\mathrm{w}}$: $\left(\mathcal{W}_{\mathrm{w}}-\frac{\varepsilon}{2},\mathcal{B}_{\mathrm{w}}\right]$, but since subsets aren't entailed, the alternative formed with the $\mathbb{X}' = \left(\mathcal{W}_{\mathrm{w}}-\frac{\varepsilon}{2},\mathcal{B}_{\mathrm{w}}\right]$ one will falsify (229bii). For any assignment for g(2) that makes (229bi) true, there will be a non-entailed subset that makes (229bii) false. (229b) is necessarily false, or else *almost* is not licensed. The analysis cannot predict truth for the *almost* comparative.

3.2.1.3 The argument for Universal Density

While $\mathcal{S}_{\text{AGE}}$ is intuitively a dense scale, that's not necessarily the case across the board. Cardinalities, for example, seem discrete: if one asks exactly how many vampires Buffy killed last night, the answer comes from the set of natural numbers. We can't say she killed 34.2 vampires. Fox & Hackl (2006) argue that in fact, $\mathcal{S}_{\text{CARDINALITY}}$ is dense as well. More generally, they propose that all measurement scales in natural language semantics are dense; this is the Universal Density of Measurement hypothesis. (UDM). We'll show now that if $\mathcal{S}_{\text{CARDINALITY}}$ were a discrete scale, we should predict examples like (230) to be acceptable, contra to fact.

(230) # Willow killed almost more vampires than Buffy did.

We'll argue, then, our analysis of *almost* provides support for the UDM in a novel way.

We need to add into our compositional semantics a means of measuring cardinalities. Following Hackl (2000, i.a.), we'll take the degree-introducing morphology in (230) to be the covert operator MANY.

(231) $[\![\textsc{many}]\!] = \lambda d \lambda A \lambda B.\ \exists x\big[x \in \mathbb{A}\ \&\ x \in \mathbb{B}\ \&\ |x| \geq d\big]$ $\langle d,\langle et,ett\rangle\rangle$

The cardinality function, let's assume, counts the number of atomic elements in a plurality and maps it an element in the set of the natural numbers $\mathbb{N}$, i.e. $\{1, 2, 3, ...\}$. This scale is discrete. For simplicity, we'll be assume that *kill* expresses a world-dependent relation between individuals, with its first argument being the killed and its second the killer.

(232) $[\![\text{kill}]\!] = \lambda w \lambda x \lambda y.\ \textsc{kill}_w(x)(y)$

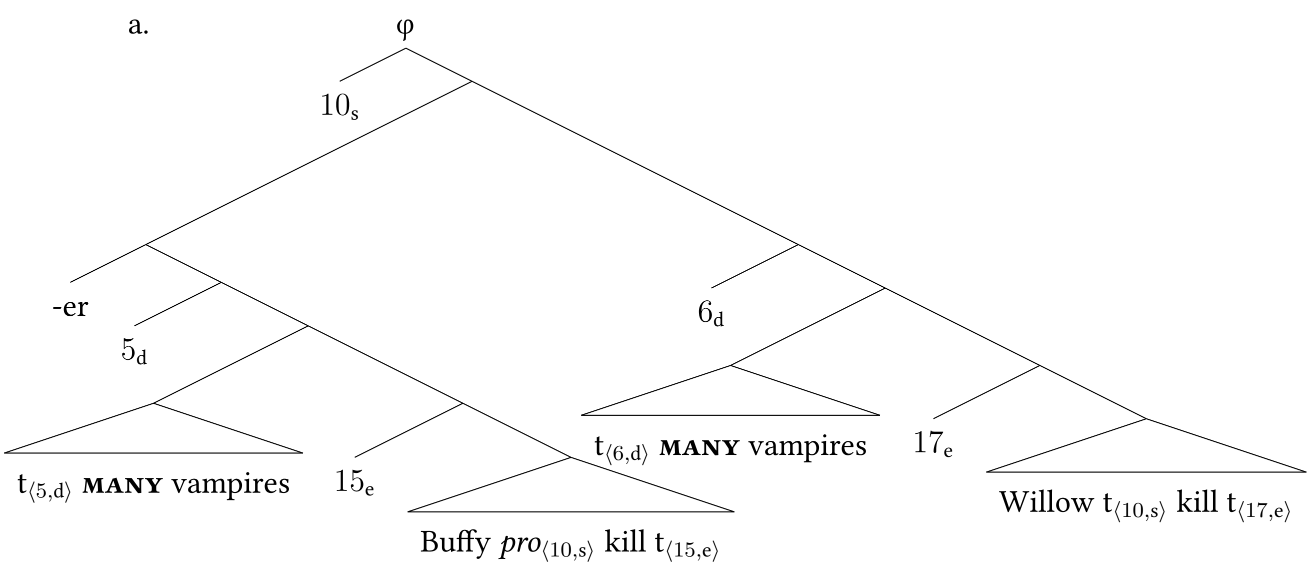
(233) a. $\mathbb{B}_w = \big\{d \mid \exists x\big[\textsc{vampire}(x)\ \&\ \textsc{kill}_w(x)(\text{Buffy})\ \&\ |x| \geq d\big]\big\}$

b. $\mathbb{W}_w = \big\{d \mid \exists x\big[\textsc{vampire}(x)\ \&\ \textsc{kill}_w(x)(\text{Willow})\ \&\ |x| \geq d\big]\big\}$

Now, putting this together with the comparative semantics assumed:

(234) Willow killed more vampires than Buffy did.

a.



b. $[\![\varphi]\!](w) = 1$ *iff* $\textsc{max}(\mathbb{B}_w) < \textsc{max}(\mathbb{W}_w)$

i.e. the number of vampires Buffy killed < the number of vampires Willow killed

The truth conditions are straightforward. Now let's add in *almost*.

(235) # Willow killed almost χ_2 more vampires than Buffy did.

a. $[\![(235)]\!]^{g,c}(w)$ is defined only if $\textbf{\textsc{max}}(\mathbb{B}_w - g(2))$ is defined, $\textbf{\textsc{max}}(\mathbb{W}_w)$ is defined, $g(2) \subseteq \mathbb{B}_w$, and $g(2) \neq \emptyset$; where defined: $[\![(235)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) = 1

i. $\textbf{MAX}(\mathbb{B}_w - g(2)) < \textbf{MAX}(\mathbb{W}_w)$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \right]$

All the now-familiar claims apply: if the number of vampires killed by Buffy is necessarily less than that killed by Willow, *almost* isn't licensed; if $g(2) = \emptyset$ and (235ai) is true, then *almost* isn't licensed, etc. Let's assume, then, that Buffy killed more vampires. For concreteness, let's say she killed one hundred and two, and Willow killed ninety-nine; for $\textbf{MAX}(\mathbb{B}_w - g(2))$ to be defined, g(2) still be a final subinterval of $\mathbb{B}_w$, even though it's not a dense set. There is, in this context, a natural candidate for g(2):

(236) $g(2) = \left\{ d \in \mathbb{N} \mid 102 \geq d \geq 99 \right\}$

(237) $\mathbb{B}_w - g(2) = \left\{ d \mid 98 \geq d \right\}$

No subset of g(2) that is a final subinterval of $\mathbb{B}_w$ is entailed, but because $\mathbb{N}$ is discrete, there are no such subsets that falisify **EXH**, either. None of them will lower $\mathbb{B}_w$ below 99. If the cardinality measure function maps to $\mathbb{N}$, a discrete scale, then we can't rule out this *almost* comparative. If the scale is dense, however, then we do predict this to be unacceptable—to yield an interval for which **MAX** is defined, g(2) must be lower-open. It must 'go below' ninety-nine by some small amount. Then, however, we'll run into the density issue again, but here, that's a good thing. Density allows us to rule out *almost* comparatives, even in intuitively discrete domains. To the extent that this analysis of *almost* is independently correct, this serves as an argument for the UDM—even cardinalities must be mapped to a dense scale[7].

3.2.1.4 Why are some *almost* comparatives okay?

Imagine that we're hosting a potluck, and every guest brings a dish, maybe even two. When all guests have arrived, and all the platters and casseroles and salads bowls are on the table, we think to ourselves that this is quite a lot—it seems there is no more room on the table. So we say to our guests:

[7]A careful discussion of Fox & Hackl's arguments in favor of the UDM would take us too far afield, but we can make some brief comments on it here. They are interested in part in explaining the contrast between (i) and (ii).

(i) # How many children does Giles not have?

(ii) How many children is Giles required to not have?

The basic idea is that *how many* questions ask for the maximally informative degree in a set. (i) is odd because there is no such degree—negation takes an upper-closed, initial interval of a scale and yields a lower-open, final interval of a scale. As a result, there is no maximally informative degree in the set of children that Giles doesn't have. (ii) is good because universal modals yield closed intervals from open ones; the resulting set has a maximally informative member, the smallest number. This is all brought up because one might reasonably ask whether modals obviate the density problem in *almost* comparatives. It's not obvious that this is relevant: the locus of the density problem here is not an open interval that can be closed by a modal, the issue is that we cannot find a unique exception. However, more careful work is needed to show this; this is left for the future.

(238) There is almost more food here than we can fit on the table.

This seems like a natural and felicitous sentence in this context, and our guests laugh and agree and the potluck starts in earnest. What we're saying is that were things different, if we had a bit more food, or a bit less space on the table, the amount of food we have wouldn't fit, but it does, in the end, fit. It appears that *almost* is modifying the comparative morpheme *more*. Other natural examples are given in (239).

(239) a. This metal rod is almost longer than it can be to fit in this gap.

b. Oz is almost taller than his father is.

(239a) conveys that, were the rod longer than it in fact is, it wouldn't fit, but as it happens, it does fit. (239b) conveys that Oz's height is close to, but not actually taller than, his father's. Again, in both cases it seems that we have *almost* modifying the comparative. Nonetheless, our theory predicts that this cannot be the case—*almost* modifying a comparative cannot yield truth on our analysis. So what's going on in these cases? Why are these *almost* comparatives acceptable?

What we'll argue is that these are not genuine instances of *almost* comparatives. They are the product of a misleading surface form. In each of these cases, we have a use of *almost* that is structurally more akin to (240b) than (240a).

(240) a. Buffy is almost as tall as Willow is.

b. Willow is almost a master of witchcraft.

Our analysis derives accurate truth conditions for (240a) through *almost* modifying, directly, the equative morpheme. We don't yet have an analysis of (240b). That said, there are good reasons to think that in (238) and (239), *almost* cannot be modifying the comparative directly, but instead, modifies something else—perhaps whatever it modifies in (240b). Certain contexts unexpectedly eliminate the possibility of *almost* modification. Tense, too, can affect the felicity of the utterance in a way that is not really expected. Perhaps most convincing is syntactic evidence suggesting that *almost* is not directly modifying these morphemes. We will discuss these in turn.

For (238) we set up a potluck context; for (239), we can clarify the context a little. Imagine we're playing a game in which we compete with metal rods. There is a range of admissible lengths for our metal rods—they must be, say, between 2 and 2.2 meters long. If the particular metal rod under consideration is 2.2 meters long, it seems that (239) is true. The metal could be malleable; we could stretch it just a hair and it would be longer than 2.2 meters. Now imagine that the metal we are using to fit in the gap is not malleable at all—once cut, it cannot be stretched without shattering. In this context, (239) is degraded, but this is unexpected if *almost* directly modifies *more*. Such modification only cares about lengths–it should be true and felicitous simply because the length of the rod is close to too long. It's unclear how, or why, malleability should play a role.

A parallel point can be made for (241a). At least one context in which it is true is if Oz is a growing boy, and it's looking like he'll end up taller than his father. It seems like the fact that he's growing is quite relevant. If we remove this kind of possibility from these examples, they too degrade. To see this, consider (241b).

(241) a. A growing boy, Oz is almost taller than his father is.

b. ?? No longer growing, Oz is almost taller than his father is.

(241a) is acceptable; *a growing boy* brings out the interpretation that Oz's height could end up exceeding his father's height. However, when we use *no longer growing*, as in (241b), the sentence is degraded. By eliminating the possibility that Oz could end up taller than he is currently through the modifier *no longer growing*, we reduce the felicity of the examples. This is unexpected, though, if *almost taller* only cares about a comparison—what bearing does the possibility or impossibility of future growth have on that? These modifiers don't disrupt *almost* across the board. Consider the felicity of *almost* equatives.

(242) a. A growing boy, Oz is almost as tall as his father is.

b. No longer growing, Oz is almost as tall as his father is.

Neither forcing nor eliminating the inference that Oz's height could change in the future affects the felicity of these examples; they do simply seem like a comparison of heights. This is what we would expect if *almost* directly modifies the equative morpheme.

The acceptability of *almost* in a comparative is also degraded through tense manipulation.

(243) a. At 16 years old, Oz is almost taller than his father is.

b. ?? At 16 years old, Oz was almost taller than his father was.

Fixing Oz's age through the modifier *at 16 years old*, (243) is acceptable. However, when we shift to the past tense, suddenly it becomes degraded. Again, though, it's not clear why this should matter, and this effect doesn't arise with the equative.

(244) a. At 16 years old, Oz is almost as tall as his father is.

b. At 16 years old, Oz was almost as tall as his father was.

Tense here doesn't affect the acceptability of *almost*. It is unclear why this should be the case with the comparative construction, but not the equative construction, if in all cases we have *almost* directly modifying a degree quantifier comparing particular intervals. It would be expected, however, if tense does affect *almost* modifying a VP—McKenzie & Newkirk (2019) show that this is borne out.

(245) a. ?? Oklahoma is almost two states.

b. Oklahoma was almost two states. McKenzie & Newkirk (2019:24-25)

Almost's contribution in these cases is restricted by tense. (245a) is odd, except perhaps on an interpretation where we're discussing the sheer size or sociopolitical divides in Oklahoma. With past tense, it improves, and it can mean that the land that is currently the state of Oklahoma could have ended up as two distinct states. This pattern is different than the one we saw above, but to the extent that its similar, it speaks to the possibility that *almost* doesn't directly modify *-er* above, but rather the entire clause. Deriving the interpretations that arise in these constructions, and showing why they are blocked through certain modifiers and tense configurations, is a topic for another chapter.

Here's a final argument. Numbers can be compared in terms of size.

(246) a. Eight is bigger than seven is.

b. Seven is almost as big as eight is.

These both seem like felicitous comparisons, but *almost* cannot appear in the comparative.

(247) # Seven is almost bigger than eight is.

This strikes us as incredibly odd. The paraphrases above for the acceptable *almost* comparatives relied on a notion of counterfactuality: were things different, the comparatives would (or could) hold. It's hard to imagine how seven could be bigger than eight. It's a mathematical fact that it's smaller! Even though it is plausibly close in size to eight, there's simply no way for there to be a world in which seven is, in fact, bigger than eight.

Implicit in the discussions above is the idea that the syntax is misleading the syntactic position of *almost*—its linear order obfuscates its structural position. Simplifying the structures, the claim is that *almost* is modifying the entire VP, rather than *more* and *too* directly.

(248) Oz is almost taller than his father is.

a. [Oz_1 [is [almost [$_{\text{VP}}$ t_1 taller than his father]]]]

b. *[Oz_1 [is [$_{\text{VP}}$ t_1 [[almost taller] than his father]]]]

Part of the problem is that the location of *almost* is obfuscated by the high position of the subject and the presence of the auxiliary. In fact, the only time *almost* seems to precede an auxiliary where it is present is when it directly modifies a quantificational determiner.

(249) Almost all the students are in the library.

Here, *almost* modifies *all*. In this position, the only interpretation that is possible is that the vast majority of the students are located in the library. This cannot mean that all of the students are nearly in the library—for example, if they are en route and are practically there. That is interesting, and it turns out that *all*-stranding patterns similarly.

(250) a. The students are almost all in the library.
b. All the students are almost in the library.
c. * Almost the students are all in the library.

(250a) can only mean that the majority of students are in the library. (250b), however, only has the interpretation that every last student is nearly in the library—exactly the interpretation (249) and (250a) lack. (250c) shows that you cannot move *almost* and the NP independent of *all*. If you move *all*, *almost* must come with; if you leave it in situ, so must *almost* remain in situ. Of course, when *almost* goes with *all*, it only allows the typical quantifier-modifier interpretation, including in SpecTP.

This fact about *almost* and SpecTP is revealing. *There*-constructions are revealing as well. Both *almost more* and *more* are permissible in *there*-constructions, but it turns out that while *more* can surface in SpecTP, *almost more* cannot.

(251) a. There are more people in my house than I invited.
b. There are almost more people in my house than I invited.

(252) a. More people are in my house than I invited.
b. # Almost more people are in my house than I invited.

Why should this be, then? Well, if *almost* directly modifies *more*, when the latter moves to SpecTP in (252), *almost* should tag along, given what we've seen. However, this isn't the case. Now, if *almost* does not and cannot directly modify *more*, it's no surprise that *almost more* cannot surface in SpecTP given the data in (250). In (251b), then, it must be modifying the VP.

Locative inversion as in (251) and (252) isn't necessary to make this point—where *more people* can occur as the subject of an accomplishment verb, *almost more people* is degraded.

(253) a. More people signed up for this course than I expected (so we'll have to kick some out).
b. ?? Almost more people signed up for this course than I expected (so we'll have to stop letting people enroll).

Here, again, *almost more people* is not a great subject, but it's unclear why this should be if *more people* can surface in this position. These data, then, are fairly striking evidence in support of the idea that *almost* doesn't directly modify *more*. While we haven't constructed similar examples yet with the comparative, and further evidence is under construction, this points to the idea that there is more than meets the eye is going on when we see *almost more*. Coupled with the effects of context and tense on the acceptability of *almost more*, we make a strong case in support of our theory, which in fact predicts that *almost* does not modify these degree operators directly, even when it appears to.

3.2.1.5 *Almost* requires subtractive alternatives, finally

In the previous chapter, we argued that *barely* necessitated the 'subtractive' alternatives proposed for exceptive *but*, as defined below.

(254) ***SUBTRACTIVE ALTERNATIVES***

For a subtractive LF φ, with an exception χ,

$\textbf{ALT}(\varphi) := \{\varphi[\chi/\chi'] \mid \mathscr{X}' \subseteq D_d \;\&\; \varphi[\chi/\chi'] \text{ is grammatical}\}$

$\textbf{ALT}(\varphi) = \{\varphi[\chi/\chi'] \mid \mathscr{X}' \subseteq D_d \;\&\; \varphi[\chi/\chi'] \text{ is grammatical}\}$

Spector (2014) and Crnič (2018), though not dealing with *barely*, assumed that the lone alternative to some φ containing *almost* was just that same φ without *almost*. That wouldn't do for *barely*; we needed these subtractive alternatives. We made the assumption that *almost* invokes the same alternatives on the basis of parsimony, but there was no positive evidence at that time. *Almost* did just fine without them.

Comparatives and equatives now provide that positive evidence. Without subtractive alternatives, we could not predict the contrast between *almost* equatives and comparatives. Suppose, then, that the alternative to an *almost* equative is just the unmodified equative, and the same for a comparative. Neither the *almost* equative nor the *almost* comparative entails their alternative, so **EXH** requires that it not be true. The truth conditions are given below (presuppositions are suppressed, but assumed).

(255) Willow is almost as old as Buffy is.

a. $\textbf{MAX}\{(0,\mathcal{B}_w] - g(2)\} \leq \mathcal{W}_w \;\&\; \textsc{small}_c(g(2)) \;\&\; \textbf{MAX}\{(0,\mathcal{B}_w]\} > \mathcal{W}_w$

(256) # Willow is almost older than Buffy is.

a. $\textbf{MAX}\{(0,\mathcal{B}_w] - g(2)\} < \mathcal{W}_w \;\&\; \textsc{small}_c(g(2)) \;\&\; \textbf{MAX}\{(0,\mathcal{B}_w]\} \geq \mathcal{W}_w$

Both sets of truth conditions are perfectly coherent. The *almost* equative requires that subtracting g(2) from Buffy's age yield a degree at least as low as Willow's age, but Willow is in fact younger than Buffy. That's fine. The *almost* comparative requires that g(2) subtract enough from Buffy's age to yield a degree lower than Willow's age, but Buffy is at least as old as Willow is. That's fine, too, or at least, it's not necessarily contradictory. It's perfectly coherent! As we've seen, though, subtractive alternatives *do* make a clear distinction between the two cases.

Another alternative might be to invoke Horn scale-mates for substitution: *almost* comparatives have *almost* equatives as alternatives and vice versa. This is essentially the "scalar" view discussed briefly in the previous chapter. On the kind of view argued for by Hitzeman (1992), Penka (2006), a.o. *almost* is a propositional operator; we know that won't work given the arguments by Gajewski (2008), et seq., and *almost* degree constructions make matters worse.

(257) $[\![\text{almost}_{\text{scalar}}]\!] = \lambda p_{\langle s,t\rangle}\lambda w.\ \neg p(w)\ \&\ \exists q\big[q \in \textbf{ALT}(p)\ \&\ q(w)\big]$

Now, the alternatives to *Willow is almost as old as Buffy is* would be generated by swapping the equative for a Horn-scale mate, i.e. the comparative. However, since the comparative is logically stronger than the equative, the negation of the equative entails the negation of the comparative, and there is no weaker alternative that can be asserted true. Such an analysis would predict that *Willow is almost as old as Buffy is* would be false necessarily. To make matters worse, it would predict that *almost* comparatives are good across the board, since the comparative has a weaker alternative—the equative—that can be substituted for it.

Amaral & Del Prete (2010) give a more flexible kind of meaning, one that is in ways more akin to ours, and reject the appeal of Horn scales in determining alternatives, but the basic components aren't all that different from Hitzeman and Penka. Theirs is a focus-sensitive operator; its first argument is a pair of the semantic value of *almost*'s associate (*P*) and the contextually determined alternatives to the second argument (*S*), and its second argument (*x*) is the semantic value of the part of the sentence not in focus.

(258) $[\![\text{almost}_{\text{A\&dP}}]\!] = \lambda\langle P_{\langle\gamma,t\rangle},\mathscr{S}\rangle\lambda x_{\gamma}.\ \neg P(x)\ \&\ \exists Q_{\gamma,t} \in \mathscr{S}\big[Q <_{S} P\ \&\ \textsc{close}_{S}(Q,P)\ \&\ Q(x)\big]$

All that this semantics does is loosen the requirements on what counts as an alternative, leaving it up to the context to determine *S*. It's not clear at all how this could be leveraged to predict the distribution our analysis does. Subtractive alternatives are, in the end, necessary for *almost.* It takes looking closely at degree constructions, and specifically comparatives and equatives, to see that.

3.2.1.6 Mix 'n' MAX

Our **MAX** operator is really quite restrictive when it comes to definedness.

(259) For any set of degrees $\mathbb{D}$ on a scale $\mathcal{S}$:
$\textbf{MAX}(\mathbb{D}) := \iota d\big[d \in \mathbb{D}\ \&\ \forall d'\big[\big[d' \in \mathcal{S}\ \&\ d' \leq d\big] \leftrightarrow d' \in \mathbb{D}\big]\big]$

This operator is only defined for a set of degrees $\mathbb{D}$ if $\mathbb{D}$ is an initial interval on a scale:

(260) For any interval $\mathbb{D}$ and scale $\mathcal{S}$ such that $\mathbb{D} \subseteq \mathcal{S}$:
$\mathbb{D}$ is an *initial* interval of $\mathcal{S}$ *iff* $\exists \mathrm{d}\big[\mathrm{d} \in \mathbb{D}\ \&\ \forall \mathrm{d}'\big[\big[\mathrm{d}' \in \mathcal{S}\ \&\ \mathrm{d}' \leq \mathrm{d}\big] \leftrightarrow \mathrm{d}' \in \mathbb{D}\big]\big]$

This is stronger than a more standard **max** operator like Rullmann's (1995).

(261) $\textbf{MAX}_{\mathrm{R(ullmann)}}(\mathbb{D}) := \iota \mathrm{d}\big[\mathrm{d} \in \mathbb{D}\ \&\ \forall \mathrm{d}'\big[\mathrm{d}' \in \mathbb{D} \rightarrow \mathrm{d} \geq \mathrm{d}'\big]\big]$

This operator doesn't require $\mathbb{D}$ to be convex or to be an initial interval of a scale, so it is defined for many sets that our **max** is not. Like our **max**, though, $\textbf{MAX}_{\mathrm{R}}$ does require is that the upper bound of $\mathbb{D}$ be included in $\mathbb{D}$.

Utilizing $\textbf{MAX}_{\mathrm{R}}$ changes how we analyze subtractive degree constructions, so it's good to briefly go over this, utilizing *almost* equatives. Although the rationale for certain kinds of exceptions is different with this operator, the end result for our analysis is, fortunately, the same. Let's take a look.

(262) Willow is almost as old as Buffy is.

a. $[\![(262)]\!]^{g,c}(w)$ is defined only if $\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{B}_{\mathrm{w}}\big]-\mathrm{g}(2)\big\}$ is defined, $\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{W}_{\mathrm{w}}\big]\big\}$ is defined, $\mathrm{g}(2) \subseteq \big(0,\mathcal{B}_{\mathrm{w}}\big]$, and $\mathrm{g}(2) \neq \emptyset$; where defined: $[\![(262)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{B}_{\mathrm{w}}\big]-\mathrm{g}(2)\big\} \leq \mathcal{W}_{\mathrm{w}}\ \&\ \textsc{small}_{c}(\mathrm{g}(2))$

ii. $\forall \chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\mathrm{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

$\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{W}_{\mathrm{w}}\big]\big\}$ is defined so long as $\mathcal{W}_{\mathrm{w}}$ is included in that interval, which it is, and the subset and non-emptiness presuppositions are familiar. $\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{B}_{\mathrm{w}}\big]-\mathrm{g}(2)\big\}$ is what we need to think about. Where before, g(2) had to be a final subinterval of $\big(0,\mathcal{B}_{\mathrm{w}}\big]$ for its subtraction from the latter to satisfy the definedness conditions on **max**, $\textbf{MAX}_{\mathrm{R}}$ imposes no such requirement. For any set $\mathbb{D}$ such that $\textbf{MAX}_{\mathrm{R}}(\mathbb{D}) < \mathcal{B}_{\mathrm{w}}$, $\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{B}_{\mathrm{w}}\big]-\mathbb{D}\big\}$ is defined. Let $\mathbb{D}^*$ be such a set. That means, then, if $\mathrm{g}(2) = \mathbb{D}^*$, $\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{B}_{\mathrm{w}}\big]-\mathbb{D}\big\}$ is just $\textbf{MAX}_{\mathrm{R}}\big\{\big(0,\mathcal{B}_{\mathrm{w}}\big]\big\}$, i.e. $\mathcal{B}_{\mathrm{w}}$, and so the truth of (262ai) depends on whether $\mathcal{B}_{\mathrm{w}} \leq \mathcal{W}_{\mathrm{w}}$, and whether $\mathbb{D}^*$ is small. Grant that both are true. Now, certainly small supersets of g(2) will be entailed, as well as any set whose maximal degree is less than or equal to the maximal element in g(2). If there are only other such sets that satisfy the presuppositions of *almost*, then (262ai) entails (262aii), and neither **exh** nor *almost* is licensed. Any set $\mathbb{X}'$ whose maximum is greater than the maximum of g(2) is not entailed, so if there are such sets $\mathbb{X}'$, they will thus falsify (262aii) above, since the subtraction of $\mathbb{X}'$ from $\big(0,\mathcal{B}_{\mathrm{w}}\big]$ cannot change the contingent fact that $\mathcal{B}_{\mathrm{w}} \leq \mathcal{W}_{\mathrm{w}}$. So (262aii) will be false, and hence (262a) will be, too. $\textbf{MAX}_{\mathrm{R}}$ doesn't really change much, in the end.

3.2.2 *Barely* needs a change

Our analysis of *almost* works beautifully so far, accurately predicting the felicity of *almost* equatives and ruling out *almost* comparatives. *Barely*'s distribution with respect to comparatives and equatives is complementary to that of *almost*: it happily goes with comparatives, but equatives are unacceptable.

(263) a. Willow is barely older than Buffy is.

b. # Buffy is barely as old as Willow is.

Now, we'll want to check if our semantics for *barely* works just as well.

(264) $[\![\text{barely}]\!]^{c} = \lambda X \lambda \mathscr{D} \lambda D \lambda D'$: $\mathbb{X} \subseteq \mathbb{D}$ & $\mathbb{X} \neq \emptyset$. $\neg\mathscr{D}(\mathbb{D} - \mathbb{X})(\mathbb{D}')$ & $\textsc{small}_c(\mathbb{X})$

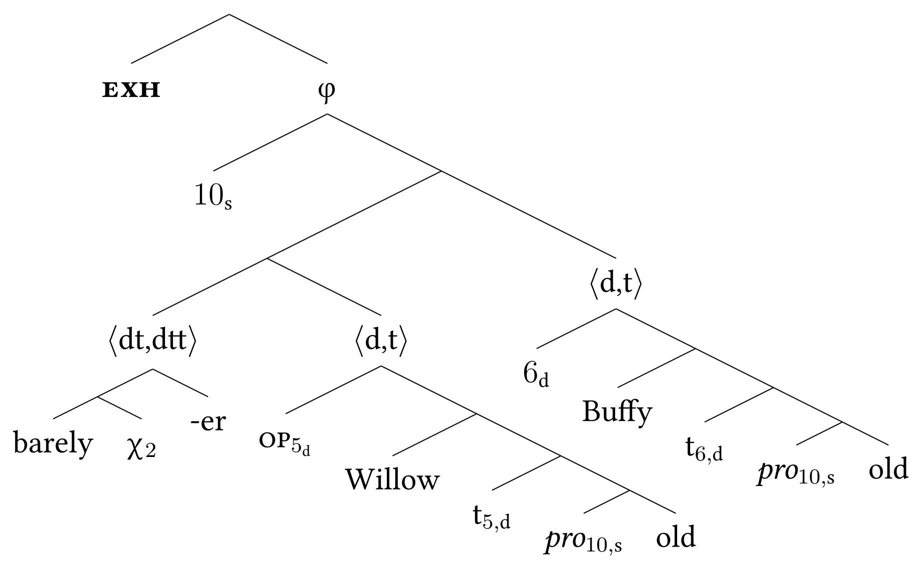
As it turns out, we run into a problem right away with comparatives: *barely* subtracts from the wrong set. §3.2.2.1 shows how, by altering the meaning of *barely* so that it subtracts from the nuclear scope of the quantifier it modifies, we derive the right results for *barely* comparatives and rule out *barely* equatives. §3.2.2.2 revisits quantificational determiners, discusses why it took degree quantifiers to reveal to us this necessary change, and shows that the revised analysis still works.

3.2.2.1 The problem and the solution

Intuitively, (265) is true where Buffy is older than Willow is, but the difference is rather small; were we to eliminate that difference, Willow would be at least as old as Buffy.

(265) Buffy is barely older than Willow is.

a.



b. $[\![(265\text{a})]\!]^{g,c}(w)$ is defined only if $\textbf{max}\{(0,\mathcal{W}_w] - g(2)\}$ is defined, $\textbf{max}\{(0,\mathcal{B}_w]\}$ is defined, $g(2) \subseteq (0,\mathcal{B}_w]$, and $g(2) \neq \emptyset$; where defined:

$[\![(265\text{a})]\!]^{g,c}(w) = 1$ *iff* (i) & (ii)

i. $\textbf{max}\{(0,\mathcal{W}_w] - g(2)\} \not< \mathcal{B}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

Right off the bat, we know these truth conditions aren't right. First, *barely* is subtracting from the wrong set—it's requiring that we lower Willow's age, rather than Buffy's. Second, (265bi) can only be true if in *w* Willow is at least as old as Buffy is, and g(2) subtracts *less* than the difference between their two ages. In a world where Buffy is older than Willow, there is no assignment for g(2) that can make (265bi) true. That's a problematic state of affairs!

Our semantics for subtractives takes them to subtract from the restrictor of the quantifier they modify; in comparatives and equatives, that means the subordinate clause. Intuitively, though, subtraction needs to apply to the matrix clause in *barely* comparatives. Let's make that change. Instead of subtracting from the restrictor, it will now subtract from the nuclear scope of the modified quantifier; the subset presupposition is changed accordingly.

(266) $[\![\text{barely}]\!]^c = \lambda X\lambda\mathscr{D}\lambda D\lambda D'\text{: } \mathbb{X} \subseteq \mathbb{D}' \;\&\; \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{D})(\mathbb{D}'-\mathbb{X}) \;\&\; \textsc{small}_c(\mathbb{X})$

Plugged into (265a), we derive the following truth conditions.

(267) $[\![(265a)]\!]^{g,c}(w)$ is defined only if $\textbf{MAX}\big\{(0,\mathcal{W}_w]\big\}$ is defined, $\textbf{MAX}\big\{(0,\mathcal{B}_w]-g(2)\big\}$ is defined, $g(2) \subseteq (0,\mathcal{B}_w]$, and $g(2) \neq \emptyset$; where defined: $[\![(265a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii)

i. $\mathcal{W}_w \not< \textbf{MAX}\big\{(0,\mathcal{B}_w]-g(2)\big\}$ & $\textsc{small}_c(g(2))$

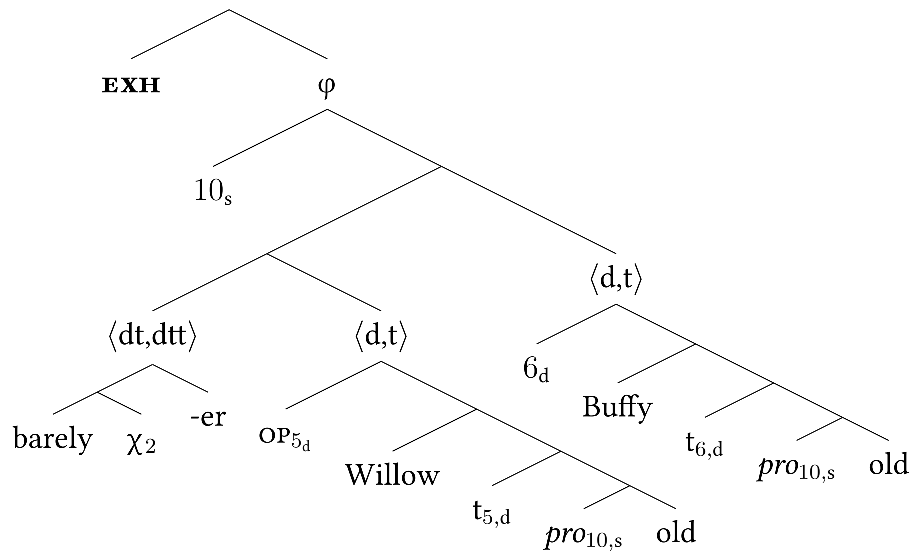
ii. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

Now the truth conditions say when we lower Buffy's age a bit, Willow's age is no longer strictly lower than Buffy's; it's at least as big. Notice that this is quite like the meaning of *Willow is almost as old as Buffy is*; that's no accident! The two sentences convey essentially the same content. In fact, they're logically equivalent now. So, the logic of *barely* comparatives on our analysis reduces to the logic of *almost* equatives. The exception must pick out the difference between Buffy and Willow's ages, i.e. $(0,\mathcal{B}_w]-(0,\mathcal{W}_w]$, for (267i) to be true in w. Anything smaller will make (267i) false, and anything larger will make (267ii) false. Once again, the exception must pick out the difference between the matrix and subordinate clauses, and that difference must be small, for (265) to be true.

Nuclear scope-*barely* also makes *barely* equatives logically equivalent to *almost* comparatives.

(268) Buffy is barely older than Willow is.

a.



b. $[\![(268a)]\!]^{g,c}(w)$ is defined only if $\textbf{MAX}\{(0,\mathcal{W}_w]\}$ is defined, $\textbf{MAX}\{(0,\mathcal{B}_w]-g(2)\}$ is defined, $g(2) \subseteq (0,\mathcal{B}_w]$, and $g(2) \neq \emptyset$; where defined:
$[\![(268a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii)

i. $\mathcal{W}_w \not\leq \textbf{MAX}\{(0,\mathcal{B}_w]-g(2)\}$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

The logic is exactly as it was for *Willow is almost older than Buffy is*; we won't rehash it here. The analysis will predict that *barely* equatives are unacceptable across the board.

3.2.2.2 Why didn't we see this before?

We proposed altering the meaning of *barely* in light of its interaction with the comparative and equative; we needed it to subtract from the nuclear scope rather than the restrictor. This move raises a natural question: why didn't we notice this before, why wasn't this change necessary earlier? It raises another natural question: does this do any violence to our analysis of *barely* with nominal quantificational determiners?

The reason we didn't see it before is simple: existential quantificational determiners are *symmetric*, and *barely* only went with (NPI) existential determiners.

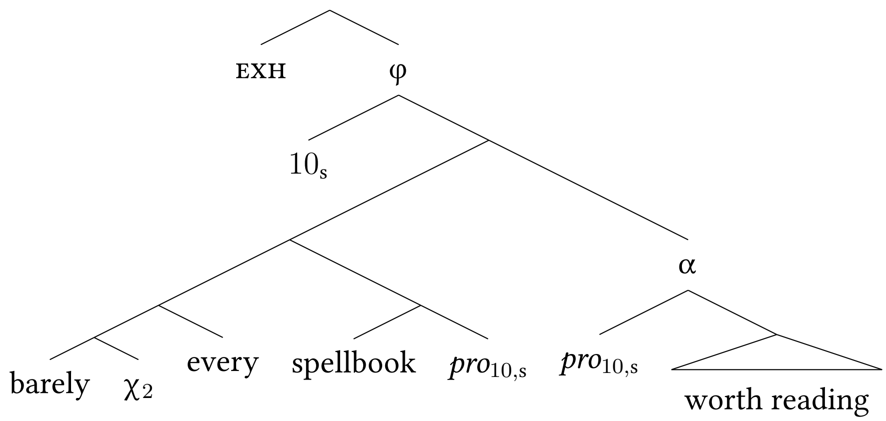
(269) ***SYMMETRY***

A type ⟨et,ett⟩ determiner $\mathscr{D}$ is *symmetric* just in case for all A, B of type ⟨e,t⟩:
If $\mathscr{D}(\mathbb{A})(\mathbb{B}) = 1$, then $\mathscr{D}(\mathbb{B})(\mathbb{A}) = 1$.

It is easy to see that *any* is symmetric. *Barely*, in modifying *any*, alters one of the arguments that it then supplies to the determiner. It didn't matter that our earlier analysis subtracted from the restrictor, and the logic of *barely any* doesn't change substantively. We are only confronted with the insufficiency of restrictor *barely* in the degree domain, specifically in the subtractive's incarnation as a modifier of the comparative morpheme, which does not express a symmetric relation. It matters which set we subtract from there.

What happens with other quantificational determiners? *No*, like *any*, is symmetric, so the logic of ruling out *barely no* is not altered in a deep way. *Every* is not symmetric, though, so we ought to revisit *barely every* to make sure that nothing goes awry. Fortunately, nothing does.

(270) # Barely every spellbook is worth reading.

a. 

b. $[\![(270)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{WR}_w$ and $g(2) \neq \emptyset$; where defined:
$[\![(270)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\mathbb{S}_w \not\subseteq \mathbb{WR}_w - g(2)$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \big]$

Note that (270bi) is logically equivalent to $\mathbb{WR}_w - g(2) \cap \overline{\mathbb{S}}_w$ & $\textsc{small}_c(g(2))$, which is the unexhaustified meaning of *Almost something worth reading is not a spellbook*:

(271) $[\![$Almost something χ_2 worth reading is not a spellbook$]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{WR}_w$ and $g(2) \neq \emptyset$; where defined:
$[\![(271)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

a. $\mathbb{WR}_w - g(2) \cap \overline{\mathbb{S}}_w$ & $\textsc{small}_c(g(2))$

b. $\forall \chi' \big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \big]$

We don't even need to go through the logic here, then. Nuclear scope *barely* cannot yield truth!

While in the degree domain density prevents us from finding a unique exception, in the nominal domain it's the arbitrariness of what we have to subtract to satisfy the truth conditions. As long as the exception contains everything in $\mathbb{WR}_w$ not in $\mathbb{S}_w$, it doesn't matter which additional element is in the exception, there just needs to be one. But since any is good enough, none is unique, and **exh** will fail. More generally, we see that this change to *barely* makes it parallel, though not identical, to how it operates in degree constructions. That's a fun thing. Something important to note, though, is that our move to nuclear scope-*barely* costs our analysis a particular understanding of the 'duality' of *almost* and *barely*. That is, *barely* is not definable as *almost not*. This evidence converges on conclusions drawn elsewhere in the literature, for example, Atlas (1997) on VP uses of *almost* and *barely*.

(272) It barely rained $\not\leftrightarrow$ It almost didn't rain.

These are not equivalent statements, but we might expect that if *almost* and *barely* are interdefinable[8]. Similar evidence is adduced by Amaral (2007), who investigates European Portuguese counterparts to *almost* and barely.

Nonetheless, there is still clearly a close link between *almost* and *barely* on our analysis: they have an exceptive semantics, and their polar implications are derived through exhaustification over alternatives invoking different exceptions coupled with their basic, subtractive truth conditions. Natural questions arise, including why these two operators should have such similar semantics. These questions must wait for their answers.

3.2.2.3 A brief note on SMALL

Barely now subtracts from the nuclear scope of the quantifier it modifies. This raises the question of whether the size of the exception is calculated with respect to size of the new minuend, i.e. the set characterized by the nuclear scope, or if it is still calculated with respect to the restrictor set. The contrast in perceived truth between the two following examples points us in the right direction[9].

(273) a. Barely any Americans are active astronauts.

b. Almost no active astronauts are Americans.

The truth conditions we derive here, given the discussion in the previous section, will be logically equivalent, yet these two do not feel equivalent. (273a) feels true, or at least more likely to be true, than (273b) does. In both cases, since we're dealing with intersective quantifiers, the exception picks out the set of active, American astronauts. (273a) requires that, discounting that set, there are no Americans that are active astronauts, and that set is small; (273) requires the same. The contrast, then, seems unexpected at first, but we can explain it by maintaining that, in spite of the fact that it subtracts from the nuclear scope of *any*, *barely* still requires that we check the size of the exception against the size of the restrictor.

[8]Of course, we have not yet given an analysis of these kinds of uses, but in Chapter 5 we'll briefly discuss them.

[9]Much obliged to Roger Schwarzschild for pointing this out and providing this data.

(274) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, μ_c be a measure function, determined by the context, mapping sets to numerical values representing their size, and n_c be a contextually determined numerical threshold for size.

$\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\mu_c(\mathbb{X}) \leq n_c$

(275) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

For any context c, an utterance of φ in c is felicitous only if $n_c < \frac{1}{2}(\mu_c(\mathbb{A}))$

The reason (273a) feels true is that the number of active American astronauts is small (there are forty-four) relative to the set of all Americans (there are roughly three hundred and thirty million); (273b) feels false because the number of active American astronauts is large relative to the total number of astronauts (there are one hundred and twenty eight)[10]. Since forty four is plausibly quite large relative to one hundred and twenty eight, (273b) is judged false.

3.2.2.4 A brief note on cross-categoriality

When we first turned to uses of *almost* and *barely* as modifiers of degree quantifiers, we simply stipulated new types. We can generalize the meanings we give, though, to any arbitrary semantic type σ.

(276) $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle \sigma t, \sigma t t\rangle} \lambda A_{\langle \sigma, t\rangle} \lambda B_{\langle \sigma, t\rangle}\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B}')\ \&\ \textsc{small}_c(\mathbb{X})$

(277) $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle \sigma t, \sigma t t\rangle} \lambda A_{\langle \sigma, t\rangle} \lambda B_{\langle d, t\rangle}\colon \mathbb{X} \subseteq \mathbb{B}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A})(\mathbb{B}-\mathbb{X})\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle \sigma t, \langle \sigma t, \langle\langle \sigma t, \sigma t t\rangle, \langle \sigma t, t\rangle\rangle$

(278) ***SUBTRACTIVE ALTERNATIVES***

For a subtractive LF φ, with an exception χ of type $\langle \sigma, t\rangle$,

$\textbf{ALT}(\varphi) := \{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_\sigma\ \&\ \varphi[\chi/\chi'] \text{ is grammatical}\}$

That these operators now are defined for arbitrary types means we could expect them to modify all sorts of quantifiers. This will be discussed more in the conclusion of the dissertation, but we should now comment on *but*. We only see this operator modifying quantificational determiners—nominal quantifiers! For us, this simply comes down to the fact that *but* is more restricted in the types of quantifiers it can modify[11]. *Almost* and *barely* are more promiscuous, as it were.

[10] Numbers from: https://tinyurl.com/327bphrt

[11] Though sometimes it does appear in degree constructions, at least for some speakers:

Perhaps, when we expand our search for subtractive operators to other domains in English and cross-linguistically, we'll see more of this: certain operators are dedicated to particular domains of quantification, others are more flexible. This investigation must be left for the future.

3.2.3 Antonymy and scale structure

Up until now, the subtractive comparatives and equatives we've focused on have featured relative adjectives, with *old* as our exemplar. Relative adjectives often come in *antonym* pairs; the truth conditional equivalence between (279a) and (279b) shows the tight connection between *young* and *old*.

(279) a. Willow is older than Buffy is.
b. Buffy is younger than Willow is.

Both of these are true just in case Willow's age exceeds Buffy's age; let's call *old* the *positive* antonym (A^{+}) and *young* the *negative* antonym (A^{-}). Other antonym pairs include *long/short*, *intelligent/unintelligent*, *safe/dangerous*, and *beautiful/ugly*; in each pair, the first is the positive antonym and the second the negative. These are diagnosable by a variety of empirical properties, but it's rather messy generally cross-linguistically and even within a single language (Seuren 1978; Kennedy 2001, i.a.). What is important is that antonyms come in pairs, and we'll make some stipulations about which is which.

Negative antonyms, just like their positive variants, occur in subtractive degree constructions.

(280) a. Willow is almost as $\{$young/short/unintelligent$\}$ as Buffy is.
b. Willow is barely $\{$younger/shorter/more intelligent$\}$ than Buffy is.

All of these are acceptable subtractive comparatives and equatives. Our analysis needs to account, then, for these.

3.2.3.1 The problem

Both (279a) and (279b) are comparatives, obviously. The matrix clause subject and the subordinate clauses are reversed, also obviously. The adjectives differ. They mean the same thing—Buffy's age is strictly less than Willow's age. How does this come about?

The semantics we've assumed so far treats positive relative adjectives as world-degree-individual relations.

(i) Willow is as old as Buffy is, but for a few days.

For speakers who accept this, it means pretty much what the *almost* equative does. That said, we haven't investigated *but for* at all, nor measure phrases like *a few days*, which have been argued to have both degree and individual senses (Brasoveanu 2008). More work is needed here.

(281) $[\![\text{old}]\!] = \lambda w \lambda d \lambda x \text{: } d \in \mathcal{S}_{\text{AGE}}.\ \text{AGE}_w(x) \geq d$

With such a semantics, the subordinate and matrix clauses of comparatives characterize initial intervals of $\mathcal{S}_{\text{AGE}}$, i.e. starting at, but not including, 0, and extending upward to an entity's age in a world. *Young*, *old*'s negative antonym, seems to give the same kind of information about an individual—it's anchored, in a sense, to an entity's age in a world, making reference to the same scale—but the perspective on that information is rather different (Seuren 1978; von Stechow 1984b; Kennedy 2001). Negative relative adjectives 'flip' the relation that holds between an entity's measure at a world and all other degrees on the scale; one way to formalize this is to give *young* the entry below.

(282) $[\![\text{young}]\!] = \lambda w \lambda d \lambda x \text{: } d \in \mathcal{S}_{\text{AGE}}.\ \text{AGE}_w(x) < d$

The entry in (282b) characterizes the complement of the initial interval that *old* characterizes. Antonym pairs share a measure function and scale on this kind of view, and the negative adjective is a *degree upward monotone* function.

(283) A function A of type $\langle s,\langle d,et\rangle\rangle$ is *degree upward monotone* just in case
$\forall w \forall x \forall d \forall d' \big[A(w)(d)(x) = 1 \ \&\ d < d' \rightarrow A(w)(d')(x) = 1\big]$

Such a semantics has roots in Cresswell (1976) and von Stechow (1984b), and is taken up in a form in Kennedy (2001). Part of the motivation concerns measure phrases, which are compatible with positive, but not negative, relative adjectives:

(284) Xander is sixteen years $\big\{$old/#young$\big\}$.

The contrast is unexpected unless there is some formal distinction between positive and negative adjectives. Negative adjectives' upward degree monotonicity is a way to cash this out. However, when put with our maximality-based semantics of the degree quantifiers. First, our **max** is defined only for *initial* intervals of a scale; the intervals derived by *young* in comparatives and equatives will be *final* intervals of $\mathcal{S}_{\text{AGE}}$.

(285) Buffy is younger than Willow is.

(286) $\textbf{MAX}\big\{\big(\mathcal{W}_w,\infty\big)\big\} < \textbf{MAX}\big\{\big(\mathcal{B}_w,\infty\big)\big\}$ BOTH **MAX** ARE UNDEFINED!

Even modifying **max** in a way so that it is defined for final intervals of a scale—for example, by having it pick out the minimal element in a final interval—still won't solve all our problems with this semantics for *young*. It's not even about the semantics of *young*. The issue is maximality. Consider these alternatives to the semantics for *young* and **max**.

(287) $[\![\text{young}_{\leq}]\!] = \lambda w \lambda d \lambda x \text{: } d \in \mathcal{S}_{\text{AGE}}.\ \text{AGE}_w(x) \leq d$

(288) For any set of degrees $\mathbb{D}$ on a scale $\mathcal{S}$:

$$\textsc{max}(\mathbb{D}) := \iota d\big[d \in \mathbb{D}\ \&\ \forall d'\big[[d' \in \mathcal{S}\ \&\ d \leq d'] \leftrightarrow d' \in \mathbb{D}\big] \vee \forall d'\big[[d' \in \mathcal{S}\ \&\ d' \leq d] \leftrightarrow d' \in \mathbb{D}\big]\big]$$

This *young* is the mirror image of *old*; in a comparative, it would derive an interval starting at an entity's height in a world and extend onward infinitely[12]. It is not quite the complement of the initial interval that *old* gives rise to, but it's close. The redefinition of **MAX** allows it to pick out the *minimal* degree of a final interval on a scale. Even these modifications don't generate the right results for comparatives and equatives with a negative relative antonym.

(289) $[\![\text{Buffy is younger than Willow is}]\!](w) = 1$ *iff* $\textsc{max}\{[\mathcal{W}_w,\infty)\} < \textsc{max}\{[\mathcal{B}_w,\infty)\}$
i.e. Willow's age in w < Buffy's age in w

What do we do about this?

3.2.3.2 Inverse scale structure

The way negative antonyms flip the 'perspective' on the information encoded by a positive adjective doesn't have to be a matter of reversing the relation between degrees encoded by adjectives. That is, even though above positive antonyms expressed the $\geq$ relation between two degrees, and negative antonyms either $<$ or $\leq$, we can flip the perspective in a different way altogether that allows us to maintain our semantics for comparatives and equatives while also making explicit exactly how antonyms are related semantically, and in such a way that we are set up to discuss the positive construction, and in particular, absolute adjectives, in §3.3. Here's the idea. Antonymy is about inverting scales: where positive adjectives invoke 'positive' scales, negative adjectives invoke their mirror images, 'inverse' scales[13].

(290) $\mathcal{S}_{\text{old}} = (0,\infty)$ (291) $\mathcal{S}_{\text{young}} = (-\infty,0)$

(292) **INVERSE DEGREES**
For any degree $d \in D_d$, there is a unique $d' \in D_d$ such that $d' = -d$; -d is the *inverse* of d.

(293) **INVERSE SCALES**
Let $\mathcal{S}_{\mu_A}$ be the scale associated with an adjective *A*, $\mathcal{S}_{-\mu_A}$ be the *inverse* of $\mathcal{S}_{\mu_A}$.
$\mathcal{S}_{-\mu_A} := \{d \mid -d \in \mathcal{S}_{\mu_A}\}$
For any d, d′ on a scale $\mathcal{S}_A$, if $d <_{\mu_A} d'$, then $-d' <_{-\mu_A} -d$.

Formally, the idea should be clear: for every age d in $\mathcal{S}_{\text{AGE}}$ there is a unique inverse age -d in $\mathcal{S}_{\text{-AGE}}$. 0, of course, is in neither scale; it serves as the non-included lower bound of $\mathcal{S}_{\text{AGE}}$ and the

[12] This kind of meaning is a degree-relational version of the proposal in Kennedy (2001).
[13] Note that we use '-' to mark an inverse or negative degree, contrasted with '−' for set subtraction.

non-included upper bound of $\mathcal{S}_{\text{-AGE}}$. Both *old* and *young* will invoke the same measure function—AGE—but *young* is defined only for degrees on the inverse scale of $\mathcal{S}_{\text{AGE}}$, and gives the inverse of an individual's age in a world.

(294) a. $[\![\text{old}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{AGE}}.\ \text{AGE}_w(x) \geq d$

b. $[\![\text{young}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{-AGE}}.\ \text{-AGE}_w(x) \geq d$

Both adjectives are degree downward monotone functions, though the intervals they will generate are rather different[14]

(295) a. $\{d \mid [\![\text{old}]\!](w)(d)(\text{Buffy}) = 1\} = (0,\mathcal{B}_w]$

b. $\{d \mid [\![\text{young}]\!](w)(d)(\text{Buffy}) = 1\} = (\text{-}\infty,\text{-}\mathcal{B}_w]$

Figure 1 illustrates.

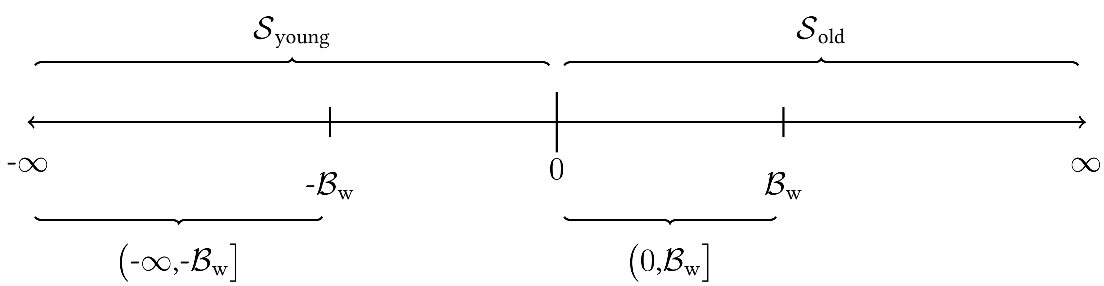


Figure 1

Comparatives and equatives of negative antonyms can now be dealt with exactly as their positive variants. **MAX** has no problem picking out maxima from the intervals generated via negative antonyms now.

(296) $[\![\text{Willow is shorter than Buffy is}]\!](w) = 1$ *iff* $\textbf{MAX}\{(\text{-}\infty,\text{-}\mathcal{B}_w]\} < \textbf{MAX}\{(\text{-}\infty,\text{-}\mathcal{W}_w]\}$
i.e. $\text{-}\mathcal{B}_w < \text{-}\mathcal{W}_w$, i.e. $\mathcal{B}_w > \mathcal{W}_w$, i.e. Buffy is taller than Willow is

(297) $[\![\text{Willow is as short as Buffy is}]\!](w) = 1$ *iff* $\textbf{MAX}\{(\text{-}\infty,\text{-}\mathcal{B}_w]\} \leq \textbf{MAX}\{(\text{-}\infty,\text{-}\mathcal{W}_w]\}$
i.e. $\text{-}\mathcal{B}_w \leq \text{-}\mathcal{W}_w$, i.e. $\mathcal{B}_w \geq \mathcal{W}_w$, i.e. Buffy is as tall as Willow is

It should be clear that this semantics validates the more general equivalence 'inverse' comparatives and equatives, since the degrees picked out by negative antonyms are the inverses of those picked out by positive antonyms.

[14]We are giving the antonyms as basic entries that stand in a clear relation to one another. We could go further, following Büring (2007a, 2007b) and Heim (2008), and take negative antonyms to be derived from their positive versions and a negative morpheme **LITTLE**. We leave a further exploration of this for the future.

(298) For any A, A^{-}, x, y:

a. $[\![$x is A-er than y is$]\!]$(w) $= 1$ *iff* $[\![$y is A^{-}-er than x is$]\!]$(w) $= 1$

b. $[\![$x is as A as y is$]\!]$(w) $= 1$ *iff* $[\![$y is as A^{-} as x is$]\!]$(w) $= 1$

While a more thorough treatment of such expressions is a subject for of the next chapter, our proposal allows us to make sense of why, at least in English, measure phrases are generally only compatible with positive adjectives. It is fairly standard to treat expressions like *six feet* as denoting degrees saturating the degree arguments of adjectives. Specifically, measure phrases could denote exclusively *positive* degrees (von Stechow 1984b), and therefore would not satisfy the presuppositions of negative adjectives[15].

(299) $[\![$six feet$]\!] = 6'$ d

a. Giles is six feet tall.

b. # Giles is six feet short.

Even if our proposal isn't right for measure phrase constructions, something along the lines of ***MEAS*** is necessary. Schwarzschild (2005) takes measure phrases to denote intervals, rather than degrees, and proposes a lexical rule that turns degree-relational semantics to an interval-relational semantics for a language-specified set of degrees. Such a view is quite compatible with a strict ordering semantics for adjectives; what would need ironing out is the meaning contributed by the measure phrase interval and what the semantic output of the lexical rule is.

Let's return to subtractives.

3.2.3.3 Subtractives again

The proposal above allows us to capture the distribution of *almost* and *barely* with respect to comparatives and equatives with negative adjectives exactly as we did for positive adjectives. Note

[15]Of course, not all positive relative adjectives permit measure phrases anyways, though their counterparts in other languages do. Furthermore, some English negative relative adjectives have counterparts that take measure phrases.

(i) # The village is two kilometers far.

(ii) Het dorp is twee kilometer ver.
the village is two kilometer far
'The village is two kilometers away.' Seuren (1978)

(1) An dieser Stelle der Garage fehlt noch ein drei Zentimeter schmales Brett.
at this place the.DEN garage is.missing still a three centimeter(s) narrow plank.
'In this part of the garage, a three-centimeter-{wide/*narrow} plank is still missing.' Hofstetter (2016:89)

The cross-linguistic distribution of measure phrases with respect to those adjectives is beyond the scope of this dissertation; see, e.g., Schwarzschild (2005) and Hofstetter (2016).

that even with negative adjectives, the comparative and equative morphemes are still Strawson Downward Entailing on their first arguments, and Strawson Upward Entailing on their second.

(300) [**EXH** [Willow is almost χ_2 as young as Buffy is]]

a. $[\![(300)]\!]^{g,c}(w)$ is defined only if $\textbf{MAX}\left\{\left(-\infty,-\mathcal{B}_w\right]-g(2)\right\}$ is defined, $\textbf{MAX}\left\{\left(-\infty,-\mathcal{W}_w\right]\right\}$ is defined, $g(2) \subseteq \left(-\infty,-\mathcal{B}_w\right]$, and $g(2) \neq \emptyset$; where defined:
$[\![(300)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}\left\{\left(-\infty,-\mathcal{B}_w\right]-g(2)\right\} \leq -\mathcal{W}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

The *almost* equative is intuitively true if Buffy is younger than Willow is, but only by a little bit; the truth conditions predict this. g(2) subtracts the difference between $\left(-\infty,-\mathcal{B}_w\right]$ and $\left(-\infty,-\mathcal{W}_w\right]$ from the former, in effect *raising* Buffy's age to Willow's. Were we to subtract any less, then Buffy would still be younger than Willow is. This is exactly how positive relative adjectives operated in *almost* equatives.

The rest of the cases operate as they did before as well; we won't walk through them, but their logic should be clear. *Almost* comparatives and *barely* equatives will face a density problem, rendering (301a) and (302a) impossible to make true. *Barely* equatives will be just like *almost* comparatives: the exception is the difference between Buffy and Willow's (negative) ages, and it must be small.

(301) # [**EXH** [Willow is almost χ_2 younger than Buffy is]]

a. $[\![(301)]\!]^{g,c}(w)$ is defined only if $\textbf{MAX}\left\{\left(-\infty,-\mathcal{B}_w\right]-g(2)\right\}$ is defined, $\textbf{MAX}\left\{\left(-\infty,-\mathcal{W}_w\right]\right\}$ is defined, $g(2) \subseteq \left(-\infty,-\mathcal{B}_w\right]$, and $g(2) \neq \emptyset$; where defined:
$[\![(301)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}\left\{\left(-\infty,-\mathcal{B}_w\right]-g(2)\right\} < -\mathcal{W}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

(302) # [**EXH** [Willow is barely χ_2 as young as Buffy is]]

a. $[\![(302)]\!]^{g,c}(w)$ is defined only if $\textbf{MAX}\left\{\left(-\infty,-\mathcal{B}_w\right]-g(2)\right\}$ is defined, $\textbf{MAX}\left\{\left(-\infty,-\mathcal{W}_w\right]\right\}$ is defined, $g(2) \subseteq \left(-\infty,-\mathcal{B}_w\right]$, and $g(2) \neq \emptyset$; where defined:
$[\![(302)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $-\mathcal{W}_w \nleq \textbf{MAX}\left\{\left(-\infty,-\mathcal{B}_w\right]-g(2)\right\}$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

(303) [**EXH** [Willow is barely χ_2 younger than Buffy is]]

a. $[\![(303)]\!]^{g,c}(w)$ is defined only if $\textsc{max}\{(-\infty,-\mathcal{B}_w]-g(2)\}$ is defined, $\textsc{max}\{(-\infty,-\mathcal{W}_w]\}$ is defined, $g(2) \subseteq (-\infty,-\mathcal{B}_w]$, and $g(2) \neq \emptyset$; where defined:
$[\![(303)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $-\mathcal{W}_w \not< \textsc{max}\{(-\infty,-\mathcal{B}_w]-g(2)\}$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' [[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1]$

Our proposal for the scale structure and meaning of negative relative adjectives allows us a tidy account of the semantics of the comparative and equative and predict the distribution of *almost* and *barely* with respect to those operators across the board.

3.2.4 Interim summary

We've done a lot in the first half of this chapter. We've taken the analysis of subtractives as modifiers of quantificational determiners developed in the previous chapter and extended it to degree constructions, showing that it is quite successful at explaining the distribution of subtractives therein. This is big! Comparatives and equatives have played little role in this way in previous literature on *almost* and *barely*[16], but it has been incredibly fruitful to do so. We've seen that without subtractive alternatives, the contrast between *almost* equatives and *almost* comparatives is entirely unexpected. We've also seen that subtractives give us evidence of the idea that all measurement scales in natural language are dense, even when we might not expect it. We've also learned that *barely* has to subtract its exception from the nuclear scope of the modified quantifier, not the restrictor; the results of the previous chapter remain, too. We also started building an analysis of antonymy that allows us to maintain our semantics for comparatives and equatives, and which sets us up for the topic to which we now turn: the positive construction.

3.3 The positive construction

There are several different classes of adjectives, and the distribution of *almost* and *barely* tracks some of the distinctions. Generally, there are adjectives that appear felicitously in constructions like the comparative, and those that do not.

(304) a. Willow is {older/taller/smarter/more powerful} than Buffy is.

b. Sunnydale is {cleaner/dirtier/safer/more dangerous} than Los Angeles is.

Adjectives that are at home in such constructions, like those in (304a) and (304b), are called GRADABLE adjectives (GAs).

[16] Though see §3.3.3.3 for a discussion of Rotstein & Winter (2004), the lone place *almost* equatives and comparatives have been discussed.

Among GAs, there are different subclasses intuitively and empirically. Intuitively, the adjectives in (304a) are vague and heavily context-dependent; what counts as *tall* or *smart* depends on what is being compared. A tall or smart child typically doesn't count as tall or smart when compared to adults, for example, since we judge them by different standards. The adjectives in (304b) lack this context-dependence, or at least have it to a much lesser degree (see Lassiter & Goodman 2013 for arguments and evidence for the latter). Intuitively, *Sunnydale is safe* is true iff there is no degree of danger there, whereas *Sunnydale is dangerous* is true iff there is some degree of danger. Note that it doesn't seem right to say that *Cordelia is tall* is true entirely because Cordelia has some amount of height; not only is that uninformative, it's trivial if that sentence is defined. For *dangerous*, though, having some degree of danger actually seems reasonably informative. Among these adjectives we also find differences empirically in terms of permissible modifiers (Rotstein & Winter 2004; Kennedy & McNally 2005).

(305) a. Willow is ??{slightly/partially/completely/totally} tall/old.
b. Los Angeles is {??slightly/??partially/completely/totally} safe.
c. Dawn's room is {??slightly/??partially/completely/totally} clean.
d. Sunnydale is {slightly/partially/??completely/??totally} dangerous.
e. Anya's room is {slightly/partially/??completely/??totally} dirty.

Adjectives like *tall* and *old* are not compatible with any of these modifiers. *Safe* and *clean* are compatible with *completely* and *totally*, but not *slightly* and *partially*; *dangerous* and *dirty* flip the pattern of *safe* and *clean*.

Adjectives that have very vague, context dependent standards and which do not permit modification by *slightly* and *completely*, are called RELATIVE adjectives (RAs). The gradable, non-relative adjectives that permit modification by *slightly* or *completely*, and which have more fixed standards of comparison are called ABSOLUTE adjectives (AAs, Unger 1978; Rusiecki 1985; Kamp & Rossdeutscher 1994). Even among AAs we can make distinctions: those that allow *completely*, for example, and whose standards are 'maximal' in a sense are called TOTAL adjectives (TAs), and those which permit *slightly* and have a 'minimal' standard are called PARTIAL adjectives (PAs) (Yoon 1996; Rotstein & Winter 2004).

The distribution of *almost* and *barely* tracks these distinctions in the adjectival domain. They are unacceptable generally with RAs; *almost* naturally modifies TAs and *barely* PAs.

(306) a. # Willow is almost {old/tall/smart/powerful}.
b. Dawn's room is {almost/#barely} clean.
c. Anya's room is {barely/#almost} dirty.

RAs are not always unacceptable with subtractives, though. They require a special kind of context, one in which a standard is explicitly set. Imagine we're in gym class, and we're trying to categorize everyone in terms of whether or not they're tall. We decide that, to count as tall, one must be five feet seven inches or greater. Willow is $5'6.5''$, Buffy is $5'7.5''$, and Cordelia is $5'9''$.

(307) a. Cordelia is (definitely) tall.
b. Buffy is barely tall.
c. Willow is almost tall.

While perhaps not perfect, *almost* and *barely* improve in this kind of context. Explicit comparison helps, as does focus on the subtractives.

(308) a. Buffy? Eh, she's $\left[\text{barely}\right]_{\text{F}}$ tall. Now Cordelia? She's definitely tall.
b. Willow? Eh, she's $\left[\text{almost}\right]_{\text{F}}$ tall. Now Cordelia? She's definitely tall.

The essential fact is that while in general RAs resist modification by *almost* and *barely*, when the standard of comparison is fixed explicitly in the context, they are permissible as modifiers[17].

Our theory of *almost* and *barely*, then, has a few important desiderata. We want to account for the acceptability of *almost* with TAs and *barely* with PAs. We also want to generally predict unacceptability with RAs, but the analysis cannot prohibit such modification across the board; contextually-established explicit standards somehow make felicitous *almost* and *barely* with RAs.

3.3.1 POS and standards

Our starting point for the semantics for adjectives remains unchanged for RAs.

(309) a. $[\![\text{old}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{AGE}}.\ \text{AGE}_{\text{W}}(x) \geq d$
b. $[\![\text{young}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{-AGE}}.\ \text{-AGE}_{\text{W}}(x) \geq d$

When bare adjectives occur in predicative position, we need something that saturates the degree argument of adjectives. Intuitively, *Cordelia is old* conveys that Cordelia's age meets or exceeds some age, or family of heights, that are contextually relevant. The contextual relevance includes facts about Cordelia; if she is a child star, the relevant ages will be lower than if she's a competitive

[17]One might wonder whether in these cases context is licensing an elided measure phrase, which also ameliorates the unacceptability of *almost* and *barely* with RAs:

(1) Buffy is {almost/barely} five feet seven inches tall.

In the next chapter, as we've mentioned before, we will address how to analyze measure phrase constructions. If we have independent reason to assume ellipsis is the right analysis of RAs in contexts with explicit standards, the solution will be found there. We'll set this possibility aside for now.

player of the card game Contract Bridge[18]. These facts can even be made salient explicitly with a *for*-PP (Klein 1980).

(310) Cordelia is old for $\{$a child star/a competitive Bridge player$\}$.

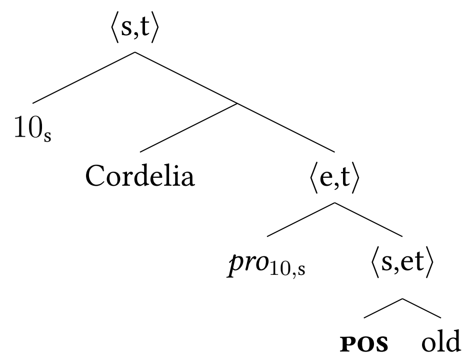
Absent such a PP, though, we need some morpheme that takes an adjective and saturates its degree argument. The standard solution is a covert morpheme **POS**, which supplies the standard of comparison via the function $\boldsymbol{STND}_c$ (see Kennedy 2007 and references therein).

(311) $[\![\textbf{POS}]\!]^c = \lambda A_{\langle s,det\rangle}\lambda w\lambda x.\ A(w)(\boldsymbol{STND}_c(A))(x)$ $\langle$sdet,set$\rangle$

This semantics aligns with a LF as in (312a)

(312) Cordelia is old.

a.



b. $[\![(312a)]\!]^c(w) = 1$ *iff* $\textsc{age}_w(\text{Cordelia}) \geq \boldsymbol{STND}_c([\![\text{old}]\!])$

How do subtractives fit into the picture?

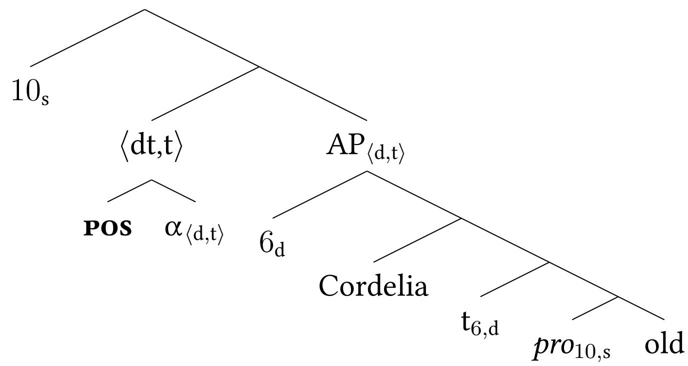
There's good reason to think that *almost* and *barely* target **POS**. Consider *Anya's room is almost clean*: this is true if the level of cleanliness Anya's room has is just below the standard of comparison for counting as 'positively' clean. If you were to lower the standard just a bit, her room would be considered positively clean. This intuition is more general, extending to RAs with explicit standards. Grant that we've established the standard for heights is 6'; *Willow is almost tall* means that if we lower the standard just a bit, Willow would count as tall, but in fact she does not. This suggests that the right account of subtractives in the positive construction involves the operators modifying **POS**. The semantics for **POS** above isn't compatible with such modification, though. It's not quantificational! What kind of shape must **POS** have, then?

The semantics for **POS** above looks something like an equative. What if **POS** itself were semantically a covert equative, expressing a relation between two sets of degrees, as in (313a)? Instead of the LF in (312a), we'd get the LF in (313b).

[18] At least in the United States in 2022, the typical Contract Bridge player is seventy-one (https://tinyurl.com/mwyc2m47).

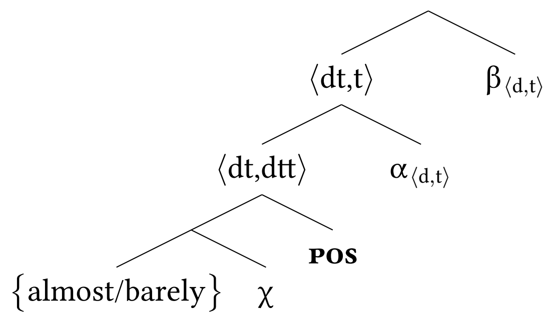
(313) a. $[\![\textsc{pos}]\!]^{c} = \lambda \mathbb{D} \lambda \mathbb{D}'.\ \textbf{MAX}(\mathbb{D}) \leq \textbf{MAX}(\mathbb{D}')$ *first version*

b.



Such a proposal would naturally allow for subtractives to compose with **POS**, at least type-wise.

(314)



Though not common by any stretch of the imagination, such a proposal is not without precedent. von Stechow (2009) treats **POS** as a generalized degree quantifier; Bylinina (2013) gives **POS** the meaning of our comparative morpheme. Now, our semantics for **POS** doesn't say anything about the standard of comparison, so how is that supplied? For the moment, let's assume the first argument of **POS** is a pronoun that picks out a degree interval. Nothing yet will actually require that the pronoun picks out the standards we want, but we'll return to this later.

To see how this analysis works, let's start with contextually explicit standards. We'll assume they pick out initial intervals of positive adjective scales, so that for some *n*, they are of the form $\big(0,\mathrm{n}\big]$. Let $pro_{4,\mathrm{dt}}$ be the interval pronoun saturating **POS**'s first argument. Then:

(315) $[\![\text{Giles is } \textbf{POS}\ pro_{4,\mathrm{dt}} \text{ tall}]\!]^{g,c}(w) = 1$ *iff* $\textbf{MAX}(g(4)) \leq \textbf{MAX}\big\{\big(0,\mathcal{G}_{w}\big]\big\}$ i.e. $6' \leq \mathcal{G}_{w}$

This seems right for the case of contextually explicit standards. Note that this semantics makes **POS** Downward Entailing on its first argument, and Upward Entailing on its second.

Let's turn to absolute adjectives.

3.3.2 Absolute adjectives

3.3.2.1 They're different!

Absolute adjectives—pairs like *clean/dirty* and *dangerous/safe*—are distinct from relative adjectives in the ways described above. They permit modification by expressions like *completely* and *slightly*. Their standards seem much more rigid than those of RAs.

(316) a. Sunnydale is clean.

b. Sunnydale is safe.

(317) a. Los Angeles is dirty.

b. Los Angeles is dangerous.

The standards associated with AAs like *clean, safe, dirty* and *dangerous* lack the kind of contextual variability that RAs had. (316a) is true if Sunnydale is free of any filth—it matters less whether we're comparing it to other cities in California or other cities of comparable size, for example. Another way to put it is that the level of cleanliness that Sunnydale has is at least as high as the 'paragon' of cleanliness. The same goes for (316b); it's true so long as Sunnydale's safety level is at least as high as the paragon of safety. (317a) and (317b) are different: as long as there is some non-zero level of threat or dirt, they seem true.

Consider the following.

(318) a. Are the toys dirty?

b. Are the toys clean? Yoon (1996:217)

A positive answer to (318a) requires that at least one toy be dirty, i.e. have some degree of filth; a positive answer to (318b) requires all to be free of filth (Yoon 1996). Kamp & Rossdeutscher (1994) call adjectives like *clean* and *safe* 'universal,' and those like *dirty* and *dangerous* 'existential' in virtue of properties like this. Note that RAs seem to pattern more like TAs with respect to this diagnostic:

(319) Are the students $\left\{\text{tall/short}\right\}$?

An affirmative answer for either question would seem to convey that *all* the students are tall, or all of them are short.

In light of this discussion, the compatibility of TAs with *almost* and PAs with *barely* is perhaps unsurprising.

(320) a. Anya's room is {almost/#barely} clean.

b. Dawn's room is {barely/#almost} dirty.

To be *almost clean* Anya's room has to be close to clean—there can't be very much filth, but if the standard were adjusted a bit, things that are just a little dirty could count as clean. Another way to phrase it: almost no part of Anya's room is dirty, or almost every part is clean. For Dawn's room to count as *barely dirty*, it has to be that removing just a little bit of filth from it would yield a clean room. There is some amount of dirt, some part if it could be dirty, but it's pretty close to being not dirty at all.

3.3.2.2 Absolute standards

Kennedy & McNally (2005, K&M) argue on the basis of the data in §3.3.1 that the scales associated with different adjective classes have distinct structures. K&M propose that the differences between AAs and RAs above comes down to whether the scales associated with those adjectives are include their bounds. The scale associated with *old*, $\mathcal{S}_{\text{AGE}}$, contains neither its upper nor a lower bound; its mirror, associated with $\mathcal{S}_{\text{-AGE}}$, similarly contains neither.

(321) $\mathcal{S}_{\text{AGE}} = \big(0,\infty\big)$ (322) $\mathcal{S}_{\text{-AGE}} = \big(\text{-}\infty,0\big)$

These are *totally open* scales, and K&M propose that RAs in general are associated with such scales.

AAs, K&M argue, are associated with scales that include at least one bound. PA scales include scalar minima: the measure function associated with *dirty*, let's call it FILTH, can map an entity *x* in its domain to 0, meaning *x* lacks any filth. There is no upper bound, though, to how dirty *x* can be; scales associated with PAs are thus lower closed and upper open. On the flipside, TAs include scalar maxima; once an entity is clean (i.e. lacks any filth), it can get no cleaner. As with *dirty*, there is no limit to how unclean clean an entity can be; TA scales are upper closed and lower open. Akin to our treatment of the scales associated with RAs, we'll treat one as a 'positive' scale, and the other as its inverse.

(323) $\mathcal{S}_{\text{FILTH}} = \big[0,\infty\big)$ (324) $\mathcal{S}_{\text{-FILTH}} = \big(\text{-}\infty,0\big]$

Since *clean* needs a scalar maximum, it will be the 'negative' adjective, and *dirty*, requiring a scalar minimum, will be the 'positive' adjective. This naturally leads to the following entries for *clean* and *dirty*.

(325) a. $[\![\text{clean}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{-FILTH}}.\ \text{-FILTH}_w(x) \geq d$

b. $[\![\text{dirty}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{FILTH}}.\ \text{FILTH}_w(x) \geq d$

For K&M (also Kennedy 2007), various modifiers, including **POS**, are sensitive to such scale structures. Essentially, they propose that **POS** in general picks out a scalar maximum or minimum if there is one to serve as the standard. Since RAs do not include maxima and minima in their scales, there can never be such a standard, and it must be context-dependent[19].

Picking out the standard and relating it to another degree is, as we've established, the purview of **POS**. So what are the standards for *clean* and *dirty*? Let's go through each in turn.

[19] There is some wiggle room for context dependence in the standards of absolute adjectives; Kennedy (2007), for example, proposes that contextual factors can override the default standards picked out by **POS**. Rotstein & Winter (2004) have a more complicated theory of scale structure than we do; we'll discuss this later on.

(326) Anya's room is [**POS** $pro_{4,dt}$] clean.

Earlier, we paraphrased this by saying the level of cleanliness Anya's room has must meet or exceed the 'paragon' of cleanliness—that is, Anya's room must be *as clean as* the paragon. The paragon will, in general, be itself clean, i.e. have no degree of filth whatsoever[20]. So, the standard must be the scalar maximum, i.e. 0. Given the semantics assumed for **POS**, the standard for *clean* should just be $\mathcal{S}_{\text{-FILTH}}$ itself[21]. so assume that g(4) picks out $\mathcal{S}_{\text{-FILTH}}$.

(327) $[\![\text{clean}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{-FILTH}}.\ \text{-FILTH}_w(x) \geq d$

(328) $[\![(326)]\!]^{g,c}(w) = 1$ *iff* $\textbf{MAX}(g(4)) \leq \textbf{MAX}\{(0,\mathcal{AR}_w]\}$
i.e. $0_{\text{-FILTH}} \leq \mathcal{AR}_w$, i.e. Anya's room lacks any filth

These truth conditions seem adequate.

Dirty and other partial adjectives require a bit more work. Note that the included bounds of $\mathcal{S}_{\text{-FILTH}}$ and $\mathcal{S}_{\text{FILTH}}$ are both 0—in fact, the very same 0. For Dawn's room to count as positively dirty, it must be *dirtier* than that same paragon of cleanliness. That is, its degree of filth must be strictly greater than 0. Suppose that the standard is simply the singleton set containing 0, and that this is the value of $pro_{4,dt}$.

(329) $[\![\text{dirty}]\!] = \lambda w \lambda d \lambda x\text{: } d \in \mathcal{S}_{\text{FILTH}}.\ \text{FILTH}_w(x) \geq d$

(330) $[\![$Dawn's room is [**POS** $pro_{4,dt}$ dirty$]\!]^{g,c}(w) = 1$ *iff* $0_{\text{FILTH}} \leq \mathcal{DR}_w$

These truth conditions hold even when Dawn's room lacks any filth at all—that is, when it's clean. That's not right. The truth conditions should be as follows:

(330′) $[\![$Dawn's room is [**POS** $pro_{4,dt}$ dirty$]\!]^{g,c}(w) = 1$ *iff* $0_{\text{FILTH}} < \mathcal{DR}_w$

These truth conditions are right, but some changes need to occur to derive them. Our **POS** won't cut it.

Let's note, though, that this is a more general issue. For example, in some of the literature, adjectives are taken to denote measure functions, rather than degree-individual relations as we've assumed, and as such in positive constructions **POS** is recruited not only to yield a predicate of individuals and supply a standard of comparison, but also to introduce the relation that holds between the degree returned by the measure function and the standard of comparison (Kennedy 2007[22], Husband 2011, e.g.).

[20]This is more of an idealization than is probably ultimately warranted, but nonetheless, one that is made in the literature, e.g. Kennedy & McNally (2005); also Rotstein & Winter (2004) and Kennedy (2007) for discussion.

[21]Note that we're suppressing the definedness presuppositions of **MAX** throughout this subsubsection, as well as those of **MIN** later on.

[22]Kennedy (2007) doesn't give a meaning for **POS** that looks like this, but he does give '$>$' as the relation expressed by $[\![\textbf{POS}\ \alpha]\!]$ where α is a PA and '$=$' where it is a TA. In a way, our analysis is, then, an attempt at compositionally deriving the desired semantics with a single **POS** morpheme.

(331) a. $[\![\text{tall}]\!] = \lambda x.\ \textsc{height}(x)$ $\langle e,d\rangle$

b. $[\![\text{clean}]\!] = \lambda x.\ \textsc{clean}(x)$ $\langle e,d\rangle$

c. $[\![\text{dirty}]\!] = \lambda x.\ \textsc{dirty}(x)$ $\langle e,d\rangle$

(332) $$[\![\textbf{POS}]\!]^c = \lambda A_{\langle e,d\rangle}\lambda x. \begin{cases} A(x) = \textbf{MAX}(A), & \text{if } \textbf{MAX}(A) \text{ is defined} \\ A(x) > \textbf{MIN}(A), & \text{if } \textbf{MIN}(A) \text{ is defined} \\ A(x) \geq \textbf{\textit{STND}}_c(A), & \text{elsewhere} \end{cases}$$

This **pos** checks the scale structure to determine the relation between degrees it encodes—let's call this a *variable relation* semantics. With a TA, it encodes $=$, with a PA $>$, and with a RA, $\geq$. This variability is necessary given the structure of scales associated with absolute adjectives and their standards of comparison. The necessity of some change when we assume adjectives express degree relations instead of measure functions alone is thus not isolated.

We can modify our **pos** so that it varies the relation depending on the input. First, let's grant that the standard for *dirty*, like the standard for *clean*, is simply the scale itself. This is, for us, one of the two hallmarks of an absolute adjective, the other being that the scale includes at least one of its bounds. Since only partial adjectives require modification to **pos**, we can give the following meaning.

(333) $$[\![\textbf{POS}]\!]^c = \lambda D\lambda D'. \begin{cases} \textbf{MIN}(\mathbb{D}) < \textbf{MAX}(\mathbb{D}'), & \text{if } \textbf{MIN}(\mathbb{D}) \text{ is defined} \\ \textbf{MAX}(\mathbb{D}) \leq \textbf{MAX}(\mathbb{D}') & \text{otherwise} \end{cases}$$

second version

(334) For any set of degrees $\mathbb{D}$ on a scale $\mathcal{S}$:
$\textbf{MIN}(\mathbb{D}) := \iota d\big[d \in \mathbb{D}\ \&\ \forall d'\big[\big[d' \in \mathcal{S}\ \&\ d' \geq d\big] \leftrightarrow d' \in \mathbb{D}\big]\big]$

Only PAs will trigger the **min** resolution of **pos**. The TA standard is a lower-open scale, and as such **min** is undefined for it. For RAs with contextually explicit standards, we assume they are initial intervals of a lower-open scale if the relevant *A* is positive, and if the adjective is negative, they still must be lower-open intervals. So,

(330) Dawn's room is [**pos** $pro_{4,dt}$] dirty.

(335) $[\![(330)]\!]^{g,c}(w) = 1$ *iff* $\textbf{MIN}(\mathcal{S}_{\textsc{filth}}) < \textbf{MAX}\big\{\big[0,\mathcal{DR}_w\big]\big\}$
i.e. iff $0_{\textsc{filth}} < \mathcal{DR}_w$

These truth conditions are right, too. This allows us to stick with the very standard assumption that all adjectives express the same relation. Now let's put this into action with subtractives.

Now let's get serious about the standard. Simply allowing the pronoun to be totally free, without a theory of how it is constrained by the context, seems undesirable, especially for absolute adjectives. If we were to put any constraints directly into **pos**, though, such constraints would, when we consider subtractives, apply to the *result* of subtractives. Consider a version of **pos**, let's call it **pos**′, that presupposes that the set characterized by its first argument is the standard of the scale associated with the second (read '$\textbf{STND}_c(\mathcal{S}_{D'})(\mathbb{D})$' as 'the contextual standard for the scale $\mathcal{S}$ associated with $\mathbb{D}'$ is $\mathbb{D}$').

(336) $$[\![\textbf{POS}']\!]^c = \lambda D \lambda D'\text{: } \textbf{STND}_c(\mathcal{S}_{D'})(\mathbb{D}).\begin{cases} \textbf{MIN}(\mathbb{D}) < \textbf{MAX}(\mathbb{D}'), & \text{if } \textbf{MIN}(\mathbb{D}) \text{ is defined} \\ \textbf{MAX}(\mathbb{D}) \leq \textbf{MAX}(\mathbb{D}') & \text{otherwise} \end{cases}$$

(337) Anya's room is [[almost χ_2 **pos**′] $pro_{4,dt}$] clean.

a. $[\![(337)]\!]^{g,c}(w)$ is defined only if $g(4) - g(2) = \mathcal{S}_{\text{-FILTH}}$.
Where defined, $[\![(337)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

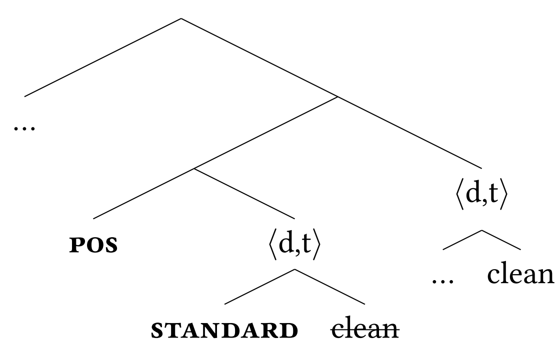
i. $\textbf{MAX}(g(4) - g(2)) \leq \mathcal{AR}_w$ & $\text{SMALL}_c(g(2))$

ii. $\forall \chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \right]$

g(4) is the contextually explicit standard, and g(2) the exception. What we want is for the exception to lower the standard just a little bit, so that Anya's room counts as clean under the modified standard. The problem is a presupposition like this would demand that the remainder of the standard pronoun minus the exception be equivalent to $\mathcal{S}_{\text{-FILTH}}$. Since *almost* modifies the first argument of **pos**, it's impossible for the presupposition to target *just* the standard pronoun.

What we can do is lean into the equative/comparative analogy even further, and take the first argument of **pos** to contain an elided copy of the adjective itself. Then, what is directly encoded into **pos** in more typical analyses—the **stnd** function—is actually encoded by a separate operator at LF, taking an adjective and returning the scale associated with that adjective should it contain a bound. As a placeholder, let's let $\textbf{STND}_c(A)$ be the otherwise case, i.e. for RAs and contextually explicit standards.

(338)



(339) $$[\![\textbf{STANDARD}]\!] = \lambda A_{\langle s,det \rangle}.\begin{cases} \mathcal{S}_{\mu_A}, & \text{if } \textbf{MIN}(\mathcal{S}_{\mu_A}) \text{ or } \textbf{MAX}(\mathcal{S}_{\mu_A}) \text{ is defined} \\ \textbf{STND}_c(A) & \text{otherwise} \end{cases}$$

$\langle\langle s,det \rangle, \langle dt \rangle\rangle$

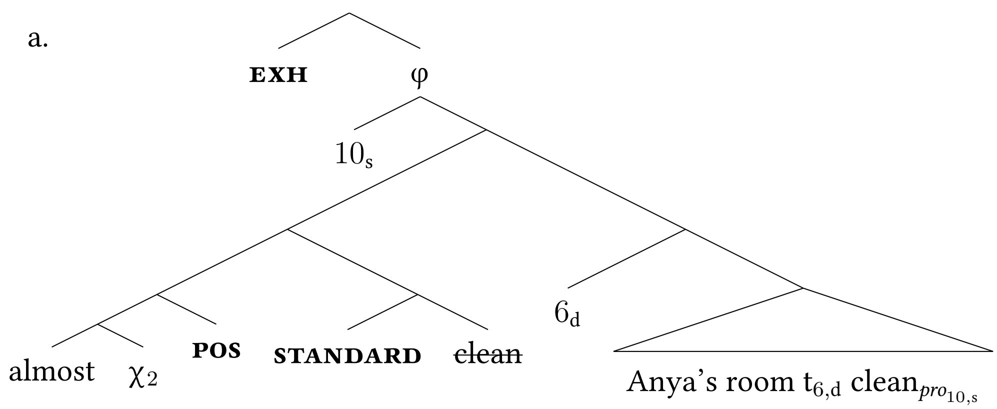
There's no need POS to carry a presupposition, since the work is done by **STANDARD**. Now *almost* will subtract from the *output* of the **STANDARD** function; the issue caused by the presupposition no longer arises. Now let's look at subtractives in earnest.

3.3.2.3 Absolute subtractives

Ultimately, the system we've developed makes TAs in the positive construction akin to equatives, and PAs comparatives. It is to be expected, then, that *almost* should be able to modify TAs and not PAs. Let's show that this is borne out; the presuppositions are once again fully fleshed out.

(340) Anya's room is almost clean.

a.



b. $[\![(340a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathcal{S}_{\text{-FILTH}}$, and $\textbf{MAX}(\mathcal{S}_{\text{-FILTH}} - g(2))$ and $\textbf{MAX}\{(-\infty, -\mathcal{AR}_w]\}$ are defined; where defined: $[\![(340a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}(\mathcal{S}_{\text{-FILTH}} - g(2)) \leq \mathcal{AR}_w$ & $\text{SMALL}_c(g(2))$

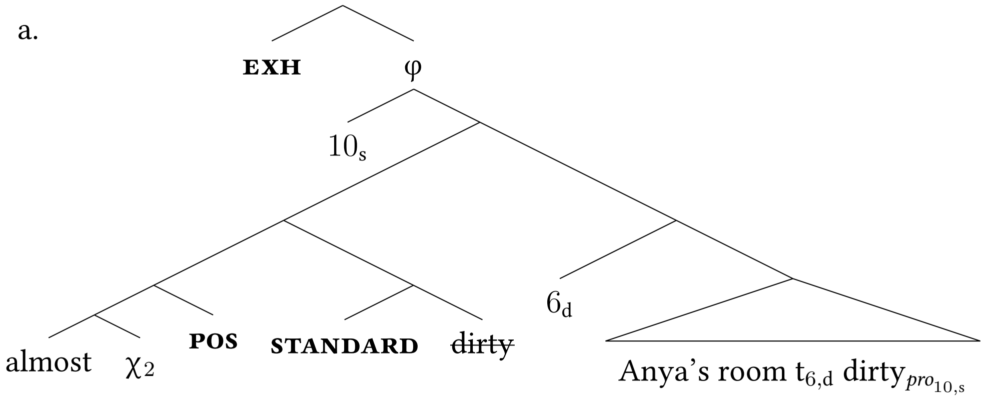
ii. $\forall \chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \right]$

This looks exactly like an *almost* equative. g(2) must be a final subinterval of $\mathcal{S}_{\text{-FILTH}}$, and g(2) must subtract the difference between $\mathcal{AR}_w$ and 0. This lowers the standard; this difference, of course, must be small. φ is Strawson Downward Entailing with respect to the first argument of **POS**, and so alternatives formed with small supersets of g(2) will be entailed; subsets won't be. So, **EXH** will require that X' denoting either supersets of g(2) that are too big or subsets of g(2) make (340bi) false or undefined when swapped for the actual exception. This is good; these truth conditions capture what we want. Anya's room is close to clean, so if we lower the threshold just a little bit, it can count as clean. Any less, and it's not.

Almost won't work with PAs on this analysis.

(341) # Anya's room is almost dirty.

a.



b. $[\![(341a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathcal{S}_{\text{FILTH}}$, and $\textbf{MIN}(\mathcal{S}_{\text{FILTH}}-g(2))$ and $\textbf{MAX}\{[0,\mathcal{AR}_w]\}$ are defined; where defined: $[\![(341a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MIN}(\mathcal{S}_{\text{FILTH}}-g(2)) < \mathcal{AR}_w$ & $\text{SMALL}_c(g(2))$

ii. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

This looks like an *almost* comparative. We'll see that density is once again the reason these truth conditions are necessarily untrue, but minimality in the first argument of **POS** means the reasoning is a bit different than what we've seen before. First, $\textbf{MIN}(\mathcal{S}_{\text{FILTH}}-g(2))$ is defined only if g(2) picks out an initial subinterval of $\mathcal{S}_{\text{FILTH}}$, and crucially is itself upper open[23]. That is, for some *n*, $g(2) = [0,n)$. For (341bi) to be true, *n* must be strictly less than $\mathcal{AR}_w$. Now, because **POS** invokes **MIN** here, and because its input is a lower closed, upper open interval, φ is actually Strawson Upward Entailing with respect to the first argument of **POS**, and so it's subsets of the exception that are entailed. (341bii) boils down to:

(342) $\forall\mathbb{X}'\big[\mathbb{X}' \not\subseteq g(2)\ \&\ \text{SMALL}_c(\mathbb{X}') \rightarrow \textbf{MIN}(\mathcal{S}_{\text{FILTH}}-\mathbb{X}') \geq \mathcal{AR}_w \vee \textbf{MIN}(\mathcal{S}_{\text{FILTH}}-\mathbb{X}') \text{ is undefined}\big]$

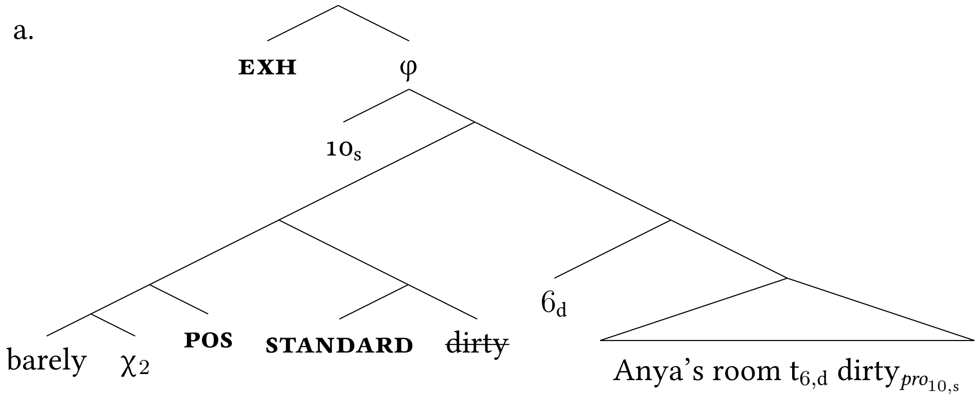
Now we can see that density makes these truth conditions impossible to satisfy. g(2) must be an upper open, proper initial subset of $[0,\mathcal{AR}_w]$ for (341bi) to be true, meaning that its upper bound is, for some small ε, $\mathcal{AR}_w-\varepsilon$. If g(2) is the largest set that makes $\text{SMALL}_c(g(2))$ true, then (341bii) is trivial. If there are supersets that are also small, there will be at least one superset of g(2) that isn't entailed but which nonetheless makes (341bii) false, e.g. the interval $[0,\mathcal{AR}_w-\frac{\varepsilon}{2})$. Density strikes again. *Almost* cannot modify PAs.

Barely can, however.

(343) Anya's room is barely dirty.

[23] We might be jumping the gun by automatically assuming **POS** yields **MIN** here. If g(2) subtracts an initial interval that is upper closed, then from the perspective of **POS**, the condition that yields **MIN** instead of **MAX** won't be met. But then **MAX** will always be undefined for $\mathcal{S}_{\text{FILTH}}$ minus such a g(2), since there is no upper bound to $\mathcal{S}_{\text{FILTH}}$. **MIN** is necessary!

a.



b. $[\![(343\text{a})]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq [0,\mathcal{AR}_w]$, and $\textbf{MIN}(\mathcal{S}_{\text{FILTH}})$ and $\textbf{MAX}\{[0,\mathcal{AR}_w]-g(2)\}$ are defined; where defined: $[\![(343\text{a})]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $0 \not< \textbf{MAX}\{[0,\mathcal{AR}_w]-g(2)\}$ & $\text{SMALL}_c(g(2))$

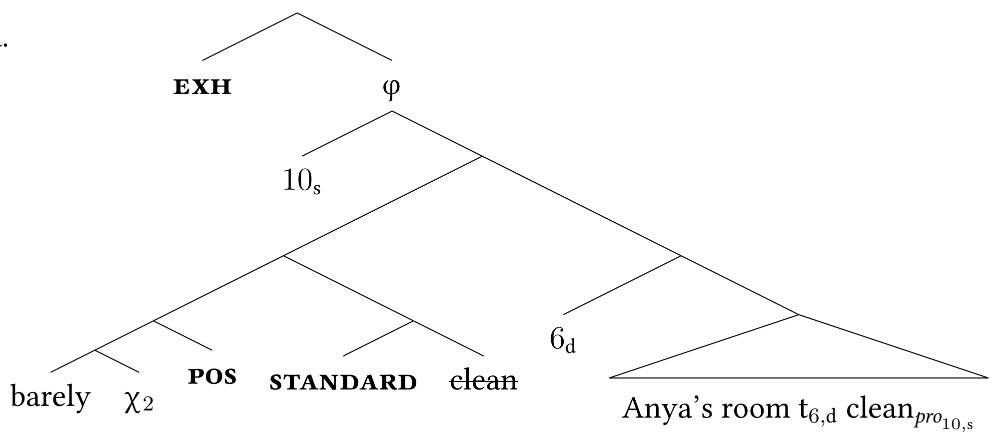
ii. $\forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

POS is generally Strawson Upward Entailing on its second argument, so φ is Strawson Downward Entailing here as a result of *barely*'s contribution. For (343bi) to be true, g(2) must be small, and remove everything that isn't 0 from $[0,\mathcal{AR}_w]$—that is, her room isn't very far at all from counting as clean. Removing any less, though, and it'd still count as positively dirty.

Since positive construction TAs are equatives, we should expect that *barely* will not be able to derive coherent, contentful truth conditions. This is borne out again.

(344) # Anya's room is barely clean.

a.



b. $[\![(344\text{a})]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq (-\infty,-\mathcal{AR}_w]$, $\textbf{MAX}\{(-\infty,-\mathcal{AR}_w]-g(2)\}$ and $\textbf{MAX}(\mathcal{S}_{\text{-FILTH}})$ are defined; where defined: $[\![(344\text{a})]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $0 \not\leq \textbf{MAX}\{(-\infty,-\mathcal{AR}_w]-g(2)\}$ & $\text{SMALL}_c(g(2))$

ii. $\forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

Barely adds nothing here. First, suppose that $0 \notin g(2)$: this means that any world *w* where (344bi) is true is a world where $0 > -\mathcal{AR}_w$, i.e. Anya's room isn't clean. In all such worlds it doesn't matter, then, what we subtract, so long as it satisfies the definedness conditions above. Then all alternatives that also meet the definedness conditions are entailed; everything else is false or yields undefinedness. If (344bi) is true, then (344bii) is true.

Now suppose that $0 \in g(2)$. This requires that $\mathcal{AR}_w = 0$, given the subset presupposition. Since φ is Strawson Downward Entailing with respect to POS's second argument, small supersets of g(2) will be entailed, then every other alternative yields undefinedness, so the truth of (344bi) guarantees the truth of (344bii). (344bi) entails (344bii)—**EXH** isn't licensed, and therefore *barely* isn't, either. That the system should predict that *barely* is useless in positive TA constructions makes sense intuitively. For Anya's room to be *barely clean* would mean that it was in fact clean, but adding a bit of gunk would make it not clean. That's already built into the semantics of TAs in the positive construction, though. *Barely* doesn't add anything!

The core proposal here is in line with general ideas in the literature about AAs and their standards (Rotstein & Winter 2004, Kennedy & McNally 2005). We proposed a variable relation, quantificational **POS** morpheme that in general has the semantics of the equative morpheme. When it combines with a partial adjective like *dirty*, **POS** contributes a comparative relation. This is a necessary modification for all theories of **POS**, it seems, once partial adjectives are taken into consideration (Kennedy 2007, Husband 2011). These ingredients together predict that *almost* can surface with total adjectives like *clean*, and *barely* with partial adjectives like *dirty*, as a product of their contributions interacting with scale structure and the semantics of **POS**. Now we'll turn to relative adjectives.

3.3.3 Relative adjectives

Relative adjectives and subtractives interact in an interesting way. When the context makes perfectly explicit a standard of comparison, both *almost* and *barely* are able to modify RAs in the positive construction. We'll see that our system can analyze *almost* RAs, but not *barely* RAs. The latter will have an analysis in the next chapter, but determining whether that's the right choice will wait. When no such explicit standard is available, RA standards are vague. Capturing this will require a more general analysis of vagueness; we'll see that neither subtractive could ever yield truth in such a situation.

3.3.3.1 Contextually explicit standards

Here's the context. We're in gym class, and we're trying to categorize everyone in terms of whether or not they're tall. We decide that, to count as positively, certainly tall, one must be five feet seven inches tall or more. Willow is $5'6.5''$, Buffy is $5'7.5''$, and Cordelia is $5'9''$.

(345) a. Cordelia is (definitely) tall.

b. Willow is almost tall.

c. Buffy is barely tall.

There are two plausible options in line with our analysis that can be used to account for the improvement in felicity of *almost* and *barely*. First, it could be that in each of these cases, there's a silent measure phrase. For example, these cases could involve ellipsis—underlying (345a) is *Willow is almost five feet seven inches tall*, and context licenses ellipsis of the measure phrase. While it's not clear how plausible such ellipsis is, this option can be made to work, but first we'll need an analysis of numerals and measure phrases. Numerals are the purview of the next chapter, though we won't fully cover measure phrases. So, the second option is to treat this not as a case of ellipsis per se, but rather that the context determines the first argument of **pos**much more strictly than is typical for RAs. Here, the context fixes it to the set corresponding to those degrees less than or equal to five feet seven inches. This will allow us to make the right predictions for *almost*, but not *barely*; more work is needed[24].

When standards are contextually explicit like this, we'll assume that **standard** function picks them out.

(346) Willow is [[almost χ_2 **pos**] [**standard** ~~tall~~]] tall.

a. $[\![(346)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \left(0{,}5'7''\right]$, and $\textbf{MAX}(\left(0{,}5'7''\right] - g(2))$ and $\textbf{MAX}\left\{\left(0{,}\mathcal{W}_w\right]\right\}$ are defined; where defined: $[\![(346)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}(\left(0{,}5'7''\right] - g(2)) \leq \mathcal{W}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

It should be obvious by now that these truth conditions are sensible, and in line with what all we've been proposing throughout. (346) is again a case of an *almost* equative; we won't rehash the logic here. What the truth conditions say is that, were we to lower the contextual standard just a bit, Willow would count as positively tall.

Barely clearly won't work, though. Since $\left(0{,}5'7''\right]$ includes no minimal element, **pos** contributes the meaning of an equative.

(347) Buffy is [[barely χ_2 **pos**] [**standard** ~~tall~~]] tall.

a. $[\![(347)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \left(0{,}\mathcal{B}_w\right]$, and $\textbf{MAX}(\left(0{,}5'7''\right])$ and $\textbf{MAX}\left\{\left(0{,}\mathcal{B}_w\right] - g(2)\right\}$ are defined; where defined: $[\![(347)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $5'7'' \not\leq \textbf{MAX}\left\{\left(0{,}\mathcal{B}_w\right] - g(2)\right\}$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\right]$

[24]Our contextually explicit standards can be understood as a particular kind of *functional* standard, in the sense of Kagan & Alexeyenko (2011) (also Solt 2012).

The problem, of course, will be density, since this is a *barely* equative. The truth conditions are necessarily untrue. What we would need is for **POS** here to encode a $<$ relation, like it does in the case when its dealing with a partial adjective, but still maintain maximality on both arguments. That is:

(347′) Buffy is [[barely χ_2 **POS**] [**STANDARD** ~~tall~~]] tall.

a. $[\![(347')]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \big(0,\mathcal{B}_w\big]$, and $\textbf{MAX}(\big(0,5'7''\big])$ and $\textbf{MAX}\big\{\big(0,\mathcal{B}_w\big] - g(2)\big\}$ are defined; where defined: $[\![(347')]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $5'7'' \not< \textbf{MAX}\big\{\big(0,\mathcal{B}_w\big] - g(2)\big\}$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

These truth conditions make the right predictions. The question is, though, how we derive them in a principled way. We would need for our **POS** to have an additional case: it encodes maximality for both arguments, but the $<$ relation between them, only when *barely* is at play with a relative adjective. This doesn't seem ideal. It seems like we might need a different route altogether; as mentioned earlier, we'll see some possible solution in the next chapter. For now, we'll leave this as an open issue.

3.3.3.2 Vague standards and relative adjectives

In general, the contextual standards for RAs are much more variable than those of AAs. The positive construction in particular has often been called vague—precisely what does and doesn't count as positively old, for example, isn't always clear. *Almost* and *barely* don't work with RAs unless the standard is contextually explicit and deterministic.

(348) # Cordelia is {almost/barely} tall.

That said, it seems like we can evaluate these kinds of sentences as false sometimes. Imagine that Cordelia is four feet tall; it just seems false to say *Cordelia is almost tall.* Here's another case.

(349) a. Boston is near New York City.

b. Los Angeles is near New York City.

(350) a. # Boston is almost near New York City.

b. ?/# Los Angeles is almost near New York City.

Considered at the national level, (349a) seems like it could be judged true, but it also could judged false. It depends. (349b) simply seems false—LA is really quite far from NYC. Adding *almost* makes things a little interesting. (350a) seems really quite odd, but (350b) seems simply false, in addition to being rather odd as far as a sentence goes. Something about the indeterminacy of cases like

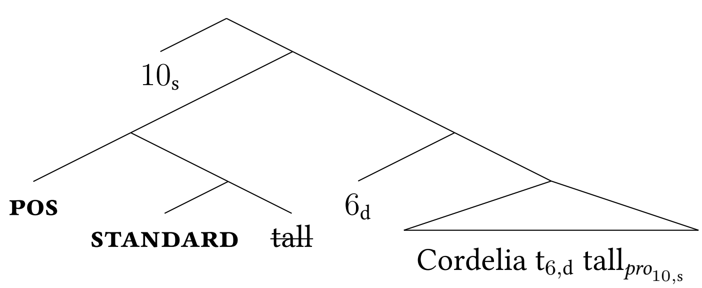
(349a) makes it even odder to have *almost* modification, whereas the latter we can outright reject because it couldn't possibly be true. We want our analysis to capture this.

The solution can be found in two facts. First, RA standards are typically taken to be vague, indeterminate, or arbitrary. In general, there is not a particular degree we can point to and say, *that* is the point at which someone is positively tall, positively bald, positively intelligent, etc. There are a host of degrees which, even in a single context, are just as good as any other degree to serve as the standard. Second, subtractives require, as we've seen, precision: the exception to a subtractive sentence has to be a unique, minimal, non-empty set. We saw that when there was no such set—e.g., with *almost some*, *barely no*, *almost* comparatives and *barely equatives*—the analysis crashed and burned. That was good, though, since we wanted to rule out such cases. Put together, the first and second facts conspire to render RAs with subtractives impossible to make true: the lack of an explicit, particular boundary with which the subtractives can operate means that we will not be able to determine a unique, minimal, non-empty exception. We will have no way to actually determine, semantically, if someone is almost or barely positively tall.

Let's turn now to the slightly more formal version of the proposal above. As it stands, the way we've structured the semantics of the positive constructions is quite precise, even for RAs. Since $\mathcal{S}_{\text{HEIGHT}}$ contains neither its upper nor its lower bound, **STANDARD** picks out a contextually-determined interval, just as it does when the standard is contextually explicit.

(351) Cordelia is tall.

a.



b. $[\![(351a)]\!]^c(w)$ is defined only if $\textbf{MAX}(\textbf{STND}_c([\![\text{tall}]\!])$ and $\textbf{MAX}\{\,(0,\mathcal{C}_w]\,\}$ are defined; where defined: $[\![(351a)]\!]^c(w) = 1$ *iff* $\textbf{MAX}(\textbf{STND}_c([\![\text{tall}]\!])) \leq \mathcal{C}_w$

RA standards are, in general, vague, though. The boundaries between the categories of positively tall, positively short, and in-between are not easy to define. Now, vagueness itself is a deep and rich topic (see van Rooij 2011, Solt 2015b, and Sorensen 2018 for overviews), and we have no horse in the race of the right analysis of vagueness. We can augment our analysis of the positive construction with a particular kind of analysis of vagueness—a supervaluationist model, as Klein (1980) applies to the positive construction—and show that can account for the way subtractives interact with RAs[25].

[25]Of course, vagueness isn't limited to the positive construction, so some model of vagueness is independently

The interpretation function $[\![]\!]$ is relativized to a context of utterance c. In general, we can think of c as one of a family of contexts that are more or less equivalent. Let's call $\mathscr{F}$(c) the family of contexts c′ that are at least as determinate as c[26]. $\mathscr{F}$(c) helps us define a rough notion of *supertruth* and *superfalsity*:

(352) **SUPERTRUTH AND SUPERFALSITY**

For some φ of type ⟨s,t⟩:

a. $[\![\varphi]\!]^{\mathscr{F}(c)}(w)$ is *supertrue* just in case $\forall c' \in \mathscr{F}(c)$, $[\![\varphi]\!]^{c'}(w)$ is true

b. $[\![\varphi]\!]^{\mathscr{F}(c)}(w)$ is *superfalse* just in case $\forall c' \in \mathscr{F}(c)$, $[\![\varphi]\!]^{c'}(w)$ is false

c. $[\![\varphi]\!]^{\mathscr{F}(c)}(w)$ is undefined otherwise

Consider RAs in the positive construction again. To assert that someone is positively tall means that in every c′, they must be positively tall. Above, we wrote that the standard is fixed in a particular c, but since we know that the standards are extremely fuzzy, for each $c' \in \mathscr{F}(c)$, the standard can be different. So for someone to be positively tall, they must count as tall with respect to every possible precisification of the standard, i.e. in every $c' \in \mathscr{F}(c)$. It must be clear that they are tall. The same goes, mutatis mutandis, for counting as positively short. Now, when we consider subtractives, they will get caught in the calculation of supertruth and superfalsity. That is, to count as positively *almost* tall, it must be the case that in all $c' \in \mathscr{F}(c)$, you count as *almost* tall. This is impossible to make true, though, assuming that across all such precisifications g maps the exception to the same value. If even one $c' \in \mathscr{F}(c)$ results in a different standard from another c″, then the *almost* sentences will relative to c′ and c″ will contradict one another. This allows an explanation, too, of why AAs and contextually explicit standards are rather different from typical RA standards: the former determine the same standard in all $c' \in \mathscr{F}(c)$, whereas the latter do not. RAs and subtractives cannot go together in the positive construction.

This also gives us a way of distinguishing the *near* cases, repeated below.

(349) a. Boston is near New York City.

b. Los Angeles is near New York City.

(350) a. # Boston is almost near New York City.

b. ?/# Los Angeles is almost near New York City.

(349a) is plausibly true, and possibly supertrue, depending on how much all the $c' \in \mathscr{F}(c)$ agree on the positive extension of nearness. (349b) is definitely superfalse. (350a) is simply undefined,

necessary more generally. As long as one's favorite analysis of vagueness can be squared with our proposal for the positive, the end result should be the same; we're not committed to any particular model of vagueness in the end.

[26] This probably needs more rigorous defining.

since it couldn't be supertrue as a result of *almost*'s contribution. We can plausibly judge (350b) as false though because it would be superfalse, both as a result of *almost*'s contribution and because (349b) is itself superfalse.

3.3.3.3 One more alternative

The vast majority of analyses of *almost* fall into camps we've discussed before, e.g. a scalar closeness semantics for the operator, and a propositional one at that. We won't worry about such analyses now. What we should discuss, though, is Rotstein & Winter (2004) (R&W), since their proposal for *almost* depends crucially on the model of the scale structure of absolute adjectives that they develop.

The core of R&W's proposal shares with ours the idea that the scales on which total and partial adjectives operate are related, but not identical. They differ from us in allowing more flexibility in the setting of the standard for both scales. Let *T* and *P* be an antonymous pair of absolute adjectives, where *T* is the total and *P* the partial adjective. The strictest requirement R&W impose comes in the relationship between the minimal degree on $\mathcal{S}_{\mathrm{P}}$ and the standard for $\mathcal{S}_{\mathrm{T}}$. To count as positively *T*, one must be totally "free of *P*-ness" (R&W:270) in their view; this accords with our view that to count as positively clean, one must lack any dirt. This means that the minimal degree on $\mathcal{S}_{\mathrm{P}}$, $\mathrm{P}_{\textsc{min}}$, must be equivalent to the standard on $\mathcal{S}_{\mathrm{T}}$, $\textbf{\textsc{stnd}}_{\mathrm{T}}$. They allow, though, that there are *T* degrees on $\mathcal{S}_{\mathrm{T}}$ both above and below $\textbf{\textsc{stnd}}_{\mathrm{T}}$. The degrees above $\textbf{\textsc{stnd}}_{\mathrm{T}}$ allow that in some contexts, we might allow that to count as positively clean needn't mean absolutely maximal cleanliness. Room below $\textbf{\textsc{stnd}}_{\mathrm{T}}$ means that *T* can map entities to a degree that doesn't mean positively *T*—this is helpful for comparatives, where *Anya's room is cleaner than Dawn's room* doesn't entail that either is clean. Furthermore, they allow for far more contextual flexibility in determining $\textbf{\textsc{stnd}}_{\mathrm{P}}$—it can be strictly greater than $\mathrm{P}_{\textsc{min}}$, though it needn't be.

(353) Where $-\infty < \mathrm{T}_{\textsc{max}} \leq \mathrm{P}_{\textsc{min}} < \mathrm{P}_{\textsc{max}}, \mathrm{T}_{\textsc{min}} \leq \infty$:

a. $\mathcal{S}_{\mathrm{P}} := \overrightarrow{(\mathrm{P}_{\textsc{min}}, \mathrm{P}_{\textsc{max}})}$

b. $\mathcal{S}_{\mathrm{T}} := \overleftarrow{[\mathrm{T}_{\textsc{max}}, \mathrm{T}_{\textsc{min}})}$

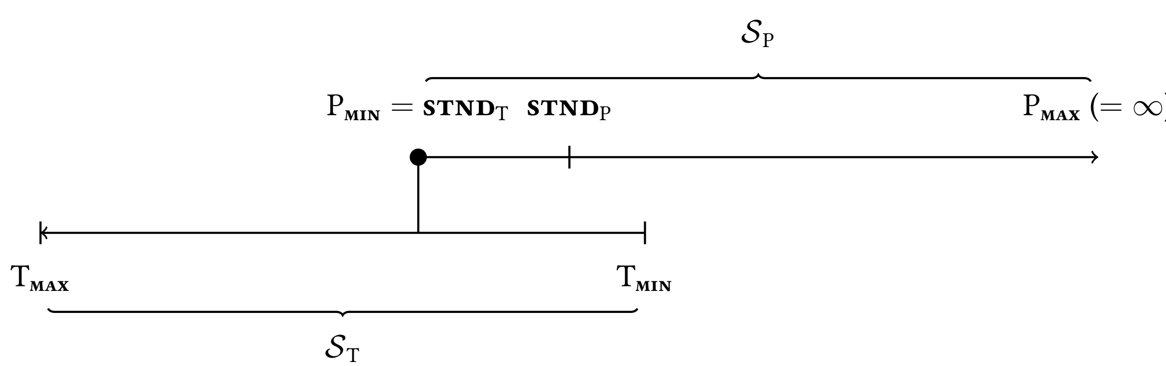


Figure 3

Positive adjectives in their analysis denote intervals on a scale.

(354) $[\![\textbf{POS}]\!]([\![T]\!](x)) = \lambda d.\ \mu_T(x) \geq \textbf{STND}_T$
i.e. $\left[T_{\textbf{MAX}},\textbf{STND}_T\right]$

(355) $[\![\textbf{POS}]\!]([\![P]\!](x)) = \lambda d.\ \mu_P(x) \geq \textbf{STND}_P$
i.e.: $\left[\textbf{STND}_P,P_{\textbf{MAX}}\right)$, if $P_{\textbf{MIN}} < \textbf{STND}_P < P_{\textbf{MAX}}$
...$\left(\textbf{STND}_P,P_{\textbf{MAX}}\right)$, if $P_{\textbf{MIN}} = \textbf{STND}_P$
...$\emptyset$, if $P_{\textbf{MAX}} = \textbf{STND}_P$

The context-dependence of $\textbf{STND}_P$ allows for the positive extension of *P* to have more variability than the positive extension of *T*.

Now, R&W propose that *almost* modifies not the **POS** operator itself, but the whole positive adjective construction. It is a function from intervals $\mathcal{I}$ on a scale $\mathcal{S}$ to the complement of $\mathcal{I}$ on $\mathcal{S}$. So for some *T*, *almost* [**POS** *T*] returns a small interval below $\textbf{STND}_T$. This is always available. *almost* [**POS** *P*] is only viable when $\textbf{STND}_P > P_{\textbf{MIN}}$, i.e. when the context makes salient a particular, non-default minimal standard. If that's not the case, then there will be no interval that *almost* [**POS** *P*] can map to.

Their analysis of AAs is really not fundamentally all that different from ours, though theirs admits flexibility in the relationship standards and contextual factors, at least in a sort of stipulative way. In general, for example, they assume that $\textbf{STND}_P = P_{\textbf{MIN}}$, but context allows the fixing of a higher degree as $\textbf{STND}_P$. It's not obvious how the context does set this different standard on their account. We would need to do a little more work to allow this, since our **POS** right now will return the maximum/minimum of a scale if there is one, and only a contextual standard if there there is neither. Provided we can come up with such a mechanism, we're set.

That said, our analysis won't readily account for the felicity of *almost P* in such cases, since that would be an *almost* comparative. However, perhaps when *almost P* is felicitous, the right analysis—as we suggested for what appeared to be superficially *almost* comparatives—is not to treat them as *almost* modifying POS, but rather the whole predicate. That is, we would analyze it structurally as R&W do, but semantically it would be quite different. We'll have to leave this to future work. The main difference is in allowing a gap between $T_{\textbf{MAX}}$ and $\textbf{STND}_T$, but that we could chalk up again to context; R&W do assume that the default value of $\textbf{STND}_T$ is $T_{\textbf{MAX}}$.

Let's discuss now how our analysis is perhaps more parsimonious than theirs when it comes to *almost* and *barely*. First, it's not clear how their analysis will extend to *barely*. Their analysis relies on *almost* picking out a small interval in the complement of $[\![\textbf{POS}\ A]\!]$. This captures, of course, the negative polar inference *almost* carries. What about *barely*? To capture its own positive polar inference, it must pick out some small interval in $[\![\textsc{pos}\ A]\!]$, but it's not clear how what we'd have to do to get the interval to be just above the standard. This comes out for free on our analysis. Their brief notes about how *almost every* works, too, leave something to be desired; they treat *almost* as a modifier of *generalized* quantifiers to generalized quantifiers, but this prevents a unified semantics of *almost every* and *almost no* (cf. Moltmann 1995), whereas we can provide a uniform semantics.

None of this is necessarily unfixable. The issues are just things that need much more ironing out, and they arise, I suspect, because they *start* from the positive construction, and want to work their way backwards. The central theoretical thrust of this dissertation is that starting from determiner quantifiers and working towards degree constructions shows us how a unified, exceptive, analysis can work for all with minimal adaptations along the way.

3.4 Onwards

The first half of this chapter analyzed comparatives and equatives modified by *almost* and *barely*, and showed that it is incredibly fruitful to extend the quantifier modifier analysis of subtractives to this domain. It is a novel extension, and importantly, the resulting theory is quite uniform. We make the meanings of *almost* and *barely* general in the following way. Let σ be an arbitrary semantic type.

(356) $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle \sigma t, \sigma tt\rangle} \lambda A_{\langle \sigma, t\rangle} \lambda B_{\langle \sigma, t\rangle}\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B}')\ \&\ \textsc{small}_c(\mathbb{X})$

(357) $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle \sigma t, \sigma tt\rangle} \lambda A_{\langle \sigma, t\rangle} \lambda B_{\langle d, t\rangle}\colon \mathbb{X} \subseteq \mathbb{B}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A})(\mathbb{B}-\mathbb{X})\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle \sigma t, \langle \sigma t, \langle \langle \sigma t, \sigma tt\rangle, \langle \sigma t, t\rangle\rangle$

Particular properties of degree constructions naturally affect the fine details of the analysis, but the general picture is very much the same as it was in analyzing quantificational determiners. The second half of this chapter is a natural extension of the first, with some embellishments. Analyzing total and partial adjectives in the positive construction as equatives and comparatives respectively naturally captures the distribution of *almost* and *barely*. Throughout the whole chapter and in the last, we've predicted the complementarity of *almost* and *barely* with respect to a given determiner. Now we'll turn to numeral constructions, where that distributional complementarity breaks down.

Chapter 4

Numeral constructions

4.1 The overview

Our analysis of *almost* and *barely* treats them as modifiers of quantifiers, and they encode set subtraction of a small *exception* from one of the modified quantifier's arguments; *barely* additionally negates the modified quantificational assertion. Let φ be an LF with exactly one subtractive Σ modifying a quantifier Δ, with an exception χ. *Almost* subtracts from the restrictor of the modified quantifier, and *barely* from the nuclear scope. Both subtractive operators must occur within the scope of the exhaustification operator EXH, and the alternatives to φ are those that are formed by substituting alternative exceptions for χ. What the analysis ultimately requires is that the exception must be the smallest, non-empty set that can be subtracted and yield truth for the modified quantification. If it impossible for there to be such a set, then the subtractives are not licensed.

What we've seen so far is complementarity in the distribution of the subtractives. When we turn to constructions with numerals modified by subtractives, though, that complementarity disappears.

(358) Almost {one hundred/half of the} vampires attacked Sunnydale High School.

(359) Barely {one hundred/half of the} vampires attacked Sunnydale High School.

(360) a. Willow is {almost/barely} a month older than Buffy is.

b. Willow is {almost/barely} sixteen years old.

c. {Almost/barely} one hundred pounds of (the) gold is on the table.

Numerals like *one hundred* and proportional quantifiers like *(a) half* are all acceptable with both *almost* and *barely*. We see this with differential measure phrases in comparatives and measure phrases constructions, too. Now, the analysis we've developed so far has been remarkably successful in predicting the distribution of the modifiers, and accurately capturing the meaning they contribute across quantifier types in with a uniformity and degree of precision not present in other extant analyses. However, it has also predicted that complementarity we have seen: if a quantifier Δ in a particular syntactic and semantic environment can be modified by *almost*, it can't be by *barely* in that same environment, and vice versa. What kind of structure and meaning for numerals allows these constructions to be modified by both operators?

A solution to the quagmire we're in is suggested by thinking a little more carefully about what *almost* and *barely* seem to do in numeral constructions. Consider the paraphrases for (358) and (359), given in (361) and (362), respectively.

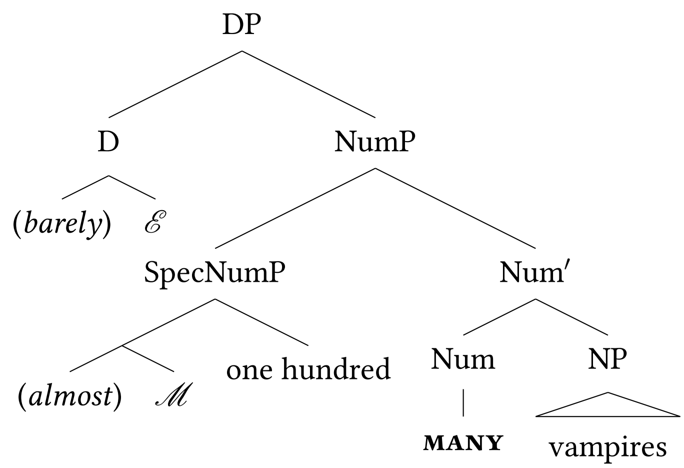
(361) *almost one hundred*: There is some set of vampire-attackers with a cardinality of just less one hundred, but it is close to one hundred.

(362) *barely one hundred*: There is some set of vampire-attackers such that if you remove a few of them, the cardinality of the set would be less than 100.

The informal paraphrases of these sentences seem to capture, roughly, the right meanings. *Almost* tells us that it's false that one hundred vampires attacked Sunnydale High School, but we've got a group of close to 100 vampire-attackers; plausibly, *almost* is lowering the threshold set by the numeral. *Barely*, on the other hand, doesn't lower the threshold; it seems to say that if we remove some small set of vampire-attackers from consideration, we wouldn't have one hundred vampire-attackers (but in fact we do). *Barely* seems like it's operating on a set of individuals, rather than the numerical threshold.

We can also simply reflect on the fact that we know what kinds of quantifiers *almost* can modify, and we know what kinds of quantifiers *barely* can modify. This leads us to the idea that both kinds of quantifiers must be present in the structure of numeral constructions. This is the ultimate proposal: that both modifiers are compatible with numeral constructions precisely because the latter involve two layers of quantification that we've already seen. First, there is covert analog of *some*, here given as $\mathscr{E}$, and second, there is a cover analog of an equative, $\mathscr{M}$. Here is the rough structure of the DP *one hundred students*.

(363)



The existential quantifier is the head of the complex DP, and *barely* can modify it, but not *almost*. Numerals themselves are not determiners at all, but rather sit in the specifier position of a dedicated projection Num. Sitting in Num is the operator **MANY**, which relates sets of individuals to degrees. Within the SpecNumP, *almost* modifies $\mathscr{M}$, a quantifier over degrees, which admits modification by *almost*, but not *barely*.

(364) $[\![\mathscr{E}]\!]^{g,c} = \lambda A \lambda B.\ \exists y \big[y \in \mathbb{A}\ \&\ y \in \mathbb{B}\big]$

(365) $[\![\mathscr{M}]\!]^{c} = \lambda D \lambda D'.\ \textbf{MAX}(\mathbb{D}) \leq \textbf{MAX}(\mathbb{D}')$

(366) For any set of degrees $\mathbb{D}$ on a scale $\mathcal{S}$:
$\textbf{MAX}(\mathbb{D}) := \iota d\big[d \in \mathbb{D}\ \&\ \forall d'\big[\big[d' \in \mathcal{S}\ \&\ d' \leq d\big] \leftrightarrow d' \in \mathbb{D}\big]\big]$

Each of these elements has precedent in the literature on the syntax and semantics of numerals in one way, shape, or form. An existential semantics for numeral DPs (henceforth numDPs) dates back to Barwise & Cooper (1981) and Keenan & Stavi (1986). **MANY**, as a way of measuring the cardinality of a plurality, is posited in many accounts including, but not limited to, Hackl (2000), Buccola & Spector (2016), and Bylinina & Nouwen (2018). Numerals have been given the semantics of maximality-based generalized quantifiers by Kennedy (2013, 2015). We take this a step further, pulling out the maximality from numerals themselves, and into an operator $\mathscr{M}$, which has the semantics of an equative. Subtractives provide an explicit argument for this kind of structure; the central goal is to show how we get there, and to show that it is the right analysis given *almost* and *barely*'s contributions.

We'll start by showing how a GQT analysis of numerals fails for both subtractives, but it is a useful starting point for *barely*. We'll then see that incorporating **MANY** is essential, and that numDPs require plural NPs; we'll briefly backtrack once more to the basic cases of quantificational determiners to discuss the implications of this. From there we'll turn to the contribution of the numeral itself, and the role that $\mathscr{M}$ plays in our analysis. We'll discuss a few areas in which issues arise for our analysis, and possible remedies; we'll also discuss a bit how $\mathscr{M}$ is related to other quantifiers that *almost* modifies. We will then conclude with a discussion of areas in which this syntax and semantics should be applied and some open issues.

4.2 The structure and meaning of numeral constructions

4.2.1 Numerals as quantificational determiners

Existential quantification has been proposed as a key component of the semantics of numeral constructions in many ways. Syntactically and semantically, one might think they are quantificational determiners; this classical, GQT view (Barwise & Cooper 1981, Keenan & Stavi 1986) is natural considering that, like *every*, *no*, and *some*, they occur prenominally, and their contributed meaning encode relations between sets. For example, *one hundred* in (367a) requires the intersection of the students and lecture-attendees to have a cardinality of at least one hundred; the lexical entry given in (367b) encodes the cardinality function '| |' which counts the number of atomic elements in a set. A typographical note: we will distinguish between numerals (1,2,...) and the corresponding English words (*one*, *two*,...). The former will appear only in the metalanguage, and generally be taken to denote degrees; the latter are expressions of the object language.

(367) a. One hundred vampires attacked Sunnydale High School..

b. $[\![\text{one hundred}]\!]^{g,c} = \lambda A \lambda B.\ |\mathbb{A} \cap \mathbb{B}| \geq 100$ $\langle et,ett \rangle$

Such a quantificational determiner semantics for numerals would seem to be along the lines of what we're looking for, given that *almost* and *barely* are quantifier modifiers. This semantics is plausible in particular for *barely*, but not *almost*: it is an existential quantifier at its core, since it requires the intersection of its arguments to be non-empty. Recall the basic, formal components of our analysis.

(368) $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle \sigma t, \sigma tt \rangle} \lambda A_{\langle \sigma,t \rangle} \lambda B_{\langle \sigma,t \rangle}{:}\ \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A} - \mathbb{X})(\mathbb{B}')\ \&\ \textsc{small}_c(\mathbb{X})$

(369) $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D}_{\langle \sigma t, \sigma tt \rangle} \lambda A_{\langle \sigma,t \rangle} \lambda B_{\langle d,t \rangle}{:}\ \mathbb{X} \subseteq \mathbb{B}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A})(\mathbb{B} - \mathbb{X})\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle \sigma t, \langle \sigma t, \langle \langle \sigma t, \sigma tt \rangle, \langle \sigma t, t \rangle \rangle$

(370) ***SUBTRACTIVE LFs***

A LF φ is a *subtractive LF* just in case φ is a LF containing exactly one subtractive operator Σ introducing an exception χ, and no other alternative-triggering expression, and φ is of type ⟨s,t⟩.

(371) ***SUBTRACTIVE ALTERNATIVES***

For a subtractive LF φ, with an exception χ of type ⟨σ,t⟩,

$\textbf{ALT}(\varphi) := \{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_\sigma\ \&\ \varphi[\chi/\chi'] \text{ is grammatical}\}$

(372) For a subtractive LF φ, with an exception χ,
$$[\![\textsc{exh}\ \varphi]\!]^{g,c}(w) = 1 \textit{ iff } [\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}(w) \neq 1 \right]$$

Let us also note that the semantics in (367b) is Upward Entailing on both of its arguments: if the intersection of $\mathbb{A}$ and $\mathbb{B}$ has a cardinality of at least one hundred, then substituting a superset of either can only increase the size of the intersection. Now, it should be obvious that that will cause us problems with *almost* simply because of the fact that this is an existential quantifier, but there's another, more pressing issue. Our analysis of *almost*, coupled with the semantics for numerals above, cannot actually lower the threshold set by the numeral given this semantics. Consider the meaning of (358) prior to application of EXH (we're suppressing the presuppositions for the moment).

(358) Almost χ_2 one hundred vampires attacked Sunnydale High School.

(373) $[\![(358)]\!]^{g,c} = \lambda w.\ |\mathbb{V}_w - g(2) \cap \mathbb{A}_w| \geq 100\ \&\ \textsc{small}_c(g(2))$

Even without EXH, the problem is evident: given the semantics of *one hundred*, the only set that *almost* has access to $\mathbb{V}_w$. As a result, once its subtractive work is done, the truth conditions still require that the intersection of the set of remaining vampires and the set of attendees is at least one hundred. Given that (358) is intuitively true where less than one hundred vampires attacked the school, this strikes us as a fundamental problem.

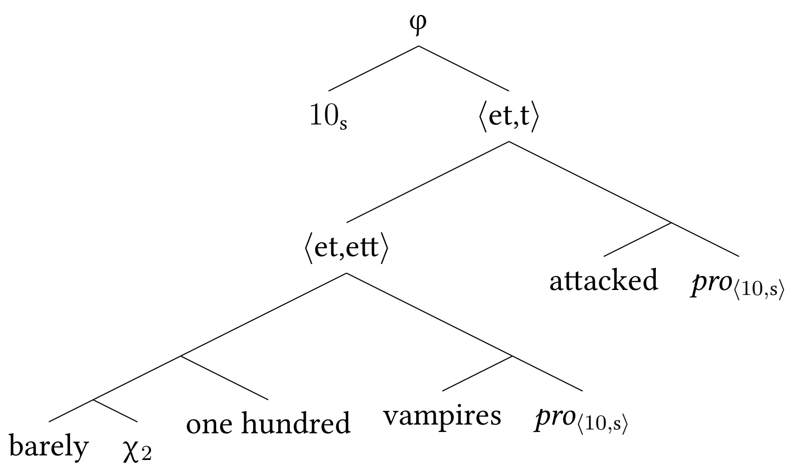
There are even problems that arise for *barely* on this classical view.

(359) Barely χ_2 one hundred vampires attacked Sunnydale High School.

(362) There is some set of vampire-attackers such that if you remove a few of them, the cardinality of the set would be less than one hundred.

What we want is for *barely* to encode that removing some small number of vampire-attackers would make it false that one hundred vampires attacked; we should also derive that one hundred vampires did in fact attack. φ is Downward Entailing with respect to the nuclear scope of *barely* χ_2 *one hundred*, so entailed alternatives to φ are formed with supersets of the exception g(2) (again, we suppress *barely*'s presuppositions).

(374)



(375) $[\![\textsc{exh}\ \varphi]\!]^{g,c}(w) = 1$ *iff* (a) & (b) = 1

a. $|\mathbb{V}_w \cap \mathbb{A}_w - g(2)| < 100\ \&\ \textsc{small}_c(g(2))$

b. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi\,[\chi_2/\chi']\,]\!]^{g,c} \rightarrow [\![\varphi\,[\chi_2/\chi']\,]\!]^{g,c}(w) \neq 1\big]$

These truth conditions, at first glance, seem to accomplish what we want. Again because of the semantics of *one hundred*, *barely*'s subtraction can only apply to sets of individuals, as was the case with *almost* (unlike *almost*, *barely* subtracts from the nuclear scope, of course). While this was a problem for *almost*, this is in fact right for *barely*. These truth conditions require that there are fewer than one hundred vampire-attackers when we remove g(2) from $\mathbb{A}_w$.

When we consider the contribution of **exh**, supersets of g(2) yield entailed alternatives, which is good. We've assumed throughout our investigation of quantificational determiners that we were trafficking in sets of atomic entities, so presumably, *barely*'s exception would be a set of atomic entities. Assuming that for numerals gets us into trouble, though. To make things simpler expositionally, let's grant that (379) is acceptable given the extensions of *vampire* and *attacked the high school* below, and grant that exceptions with a cardinality greater than than two do not count as small in the context of evaluation.

(376) $\mathbb{V}_w = \{a, b, c, d, e, f\}$ (377) $\mathbb{A}_w = \{a, b, c, d, e, g\}$ (378) $\mathbb{V}_w \cap \mathbb{A}_w = \{a, b, c, d, e\}$

(379) Barely four vampires attacked Sunnydale High School.

(380) $[\![$**exh** [*barely four vampires attacked Sunnydale High School*]$]\!]^{g,c}(w) = 1$ *iff* (a) & (b) = 1

a. $|\mathbb{V}_w \cap \mathbb{A}_w - g(2)| < 4\ \&\ \textsc{small}_c(g(2))$

b. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi\,[\chi_2/\chi']\,]\!]^{g,c} \rightarrow [\![\varphi\,[\chi_2/\chi']\,]\!]^{g,c}(w) \neq 1\big]$

There are many possible values for g(2) that predict truth: so long as g(2) contains at most two entities in $\mathbb{V}_w \cap \mathbb{A}_w$, e.g. $\{a, b\}$, $\{b, d\}$, and so on, (380a) will be true in w. This is the start of the problem: there are many sets that are possible values for g(2). **exh** is the rest of the problem.

We know that for a given $g(2)$, alternatives to φ formed with small supersets of $g(2)$ will be entailed, but since we're dealing with an existential quantifier, we know that non-supersets won't be entailed[1]. So, **EXH** requires that for all small non-supersets $\mathbb{X}'$ of $g(2)$, subtracting $\mathbb{X}'$ from the set of attackers leaves at least four vampire-attackers. For any possible value of $g(2)$ that makes (380a) true, there will be a set $\mathbb{X}'$ disjoint from $g(2)$ such that $\mathbb{X}' \subseteq (\mathbb{S}_w \cap \mathbb{A}_w)$, with the same cardinality as $g(2)$, and which makes (380a) true. It won't be entailed, though, so **EXH** will require that it makes $|\mathbb{V}_w \cap \mathbb{A}_w - X'| \geq 4$ true. This will generate falsity. The essential problem here is that there is no value for $g(2)$ that uniquely entails all other true values—there are many true values that don't stand in any asymmetric entailment relationships with one another. That there is no ordering relationship between elements in these sets creates a novel issue with numerals that is absent from the more basic quantificational determiners we investigated before[2].

There are other, independent issues with assuming a classical, GQT-style semantics for numerals. For example, other determiners can sit on top of numerals, but the quantifier semantics above doesn't make clear how they could compose (Bylinina & Nouwen 2020).

(381) Every two houses come with one parking space.

(382) The twelve students that came to the party had a nice time.

Bylinina & Nouwen (2020:3)

Still, the classical view is a useful starting point for us. It makes quite clear the issue that *almost* would face with any existential quantifier semantics for the numeral: on such a view, it could never modify the threshold set by the numeral. That contribution must be separated out from any existential component to the semantics, at least in such a way that gives *almost* access to it. Let's grant, though, that an existential semantics is a core component of numeral constructions; how does it come in?

4.2.2 MANY pluralities

Hackl (2000) introduced an analysis of numeral constructions in terms of two parts: a numeral, taken to denote a degree, and **MANY**, a parameterized existential determiner that quantifies over individuals with particular cardinalities.

(383) $[\![\textbf{MANY}]\!]^{g,c} = \lambda d \lambda A \lambda B.\ \exists x \big[x \in \mathbb{A}\ \&\ x \in \mathbb{B}\ \&\ |x| \geq d \big]$ $\langle d, \langle et, ett \rangle \rangle$

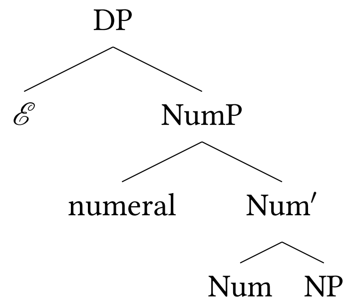
(384) $[\![\text{one hundred}]\!]^{g,c} = 100$ d

[1]See the appendix for a proof that such sets aren't entailed.

[2]On a GQT analysis, *one* and NPI *any* are plausibly equivalent, and as such should pattern alike. That this is not the case is an argument against a GQT analysis.

This proposal is for our purposes a variant of the denotation in (367b). What we want is to sever the quantification from this semantics, and align its contribution with the syntax below.

(385)



We'll take $\mathscr{E}$ to denote a standard existential quantifier:

(386) $[\![\mathscr{E}]\!]^{g,c} = \lambda A \lambda B.\ \exists x \big[x \in \mathbb{A}\ \&\ x \in \mathbb{B}\big]$

Following proposals by Ionin & Matushansky (2006), Solt (2015a), and Mendia (2018), we'll take Num to be the locus of **MANY**; its semantics is given below[3].

(387) $[\![\textbf{MANY}]\!]^{g,c} = \lambda A \lambda d \lambda x.\ x \in \mathbb{A}\ \&\ |x| \geq d$ $\qquad \langle et,\langle d,et\rangle\rangle$

This semantics allows an entity to be measured in terms of the number of its atomic parts; if there is only one atom, then the cardinality is one. What happens if the saturating degree is greater than one? We'll incorporate *pluralities* into our semantics.

Consider the set $\mathbb{S}$:

(388) $\mathbb{S} = \big\{a, b, c\big\}$

This set consists of three atoms, to be defined momentarily; pluralities are constructed from, or minimally related to, atomic elements in one way, shape, or form. An influential theory, developed by Link (1983), et seq., is what we will adopt here, though nothing crucial hinges on this assumption[4]. Link proposed that plural individuals are constructed through a summation operation, $\oplus$, and that atomic and plural individuals are no different type-theoretically. That is, D_e consists of atomic individuals and plural individuals. The $\oplus$ operator is associative, commutative, and idempotent:

(389) a. $(a \oplus b) \oplus c = a \oplus (b \oplus c)$ *associative*

[3]An alternative would be to treat it as a gradable adjective, with only a degree and an individual argument (as in, e.g., Rett (2008), Wellwood (2014), Solt (2015a), Buccola & Spector (2016), Bylinina & Nouwen (2018), a.o.). That would be equivalent for our purposes semantically, but would involve a different syntactic structure: our SpecNumP, the numeral, would have to be within the Num head, i.e. sit as the first argument of **MANY**. Nothing deeper hinges on the distinction.

[4]This discussion follows Nouwen (2016).

b. $a \oplus b = b \oplus a$ *commutative*

c. $a \oplus a = a$ *idempotent*

This allows us to define a partial order $\leq$ on D_e, and we can then define atoms in terms of $\leq$.

(390) $a \leq b$ *iff* $a \oplus b = b$.

(391) ATOM(a) *iff* $\forall b \big[b \leq a = a \big]$.

Now we take from Link an operator, $*$, which takes a set of atomic individuals and returns a set of atomic and plural individuals. Applied to $\mathbb{S}$ above, it derives the set consisting of a, b, c, and all their joins. The structure of this set is a join semi-lattice.

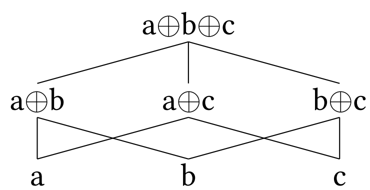
(392) $*\mathbb{A}$ is the smallest set such that:

a. $\mathbb{A} \subseteq *\mathbb{A}$, and

b. $\forall x \forall y \big[x \in *\mathbb{A} \ \& \ y \in *\mathbb{A} \rightarrow x \oplus y \in *\mathbb{A} \big]$

(393) $*\mathbb{S} = \{a, b, c, a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c\}$

(394)



$*\mathbb{S}$—call it the plural set of $\mathbb{S}$—contains all the elements in $\mathbb{S}$, plus all their possible joins. It is common to assume that plural morphology—e.g., *-(e)s* on many nouns in English—is the morphosyntactic correlate to the $*$ operator. That is:

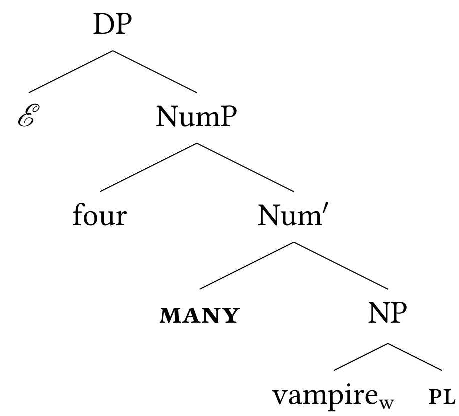
(395) $[\![\textsc{pl(ural)}]\!]^{g,c} = \lambda A \lambda x.\ x \in *\mathbb{P}$ $\langle et,et \rangle$

The English expression *vampires* can be taken to comprise two functions at LF.

(396) $[\![\textsc{pl}]\!]^{g,c}([\![\text{vampire}]\!]^{g,c}(w)) = \lambda x.\ x \in *\mathbb{V}_w$

The plural *students*, then, is a set of atomic and plural individuals. Now let's put the pieces together for *four vampires.*

(397) four vampires

a. 

b. $[\![(397a)]\!]^{g,c} = \lambda B.\ \exists x\big[x \in *\mathbb{V}_w \ \&\ x \in \mathbb{B} \ \&\ |x| \geq 4\big]$

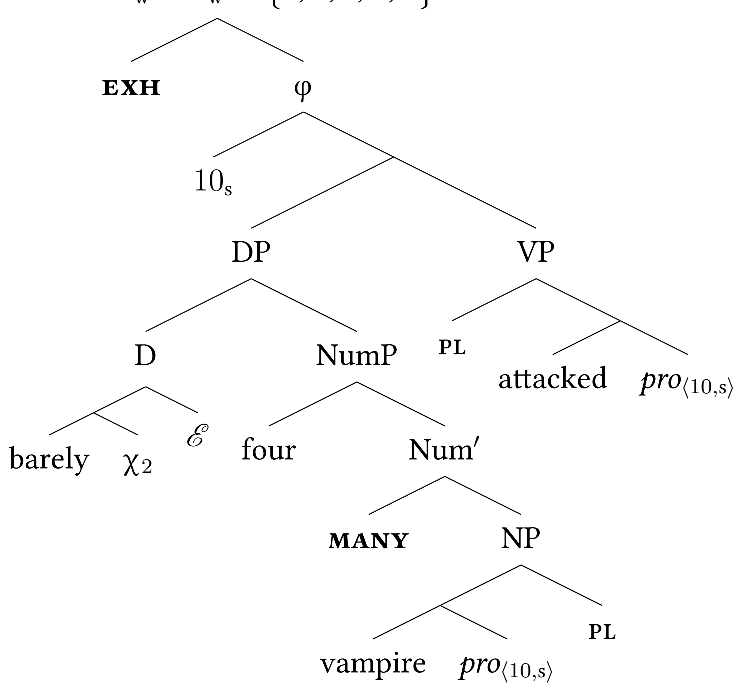
This semantics makes *four vampires* a generalized quantifier. We'll need to assume that the nuclear scope of this quantifier is also pluralized. We'll just stick a PL there, too.

4.2.3 *Barely* back in the mix

This semantics composes well with *barely*. Let's return again to a context in which five vampires attacked Sunnydale High, granting that (398) is felicitous for the sake of simplicity in explication. *Barely*'s presupposition is put back in.

(398) Barely four vampires attacked Sunnydale High School.

a. Let w be such that $\mathbb{V}_w \cap \mathbb{A}_w = \big\{a, b, c, d, e\big\}$

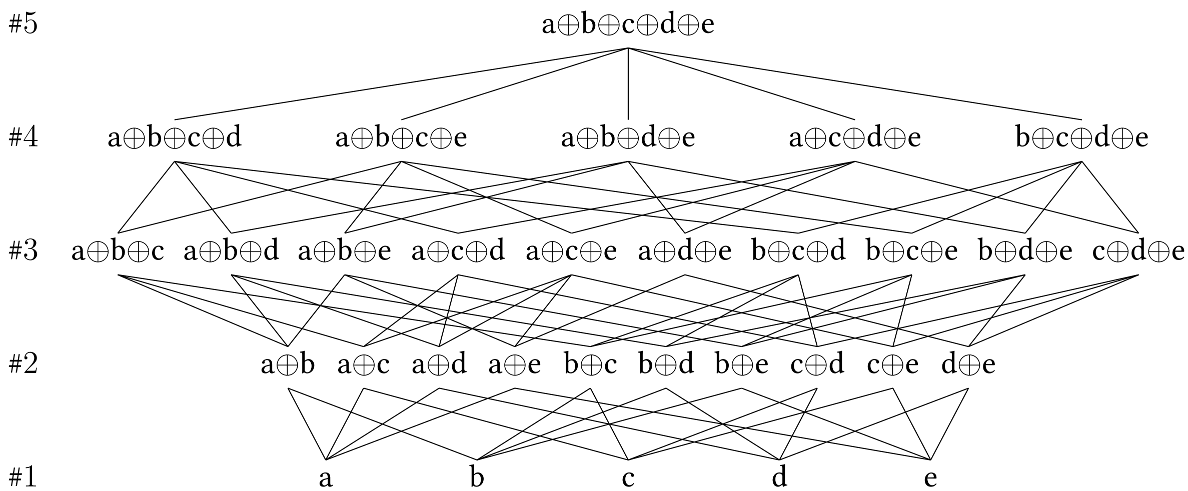
b. 

c. $[\![(398b)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{A}_w \ \&\ g(2) \neq \emptyset$; where defined:
$[\![(398b)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\neg\exists y\big[y \in *\mathbb{V}_w \;\&\; y \in *\mathbb{A}_w - g(2) \;\&\; |y| \geq 4\big] \;\&\; \textsc{small}_c(g(2))$

ii. $\forall \chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c} \neq 1\big]$

The truth conditions are very different than what we derived before. (398bi) requires that g(2) is small and removes all pluralities of four+ vampire-attackers. Consider the set of pluralities of vampire-attackers in w.



$*\mathbb{V}_w \cap *\mathbb{A}_w$

Figure 1

Because $\mathscr{E}$, and by extension *barely* $\mathscr{E}$, is an intersective quantifier, only pluralities that are in the intersection of $*\mathbb{V}_w$ and $*\mathbb{A}_w$ matter, i.e. those in Fig. 1; (398bi) can only be true if g(2) contains all pluralities in #4 and #5, i.e. those with a cardinality of at least four.

(399) $\big\{a{\oplus}b{\oplus}c{\oplus}d,\ a{\oplus}b{\oplus}c{\oplus}e,\ a{\oplus}b{\oplus}d{\oplus}e,\ a{\oplus}c{\oplus}d{\oplus}e,\ b{\oplus}c{\oplus}d{\oplus}e,\ a{\oplus}b{\oplus}c{\oplus}d{\oplus}e\big\}$

If g(2) does not contain *at a minimum* all pluralities in (399), then (398bi) is false, since there is a plurality of vampire-attackers with a cardinality of at least four in $*\mathbb{V}_w \cap *\mathbb{A}_w - g(2)$.

What about (398bii)? The additional ordering imposed by pluralization helps constrain relations between possible exceptions. φ is Strawson Downward Entailing with respect to the VP, and so alternatives to φ formed with supersets of g(2) that are small and subsets of $*\mathbb{A}_w$ are entailed[5]. Alternatives with $\mathbb{X}'$ that are not subsets of $*\mathbb{A}_w$ or which are empty are undefined, so they make the consequent of **exh** true. Alternatives where $\mathbb{X}'$ are not small are false, and also make the consequent of **exh** true. (398) thus amounts to the following.

(398′) a. $[\![(398b)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{A}_w \;\&\; g(2) \neq \emptyset$; where defined:

[5]If you're wondering how we determine SMALLness when it comes to sets of pluralities, hold on to your hats; we'll discuss this in §4.2.3.2.

$$\llbracket(398\text{b})\rrbracket^{g,c}(w) = 1 \textit{ iff } (i) \;\&\; (ii) = 1$$

$$\text{i.}\;\; \neg\exists y\big[y \in *\mathbb{V}_w \;\&\; y \in *\mathbb{A}_w - g(2) \;\&\; |y| \geq 4\big] \;\&\; \textsc{small}_c(g(2))$$

$$\text{ii.}\;\; \forall \mathbb{X}'\Big[\big[\mathbb{X}' \subseteq *\mathbb{A}_w \;\&\; \mathbb{X}' \neq \emptyset \;\&\; \textsc{small}_c(\mathbb{X}') \;\&\; g(2) \not\subseteq \mathbb{X}'\big] \rightarrow \exists y\big[y \in *\mathbb{V}_w \;\&\; y \in *\mathbb{A}_w - \mathbb{X}' \;\&\; |y| \geq 4\big]\Big]$$

It should be easy to see that these truth conditions yield truth if g(2) = (399).

Now let's imagine g(2) contains everything in (399), plus one additional plurality; for concreteness, a⊕b⊕c.

(400) $\{a{\oplus}b{\oplus}c{\oplus}d,\ a{\oplus}b{\oplus}c{\oplus}e,\ a{\oplus}b{\oplus}d{\oplus}e,\ a{\oplus}c{\oplus}d{\oplus}e,\ b{\oplus}c{\oplus}d{\oplus}e,\ a{\oplus}b{\oplus}c{\oplus}d{\oplus}e,\ \underbrace{a{\oplus}b{\oplus}c}_{\text{new!}}\}$

This set would make (398bi) true. **EXH** then does its work, but predicts falsity. Since it quantifies over non-supersets of g(2), any subset of g(2) is caught within its grasp, and all such subsets must leave at least one four+ plurality in $*\mathbb{V}_w \cap *\mathbb{A}_w - \mathbb{X}'$. The problem is that (399) is a subset of (400), which means not all non-supersets of (400) satisfy **EXH**; we predict falsity here. Taken together, then, the truth conditions require that g(2) be the smallest exception that removes all pluralities of vampire-attackers with a cardinality of at least four.

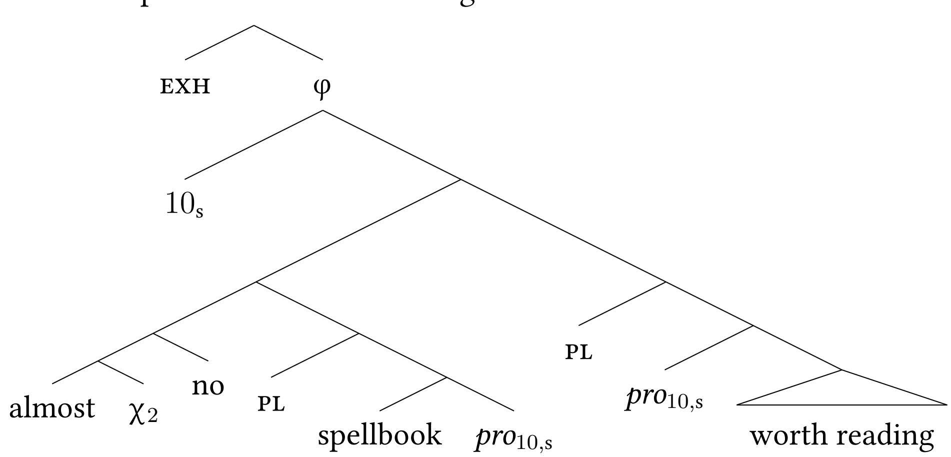
4.2.3.1 Pluralities and quantificational determiners

A natural question that arises is whether pluralities affect how we analyze quantificational determiners. There are two major points. First, with our intersective quantifier *no*, pluralities don't change anything foundational about the semantics. When we turn to *some*, we see potential problems. Pluralization does *not* obviously work with *every*; that said, it's entirely reasonable to think that *every* itself either "undoes" pluralization, if present, or blocks its presence, given that it is standardly taken to be a *distributive* determiner (Vendler 1967, i.a.).

Let's look at *no*, with *almost* as our focus.

(401) Almost no spellbook is worth reading.

a.



b. $[\![(401)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{S}_w$ & $g(2) \neq \emptyset$; where defined:
$[\![(401)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $*\mathbb{S}_w - g(2) \cap *\mathbb{WR}_w = \emptyset$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \right]$

Prior to pluralization, g(2) had to contain all the elements in $\mathbb{S}_w$ not in $\mathbb{WR}_w$, which were atoms. Consider this toy model for w.

(402) $\mathbb{S}_w = \{a, b, c, d, e\}$

(403) $\mathbb{WR}_w = \{d, e, f, g, h\}$

For (401) to be true in w, g(2) must pick out all the spellbooks worth reading (ignore concerns about size), i.e. the intersection of these two sets, $\{d, e\}$. If we were to add in pluralization, the intersection of the plural sets would contain all and only those atoms in $\mathbb{S}_w$ and $\mathbb{WR}_w$ alongside any pluralities formed with those atoms.

(404) $*\mathbb{S}_{w'} \cap *\mathbb{WR}_{w'} = \{d, e, d \oplus e\}$

This is all g(2) needs to pick out; the rest of the logic proceeds as if we were considering only sets of atomic entities. Alternatives formed with small supersets of g(2) will be entailed, and nothing else will be; **EXH** requires that other possible exception leave in $*\mathbb{S}_{w'}$ a spellbook worth reading. This is true for all non-supersets of g(2), so we're golden. *Almost some* is a little different.

(405) # Almost some spellbook is worth reading.

a. $[\![(405)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{S}_w$ & $g(2) \neq \emptyset$; where defined:
$[\![(405)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $*\mathbb{S}_w - g(2) \cap *\mathbb{WR}_w \neq \emptyset$ & $\textsc{small}_c(g(2))$

ii. $\forall\chi' \left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \right]$

As long as $|*\mathbb{S}_w \cap *\mathbb{WR}_w| > 1$, there is no value for g(2) such that both (405ai) and (405aii) are true. If $|*\mathbb{S}_w \cap *\mathbb{WR}_w| = 1$, we can find a unique exception, we have to rely on SMALL to rule it out. Here's such a world.

(406) $\mathbb{S}_w = \{a, d\}$

(407) $\mathbb{WR}_w = \{d, e, f, g, h\}$

(408) $*\mathbb{S}_w = \{a, d, a \oplus d\}$

(409) $\mathbb{S}_w \cap \mathbb{WR}_w = *\mathbb{S}_w \cap *\mathbb{WR}_{w'} = \{d\}$

When we start pluralizing, for (405ai) to be true, g(2) still must not contain d, and must be a subset of $*\mathbb{S}_w$. There are three such sets.

(410) $\{a{\oplus}d\}$ (411) $\{a\}$ (412) $\{a, a{\oplus}d\}$

$\{a{\oplus}d\}$ entails neither of the other sets, so they will make (405aii) false, but the other two sets are different. First, $\{a, a{\oplus}d\}$: it entails all other sets that make (405ai) true because subsets of any true value are entailed. Every other set $\mathbb{X}'$ that is a subset of $*\mathbb{S}_w$ contains d, and could not be entailed, and they are correctly negated by **EXH**. So, we should predict this to be true. We have recourse here to SMALL: such an exception couldn't be considered small. Perhaps that's sufficient to rule this out.

$\{a\}$ is the real problem. It entails both $\{a{\oplus}d\}$ and $\{a, a{\oplus}d\}$, even though they aren't subsets, as a result of how the $*$ operator works. To show they are entailed, consider the following. If $\{a\}$ didn't entail $\{a{\oplus}d\}$, we should be able to find a world w′ where $*\mathbb{S}_{w'} - \{a\} \cap *\mathbb{WR}_{w'} \neq \emptyset$ is true and $*\mathbb{S}_{w'} - \{a{\oplus}d\} \cap *\mathbb{WR}_{w'} = \emptyset$, as well as a world w″ where $*\mathbb{S}_{w''} - \{a\} \cap *\mathbb{WR}_{w'} \neq \emptyset$ is true and $*\mathbb{S}_{w''} - \{a, a{\oplus}d\} \cap *\mathbb{WR}_{w'} = \emptyset$[6].

Assume there is a world w′ such that $*\mathbb{S}_{w'} - \{a\} \cap *\mathbb{WR}_{w'} \neq \emptyset$ is true and $*\mathbb{S}_{w'} - \{a{\oplus}d\} \cap *\mathbb{WR}_{w'} = \emptyset$. This means that $*\mathbb{S}_{w'} \cap *\mathbb{WR}_{w'} \subseteq \{a{\oplus}d\}$. But then since $a{\oplus}d \in *\mathbb{S}_{w'} \cap *\mathbb{WR}_{w'}$, all its constituent atoms must be, so $d \in *\mathbb{S}_{w'} \cap *\mathbb{WR}_{w'}$[7]. This means that $d \in *\mathbb{S}_{w'} - \{a{\oplus}d\} \cap *\mathbb{WR}_{w'} \neq \emptyset$, which contradicts our assumption. $\{a\}$ entails $\{a{\oplus}d\}$. The same proof works for $\{a\}$ entailing $\{a, a{\oplus}d\}$. This means that $\{a\}$ entails all true exceptions, too. That's a problem!

There are a few ways we could work around this issue. We could make the size constraint stronger: the exception cannot be greater than $\frac{1}{3}$ the size of the restrictor. Perhaps that's fine and intuitive to do (personally, the $\frac{1}{2}$ mark seems a bit high anyways). If this isn't a tolerable route, we could abandon appeal to SMALL and follow Gajewski (2013). In contexts like this, where there is no *unique* exception, Gajewski proposes that an additional layer of exhaustification, negating higher order alternatives. We won't push that here. Finally, we could maintain that in QPs like this, there isn't pluralization, and maintain the $\frac{1}{2}$ threshold set previously. Deciding definitively between these options is left for the future; for concreteness, though, we'll adopt the stronger constraint on SMALL.

(413) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, μ_c be a measure function, determined by the context, mapping sets to numerical values representing their size, and n_c be a contextually determined numerical threshold for size.

$\text{SMALL}_c(\mathbb{X}) = 1$ *iff* $\mu_c(\mathbb{X}) \leq n_c$

(414) ***MAXIMAL SIZE OF*** n_c

[6] I'm assuming they all satisfy (405)'s definedness conditions, i.e. they are all subsets of $*\mathbb{S}$ in the relevant worlds.

[7] Note: this isn't true for predicates in general, but distributive predicates. Collective predicates, which we do not discuss here, can have pluralities in their extension and not the atoms comprising those pluralities

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.
For any context c, an utterance of φ in c is felicitous only if $n_c < \frac{1}{3}(\mu_c(\mathbb{A}))$

Now let's briefly turn to *every*.

Pluralization is a real problem for our universal quantifier.

(415) Almost every spellbook is worth reading.

a. $[\![(415)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{S}_w$ & $g(2) \neq \emptyset$; where defined:
$[\![(415)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $*\mathbb{S}_w - g(2) \subseteq *\mathbb{WR}_w$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' [[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1]$

Here, as it did when we were only trafficking in atoms, g(2) must contain all the elements of $*\mathbb{S}_w$ not in $*\mathbb{WR}_w$. However, the problem is g(2) now has to be much, much larger—it not only has to contain every atom x in $*\mathbb{S}_w$ not in $*\mathbb{WR}_w$, but every plurality containing x. For any such x, at least half the elements in $*\mathbb{S}_w$ contain x. Here's a toy example:

(416) $\mathbb{S}_w = \{a, b, f\}$

(417) $\mathbb{WR}_w = \{a, b, c, d, e\}$

(418) $*\mathbb{S}_w = \{a, b, f, a \oplus b, a \oplus f, b \oplus f, a \oplus b \oplus f\}$

Without pluralization, g(2) would have to be $\{f\}$, which is small relative to $\mathbb{S}_w$ on our original definition. With pluralization, though, g(2) has to contain f and all the pluralities in $*\mathbb{S}_w$ of which it is a subpart; that's four of the seven. Even on our original definition that's too much! The problem only gets worse if there is more than one element in the exception. One thing to think about is *how* we calculate the size of the exception—right now, we're just counting the number of elements in the exception and restrictor. There might be other ways to do this, and we'll turn to this in the next section. Another angle is to claim that *every* is a *distributive* determiner (Vendler 1967, i.a.): when a predicate is necessarily collective, like *gather* is, *every* is unacceptable:

(419) # Every vampire gathered at dusk.

For present purposes, we could take this to mean one of two things: either *every* never composes with a plural restrictor, perhaps through a presupposition, or its semantics 'undoes' any pluralization. Below are two quick-and-dirty entries to circumvent the issue that pluralization poses for *every* here.

(420) a. $[\![\text{every}_1]\!]^{g,c} = \lambda A \lambda B{:}\ \forall y[y \in \mathbb{A} \rightarrow \textsc{atom}(y)].\ \mathbb{A} \subseteq \mathbb{B}$

b. $[\![\text{every}_2]\!]^{g,c} = \lambda A \lambda B.\ \forall y \big[y \in \mathbb{A}\ \&\ \textsc{atom}(y) \rightarrow y \in \mathbb{B}\big]$

We'll set making a firm decision aside for now, and where *every* comes up, we'll simply not include pluralization.

It's important to note that *all*, famously compatible with *almost*, does permit collective quantification with surface-plural restrictors.

(421) Almost all of the vampires gathered at dusk.

Predicates like *gather* are not distributive, nor can they be used with singular subjects:

(422) # Willow gathered at dusk.

All can be analyzed as a non-distributive counterpart to *every*, but subtractives tell us that can't be right. If *all* were indeed just a non-distributive version of *every*, the counting problem discussed above should arise, and *almost* should not be able to modify *all*. There's plenty of literature that discuss the contribution of *all* as non-quantificational (e.g. Brisson 1998, 2003; Križ 2015, 2016; Bar-Lev 2018) or as a quantifier over candidate interpretations rather than entities (Križ & Spector 2021). Whichever route is right, there's still a puzzle as to how *almost* composes (what would it be, for example, to subtract 'candidate interpretations'?). At the very least, subtractives push us towards those investigating those puzzles.

Now we'll turn to the promised discussion of how we calculate the size of an exception when pluralities are considered.

4.2.3.2 Thinking about size again

As is usual, an exception must count as small in the context. We defined the size constraint in the following way.

(423) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, and n_c be a contextually determined numerical threshold for size.

Where $\mathbb{X}$ is a dense set of degrees, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\textsc{length}(\mathbb{X}) \leq n_c$;

Otherwise, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $|\mathbb{X}| \leq n_c$

(424) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = \big[\Delta\ \big[\alpha\ \big[\Sigma\ \chi\big]\big]\big]$ or $\big[\big[\big[\Sigma\ \chi\big]\ \Delta\big]\ \alpha\big]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

An utterance of φ in c is felicitous only if, for any context c:

Where $\mathbb{A}$ is a dense set of degrees, $n_c < \frac{1}{3}(\textsc{length}(\mathbb{A}))$;
Otherwise, $n_c < \frac{1}{3}|\mathbb{A}|$

In the context of quantificational determiners like *every* and *no* when their arguments are sets of atomic elements, it makes sense intuitively to treat μ_c as a cardinality function. The exception cannot contain more than a third of the elements in the set from which it is subtracted. When we turn to pluralities, we might try to simply apply this definition: we need simply determine if the number of pluralities in the exception is small relative to the number of pluralities in the restrictor. Our toy example is repeated below.

(425) Barely four vampires attacked Sunnydale High School.

a. In w, $\mathbb{V}_w \cap \mathbb{A}_w = \{a, b, c, d, e\}$

b. $[\![(425b)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{A}_w$; where defined:
$[\![(425b)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\neg\exists y\big[y \in *\mathbb{V}_w \;\&\; y \in *\mathbb{A}_w - g(2) \;\&\; |y| \geq 4\big] \;\&\; \textsc{small}_c(g(2))$

ii. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi\,[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi\,[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

Here, $g(2)$ is not a dense set, so we are presumably counting pluralities. It contains six pluralities, roughly 20% of the thirty-one pluralities in $*\mathbb{V}_w \cap *\mathbb{A}_w$, so we know that $g(2)$ is no more than 20% of the elements in $*\mathbb{V}_w$. Considering that *barely n vampires attacked* generally conveys that the actual number of attacking vampires was much lower than the number we expected, it seems plausible to assume that the size of $g(2)$ is dwarfed by the size of $*\mathbb{V}_w$.

One might wonder whether this way of calculating the size of the exception when pluralities are concerned is reasonable. That is, when using *barely* as a modifier in a numeral construction in the way proposed above, does the judgement proceed by counting pluralities? If we were to deny this, what could we do? If we want SMALL to circumvent the supplied pluralities, here's a modification to our proposal above.

(426) ***THE SIZE CONSTRAINT***

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, and n_c be a contextually determined numerical threshold for size.
Where $\mathbb{X}$ is a dense set of degrees, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\textsc{length}(\mathbb{X}) \leq n_c$;
Otherwise, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $|\{y \mid y \in \mathbb{X} \;\&\; \textsc{atom}(y)\}| \leq n_c$

(427) ***MAXIMAL SIZE OF*** n_c

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

An utterance of φ in c is felicitous only if, for any context c:
Where $\mathbb{A}$ is a dense set of degrees, $n_c < \frac{1}{3}(\textsc{length}(\mathbb{A}))$;
Otherwise, $n_c < \frac{1}{3}|\{y \mid y \in \mathbb{A} \; \& \; \textsc{atom}(y)\}|$

This seems sufficient for our purposes: we don't count pluralities, but rather, the number of unique atoms across the pluralities we remove.

4.2.4 The numeral itself

We've assumed so far that numerals denote degrees, and saturate the degree argument of **MANY**. There is nowhere, yet, that *almost* can fit in. Given that we know that *almost* can modify universal quantifiers and equatives, we want to look for such a quantifier within numeral constructions, and in particular, one that is intimately connected to the numeral itself. We'll propose it's a quantifier over degrees, which we'll call $\mathscr{M}$, with the semantics of an equative.

(428) $[\![\mathscr{M}]\!]^{g,c} = \lambda D \lambda D'.\ \textbf{MAX}(\mathbb{D}) \leq \textbf{MAX}(\mathbb{D}')$

We'll further posit that numerals denote *sets* of degrees, in line with our more general enterprise of set subtraction. Others have assumed that degrees denote $\langle d,t \rangle$ functions, e.g. Anderson (2015) and Mendia (2018); for Mendia, they have the following meaning:

(429) $[\![\text{four}]\!]^{g,c} = \lambda d.\ d = 4$

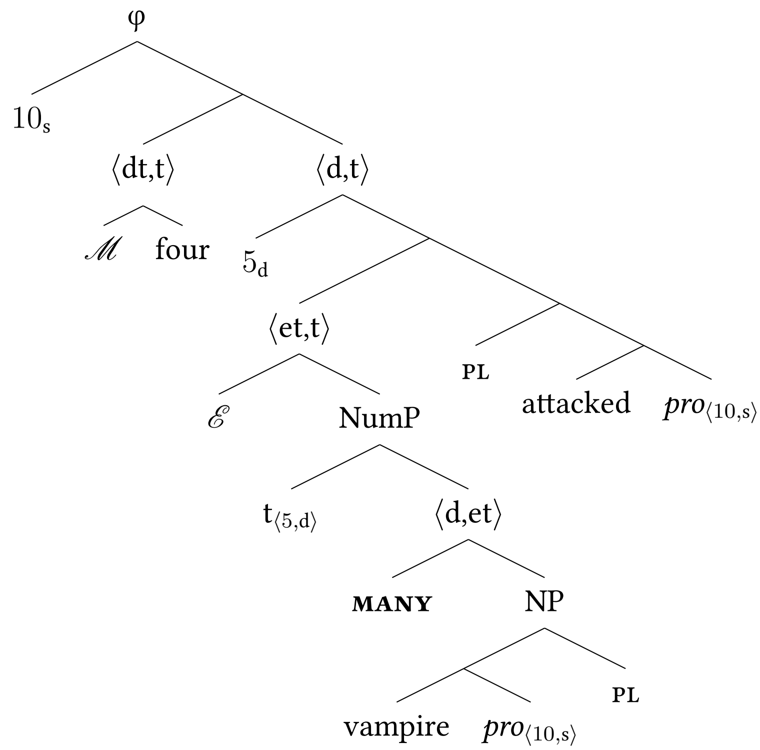
Such a semantics won't work for our subtraction: this is a singleton set, and won't allow us to subtract anything more than that singleton, so, we'll do things slightly differently.

(430) $[\![\text{four}]\!]^{g,c} = \lambda d.\ 4 \geq d$

Putting the pieces together, we get the following LF.

(431) Four vampires attacked Sunnydale High School.

a.

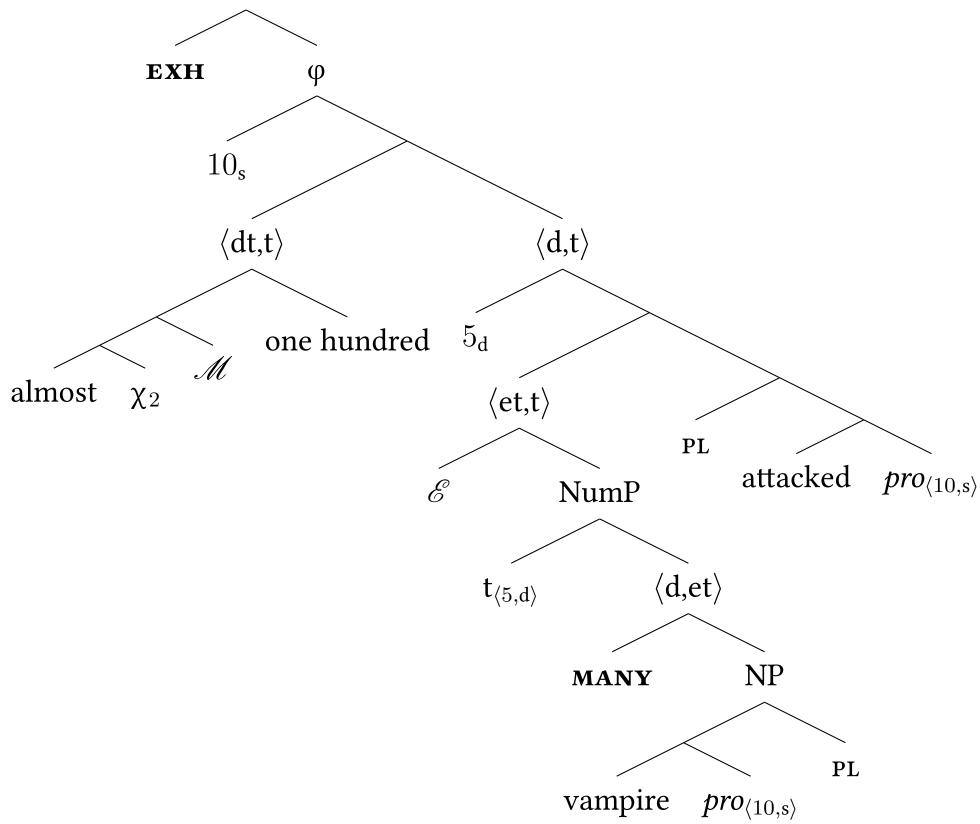


b. $[\![(431a)]\!]^{g,c}(w)$ is defined only if $\textsc{max}\{d \mid 4 \geq d\}$ is defined; where defined:
$[\![(431a)]\!]^{g,c}(w) = 1$ *iff* $\textsc{max}\{d \mid 4 \geq d\} \leq \textsc{max}\{d \mid \exists x[x \in *\mathbb{V}_w \;\&\; x \in *\mathbb{A}_w \;\&\; |x| \geq d]\}$
i.e. there is a vampire-attacker plurality consisting of at least four atomic vampires

The maximal degree in $\{d \mid \exists x[x \in *\mathbb{V}_w \;\&\; x \in *\mathbb{A}_w \;\&\; |x| \geq d]\}$ is the size of the largest plurality of vampire-attackers, generating an at-least interpretation of the numeral. *Almost* now has a foothold. It modifies $\mathscr{M}$; its integration is straightforward. φ is Strawson Downward Entailing on the restrictor of $\mathscr{M}$, small supersets to any exception will be entailed. For concreteness, assume ninety-eight vampires attacked Sunnydale High School.

(432) Almost one hundred vampires attacked Sunnydale High School.

a.



b. $[\![(432a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \{d \mid 100 \geq d\}$, $g(2) \neq \emptyset$, and $\textbf{MAX}\big(\{d \mid 100 \geq d\} - g(2)\big)$ and $\textbf{MAX}\{d \mid \exists x[x \in *\mathbb{V}_w \ \&\ x \in *\mathbb{A}_w \ \&\ |x| \geq d]\}$ are defined; where defined: $[\![(432a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\textbf{MAX}\big(\{d \mid 100 \geq d\} - g(2)\big) \leq \textbf{MAX}\{d \mid \exists x[x \in *\mathbb{V}_w \ \&\ x \in *\mathbb{A}_w \ \&\ |x| \geq d]\}$ & $\textsc{small}_c(g(2))$

ii. $\forall \chi' \big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

The logic of this example is the logic of the equative, and so it should be familiar. $\mathscr{M}$ will only be defined if g(2) subtracts a non-empty, final, lower-open subinterval of $\{d \mid 100 \geq d\}$. Small supersets of g(2) will be entailed, subsets of g(2) aren't, and everything else makes $\mathscr{M}$ undefined. g(2) must therefore contain everything in $\{d \mid 100 \geq d\}$ not in $\{d \mid \exists x[x \in *\mathbb{S}_w \ \&\ x \in *\mathbb{A}_w \ \&\ |x| \geq d]\}$, i.e. the difference between 100 and the cardinality of the largest plurality of vampire-attackers. That difference, of course, must be small.

Our $\mathscr{M}$ is not standard in the literature on numeral constructions, but there are precedents for maximality in the semantics of numerals; for example, Kennedy (2013, 2015) proposes that numerals denote generalized quantifiers over degrees.

(433) $[\![\text{four}_{\text{Kennedy}}]\!]^{g,c} = \lambda D.\ \textbf{MAX}(\mathbb{D}) = 4$

What we're doing with $\mathscr{M}$ is pulling out maximality from the numeral itself, so that we have an *almost* can modify. Of course, our $\mathscr{M}$ expresses a $\geq$ relation, whereas Kennedy's numerals express =; if our $\mathscr{M}$ were instead an equality operator, e.g. $\mathscr{M}'$, in tandem with our semantics for numerals, we'd derive his semantics.

(434) $[\![\mathscr{M}']\!]^{g,c} = \lambda D \lambda D'.\ \textbf{MAX}(\mathbb{D}) = \textbf{MAX}(\mathbb{D}')$

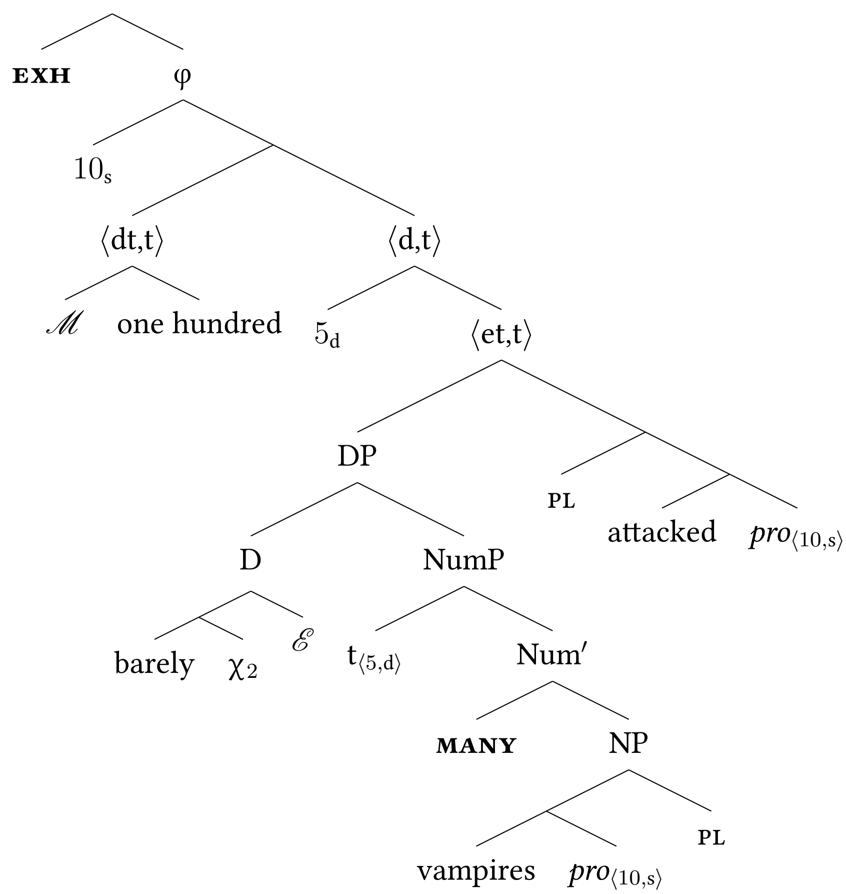
(435) $[\![\mathscr{M}]\!]^{g,c}([\![\text{four}]\!]^{g,c}) = \lambda D'.\ \textbf{MAX}\{d \mid 4 \geq d\} = \textbf{MAX}(\mathbb{D}')$
i.e. $\lambda D'.\ 4 = \textbf{MAX}(\mathbb{D}')$

Of course, $\mathscr{M}'$ is non-monotonic, which mucks up entailments when we consider subtracted exceptions and entailed alternatives, so we stick with our equative $\mathscr{M}$.

Now that we have the full picture of numeral constructions, we ought to put *barely* back in. The generalized quantifier $[\mathscr{M}$ one hundred$]$ must scope over $\mathscr{E}$, but this causes an immediate problem with definedness.

(436) Barely one hundred students attacked Sunnydale High School.

a.



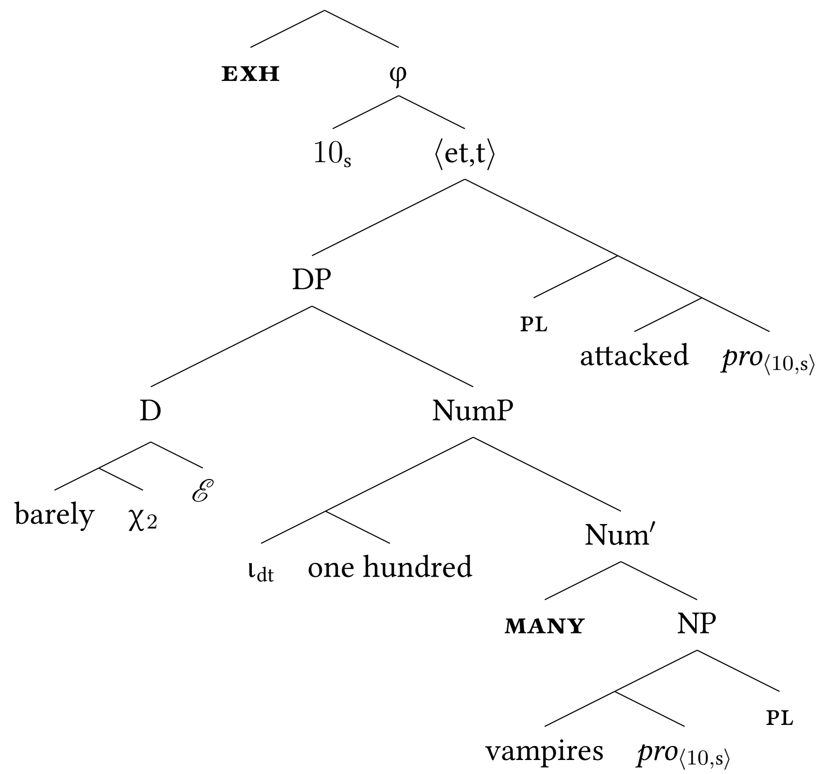
b. $[\![(436a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{A}_w$, $g(2) \neq \emptyset$, $\textbf{MAX}\{d \mid 100 \geq d\}$ and $\textbf{MAX}\{d \mid \neg\exists x[x \in *\mathbb{V}_w \ \&\ x \in *\mathbb{A}_w - g(2) \ \&\ |x| \geq d \ \&\ \textsc{small}_c(g(2))]\}$ are defined.

The negation contributed by *barely* within the scope of $\mathscr{M}$ *one hundred* turns the set of degrees into an interval starting at, but not including, the largest number of vampire-attackers, once the exceptions are removed, and extending upwards ever on. Such a set has no maximal element! **MAX** is undefined, and this example can never be true[8]. How, then, do we rectify this issue?

In spite of the support we've mustered for $\mathscr{M}$, it is not strictly necessary, at least not always. Its utility for us is as a hook on which to hang *almost*, but independent motivation comes from the need for numerals to have scope independent of quantifiers they are a part of (Hackl 2000, Kennedy 2013, i.a.). The *at least* reading discussed above could be derived without appeal to a scope-taking operator—instead of $\mathscr{M}$, we could utilize an operator that picks out the maximal degree in an interval. This allows interpretation of the numeral in situ.

(437) Barely one hundred students attacked Sunnydale High School.

a.



b. $[\![(437a)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq *\mathbb{A}_w$, $g(2) \neq \emptyset$, $\textbf{MAX}\{d \mid 100 \geq d\}$ is defined; where defined: $[\![(437a)]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\neg\exists x\big[x \in *\mathbb{V}_w \;\&\; x \in *\mathbb{A}_w - g(2) \;\&\; |x| \geq 100 \;\&\; \textsc{small}_c(g(2))\big]$

ii. $\forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\varphi[\chi_2/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi_2/\chi']]\!]^{g,c}(w) \neq 1\big]$

Now the truth conditions work: the logic proceeds as before. We can take these two operators—ι_{dt} and $\mathscr{M}$—to be in free variation generally speaking with one another, but a few things will lead us to conclude one is in play and not the other. First, neither *almost* nor *barely* can modify ι_{dt} for

[8]Note that on a more standard definition of **MAX** à la Rullman (1995) this is still undefined; if we tried to appeal to maximal informativity à la von Fintel et al. (2015)., we'd still get undefinedness since this totally open interval contains no maximally informative degree.

type reasons. That said, the presence of *almost* requires the presence of $\mathscr{M}$: if it tried to modify $\mathscr{E}$, we'd get unacceptability across the board for the same reasons *almost some* is unacceptable. Similarly, the presence of *barely* requires ι_{dt} rather than $\mathscr{M}$, which makes the numeral construction an equative. *Barely* can't modify equatives, and $\mathscr{M}$'s movement to a position above *barely* undefinedness. Second, and separate from subtractives, if a scope ambiguity is detected in a numeral construction, both operators must be possible at LF; our $\mathscr{M}$ allows for numerals to take variable scope, separate from the individual quantifier. Such configurations have been discussed for more complex examples in the literature (see Kennedy 2015 and Bylinina & Nouwen 2020, for example), and while we don't address these topics in any detail, it's good to have in our back pocket[9].

4.2.5 What about $\mathscr{M}$?

As stated at the outset, $\mathscr{M}$ is exactly like the equative morpheme. It could, like $\mathscr{E}$ is to *some*, be taken as a covert analog. One might also wonder if we could avoid positing $\mathscr{M}$ by replacing it with one of the other quantifiers that *almost* does modify: **POS**. Such a proposal would be novel—unifying numerals and the positive construction. In what follows, we'll discuss this possibility and what it requires in more detail. The pros and cons, it seems, probably shake out about the same, but it's worth a brief tour.

POS has the following semantics.

(438) $[\![\textbf{POS}]\!]^c = \lambda D \lambda D'. \begin{cases} \textbf{MIN}(\mathbb{D}) < \textbf{MAX}(\mathbb{D}'), & \text{if } \textbf{MIN}(\mathbb{D}) \text{ is defined} \\ \textbf{MAX}(\mathbb{D}) \leq \textbf{MAX}(\mathbb{D}') & \text{otherwise} \end{cases}$

(439) For any set of degrees $\mathbb{D}$ on a scale $\mathcal{S}$:

a. $\textbf{MAX}(\mathbb{D}) := \iota d\big[d \in \mathbb{D} \;\&\; \forall d'\big[\big[d' \in \mathcal{S} \;\&\; d' \leq d\big] \leftrightarrow d' \in \mathbb{D}\big]\big]$

b. $\textbf{MIN}(\mathbb{D}) := \iota d\big[d \in \mathbb{D} \;\&\; \forall d'\big[\big[d' \in \mathcal{S} \;\&\; d' \geq d\big] \leftrightarrow d' \in \mathbb{D}\big]\big]$

At least sometimes, **POS** has the meaning of the equative morpheme and our $\mathscr{M}$, i.e. the 'otherwise' condition. Given our definition of **MIN** above, in numeral constructions, the first condition for **POS** will never be met—numerals denote degree sets of degrees that include their upper bound.

[9]One might be irked at the presence of both $\mathscr{M}$ and ι_{dt} in our theory; here's an alternative, though I'm not sure it's much better. Instead of assuming that numerals denote sets of degrees, necessitating both $\mathscr{M}$ and ι_{dt}, we can take them to denote degrees, and lift them to a set with an operator like the following:

(i) $[\![\textsc{lift}]\!]^{g,c} = \lambda d \lambda d'.\, d \geq d'$

$\mathscr{M}$ would thus necessitate lifting of a degree to a set like this, but we wouldn't need ι_{dt}. It's really a pick-your-poison moment, it would seem, since we would need to posit two operators on both approaches. We'll stick with the one above.

MIN requires that they contain every degree on the scale above their minimal degree, so we'd get undefinedness. This means that only the 'equative' meaning, the otherwise condition, would be triggered, regardless of whether *almost* or *barely* were modifying it. That works beautifully for *almost*, and not so well, fortunately, for *barely*. The latter still requires that **POS** alternate with ι, though, for the same reasons we needed it with $\mathscr{M}$. So, to the extent that that is a detriment to the $\mathscr{M}$ analysis, it's a detriment to the POS analysis.

That said, maybe $\mathscr{M}$ does win out in the end. Our definitions for **MAX** and **MIN** are very restrictive; they were made so largely for expository purposes. Such definitions allowed us in the previous chapter to set aside a lot of exceptions quickly, because they yielded undefinedness. If we opt to go with more standard definitions, i.e. those that pick out the biggest/smallest degree in a set, we'd have to discuss a lot more possible exceptions. Ultimately, though, as we saw in the previous chapter, such exceptions didn't matter anyways—the analysis predicted they weren't good. The definitions are substantive in a sense, but they are for our purposes just a simplification. Now, however, our restrictive definitions play a *crucial* role.

A less restrictive, Rullman-style meaning for **MIN** would be as follows.

(440) $\textbf{MIN}(\mathbb{D}) := \iota d\big[d \in \mathbb{D} \;\&\; \forall d'\big[d' \in \mathbb{D} \rightarrow d' \geq d\big]\big]$

Whether this meaning could be triggered given a numeral depends on the scale associated with numerals. We've tacitly assumed that cardinalities *include* 0 and do not extend into negatives; that is, $\{d \mid 100 \geq d\}$ can be written as $[0,100]$ in our interval notation. If such a set is the first argument of **POS**, and the active **MIN** is the weaker meaning just above, then we will trigger the non-equative meaning, and it's not clear how we could circumvent that. It seems like we'd have two options. First, we could deny that 0 is included on the scale of cardinalities, so that $\{d \mid 100 \geq d\}$ does not contain 0, i.e. its interval notation is $(0,100]$. The issues surrounding 0 are complex and, as far as I can tell, not deeply studied, though; Bylinina & Nouwen (2018) and Chen (2018) are two investigations into zero. An option that circumvents this entirely would be to include 0, but grant that the scale continues on down the real numbers to $-\infty$; then $\{d \mid 100 \geq d\}$ would translate in interval notation to $(-\infty,100]$. Neither definition of **MIN** would apply to such an interval, and so we'd expect only the equative interpretation of **POS** to surface, and then the derivation would proceed as expected. We'll leave a final decision between these two options and all their entailments and choice points up to the reader.

4.3 Wrapping up

The fact that both *almost* and *barely* are compatible with numeral constructions suggests, at first, a problem for our analysis: given that we predict complementarity, why are we seeing overlapping distributions? The solution comes from a closer look at how numeral constructions are built. The solution to the apparent paradox comes in the form of a bipartite quantificational structure for numeral constructions: (i) an existential quantifier over individuals, $\mathscr{E}$, which is modified by *barely*; and (ii) a maximality operator, $\mathscr{M}$, which is modified by *almost*. In the case of *barely*, we saw that pluralization was a solid means of getting the ordering we needed in order to properly subtract; extending this to other determiners has some repercussions that we explored a bit. *Almost* gave us evidence for an additional piece of degree-semantic machinery that has fewer precursors in the literature, but which gets the right results and in fact makes good predictions here. As we close out, we'll briefly discuss some possible extensions and open issues.

Numerals appear in one way or another in other kinds of constructions, and we see that the distribution of *almost* and *barely* again overlaps here. Both seem to modify measure phrases in predicative sentences like (441).

(441) a. Xander is almost six feet tall.
b. Giles is barely six feet tall.

(441a) is true if Xander's height is close to, but does not reach, six feet tall; (441b) is true if Giles's height is at least six feet tall, but if it it exceeds that, it's not by much at all. We briefly discussed measure phrases with *almost* and *barely* in the previous chapter, but as mentioned there, a comprehensive treatment is necessary.

Differential comparatives permit both *almost* and *barely*, whereas unmodified comparatives only permit *barely*.

(442) a. Willow is almost *(two months) older than Buffy is.
b. Anya is barely (two months) older than Tara is.

Density prevents us from finding a non-empty, minimal exception when *almost* composes with a comparative, and they can never be true. (442a) is perfectly natural, though, and true if Willow is older than Buffy is, and the difference between their ages is close to, but isn't quite, two months. *Barely* and comparatives already went hand-in-hand, and that is no different with a measure phrase. (442b) is true if Anya is two months older than Tara is, but not much more than that. Equatives can be modified as well, and once more both *almost* and *barely* can modify such modified equatives.

(443) a. Giles is almost (twice) as old as Willow is.
b. Giles is barely *(twice) as old as Buffy is.

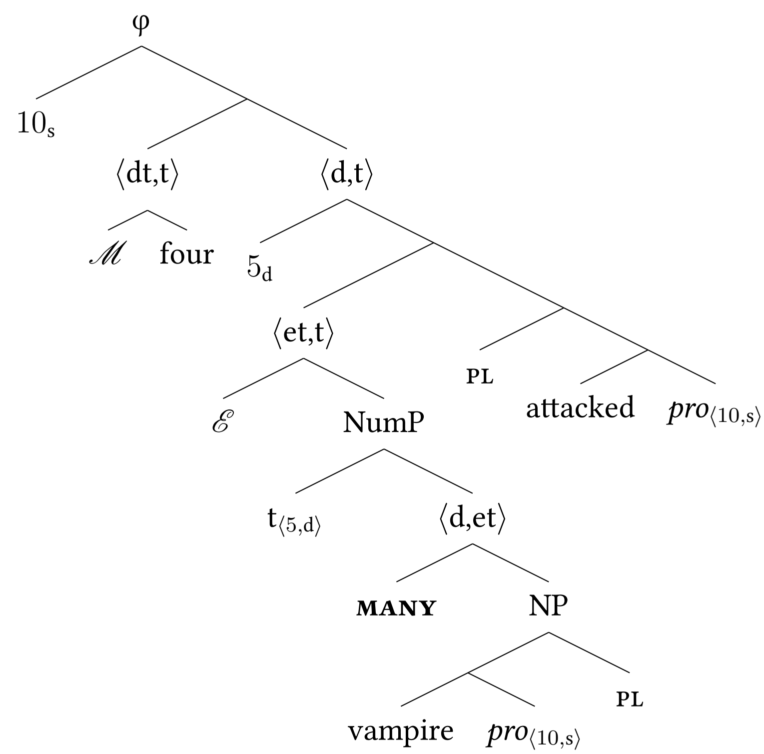
(443a) is true if Giles's age is just under two times Willow's age; (443b) is true if it's just above Buffy's age. Without *twice*, the *barely* equative is unacceptable. These too need analyzing in the context of subtractives.

Typically, measure phrases are analyzed as denoting degrees, saturating degree arguments. That won't obviously be sufficient to account for these data. If we've learned anything from this chapter, though, it's that there must be more than meets the eye to measure phrases. They've got numerals in them, so maybe they invoke more structure internally, even if they ultimately denote degrees—there must be some operators that *almost* and *barely* each latch onto. Precisely what the compositional shape of these constructions awaits future research.

We've discussed a bit that our analysis of numeral constructions yields *at least* readings, but not *exactly* readings. That is, the truth conditions we derive for (431), repeated below, allow that more than four vampires attacked, when generally one infers that exactly four did.

(431) Four vampires attacked Sunnydale High School.

a.



b. $[\![(431a)]\!]^{g,c}(w) = 1$ *iff* $\exists x\big[x \in *\mathbb{V}_w \;\&\; x \in *\mathbb{A}_w \;\&\; |x| \geq 4\big]$

One way that exactly readings can be derived is through **EXH** (Bylinina & Nouwen 2020). $[\![(431a)]\!]^{g,c}$ doesn't Strawson entail, for any degree $d' > 4$, that there is a vampire-attacker plurality with d' atoms that attacked the school; **EXH** applied to φ could require that there be no such pluralities, depending on what we assume about the alternatives to φ. This would mean that at least four, and no more than four, vampires attacked Sunnydale High School, an *exactly* reading. It should be noted that, in its way, **EXH** forces an *exactly* reading when subtractives are in play. The boundary set by *one hundred* in *almost one hundred* is set at exactly one hundred; the logic requires it. There is no need for any additional mechanism for deriving *exactly* interpretations when we've got subtractives and their own **exh**—this might be taken as support for **exh** as the locus more generally of *exactly* readings of numerals. That said, a full discussion of the complexities of bare numeral constructions and their implicatures is beyond our scope; see Spector (2013) for a discussion of factors leading to an **exh**-based view, and see Kennedy (2015) for problems for such a view.

The discussion above segues into the fact that, as it stands, we predict no contrast between (444a) and (444b).

(444) a. Almost one hundred vampires attacked Sunnydale High School.

b. ?? Almost one hundred and one vampires attacked attacked Sunnydale High School.

The reason we predict no contrast between the two is that both *one hundred* and *one hundred and one*, or more accurately $[\mathscr{M}$ *one hundred*$]$ and $[\mathscr{M}$ *one hundred and one*$]$ set a clear boundary from which *almost* can subtract. It is intuitive that *almost* requires such clear boundaries, and our analysis delivers the same truth conditions for (444a) as it does for (444b), modulo the increased number. What's the deal? 'Round' numbers like *one hundred* tend to have approximate interpretations, whereas 'non-round' numbers like *one hundred and one* tend to have precise interpretations (Krifka 2007). It's not clear, though, in what sense *we* can say *one hundred* is round and imprecise, since we just said our system delivers an exact reading of the numeral when modified by *almost*! Perhaps the nature of precise numerals is different. At this point, we have to leave this as an open issue for our theory.

We have one additional problem that needs solving. We've posited that the quantificational determiner invoked in numeral constructions is a covert existential operator $\mathscr{E}$, allowing modification by *barely*. How, then, do we account for the following?

(445) The barely one hundred vampires attacked Sunnydale High School.

Admittedly, this is not my favorite sentence, but it seems alright. *The* is presumably the head of the DP *the barely one hundred vampires*, taking the place of the existential $\mathscr{E}$, but the definite article isn't modifiable by *barely* (or *almost*, for that matter). How is *barely* composing if this is the right view? Other constructions raise similar questions: why can we not compose subtractives with other kinds of modified numerals, as in (446)?

(446) a. # $\{$Almost/barely$\}$ at least one hundred vampires attacked Sunnydale High School.

b. # $\{$Almost/barely$\}$ at most one hundred vampires attacked Sunnydale High School.

c. # $\{$Almost/barely$\}$ no fewer than one hundred vampires attacked Sunnydale High School.

d. # $\{$Almost/barely$\}$ between one hundred and two hundred vampires attacked Sunnydale High School.

Unfortunately, we don't have the time to analyze these in detail. We'll have to leave accounting for these data for the future. In any case, our analysis of numerals needs testing in a host of domains, but subtractives do give us good reason to think it's right, so that testing will hopefully be fruitful.

Chapter 5

The polar inference revisited

In this chapter, we'll discuss a bit more the relationship between **exh**, the semantic means with which we derive the polar inference, and subtractives. As we'll see, some changes might be warranted. We'll start by looking at multiple quantifier constructions in which a subtractive modifies a quantifier within the scope of another. We've seen some of these before, actually, when we discussed Gajewksi's (2008) and Crnič's (2018) arguments for **exh** as the source of the polar inference. What we'll focus on is multiple quantificational determiners in a single clause, and what these configurations reveal about the scope of **exh** with respect to subtractives.

(447) a. No witch read any spellbook but *The Brekenkrieg Grimoire.*
b. # Every witch read some spellbook but *The Brekenkrieg Grimoire.*
c. No witch read every spellbook but *The Brekenkrieg Grimoire.*

The first two examples lead us to the view that **exh** must be *able* to scope as close as possible to the subtractive with which it is associated; that proximity is tied to the licensing of those quantifiers modified by the subtractive. We'll run into a problem with the third sentence, though, and we'll take a detour. We'll discuss cases mentioned at the very beginning of our journey with *almost* and *barely*, namely those in which the polar inference seems to be invisible to embedding operators like *good news!* and *because*. We'll sketch what our **exh**-based theory would have to say, but then, finding such a view unsatisfactory, we'll argue that all this data supports a view on which the main contribution of **exh** is not truth conditional, but rather presuppositional, as proposed by Bassi et al. (2021). The success of our implementation can serve as an argument in favor of such a view. What is really quite neat is that when we return to the problematic multiple quantifier cases, we have the means with which to explain them in terms of this presuppositional **exh**. The resulting theory allows a unified analysis of the locality of presuppositional **exh**, or **pex**, with respect to subtractives, accounting for the data above, and the invisibility of the polar inference in certain environments by the same mechanism. That is, we get to make good on the promise to properly analyze the polar inference of subtractives in a neat and uniform way. Let's get started.

5.1 Locality and multiple quantifier constructions

In general, we have stayed away from discussion of multiple quantifier constructions in the context of subtractives[1].

(448) a. No witch read any spellbook but *The Brekenkrieg Grimoire*.

b. No witch read almost any spellbook.

As we discussed in Chapter 2, this kind of example—where subtractive *but* modifies an existential determiner in a Downward Entailing environment—is an argument for severing the derivation of the polar inference from the meaning of the subtractive operator proper, as in Gajewski (2008). The classical, Fintelian semantics for *but*, as in (449), predicts triviality for such cases.

(449) $$[\![\text{but}]\!]^{g,c} = \lambda X \lambda P \lambda \mathscr{D} \lambda Q.\ \underbrace{\mathscr{D}(\mathbb{P}-\mathbb{X})(\mathbb{Q})}_{\textbf{proximal inference}}\ \&\ \underbrace{\forall \mathbb{X}'\big[\mathbb{X} \not\subseteq \mathbb{X}' \rightarrow \neg\mathscr{D}(\mathbb{P}-\mathbb{X}')(\mathbb{Q})\big]}_{\textbf{polar inference}}$$

(450) $$[\![(448\text{a})]\!]^{g,c}(w) = 1 \textit{ iff}$$
$$\mathbb{W}_w \cap \Big\{ y \mid \mathbb{S}_w - \big\{BG\big\} \cap \big\{ z \mid \textbf{READ}_w(y)(z) \big\} \neq \emptyset\ \& \\ \forall \mathbb{X}'\big[\big\{BG\big\} \not\subseteq \mathbb{X}' \rightarrow \mathbb{S}_w - \mathbb{X}' \cap \big\{ z \mid \textbf{READ}_w(y)(z)\big\} = \emptyset\big]\ \Big\} = \emptyset$$

The set intersected with $\mathbb{W}_w$ is necessarily empty; the truth conditions as trivially true. This is a function of von Fintel's semantics generating falsity necessarily when *but* combines with an Upward Entailing determiner, as *any* is. The same goes for a Fintelian semantics for *almost*; from here on out, we'll utilize the *but* examples to keep intuitions about truth conditions sharp.

Gajewski's solution is the one we've assumed all along: EXH is responsible for the derivation of the polar inference, and it can scope above *no witch*. This derives non-trivial, and correct, truth conditions[2].

(451) $$[\![\text{but}]\!]^{g,c} = \lambda X \lambda P \lambda \mathscr{D} \lambda Q\colon \mathbb{X} \subseteq \mathbb{P}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{P}-\mathbb{X})(\mathbb{Q})$$

(452) $$[\![\textbf{EXH}\ \varphi]\!]^{g,c}(w) = 1 \textit{ iff } [\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\psi\big[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\big]$$

(453) EXH *AND THE LICENSING OF SUBTRACTIVES OPERATORS* *first version*

Subtractive operators are only licensed in the scope of an EXH operator.

[1]Note that we're still restricting ourselves to one subtractive, and we'll only consider our *subtractive* alternatives for simplicity.

[2]Throughout we'll suppress the size constraint for space and readability, but its inclusion is ultimately necessary.

(454) $[\![(448a)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$ & $\{BG\} \neq \emptyset$; where defined:
$[\![(448a)]\!]^{g,c}(w) = 1$ *iff* (a) & (b) $= 1$

a. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{BG\} \cap \{z \mid \textbf{READ}_w(y)(z)\} \neq \emptyset\} = \emptyset$

b. $\forall \mathbb{X}' \Big[[\{BG\} \not\subseteq \mathbb{X}' \;\&\; \mathbb{X}' \subseteq \mathbb{S}_w \;\&\; \mathbb{X}' \neq \emptyset] \rightarrow \mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \mathbb{X}' \cap \{z \mid \textbf{READ}_w(y)(z)\} \neq \emptyset\} \neq \emptyset \Big]$

This is, then, an argument for allowing that **EXH** to have non-local scope with respect to subtractive operators Sometimes, though, it seems like we need to require that **EXH** takes scope as locally as possible to *but*; consider (455).

(455) # Every witch read some spellbook but *The Brekenkreig Grimoire.*

a. *[every witch [2 [**EXH** [$_\varphi$ some spellbook but BG [3 [t_2 read t_3]]]]]]
local contradiction

b. *[**EXH** [$_\varphi$ some spellbook but BG [3 [every witch [2 [t_2 read t_3]]]]]]
local contradiction

c. [**EXH** [$_\varphi$ every witch [2 [some spellbook but BG [3 [t_2 read t_3]]]]]]
well-formed (bad!)

In all three LFs, φ is Upward Entailing with respect to the restrictor of *some*. Where no quantifier intervenes between **EXH** and *some spellbook but BG*, we predict necessary non-truth, so (455a) and (455b) are ruled out as possible LFs of (455)[3]. The problem is (455c). φ is still Upward Entailing with respect to the restrictor of *some*, and alternatives formed with subsets of $\{BG\}$ are entailed by $[\![\varphi]\!]^{g,c}$.

(456) $[\![(455c)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$ & $\{BG\} \neq \emptyset$; where defined:
$[\![(455c)]\!]^{g,c}(w) = 1$ *iff*
$\mathbb{W}_w \subseteq \{y \mid \mathbb{S}_w - \{BG\} \cap \{z \mid \textbf{READ}_w(z)(y)\} \neq \emptyset\}$ &
$\forall \mathbb{X}' \Big[[\mathbb{X}' \not\subseteq \{BG\} \;\&\; \mathbb{X}' \subseteq \mathbb{S}_w \;\&\; \mathbb{X}' \neq \emptyset] \rightarrow \mathbb{W}_w \not\subseteq \{y \mid \mathbb{S}_w - \mathbb{X}' \cap \{z \mid \textbf{READ}_w(z)(y)\} \neq \emptyset\} \Big]$

This is true iff every witch read at least one non-BG spellbook, and for every spellbook, at least one witch failed to read it. Here's a context that makes this true. Tara read DLM and WD but failed to read BG; Willow read WD and BG, but failed to read DLM; Amy read BG and DLM, but failed to read WD. Each has read a non-BG spellbook, and every spellbook was left unread by at least one witch. Unfortunately, these truth conditions are coherent, though the sentence is not. Furthermore, it's not the case that the exhaustified meaning is entailed by its counterpart without **EXH**.

[3]Of course, to actually do this, we'd need SMALL. We trust the readers can grant that.

(457) $\mathbb{W}_w \subseteq \{y \mid \mathbb{S}_w - \{BG\} \cap \{z \mid \textbf{READ}_w(z)(y)\} \neq \emptyset\}$

If (457) is true, can (456) be false? Yes: if Tara, Willow, and Amy all read the same non-BG spellbook, say WD, then (457) and the first conjunct of (456) are true, but $\{DLM\}$ is an $\mathbb{X}'$ that falsifies **exh**, since it is not a subset of $\{BG\}$ and every witch did indeed read at least one of non-DLM spellbook. Exhaustification yields a stronger meaning. This is unfortunate, since this sentence is reported to be unacceptable across the board.

Here are the facts we need to account for at this point.

1. **exh** must be able to scope above *no* to predict the acceptability of *no* > *any...but*

2. **exh** must take the closest scope possible to *but* in *every* > *some...but* in order to rule it out

We can account for this by recognizing an important difference between *some* and *any*, and by modifying the constraint by which we tie the licensing of subtractives with **exh**. Gajewski (2013), discussing *no* > *any...but*, notes that the non-local scope position of **exh** in (458) has as its sister the smallest node φ in which *any* is licensed.

(458) []No witch read any spellbook but the Brekenkrieg Grimoire.

a. *[$_\varphi$ no witch [2 [**exh** [$_\psi$ any spellbook but BG [3 [t_2 read t_3]]]]]]
local contradiction

b. [**exh** [$_\varphi$ no witch [2 [$_\psi$ any spellbook but BG [3 [t_2 read t_3]]]]]]
well-formed (good!)

(458a) is the LF that generates a contradiction, so that couldn't be the right LF for (458a), but it's also the case that in ψ, *any* is not licensed. It is only in φ that *any* is licensed. **exh**, then, is as close to *any spellbook but BG* as it can get while *any* is still licensed. We can alter the licensing condition on subtractives, the constraint that ties the distribution of subtractives to the presence of an associated **exh**, in a way that reflects this.

(459) ***exh and the licensing of subtractives operators*** *revised*

Subtractive operators are only licensed as modifiers of some quantifier Δ if:

a. They are in the scope of an **exh** operator, and
b. That **exh**'s sister is the minimal expression of type ⟨s,t⟩ in which Δ is licensed.

These are taken to be necessary, not sufficient, conditions for the licensing of subtractive operators, of course. This allows us to rule out the problematic scopal configuration for *every* > *some...but* too!

(455) # Every witch read some spellbook but *The Brekenkreig Grimoire.*

a. *[$_{\varphi}$ every witch [2 [**EXH** [$_{\psi}$ some spellbook but BG [3 [t_2 read t_3]]]]]]

local contradiction

b. *[**EXH** [$_{\psi}$ some spellbook but BG [3 [every witch [2 [t_2 read t_3]]]]]]

local contradiction

c. *[**EXH** [$_{\varphi}$ every witch [2 [$_{\psi}$ some spellbook but BG [3 [t_2 read t_3]]]]]]

barred by locality of licensing condition

The problematic LF was (455c): we were able to derive coherent truth conditions, even though the sentence itself is garbage. Now, though, that LF is not possible, given (459): φ is not the most local position for **EXH**, since *some* is licensed in ψ. We can, therefore, predict that there is no possible LF for (455) that generates coherent truth conditions[4].

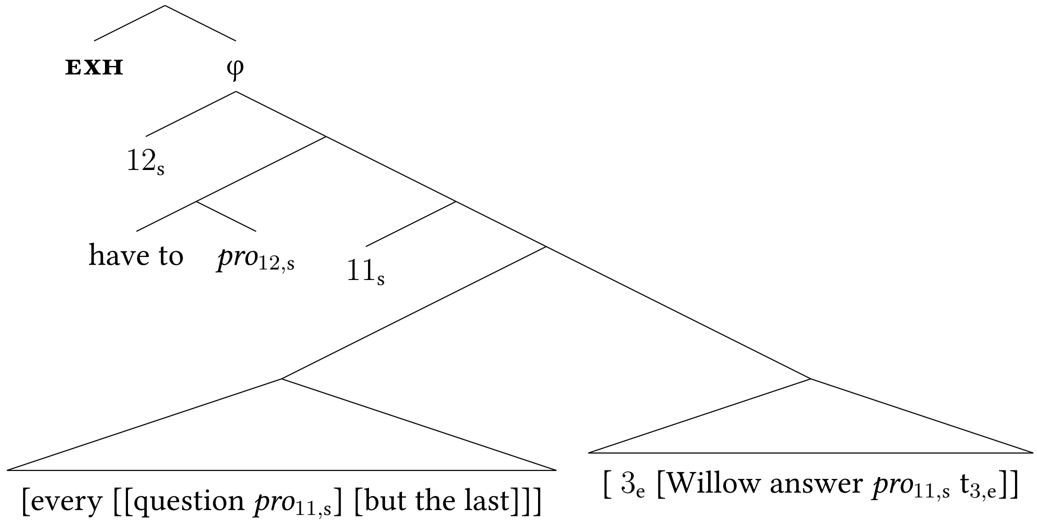
5.2 A problem or two

Multiple quantifiers still cause us some problems, unfortunately. First, some of the original cases in favor of the 'distributed,' **EXH**-based theory of subtractives run into an obvious problem: our adoption of such a strict locality constraint only permits the an LF for (460) that derives the strongest reading of (460). The LF in (460a), the one which is necessary to get the weaker interpretation of (460) that is most accessible, i.e. that Willow wasn't required to *skip* the last question, is no longer permitted.

(460) Willow answered every question on the exam but the last. To get an A, she had to ~~answer every question but the last.~~

[4]Gajewski (2013) provides a richer explanation of our constraint in terms of relativized minimality: following Chierchia (2006, 2010), he takes the licensing of *any* to be tied to its own **EXH** operator, distinct from the **EXH** associated with *but* (or the whole DP *any spellbook but BG*). The **EXH** operator associated with *any* must adhere to our constraint in (459), but were the **EXH** associated with *but* to sit below the other, there would be crossing dependencies between them, and this is independently banned.

a.



b. $[\![(460\text{a})]\!]^{g,c}(w)$ is defined only if $\{L\} \subseteq \mathbb{Q}_w$ & $\{L\} \neq \emptyset$; where defined, $[\![(460\text{a})]\!]^{g,c}(w) = 1$ *iff* (i) & (ii) $= 1$

i. $\{w' \mid \textbf{REQ}_w(w') \text{ \& } \text{GAA}_{w'}(\text{the exam})(\text{Willow})\} \subseteq \{w' \mid \mathbb{Q}_{w'} - \{L\} \subseteq \mathbb{A}_{w'}\}$

ii. $\forall \mathbb{X}'\Big[\big[\{L\} \not\subseteq \mathbb{X}' \text{ \& } \mathbb{X}' \neq \emptyset \text{ \& } \mathbb{X}' \subseteq \mathbb{Q}_w\big] \rightarrow$
$\{w' \mid \textbf{REQ}_w(w') \text{ \& } \text{GAA}_{w'}(\text{the exam})(\text{Willow})\} \not\subseteq \{w' \mid \mathbb{Q}_{w'} - \mathbb{X}' \subseteq \mathbb{A}_{w'}\}\Big]$

That such a configuration is, prima facie, barred given our licensing condition for subtractives is deeply irksome. Note that this is entirely our own doing: Crnič (2018) faces no problem with (460) because there are no locality constraints in his theory.

Perhaps what we can do is the following:

(461) **EXH *AND THE LICENSING OF SUBTRACTIVES OPERATORS*** *revised again*

Subtractive operators are only licensed as modifiers of some quantifier Δ if:

a. They are in the scope of an **EXH** operator, and
b. **EXH** is licensed where its sister is the minimal expression of type ⟨s,t⟩ in which Δ is licensed.

In a way, we're setting up a cyclical, bottom-up story of how we get to candidate LFs for a given sentence: we derive the LF in (460a) first by checking if the more local scope is one in which it is licensed, i.e. within the scope of the modal *have to*. It is licensed in this position, so we can consider more distant locales for **EXH**, now permitting the LF in (460a). This formulation still requires that **EXH** scope above *no* in *no* > *any...but* configurations, since it's still the closest it can get to *but* while *any* is licensed. This formulation also still rules out the problematic LF for *every* > *some...but*: **EXH** cannot take scope between *every* and *some...but*, since this derives a contradiction, so it's therefore not possible for it to scope above *every*. That's good. This formulation also makes a welcome prediction: we cannot utilize a high-scope **EXH** to generate coherent truth conditions for (462).

(462) # Willow had to answer some question on the exam but the last one.

a. *[had to [**EXH** [$_\varphi$ some question but the last [3 [Willow answer t_3]]]]]

local contradiction

b. *[**EXH** [$_\varphi$ some question but the last [3 [had to [Willow answer t_3]]]]]

local contradiction

c. *[**EXH** [$_\varphi$ had to [some question but the last [3 [Willow answer t_3]]]]]

barred by locality of licensing condition

This is just the same configuration as *every* > *some...but*: highest scope **EXH** here would predict coherent truth conditions; the other two LFs will, naturally, yield necessary non-truth. Note that *without* our kind of licensing condition, Crnič's theory does predict the availability of the problematic LF in (462c) for (462); the very same reason he avoids the problem in (460) causes a problem here. It would seem, then, that allowing **EXH** to have non-local scope with respect to a determiner-subtractive combination *only if* it is licensed in locally is a good constraint.

Well, until we consider the following.

(463) CONTEXT: Willow, Tara, and Amy are all the witches, and *The Brekenkrieg Grimoire*, *The Du Lac Manuscript*, and *Writings of Dramius* (BG, DLM, and WD) are all the spellbooks. Willow read all three spellbooks. Tara read BG and DLM, but not WD. Amy read BG and WD, but not DLM.

a. No witch read every spellbook but the Brekenkrieg Grimoire.

(463a) is acceptable and true in this context. An intuitive, semi-formal paraphrase of the truth conditions can be stated as follows:

(464) If y is a witch, then either y read every spellbook or y failed to read at least one spellbook other than *The Brekenkrieg Grimoire*.

Our analysis faces problems deriving this. Given the constraint placed on the locality of **EXH** with respect to *but*, (465) is the first LF we ought to consider.

(465) [no witch [2 [**EXH** [$_\varphi$ every spellbook but BG [3 [t_2 read t_3]]]]]]

φ is Downward Entailing with respect to the restrictor of *every* in this configuration. Alternatives formed with supersets of $\{\text{BG}\}$ will be entailed within the scope of **EXH**; so, **EXH** negates alternatives to φ that are formed with non-*supersets* of $\{\text{BG}\}$. (466) are the truth conditions derived from (465).

(466) $[\![(465)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$ & $\{BG\} \neq \emptyset$; where defined:

$$[\![(465)]\!]^{g,c}(w) = 1 \textit{ iff } \mathbb{W}_w \cap \Big\{y \mid \mathbb{S}_w - \{BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\} \ \& \ \forall \mathbb{X}'\Big[\big[\{BG\} \not\subseteq \mathbb{X}' \ \& \ \mathbb{X}' \subseteq \mathbb{S}_w\big] \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \{z \mid \textbf{READ}_w(z)(y)\}\Big]\Big\} = \emptyset$$

The set with which $\mathbb{W}_w$ is intersected is the set of individuals who read both DLM and WD and who didn't BG. Let's call that set $\mathbb{RB}_w$ for short—those who Read some Books. That is:

(467) $\mathbb{RB}_w = \Big\{y \mid$

a. $\mathbb{S}_w - \{BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}$ &

b. $\mathbb{S}_w - \{DLM\} \not\subseteq \{z \mid \textbf{READ}_w(z)(y)\}$ &

c. $\mathbb{S}_w - \{WD\} \not\subseteq \{z \mid \textbf{READ}_w(z)(y)\}$ &

d. $\mathbb{S}_w - \{DLM,WD\} \not\subseteq \{z \mid \textbf{READ}_w(z)(y)\} \Big\}$

$\mathbb{RB}_w$ is the set of people who read both DLM and WD and not BG. (466) is true iff $\mathbb{RB}_w$ is disjoint from $\mathbb{W}_w$—that is, if no witch read all and only DLM and WD. Note that this allows the witches to have read all three spellbooks, or a subset thereof, just not the subset $\{DLM, WD\}$. These truth conditions, then, map onto the intuitive meaning of (463a) and the informal truth conditions above.

Unfortunately, **EXH** is in a Downward Entailing environment in the nuclear scope of *no*, and eliminating **EXH** from (465) yields a meaning that entails the meaning with **EXH**.

(468) $\mathbb{W}_w \cap \big\{y \mid \mathbb{S}_w - \{BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\big\} = \emptyset$

This is a problem for our constraint restricting the distribution of **EXH**.

(469) ***ECONOMY CONSTRAINT ON THE DISTRIBUTION OF* EXH**

Let φ be an arbitrary expression containing an occurrence of the **EXH** operator. **EXH** is not licensed in φ if eliminating this occurrence leads to an expression ψ such that for any admissible g, c, $[\![\psi]\!]^{g,c} \Rightarrow [\![\varphi]\!]^{g,c}$.

The issue is that (the intension of) (468) Strawson entails (the intension of) (466); this constraint means that **EXH** is not licensed, and since the licensing of subtractives is tied intimately to the licensing of **EXH**, this LF won't work with the tools at hand.

Other LFs don't fare any better, really. Given the constraints placed on the locality of **EXH** with respect to subtractives, (470) is *not* a possible LF for (463a). That said, even if it were, (471) are the incorrect truth conditions.

(470) [**EXH** [$_\varphi$ no witch [2 [every spellbook but BG [3 [t_2 read t_3]]]]]]

(471) $[\![(470)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$ & $\{BG\} \neq \emptyset$; where defined:
$[\![(470)]\!]^{g,c}(w) = 1$ *iff* (a) & (b) $= 1$

a. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} = \emptyset$

b. $\forall \mathbb{X}' \Big[[\mathbb{X}' \not\subseteq \{BG\}$ & $\mathbb{X}' \subseteq \mathbb{S}_w$ & $\mathbb{X}' \neq \emptyset]$
$\rightarrow \mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \mathbb{X}' \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} \neq \emptyset \Big]$

Assume that $\mathbb{S}_w = \{BG, DLM, WD\}$. (471a) is true if the set of witches intersected with the set of readers of both DLM and WD is empty. That is, no witch read both DLM and WD. (471b) requires all the following to be true in w.

(472) a. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{DLM\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} \neq \emptyset$

b. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{WD\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} \neq \emptyset$

c. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{DLM, WD\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} \neq \emptyset$

d. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{BG, WD\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} \neq \emptyset$

e. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{DLM, BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} \neq \emptyset$

f. $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{DLM, WD, BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} \neq \emptyset$

(472a) is true if some witch read both WD and BG. (472b) is true if some witch read both DLM and BG. (472c) is true if some witch read BG. (472d) is true if some witch read DLM. (472e) is true if some witch read WD. (472f) is true if some witch is a reader, but assuming that $[\![\text{every}]\!]^{g,c}$ presupposes that its restrictor is non-empty, we could claim that (472f) is undefined. Taken together, (471) is true if no witch read both DLM and WD, but all books were read by some witch. Were it a possible LF, it would still get us only part of the way to the truth conditions we want, but they predict falsity if some witch read every spellbook. That's not good.

One might think that inverse scope, i.e. *every...but* scoping above *no witch* would solve the issue. First, inverse scope does not appear possible in this configuration—*no* seems to prevent, in some sense, quantifiers from scoping across it.

(473) No witch read every spellbook.

a. no witch > every spellbook

b. # every spellbook > no witch

This is independent, as we can see, of subtractives. However, even if *every spellbook but BG* could scope above *no witch*, we still wouldn't derive the right results. Such a logical form derives the following.

(474) [**EXH** [$_\varphi$ [every spellbook but BG [3 [no witch [2 [t_2 read t_3]]]]]]]

(475) $[\![(474)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$ & $\{BG\} \neq \emptyset$; where defined:

$[\![(474)]\!]^{g,c}(w) = 1$ *iff* (a) & (b) $= 1$

a. $\mathbb{S}_w - \{BG\} \subseteq \{z \mid \mathbb{W}_w \cap \{y \mid \textbf{READ}_w(z)(y)\} = \emptyset\}$

b. $\forall \mathbb{X}' \Big[\big[\{BG\} \not\subseteq \mathbb{X}' \;\&\; \mathbb{X}' \subseteq \mathbb{S}_w \;\&\; \mathbb{X}' \neq \emptyset\big] \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \{z \mid \mathbb{W}_w - \cap \{y \mid \textbf{READ}_w(z)(y)\} = \emptyset\} \Big]$

Every spellbook but BG is now in an Upward Entailing environment, so the restrictor of *every* is a Downward Entailing envoronment, and so supersets of $\{BG\}$ give rise to entailed alternatives. The truth conditions in (474) require that both WD and DLM are in the set of things that no witch read, and any non-superset of $\{BG\}$ is such that some witch read an element of $\mathbb{S}_w$. This second conjunct amounts to the following.

(476) a. $\mathbb{S}_w - \{WD\} \not\subseteq \{z \mid \mathbb{W}_w \cap \{y \mid \textbf{READ}_w(z)(y)\} = \emptyset\}$

b. $\mathbb{S}_w - \{DLM\} \not\subseteq \{z \mid \mathbb{W}_w \cap \{y \mid \textbf{READ}_w(z)(y)\} = \emptyset\}$

c. $\mathbb{S}_w - \{WD, DLM\} \not\subseteq \{z \mid \mathbb{W}_w \cap \{y \mid \textbf{READ}_w(z)(y)\} = \emptyset\}$

(476a) requires that some witch read either DLM or BG, (476b), that some witch read either WD or BG, and (476c) that some witch read BG. (474) is true iff no witch read either DLM or WD, and some witch read BG. This is still not the reported interpretation. Alas. This is a problem we'll have to put on hold for a moment. Importantly, though, it's not a problem unique to this exact configuration. Subtractives can occur in (Strawson) Downward Entailing environments more generally, as in the restrictors of conditionals (Kilbourn-Ceron 2016) and relative clauses inside the restrictor *every* (Spector 2014), both of which are islands to movement.

(477) a. If Willow reads every spellbook but *The Brekenkrieg Grimoire*, Giles will reward her.

b. If Willow reads almost every spellbook, Giles will reward her.

(478) a. Every witch who read every spellbook but *The Brekenkrieg Grimoire* has become quite powerful.

b. Every witch who read almost every spellbook has become quite powerful.

These are both weak readings of the subtractives (Willow can read every spellbook to get her reward from Giles, and reading every single one doesn't seem incompatible with witches becoming powerful) and strong readings (some spellbooks are just detrimental). The tension between the subtractive licensing conditions, requiring that **EXH** be licensed within the restrictor, and the economy conditions on **EXH**, blocking that very licensing, prove problematic.

In the next section, we'll discuss some seemingly unrelated data concerning the 'invisibility' of the polar inference, and then we'll propose an analysis that seems to account for both the problem above *and* the invisibility.

5.3 The invisible polar inference

Sometimes the polar inference is, while present in its classical form, invisible to certain operators. As we've seen, the inference doesn't appear to be cancellable.

(479) a. # Buffy killed almost every vampire. In fact, she killed every one.

b. # Xander killed barely any vampires. In fact, he killed none.

This is a byproduct of EXH. The first sentence (479a) entails that some vampires weren't killed by Buffy, and the speaker is thus committed to its truth; the continuation is inconsistent with that. The same goes, mutatis mutandis, for (479b). Sometimes this disappears, though. We saw earlier that certain expressions are blind to the polar inferences of *almost* and *barely* (Horn 2002).

(480) a. Good news! Almost everyone has arrived.

b. Good news! Barely anyone has arrived.

(481) a. Bad news! Almost everyone has arrived.

b. Bad news! Barely anyone has arrived.

It's not that the good news in (480a) is that some people have yet to arrive; the good news is that most everyone has indeed arrived. (480b) requires a particular kind of context to be felicitous—say, we're scrambling to prepare for our party and at the appointed time, most everyone has not yet arrived, saving us—but then the good news is still that close to no one has arrived, not that some have. The data in (481) are similar, except is the bad news is that close to everyone has arrived in (481a), and that close to no one has in (481b). Other high modifiers, like *fortunately/unfortunately*, pattern similarly. In all these cases, the modifiers are blind to the polar inference, which is indeed present in all cases, only picking up on the proximal.

Expressions like *good news!* appear to be speaker-oriented, or 'expressive' in the sense of Potts (2005, 2007; Karttunen & Peters 1979, i.a.): they encode the speaker's perspective or opinion on the assertion. In general, their contributions seem to operate on a distinct semantic-pragmatic level from truth conditions. *Damn* in (482) is another such expression (Cruse 1986; Löbner 2002, i.a.).

(482) That *damn* vampire bit Buffy.

There's nothing obviously true or false about the vampire being described with *damn* here. Rather, the utterer of (482) is expressing a very negative view of the offending vampire.

The view that expressions like *damn*, *good news!*, and other speaker-oriented expressions operate on a distinct level of meaning than the truth conditional goes back to Grice (1975); they give rise to *Conventional Implicatures* (CIs). Potts (2005:11) characterizes CIs, following Grice, as having the following properties (italics and scare quotes original).

1. CIs are part of the conventional meanings of words.

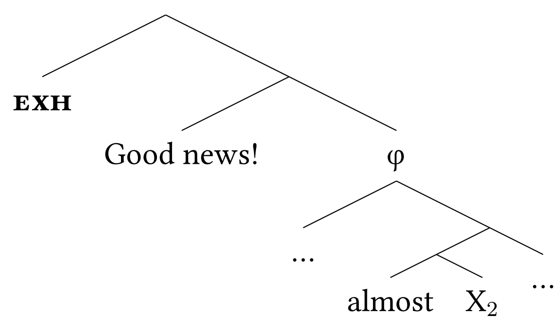
2. CIs are commitments, and thus give rise to entailments.

3. These commitments are made by *the speaker of the utterance* "by virtue of the meaning of" the words [they choose].

4. CIs are logically and compositionally independent of what is "*said* (in the favored sense)," i.e., independent of the at-issue entailments.

Expressive content like that encoded by *good news* is, we assume, a CI. Since the meaning contribution of *good news* and other expressions is such expressive content, i.e. a conventional implicature, it is logically and compositionally independent of truth conditions, i.e. at-issue entailments. Nothing in our proposal so far suggests that EXH has direct access to not-at-issue content, save any definedness conditions on its prejacent. So, one might think that if $[\![\varphi]\!]^{g,c}$ carries expressive content, then it stands to reason that EXH couldn't see the contribution of conventional implicatures, and so $[\![\textsc{exh}\ \varphi]\!]^{g,c}$ is calculated without direct reference to CIs triggered within its scope. CIs would sneak on by, as it were. We'll assume that the data in (480)-(481) are all 'monoclausal' in the sense that, despite the exclamation point, *good news!* and *bad news!* are part of the same syntactic and semantic unit. In all these cases, EXH *must* scope above the speaker-oriented expressions.

(483)



The result of such a configuration, once the actual compositional details of expressive content are incorporated, $[\![\textsc{exh}\ [\text{good news! } \varphi]]\!]^{g,c}$ is truth conditionally equivalent to $[\![\textsc{exh}\ \varphi]\!]^{g,c}$, but within the scope of EXH, *good news!* only has access to the non-exhaustified meaning of φ. In the context of subtractives, this means it only has access to the proximal inference encoded by the modifiers; on such a proposal, it could not be *good news* that φ is false, only that φ is close to being true.

It's clear, even with this sketchy analysis, that to get the facts right EXH must scope over *good news* and other expressive content. Our view of EXH treats its contribution as at-issue, truth conditional content, so if *good news* were to structurally appear higher than EXH, all of the latter's contribution should be good news. As we've seen, though, in (480)-(481), this is not a possible interpretation. So, we'd need to justify why this position is necessary, i.e. what distinguishes above and below *good news!* such that above it is the most local **EXH** can be.

If we do not wish to adhere to the highest-scope proposal for **EXH**, we'll need to make some other claims about its contribution such that *good news* doesn't have access to it.

Granting that we could make such an analysis work, it might extend to other constructions where the polar inference is invisible to embedding operators. Recall that it seemed the polar inference could not be construed as a result of the content of the *because*-clauses below (Ziegeler 2000).

(484) Almost every senior at Sunnydale High School attended the pep rally...

a. because there was pizza there.

b. # because they wanted to ditch.

(485) Barely any seniors at Sunnydale High School attended the pep rally...

a. # because there was pizza there.

b. because they wanted to ditch.

That from (484) we ultimately infer that there were some seniors who ditched the rally, the *because*-clause cannot pick up on this. The ditcher's reasons cannot serve as causal justification. The presence of pizza at the rally, though, can justify why the vast majority attended. Similarly, the inference drawn from (485) that some students did attend cannot be picked up on by the *because*-clause, with the presence of pizza justifying those few who did attend; the desire to ditch does justify (or at least, explains!) why close to none came.

Let's give an extremely rough meaning for *because*: we'll assume it relates two propositions, asserts the truth of both, and asserts that the proposition expressed by the first 'caused' the second.

(486) $[\![\text{because}]\!]^{g,c} = \lambda p_{st}\lambda q_{st}\lambda w.\ p(w)\ \&\ q(w)\ \&\ \textbf{CAUSE}_{w}(p)(q) = 1$

I would imagine this meaning won't stand up to scrutiny, but let's grant that as a rough and ready denotation, it's sufficient for our purposes. If *because* takes scope above **EXH**, then the inference should be, or at least could be, that the pizza was the reason that most everyone came *and* the reason that not everyone came.

(487) $[\![\text{because}]\!]^{g,c}([\![\text{there was pizza}]\!]^{g,c})([\![\textbf{EXH}\text{ almost }\chi_2\text{ every student attended}]\!]^{g,c})(w) = 1$ *iff* $\textbf{CAUSE}_{w}(\text{there was pizza})([\![\textbf{EXH}\text{ almost }\chi_2\text{ every student attended}]\!]^{g,c})(w)\ \&$ $[\![\textbf{EXH}\text{ almost }\chi_2\text{ every student attended}]\!]^{g,c}(w)\ \&\text{ there was pizza in w}$

The fact that such an inference is unavailable would suggest that *because* cannot be taking scope above **EXH**. What happens if **EXH** scopes over everything? In the derivation of the meaning, we'll use φ to refer to the constituent *almost every student attended*, ψ for *there was pizza*, and δ for their conjunction via *because*. The latter is the argument of **EXH** (all presuppositions suppressed here).

(488) $[\![\textbf{EXH}\ [\underbrace{[\![\text{because}]\!]^{g,c}(\underbrace{[\![\text{there was pizza}]\!]^{g,c}}_{\psi})(\underbrace{[\![\text{almost } \chi_2 \text{ every student attended}]\!]^{g,c}}_{\varphi})}_{\delta}]]\!]^{g,c}(w)$

$= 1$ *iff* (a) & (b) $= 1$

a. $[\![\varphi]\!]^{g,c}(w)\ \&\ [\![\psi]\!]^{g,c}(w)\ \&\ \textbf{CAUSE}_w([\![\psi]\!]^{g,c})([\![\varphi]\!]^{g,c})$

b. $\forall\chi'\Big[[\![\delta]\!]^{g,c} \not\Rightarrow_{ST} [\![\delta[\chi_2/\chi']]\!]^{g,c} \rightarrow \text{there's no pizza in w} \vee [\![\delta[\chi_2/\chi']]\!]^{g,c}(w) \neq 1 \vee$
$\neg\textbf{CAUSE}(\text{there's pizza in w})([\![\text{almost } \chi' \text{ every student attended}]\!]^{g,c})\Big]$

The problem with these truth conditions is that we don't predict that there should be any ditchers at all; at best, it seems that what we predict is that pizza didn't directly cause any members of the exception set to attend the rally. These are far too weak. *Because* clauses are a problem!

It's worth noting, at this juncture, that some account of the interactions between the contribution of **EXH** and expressive content is necessary independent of *almost* and *barely*. Other kinds of inferences are 'invisible' to such expressive content. Consider the well-known fact that (489) typically implies that Buffy didn't kill all the vampires.

(489) Buffy killed some of the vampires.
$\rightsquigarrow$ Buffy killed some but not all of the vampires.

This is a scalar implicature, and like our polar inferences, these scalar implicatures are often taken to be derived via **EXH**. This can be done by taking the alternatives to (489) to be formed via substituting *some* for other elements on its Horn scale (other quantificational determiners, e.g. *all*) and negating those alternatives that aren't entailed. Since (the proposition expressed by) (489) doesn't entail (the proposition expressed by) *Buffy killed all of the vampires*, **EXH** requires this alternative to be false. This generates the inference that Buffy didn't finish the job, so to speak. Importantly, this implicature is invisible to expressions like *good news*.

(490) a. Good news! Buffy killed some of the vampires.
$\rightsquigarrow$ Buffy killed some but not all of the vampires

b. Bad news! Buffy killed some of the vampires.
$\rightsquigarrow$ Buffy killed some but not all of the vampires

The good news in (490a) and the bad news in (490b) is the same news, just with a different perspective: it's that Buffy killed some of the vampires. Importantly, the good and bad news is *not* that she failed to kill all of the vampires. Our analysis of *almost* and *barely*, then, needs to be a part of a more general discussion of the precise nature of the inferences derived by **EXH**, and of **EXH**'s relationship with other operators.

In the next section, we'll turn to a variant of **EXH** that seems like a reasonable candidate for solving the issues noted here and in the previous section.

5.4 Presuppositional EXH

What if, instead of being truth conditional, the content of **EXH** had a different status? Bassi et al. (2021) propose an alternative to **EXH** that would accomplish just that—theirs is a presuppositional variant of **EXH**, called presuppositional **EXH**, or **PEX** for short. Instead of asserting the non-truth of alternatives to its prejacent, it *presupposes* their non-truth, and simply asserts the truth of its prejacent[5].

(491) $[\![\textbf{PEX}\ \varphi]\!]^{g,c}(w)$ is defined only if $\forall\psi\left[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{ST} [\![\psi]\!]^{g,c} \rightarrow [\![\varphi]\!]^{g,c}(w) \neq 1\right]$;
where defined: $[\![\textbf{PEX}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1$

Think about it in relation to the data from the previous section: presuppositions appear to project through the contribution of CIs, but expressive content like *good news* still doesn't directly comment on the presuppositions themselves. Here we're not even invoking **EXH**!

(492) a. Good news! The principal of Sunnydale High School was eaten by the mayor.
b. Good news! Willow stopped abusing magic.

(493) a. The principal of Sunnydale High School was eaten by the mayor because the mayor became a demon.
b. Willow stopped abusing magic because she feared its consequences.

The definite descriptions in (492a) presuppose the existence of a principal and a mayor; the verb *stopped* in (492b) presupposes Willow had been abusing magic. *Good news!* doesn't comment on this presupposed content, but rather on the truth conditional contributions of (492a) and (492b), the fact of the mayor's devouring the principal and the cessation of magical abuse, respectively. The presuppositions pass through *because*, too, in (493a) and (493b). It seems reasonable, then, to assume that expressive content does not apply to presuppositional content. If we adopt **PEX**, the

[5]We're simply making our semantics for **EXH** presuppositional. Their theory differs in a few ways that are, as far as I can tell, orthogonal to this endeavor. We assume that no harm comes of this, but a fuller exploration of their theory in the context of subtractives is warranted in future work.

relative scope of the operator with respect to *good news* and other speaker-oriented expressions doesn't matter, then. Regardless of configuration, neither operator can see the contribution the other makes. The 'invisibility' of the polar inference to speaker-oriented expressions on this view is a conspiracy between the distinct not-at-issue status of expressive content and presuppositions. This means that the locality constraint for **EXH** in relation to subtractives can be maintained without issue. Now, the questions are whether any issues arise with our other examples, and whether this solves the problem with *no* > *every...but.*

The issue with *no* > *every...but* for us, recall, is that the most local position for **EXH** is within the nuclear scope of *no*, and this is the position in which the derived truth conditions are correct.

(461) ***EXH AND THE LICENSING OF SUBTRACTIVES OPERATORS***

Subtractive operators are only licensed as modifiers of some quantifier Δ if:

a. They are in the scope of an **EXH** operator, and

b. **EXH** is licensed where its sister is the minimal expression of type ⟨s,t⟩ in which Δ is licensed.

(465) [no witch [2 [**EXH** [$_{\varphi}$ every spellbook but BG [3 [t_2 read t_3]]]]]]

(466) $[\![(465)]\!]^{g,c}(w)$ is defined only if $\{BG\} \subseteq \mathbb{S}_w$ & $\{BG\} \neq \emptyset$; where defined:

$$[\![(465)]\!]^{g,c}(w) = 1 \textit{ iff } \mathbb{W}_w \cap \Big\{ y \mid \mathbb{S}_w - \{BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\} \;\&\; \forall \mathbb{X}' \Big[\big[\{BG\} \not\subseteq \mathbb{X}' \;\&\; \mathbb{X}' \subseteq \mathbb{S}_w \;\&\; \mathbb{X}' \neq \emptyset \big] \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \{z \mid \textbf{READ}_w(z)(y)\} \Big] \Big\} = \emptyset$$

The problem is, though, that **EXH** is in a Downward Entailing environment, and therefore isn't licensed, so we're a bit stuck.

(469) ***ECONOMY CONSTRAINT ON THE DISTRIBUTION OF* EXH**

Let φ be an arbitrary expression containing an occurrence of the **EXH** operator. **EXH** is not licensed in φ if eliminating this occurrence leads to an expression ψ such that for any admissible g, c, $[\![\psi]\!]^{g,c} \Rightarrow [\![\varphi]\!]^{g,c}$.

An economy condition like (469) is needed on standard theories of **EXH** to block the calculation of scalar implicatures under negation (Fox & Spector 2018). Where *or* triggers a strong, exclusive inference in (494), under negation, as in (495), this isn't generated.

(494) Willow talked to Buffy or Xander.

$\rightsquigarrow$ *Willow didn't talk to both*

(495) Willow didn't talk to Buffy or Xander.

$\not\rightsquigarrow$ *Willow talked to both*

Only a constraint prohibiting **EXH** from occurring in Downward Entailing environments would block the LF for (495) that would generate the unattested inference that Willow talked to both Buffy and Xander. However, Bassi et al.'s **PEX** doesn't require such an economy constraint. For starters, they take the presuppositions encoded by **PEX** to be automatically globally accommodated, provided they are consistent with the context—they are more like the presuppositions encoded by possessives and *only*.

(496) Buffy is bringing her sister to the Magic Box.
presupposition: Buffy has a sister, x
at-issue content: Buffy is bringing x to the Magic Box

(497) Only Willow went to Cordelia's party.
presupposition: Willow went to Cordelia's party
at-issue content: No one else went to Cordelia's party

Barring incompatibility with the common ground, these kinds of presuppositions are readily accommodated, and not necessarily detected in the same way. Now, if **PEX** is in the scope of a Downward Entailing operator, what results is a presupposition that projects through that operator, but it is a presupposition that is entailed by the truth conditional, at-issue content of the entire sentence.

(498) [not [**PEX** [Willow talked to Buffy or Xander]]]

a. *predicted presupposition*: it's not the case that Willow talked to both Buffy and Xander
b. *at-issue content*: Willow didn't talk to Buffy and Willow didn't talk to Xander

The at-issue content asymmetrically entails the presupposition, which would be globally accommodated anyways, so its contribution is not detected. There is, in their theory, no need for the the economy constraint in (469). Now let's bring this back to *no* > *every...but*.

Since **PEX** can occur within the scope of *no* now, let's see what we can do with an LF that substitutes **EXH** for **PEX**. The truth conditions we want, recall, are those in (466).

(466) $$\mathbb{W}_w \cap \Big\{ y \mid \mathbb{S}_w - \{BG\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\} \;\&\; \forall \mathbb{X}' \Big[\big[\{BG\} \not\subseteq \mathbb{X}' \;\&\; \mathbb{X}' \subseteq \mathbb{S}_w \;\&\; \mathbb{X}' \neq \emptyset \big] \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \{z \mid \textbf{READ}_w(z)(y)\} \Big] \Big\} = \emptyset$$

The newly **PEX**ed LF is in (499).

(499) [no witch [2 [**PEX** [$_\varphi$ every spellbook but BG [3 [t_2 read t_3]]]]]]

(500) $[\![(499)]\!]^{g,c}(w)$ is defined only if (a) & (b) $= 1$

a. $\{BG\} \subseteq \mathbb{S}_w \;\&\; \mathbb{BG} \neq \emptyset$

b. $\forall \mathbb{X}' \Big[\big[\{\text{BG}\} \not\subseteq \mathbb{X}' \ \&\ \mathbb{X}' \neq \emptyset \ \&\ \mathbb{X}' \subseteq \mathbb{S}_w \big] \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \{z \mid \textbf{READ}_w(z)(g(2))\} \Big]$

Where defined, $[\![(499)]\!]^{g,c}(w) = 1$ *iff* $\mathbb{W}_w \cap \{y \mid \mathbb{S}_w - \{\text{BG}\} \subseteq \{z \mid \textbf{READ}_w(z)(y)\}\} = \emptyset$

At the point at which **PEX** calculates entailments, the restrictor of *every* is in a Strawson Downward Entailing environment, and so all and only supersets of $\{\text{BG}\}$ will be entailed (ignoring SMALL, of course), so the presupposition in (500b) requires that g(2) not have read BG. That's a free pronoun, though! To deal with quantified presuppositions like this, we can accommodate it locally, i.e. within the scope of the binding quantifier *no witch*. The necessity of this kind of solution for precisely this issue is within the tradition of Heim (1983), Van der Sandt (1992), and Beaver & Krahmer (2001). Following Bassi et al., we will use the $\mathscr{A}$ operator; this operator collapses falsity and undefinedness.

(501) $[\![\mathscr{A}\ \varphi]\!]^{g,c}(w) =$

$1,\ \text{if } [\![\varphi]\!]^{g,c}(w) = 1$

$0,\ \text{if } [\![\varphi]\!]^{g,c}(w) = 0 \text{ or } \#$

This means that the presuppositions of φ become part of its truth conditional content. Applying $\mathscr{A}$ below *no witch* and above **PEX** yields the LF (502) and so derives the truth conditions in (503).

(502) [no witch [2 [$\mathscr{A}$ [**PEX** [$_\varphi$ every spellbook but BG [3 [t_2 read t_3]]]]]]]

(503) $[\![(502)]\!]^{g,c}(w) = 1$ *iff*

$\{\text{BG}\} \neq \emptyset \ \&\ \{\text{BG}\} \subseteq \mathbb{S}_w \ \&\ \mathbb{W}_w \cap \Big\{ z \mid \mathbb{S}_w - \{\text{BG}\} \subseteq \{y \mid \textbf{READ}_w(y)(z)\} \ \&$

$\forall \mathbb{X}' \Big[\big[\{\text{BG}\} \not\subseteq \mathbb{X}' \ \&\ \mathbb{X}' \subseteq \mathbb{S}_w \ \&\ \mathbb{X}' \neq \emptyset \big] \rightarrow \mathbb{S}_w - \mathbb{X}' \not\subseteq \{y \mid \textbf{READ}_w(y)(z)\} \Big] \Big\} = \emptyset$

These truth conditions are just what we set out to derive: they require that no witch have read all and only those non-BG spellbooks. They are weaker than the truth conditions derived without $\mathscr{A}$ and **PEX**, of course, but that's what we wanted, of course.

A consequence of adopting the **PEX** based theory is that the economy constraint, which blocks **EXH** from being within the scope of *no*, is no longer necessary, at least in its current incarnation. **PEX**, in Downward Entailing environments, doesn't generate the wrong inferences, as **EXH** does—it is simply not detected because its contributed presupposition is asymmetrically entailed by the assertion. This allows us to maintain the licensing conditions on subtractives with respect to **PEX**, which we revise a final time for **PEX**.

(504) ***PEX AND THE LICENSING OF SUBTRACTIVES OPERATORS*** *final version*

Subtractive operators are only licensed as modifiers of some quantifier Δ if:

a. They are in the scope of a **PEX** operator, and

b. **pex** is licensed where its sister is the minimal expression of type $\langle s,t \rangle$ in which Δ is licensed.

It appears, then, that what subtractives do is provide an explicit argument in favor of a **pex**-type theory, and this in turn gives us a special insight into the polar inference subtractives give rise to: it is, generally speaking, a presupposition, albeit a readily accommodated one.

Now, there are a couple things we should clear up before concluding. The first point of order is the economy constraint. Although the economy constraint is no longer needed for the purposes of barring **pex** from occurring in Downward Entailing environments, we did make use of it to rule out cases where subtractives occurred in necessarily true quantificational claims, as in (505).

(505) # Almost every spellbook is a spellbook.

In our analysis, the pre-exhaustified meaning of (505) entailed the exhaustified meaning; if we don't appeal to the economy constraint when we switch to the **pex** theory, we don't have an immediate way of ruling such cases. Since we already avail ourselves of a cyclical sort of model for the licensing of subtractives—an LF with a particular shape must be permissible before considering other LFs—perhaps there are constraints imposed on **pex** that can feed into that cycle. For example, we can posit a novel economy condition on **pex** that is related to, but distinct from, the economy condition we abandoned above.

(506) ***economy constraint on the distribution of* pex** *to be revised*
Let φ be an arbitrary expression, and let $\psi = [\textbf{pex}\ \varphi]$.
For any g, c, if $[\![\varphi]\!]^{g,c} \Rightarrow [\![\psi]\!]^{g,c}$, then **pex** is not licensed in φ.

This condition would ensure that **pex** is not licensed in (505), since the un-**pex**haustified meaning entails the **pex**haustified meaning. [6] It also wouldn't allow a higher operator to save the derivation of an otherwise illicit LF. That is, we should still expect (507) to be odd because **pex** is not licensed because $[\![\varphi]\!]^{g,c} \Rightarrow_{ST} [\![\psi]\!]^{g,c}$.

(507) # [not [$_\psi$ **pex** [$_\varphi$ almost every spellbook is a spellbook]]]

Note, though, that the condition won't incorrectly rule out those cases we discussed above where **pex** is present in a Downward Entailing environment, and yet not detected. This is good!

(498) [not [$_\psi$ **pex** [$_\varphi$ Willow talked to Buffy or Xander]]]

a. *predicted presupposition*: it's not the case that Willow talked to both Buffy and Xander

b. *at-issue content*: Willow didn't talk to Buffy and Willow didn't talk to Xander

[6] Note that this constraint makes reference to strict entailment, not Strawson entailment. Because **pex** only adds presuppositions, for any φ, $[\![\varphi]\!]^{g,c} \Rightarrow_{ST} [\![\textbf{pex}\ \varphi]\!]^{g,c}$; the economy constraint needs to be stronger, then.

Here, although the global at-issue content entails the presupposition contributed by **PEX**, it's not the case that $[\![\psi]\!]^{g,c}$ is entailed by $[\![\varphi]\!]^{g,c}$, so **PEX** is licensed here even with our new economy condition. The basic role of the constraint is to rule out cases where **PEX** is locally vacuous; it succeeds in that. Perhaps there is a more sophisticated way in which we can rule out these cases, and maybe there are issues that will be uncovered with our constraint that point in new directions. This is left open for investigation, but, as far as I can tell, works for our theory.

The second point of order is to note that contradictions derived by **PEX** are still contradictions. The presuppositions contributed by **PEX** in (508) will contradict themselves and the truth conditions—before, it was just all truth conditional contradiction.

(508) # Almost some spellbook is worth reading.

Still, we likely need to add to the theory a means of explicitly stating that this LF is deviant, and that **PEX** is not licensed, and hence that *almost* is not licensed. Bassi et al. (2021:27) incorporate an *oddness filter* into their theory in order to rule out these kinds of LFs on the basis of their contradictions, e.g. (509) from Bassi et al. (2021:27).

(509) ***ODDNESS FILTER***

Let φ be a sentence with presupposition p and assertion q. For any context c, asserting φ in c is *odd* if:

a. p is inconsistent with c, **or**
b. p and q are individually consistent with c, but p & q is inconsistent with c

Their purpose in adopting such a filter is to explain why (510) is odd on natural assumption that, if all Italians are from a warm country (see also Magri (2014)).

(510) # Some Italians are from a warm country.

a. *predicted presupposition*: not all Italians are from a warm country
b. *at-issue content*: some but not all Italians are from a warm country

Here, the presupposition is in conflict with common knowledge, and hence a normal context c in which (510) is uttered. The oddness filter is an explicit way to say that this sentence is odd in typical contexts. The conjunction of the presuppositions and the assertion (508) are is inconsistent necessarily for us; note that we've added back in the size constraint to make sure this is crystal clear.

(511) $\big[$**PEX** $\big[$almost χ_2 some spellbook is worth reading$\big]\big]$

(512) $[\![(511)]\!]^{g,c}(w)$ is defined only if $g(2) \subseteq \mathbb{S}_w$ & $g(2) \neq \emptyset$ &
$\forall \mathbb{X}' \big[\big[\mathbb{X}' \not\subseteq g(2)$ & $\mathbb{X}' \subseteq \mathbb{S}_w$ & $\textsc{small}_c(\mathbb{X}')$ & $\mathbb{X}' \neq \emptyset\big] \rightarrow \mathbb{S}_w - \mathbb{X}' \cap \mathbb{WR}_w = \emptyset\big]$; where defined,
$[\![(511)]\!]^{g,c}(w) = 1$ *iff* $\mathbb{S}_w - g(2) \cap \mathbb{WR}_w = \emptyset$ & $\textsc{small}_c(g(2))$

The conjunction of the truth conditions and presupposition is contradictory, and so cannot be consistent with the context (or any non-trivial context). This LF is caught by the oddness filter. Now, for concreteness, we'll add into the economy constraint on **PEX** a condition barring it from being licensed if it yields oddness.

(513) ***ECONOMY CONSTRAINT ON THE DISTRIBUTION OF*** **PEX** *final version*
Let φ be an arbitrary expression, and let $\psi = [\textbf{PEX}\ \varphi]$.
For any g, c, if $[\![\varphi]\!]^{g,c} \Rightarrow [\![\psi]\!]^{g,c}$, or if $[\![\psi]\!]^{g,c}$ is odd, then **PEX** is not licensed in φ.

This allows us to ensure that embedding an odd LF won't save things, either.

(514) # Every witch read some spellbook but *The Brekenkreig Grimoire.*
a. *[every witch [2 [$_\psi$ **PEX** [$_\varphi$ some spellbook but BG [3 [t_2 read t_3]]]]]]
local contradiction
b. *[$_\psi$ **PEX** [$_\varphi$ some spellbook but BG [3 [every witch [2 [t_2 read t_3]]]]]]
local contradiction
c. *[**PEX** [every witch [2 [some spellbook but BG [3 [t_2 read t_3]]]]]]
barred by locality of licensing condition

PEX derives assertions and presuppositions that conflict necessarily in the LFs (514a) and (514b). These are odd LFs, and given the updated economy constraint, **PEX** is not licensed in these LFs, and neither is the subtractive. (514c) is not licensed by the locality of licensing condition, just as it was before.

In the end, we have quite a neat theory in **PEX**: it allows us to both solve the puzzles we set out to in multiple quantifier constructions and to predict why the polar inference disappears in certain environments. More work is needed to ensure that the theory we've ended up with is on the right track. Some questions will be inherited from Bassi et al. (2021), simply through assuming their theory. For example, Bassi et al. take the presuppositions triggered by **PEX** to be globally accommodated wherever possible—is this the right explanation for their status? When we turn to subtractives, does global accommodation really capture the nature of the polar inference? We have also proposed a sort of funny licensing condition on subtractives in claiming that the licensing of one particular LF permits the licensing, or blocks the licensing, of another. More work is needed to flesh out the details and consequences of such a view. These are essential issues, and this chapter does not answer them. It is, in the end, more of a sketch of what the theory might need to look like, but hopefully, it is clear that we're on the right track.

Chapter 6

Outlook

This chapter concludes the dissertation. In the first part of the chapter, we discuss a bit data that we've not yet gone through in detail, and discusses how we ought to analyze those constructions in light of the analysis developed and defended in the previous chapters. This will, by necessity, be a much looser discussion than in the previous chapters; unfortunately, doing justice to the distribution and interpretation of *almost* and *barely* along these lines requires more time than we have. However, the hope is that this discussion will show that there's good reason to think that the theory of the modifiers as subtractives can, and should be, extended beyond. The second part summarizes the dissertation and concludes.

6.1 Unanalyzed data

In the previous three chapters, we discussed in detail the distribution of *almost* and *barely* with respect to quantificational determiners, degree quantifiers, adjectives, and numeral constructions. We saw, generally speaking, that *almost* and *barely* were in complementary distribution up until we reached numeral constructions, where both subtractives on the surface appear to modify the same expression. There are plenty of other domains in which we haven't applied our analysis, and there's plenty of data for the theory to cover. We're going to focus on verbal uses for now. Verbs aren't typically understood as quantificational constructions, and as such, it's much harder to see at first how we can reconcile such uses with our theory. That our theory only reaches at the finale the data that other theories begin with is possibly a problem, but we'll suggest that it's not. If we're right about the subtractives, then we're committed to the presence of quantification in hiding, and there's perhaps good reason to think it is present. While we won't be able to provide a full account of these data, we'll do our best to suggest some tentative analyses.

What I call *verbal* uses of *almost*—those in which *almost* seems to modify a verb phrase—are probably the most discussed uses in the literature. They have received quite a lot of attention

(Morgan 1969; McCawley 1972; Dowty 1979; Rapp & von Stechow 1999; Eckardt 2007; McKenzie & Newkirk 2019). These are uses that can also reasonably be seen as *clausal* or *propositional*. Indeed, oftentimes we see analyses of *almost* as a proposition-modifying operator in these uses, with appeals to the modal or scalar closeness semantics discussed in Chapter 2. *Barely*, as we've mentioned, has received less attention in general than *almost*, but it is its verbal uses that have largely received attention (Ducrot 1973; Horn 1996; Atlas 1997; Sevi 1998; Amaral 2007). Think about (515): it seems to admit two, maybe three interpretations (Morgan 1969; McCawley 1972).

(515) Giles almost killed Angel.

a. Giles almost did something that would have had the effect of Angel's dying
b. Giles did something that almost had the effect of Angel's dying
c. Giles did something which had the effect of Angel's becoming almost not alive (i.e. becoming nearly dead)

One interpretation is (515a); let's call this the *counterfactual* (CF) interpretation. It's true even where Giles made no attempt to kill Angel, but he seriously considered doing it. In another world, where facts aren't too different from this one, he did kill Angel. The second and the third are quite similar (though perhaps indistinct, cf. Dowty 1979)—both entail that Giles did something that is related to Angel's near death. The second, what we can call the *scalar* (SC) interpretation, doesn't require that Angel came close to dying, but rather, had the facts been different, he would have. The third is the *resultative* (RS) interpretation. Here, Angel really came close to death. Ultimately he is alive, but it was a close one. Note that in all three, there's a way to understand the interpretations as saying that some event or state of affairs didn't come about, but it was close to coming about. Just what closeness amounts to does seem to be interpretation-specific.

Morgan (1969) and McCawley (1972) argue these kinds of data support a decompositional view of verbs like *kill* into three parts, each of which can be targeted by *almost*.

(516) Giles almost killed Angel.

a. [almost [**ACT**(Giles) [**CAUSE** [**BECOME** [Angel dead]]]]] COUNTERFACTUAL
$\rightsquigarrow$ Giles almost did something that would have had the effect of Angel's dying
b. [**ACT**(Giles) [almost [**CAUSE** [**BECOME** [Angel dead]]]]] SCALAR
$\rightsquigarrow$ Giles did something that almost had the effect of Angel's dying
c. [**ACT**(Giles) [**CAUSE** [**BECOME** [almost [Angel dead]]]]] RESULTATIVE
$\rightsquigarrow$ Giles did something which had the effect of Angel's becoming almost not alive (i.e. becoming nearly dead)

Of course, it's not settled that decomposition along these lines is the right call—Dowty (1979), for example, suggests that the different "interpretations" arise out of vagueness or ambiguity rather than such a complex structure.

Rapp & von Stechow (1999) similarly suggest based on German *fast* "almost" that decomposition might play some role, though it can't be the whole story. First, they note that CF interpretations are connected to the presence or absence of particular verbal morphology, the Konjunktiv II (glossed here as K2), in some dialects, and its presence blocks SC interpretations.

(517) a. weil David fast seinen Hasen erwürgt hätte
because David almost his rabbits strangled have-K2
'Because David almost strangled his rabbits' CF, *SC

b. weil David seinen Hasen fast erwürgt hätte
because David his rabbits almost strangled have-K2
'Because David almost strangled his rabbits' CF, SC as 'indirect discourse'

c. ?? weil David fast seinen Hasen erwürgte
because David almost his rabbits strangled-PRET
'Because David almost strangled his rabbits' ??CF, *SC

d. weil David seinen Hasen fast erwürgte
because David his rabbits almost strangled-PRET
'Because David almost strangled his rabbits' *CF, SC

Rapp & von Stechow (1999:156-157)

That there must be, for some dialects, dedicated counterfactual morphology for the CF interpretation of *fast* to appear is rather interesting. It could be telling us that *fast* is modifying the semantic correlate of that morphology, and when it isn't present, *almost* must be modifying something else. Other languages illustrate verbal morphology-based restrictions on the availability of CF vs. SC interpretations, tied to aspect. In European Portuguese, the Imperfeito do Indicativo, an imperfective past tense, allows only a CF interpretation, but the Pretérito Perfeito Simples, a perfective past tense, allows both CF and SC interpretations (Amaral 2007).

(518) a. O João quase ganhava a corrida.
The João almost win-PAST.IMPFV.3SG the race
'João almost won the race.' CF, *SC

b. O João quase ganhou a corrida.
The João almost win-PAST.PFV.3SG the race
'João almost won the race.' CF, SC

Amaral (2007:137)

Sometimes *almost* seems to be commenting on the manner of performing some action; consider the following contrast.

(519) a. ?? Rafi aß den Kuchen fast.
Rafi ate the cake almost
'Rafi almost ate the cake.'

b. Rafi fraß den Kuchen fast.
Rafi devoured the cake almost
'Rafi almost devoured the cake.'

Rapp & von Stechow (1999:182-183)

Here, the use of the more forceful, 'colorful' verb *fraß* with *fast* brings out an interpretation in which the manner is commented on: the way Rafi ate the cake approached a devouring. There was no moderation!

The data above is a lot of ground to cover, and it is, of course, not the full extent of things. That said, if we've learned anything from our investigation into subtractives, it's that starting from their uses as modifiers of quantificational determiners and generalizing from there is incredibly fruitful. The appearance of subtractive operators in the constructions above tells us we need to be looking for some quantification. We saw in our investigation of numeral constructions that sometimes we have to dig around and pull things apart in order to find the quantifiers; perhaps they've been hiding in plain sight in verbal uses.

Think about the counterfactual interpretations above in this light.

(520) Giles almost killed Angel.

As discussed above, this is true in a context where Giles came close to, but did not actually, kill Angel. Giles could have merely contemplated the murder, he could have only just missed the heart, whatever. Had the facts been different, Angel would be dead. Perhaps this interpretation arises as a result of *almost* modifying a particular kind of universal modal which trafficks in those kinds of facts, and *almost* subtracts a set of facts relevant to Giles's killing of Angel from that modal's base. **PEX** would then say that that's the smallest, non-empty set of facts in the context that you can subtract from the modal base and still yield Angel's death.

The fact that in some languages we see counterfactual interpretations of *almost* cognates correlating with particular morphology could lend support to this sketch of a proposal. The presence of K2 morphology, at least for some speakers of German, is the trigger for counterfactual interpretations of *fast* because K2 is the host of a counterfactual-type modal. The absence of K2 for such speakers would mean that there is *no* such counterfactual modal present, and *fast* would have to modify something else.

(521) weil David fast seinen Hasen erwürgt hätte
because David almost his rabbits strangled have-K2
'Because David almost strangled his rabbits' CF, *SC

(522) weil David seinen Hasen fast erwürgte
because David his rabbits almost strangled-PRET
'Because David almost strangled his rabbits' *CF, SC

The idea that morphology like K2 is utilized in constructions to contribute a sort of modal displacement is not new—it's familiar from the literature on *X-marking*, as in von Fintel & Iatridou (2020)—but this idea needs to be better grounded in the rather complex set of facts surrounding X-marking cross-linguistically, and a careful investigation goes beyond what we have time for.

What about non-counterfactual uses of *almost*, and what about *barely*? How do we incorporate them into the theory? The European Portuguese facts might be giving us a clue. Perhaps we can think of *almost* and *barely* in these cases as modifiers of aspectual morphology. Aspect is typically taken to relate events to times, and introduce existential quantification over events (Klein 1994, i.a.). Let v be the semantic type of events, t be the semantic type of times, and τ be a function that maps events to their run times; we'll assume that times are intervals for simplicity here.

(523) $[\![\textsc{perfective}]\!]^{g,c} = \lambda A_{v,t}\lambda t.\ \exists e\big[e \in \mathbb{A}\ \&\ \tau(e) \subseteq t\big]$

(524) $[\![\textsc{imperfective}]\!]^{g,c} = \lambda A_{v,t}\lambda t.\ \exists e\big[e \in \mathbb{A}\ \&\ t \subseteq \tau(e)\big]$

Already the existential quantification could be a starting point for *barely* to latch onto. *Almost* would need some work. That said, perfective aspect conveys that an event's runtime is contained in some other time t; perhaps this boudnedness gives us access to endpoints from which *almost* can subtract, whereas imperfective doesn't give such access. At least for European Portuguese *quase*, this might be the start of an account of the absence of 'scalar' interpretations when the verb is marked with imperfective, but their availability when the verb is marked with perfective. *Quase* could be subtracting, for example, subintervals of the runtime of a João winning-the-race event in (518b) saying that if we took a small amount of time away from the runtime of such an event, it would be complete, but in fact, that's not the state of affairs. Perhaps something like this is on the right track.

There are still a host of other areas to explore with subtractives, and we've only just scratched the surface of the cross-linguistic picture. Pushing the theory developed and defended in this dissertation and testing it against in sorts of novel territory is left for future work.

6.2 Wrapping up

This dissertation has been about *almost* and *barely*. We began our journey in Chapter 2 by looking at uses of the subtractives as modifiers of determiners. Building on work by von Fintel (1993), Gajewski (2013), and Crnič (2018) on exceptive *but*, we took all three operators to encode set subtraction, removing an exception from the restrictor of the modified quantifier. We added presuppositions that required the exception to be a subset of the set from which it's subtracted, and non-empty. We also argue that all three operators require their exception to be small in the context; this is the *size* constraint, and in particular, it is an essential part of the derivation of the *proximal* inference in subtractives.

(525) $[\![\text{but}]\!]^{g,c} = \lambda X \lambda A \lambda \mathscr{D} \lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

(526) $[\![\text{almost}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

(527) $[\![\text{barely}]\!]^{g,c} = \lambda X \lambda \mathscr{D} \lambda A \lambda B\text{: } \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

$\langle \text{et},\langle\langle \text{et},\text{ett}\rangle,\langle \text{et},\langle \text{et},\text{t}\rangle\rangle$

(528) ***THE SIZE CONSTRAINT*** *final version*

Let c be a context of utterance, $\mathbb{X}$ be an arbitrary set, and n_c be a contextually determined numerical threshold for size.

Where $\mathbb{X}$ is a dense set of degrees, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $\textsc{length}(\mathbb{X}) \leq n_c$;

Otherwise, $\textsc{small}_c(\mathbb{X}) = 1$ *iff* $|\{y \mid y \in \mathbb{X}\ \&\ \textsc{atom}(y)\}| \leq n_c$

(529) ***MAXIMAL SIZE OF*** n_c *final version*

Let φ be a constituent with subconstituent γ such that $\gamma = [\Delta\ [\alpha\ [\Sigma\ \chi]]]$ or $[[[\Sigma\ \chi]\ \Delta]\ \alpha]$, where Σ is a subtractive operator, Δ is quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively.

An utterance of φ in c is felicitous only if, for any context c:

Where $\mathbb{A}$ is a dense set of degrees, $n_c < \frac{1}{3}(\textsc{length}(\mathbb{A}))$;

Otherwise, $n_c < \frac{1}{3}|\{y \mid y \in \mathbb{A}\ \&\ \textsc{atom}(y)\}|$

All three operators are required to be within the scope of an operator that derives scalar implicatures via negating non-Strawson entailed alternatives of an LF φ. Until Chapter 5, this operator was **exh**; for expository purposes we restricted ourselves to 'subtractive' LFs and to 'subtractive' alternatives, as defined below, which allow us to more succinctly state the contribution of **exh** given such LFs and alternatives.

(530) $[\![\textbf{exh}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\psi\,[\psi \in \textbf{alt}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{\textsc{st}} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1]$

(531) ***SUBTRACTIVE LFS***

A LF φ is a *subtractive LF* just in case φ is a LF containing exactly one subtractive operator Σ introducing an exception χ, and no other alternative-triggering expression, and φ is of type $\langle s,t\rangle$.

(532) **SUBTRACTIVE ALTERNATIVES**

For a subtractive LF φ, with an exception χ,

$\textsc{alt}(\varphi) := \left\{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_d \;\&\; \varphi[\chi/\chi'] \text{ is grammatical}\right\}$

(533) For a subtractive LF φ,
$[\![\textsc{exh}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1 \;\&\; \forall\chi'\left[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}(w) \neq 1\right]$

Entailment between alternatives is tied to monotonicity.

(534) ***SUBTRACTIVE ENTAILMENT THEOREMS FOR DETERMINERS***
Let Σ be a subtractive operator, Δ a quantificational determiner, and α and χ be expressions whose denotations characterize $\mathbb{A}$ and $\mathbb{X}$, respectively. Let φ be a constituent with subconstituent γ such that $\gamma = \left[\Delta\ \left[\alpha\ \left[\Sigma\ \chi\right]\right]\right]$ or $\left[\left[\left[\Sigma\ \chi\right]\ \Delta\right]\ \alpha\right]$. Let g be an assignment function in context c.
Let $\textsc{alt}(\varphi) = \left\{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_e \;\&\; \varphi[\chi/\chi'] \text{ is grammatical}\right\}$.

a. If φ is Upward Entailing (and not Downward Entailing) with respect to the constituent α, then for all χ′ such that $\mathbb{X}' \subseteq \mathbb{X}$, then $[\![\varphi]\!]^{g,c} \Rightarrow_{\text{ST}} [\![\varphi[\chi/\chi']]\!]^{g,c}$.
b. If φ is Downward Entailing (and not Upward Entailing) with respect to the constituent α, then for all χ′ such that $\mathbb{X} \subseteq \mathbb{X}'$ & $\mathbb{X}'$ is small in c, then $[\![\varphi]\!]^{g,c} \Rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}$.

We saw that it is only when, for a given exception χ in a subtractive LF φ, alternatives to φ formed with supersets of $\mathbb{X}$ (the set characterized by $[\![\chi]\!]^{g,c}$) are entailed are subtractives possible. This is a necessary, not sufficient, condition, but it's an important one. This occurs when φ is Downward Entailing with respect to the restrictor of the subtractive-modified determiner. This captures the distribution of subtractives with respect to the quantificational determiners explored in Chapter 2. **EXH** is also, importantly, the locus of the polar inference—its contribution entails that the unmodified quantificational claim is false in the case of *almost*, and true in the case of *barely*.

In Chapter 3, we extended the coverage of this theory to degree constructions, i.e. comparatives, equatives, and the positive. There, we learned a few important things. First, subtractives provided evidence in favor of the idea that all natural language scales are dense, even when they aren't intuitively. Second, we learned from *barely*-modified comparatives that *barely* must subtract not from the restrictor of the quantifiers it modifies, but the nuclear scope. This revision, fortunately, doesn't change the predictions made in the previous chapter. The extension of the theory developed in Chapter 2 to degree constructions is a central contribution of the dissertation. We see that by starting from the determiner-modifier uses of subtractives, we naturally and readily capture data that have been largely ignored in the literature. This is big, and set the stage for the investigations in the following chapter.

In Chapter 4, we probed numeral constructions, where we saw that the complementarity in distribution between *almost* and *barely* we'd seen so far disappears. We argued, though, on the basis of the theory developed in Chapters 2 and 3 that there is structurally more than meets the eye in numeral constructions. We posited that, in numeral constructions, there is a covert existential quantifier over individuals $\mathscr{E}$, which *barely* can modify, and a degree quantifier $\mathscr{M}$ with the semantics of the equative, which we know *almost* can modify. We explored how the composition allows us to explain why both subtractives are licensed in these constructions. This chapter is also important: it taught us that it's fruitful to use subtractives as evidence for the presence of quantifiers, even when we don't expect them. The ability of both *almost* and *barely* to surface in numeral constructions is crucial for deducing the presence of quantifiers of different kinds.

Chapter 5 returned to quantificational determiners and the nature and role of **EXH**. We complicated matters by considering structures in which subtractive-modified quantifiers are embedded under other quantificational elements, and cases where the polar inference associated with subtractives is invisible to other operators. We were led to place constraints on the locality of **EXH** with respect to subtractive operators, which are only licensed in the scope of **EXH**, and we were led to supplant **EXH** with a presuppositional version, **PEX** (Bassi et al. 2021).

(535) $[\![\textbf{PEX}\ \varphi]\!]^{g,c}(w)$ is defined only if $\forall\psi\left[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{\mathrm{ST}} [\![\psi]\!]^{g,c} \rightarrow [\![\varphi]\!]^{g,c}(w) \neq 1\right]$;
where defined: $[\![\textbf{PEX}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1$

(536) **PEX** ***AND THE LICENSING OF SUBTRACTIVES OPERATORS*** *final version*
Subtractive operators are only licensed as modifiers of some quantifier Δ if:

a. They are in the scope of a **PEX** operator, and
b. **PEX** is licensed where its sister is the minimal expression of type $\langle s,t\rangle$ in which Δ is licensed.

The resulting picture is one in which we neatly capture the invisibility of the polar inference to certain operators—the presuppositional polar inference escapes the notice of 'expressive' content—and accurately capture the meaning and distribution of subtractives embedded under other quantificational operators.

The first half of this chapter briefly discussed some areas in which our theory should be applied, and sketched what such an extension might look for and look like. If this theory is right, we expect more quantifiers in more places; subtractives give us evidence of their presence.

Appendix A

A.1 Denotations, definitions, and a theorem

This appendix is formulated for the discussion in Chapter 2, so the denotation of *barely* is nuclear scope-subtracting. As discussed in Chapter 3, arguments don't really change too much for nuclear scope *barely*.

(537) ***DENOTATIONS OF SUBTRACTIVE OPERATORS***

a. $[\![\text{but}]\!]^c = \lambda X_{\langle e,t\rangle}\lambda A_{\langle e,t\rangle}\lambda \mathscr{D}_{\langle et,ett\rangle}\lambda B_{\langle e,t\rangle}\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

b. $[\![\text{almost}]\!]^c = \lambda X_{\langle e,t\rangle}\lambda A_{\langle e,t\rangle}\lambda \mathscr{D}_{\langle et,ett\rangle}\lambda B_{\langle e,t\rangle}\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

c. $[\![\text{barely}]\!]^c = \lambda X_{\langle e,t\rangle}\lambda A_{\langle e,t\rangle}\lambda \mathscr{D}_{\langle et,ett\rangle}\lambda B_{\langle e,t\rangle}\colon \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B})\ \&\ \textsc{small}_c(\mathbb{X})$

(538) ***CONJOINABLE TYPES***

a. t is a conjoinable type.

b. If τ is a conjoinable type, then for all types σ, $\langle\sigma,\tau\rangle$ is a conjoinable type.

(539) ***CROSS-CATEGORIAL ENTAILMENT*** '$\Rightarrow$'

a. For p, q of type t: $p \Rightarrow q$ *iff* $p = 1 \rightarrow q = 1$.

b. For f, g of conjoinable type $\langle\sigma,\tau\rangle$: $f \Rightarrow g$ *iff* for every x of type σ, $f(x) \Rightarrow g(x)$.

(540) ***CROSS-CATEGORIAL STRAWSON ENTAILMENT*** '$\Rightarrow_{ST}$' von Fintel (1999)

a. For p, q of type t: $p \Rightarrow_{ST} q$ *iff* $\left[p = 1\ \&\ q \neq \#\right] \rightarrow q = 1$.

b. For f, g of conjoinable type $\langle\sigma,\tau\rangle$: $f \Rightarrow_{ST} g$ *iff* for every x of type σ, $f(x) \Rightarrow_{ST} g(x)$.

(541) ***SUBSTITUTION-BASED STATEMENT OF (STRAWSON) ENTAILMENT*** Gajewski (2011)

For a constituent φ and some subconstituent α:

a. φ is (Strawson) Downward Entailing with respect to α *iff*
$\forall\beta\left\{[\![\beta]\!]^{g,c} \Rightarrow_{(ST)} [\![\alpha]\!]^{g,c} \rightarrow [\![\varphi]\!]^{g,c} \Rightarrow_{(ST)} [\![\varphi[\alpha/\beta]]\!]^{g,c}\right\}$

b. φ is (Strawson) Upward Entailing with respect to α *iff* $\forall\beta\{[\![\alpha]\!]^{g,c} \Rightarrow_{(ST)} [\![\beta]\!]^{g,c} \rightarrow [\![\varphi]\!]^{g,c} \Rightarrow_{(ST)} [\![\varphi[\alpha/\beta]]\!]^{g,c}\}$

(542) ***SET SUBTRACTION*** '$-$'

a. For arbitrary sets $\mathbb{A}$, $\mathbb{B}$: $\mathbb{A}-\mathbb{B} := \{x \mid x \in \mathbb{A} \;\&\; x \notin \mathbb{B}\}$.
($\mathbb{A}$ is the *minuend* and $\mathbb{B}$ the *subtrahend*; the set $\mathbb{C}$ such that $\mathbb{A}-\mathbb{B} = \mathbb{C}$ is the *difference* or *remainder.*)

b. For any arbitrary expressions α, β, of type $\langle e,t\rangle$, let $\mathbb{A}$ and $\mathbb{B}$ be the sets characterized by $[\![\alpha]\!]^{g,c}$ and $[\![\beta]\!]^{g,c}$, and let '$\alpha - \beta$' be a wellformed expression of type $\langle e,t\rangle$ such that $[\![\alpha - \beta]\!]^{g,c}$ is a function characterizing the set $\mathbb{A}-\mathbb{B}$

(543) ***SUBTRACTIVE LFs***

A LF φ is a *subtractive LF* just in case φ is a LF containing exactly one subtractive operator Σ introducing an exception χ, and no other alternative-triggering expression, and φ is of type $\langle s,t\rangle$.

Now we add one theorem for our proofs.

(544) ***SUBSTITUTION EQUIVALENCE FOR SUBTRACTIVES (S.E.S.)***

Let Σ be a subtractive operator, Δ be a quantificational determiner, and α and χ be arbitrary expressions denoting functions of type $\langle e,t\rangle$. Let φ be an LF that contains a constituent γ such that, for particular occurrences of α and χ, call them α^* and χ^*, $\gamma = [[[\Sigma\ \chi^*]\ \Delta]\ \alpha^*]$ or $[\Delta\ [\alpha^*\ [\Sigma\ \chi^*]]]$.
For any α', χ', if $[\![\alpha^*]\!]-[\![\chi^*]\!] = [\![\alpha']\!]-[\![\chi']\!]$, then $[\![\varphi]\!] = [\![\varphi[\alpha^*/\alpha', \chi^*/\chi']]\!]$.

This theorem comes in handy in proving entailments between exceptions. See below.

A.1.1 Subtractive equivalences

(1) a. $[\![\text{but}]\!]^c = \lambda X_{\langle e,t\rangle}\lambda A_{\langle e,t\rangle}\lambda\mathscr{D}_{\langle et,ett\rangle}\lambda B_{\langle e,t\rangle}\!: \mathbb{X} \subseteq \mathbb{A} \;\&\; \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B}) \;\&\; \textsc{small}_c(\mathbb{X})$

b. $[\![\text{almost}]\!]^c = \lambda X_{\langle e,t\rangle}\lambda A_{\langle e,t\rangle}\lambda\mathscr{D}_{\langle et,ett\rangle}\lambda B_{\langle e,t\rangle}\!: \mathbb{X} \subseteq \mathbb{A} \;\&\; \mathbb{X} \neq \emptyset.\ \mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B}) \;\&\; \textsc{small}_c(\mathbb{X})$

c. $[\![\text{barely}]\!]^c = \lambda X_{\langle e,t\rangle}\lambda A_{\langle e,t\rangle}\lambda\mathscr{D}_{\langle et,ett\rangle}\lambda B_{\langle e,t\rangle}\!: \mathbb{X} \subseteq \mathbb{A} \;\&\; \mathbb{X} \neq \emptyset.\ \neg\mathscr{D}(\mathbb{A}-\mathbb{X})(\mathbb{B}) \;\&\; \textsc{small}_c(\mathbb{X})$

(2) a. $[\![\text{every}]\!] = \lambda A\lambda B.\ \mathbb{A} \subseteq \mathbb{B}$

b. $[\![\text{no}]\!] = \lambda A\lambda B.\ \mathbb{A} \cap \mathbb{B} = \emptyset$ $\quad(= \lambda A\lambda B.\ \mathbb{A} \subseteq \overline{\mathbb{B}})$

c. $[\![\text{some}]\!] = [\![\text{any}_{NPI}]\!] = \lambda A\lambda B.\ \mathbb{A} \cap \mathbb{B} \neq \emptyset$

(3) Let α, β, and χ, be arbitrary expressions whose denotations characterize the sets $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{X}$, respectively; let g be such that $g(i) = \mathbb{X}$

a. $[\![[\text{every } [\alpha\ [\text{but } \chi]]]]\!]^{g,c} = [\![[[[\text{almost } \chi_i] \text{ every}]\ \alpha]]\!]^{g,c} =$
$\lambda B{:}\ \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ (\mathbb{A}-\mathbb{X}) \subseteq \mathbb{B}\ \&\ \textsc{small}_c(\mathbb{X})$

b. $[\![[\text{some } [\alpha\ [\text{but } \chi]]]]\!]^{g,c} = [\![[[[\text{almost } \chi_i] \text{ some}]\ \alpha]]\!]^{g,c} = [\![[[[\text{barely } \chi_i] \text{ no}]\ \alpha]]\!]^{g,c} =$
$\lambda B{:}\ \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ (\mathbb{A}-\mathbb{X}) \cap \mathbb{B} \neq \emptyset\ \&\ \textsc{small}_c(\mathbb{X})$

c. $[\![[\text{no } [\alpha\ [\text{but } \chi]]]]\!]^{g,c} = [\![[[[\text{almost } \chi_i] \text{ no}]\ \alpha]]\!]^{g,c} = [\![[[[\text{barely } \chi_i] \text{ any}_{\text{NPI}}]\ \alpha]]\!]^{g,c} =$
$\lambda B{:}\ \mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset.\ (\mathbb{A}-\mathbb{X}) \cap \mathbb{B} = \emptyset\ \&\ \textsc{small}_c(\mathbb{X})$

d. $[\![[[[[\text{almost } \chi_i] \text{ some}]\ \alpha]\ \beta]]\!]^{g,c}$ and $[\![[[[[\text{barely } \chi_i] \text{ every}]\ \alpha]\ [\text{not } \beta]]]\!]^{g,c}$ are defined only if $\mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset$; where defined,
$[\![[[[[\text{almost } \chi_i] \text{ some}]\ \alpha]\ \beta]]\!]^{g,c} = [\![[[[[\text{barely } \chi_i] \text{ every}]\ \alpha]\ [\text{not } \beta]]]\!]^{g,c}$
$= 1$ *iff* $(\mathbb{A}-\mathbb{X}) \cap \mathbb{B} \neq \emptyset\ \&\ \textsc{small}_c(\mathbb{X})$

e. $[\![[[[[\text{almost } \chi_i] \text{ no}]\ \alpha]\ \beta]]\!]^{g,c}$ and $[\![[[[[\text{almost } \chi_i] \text{ every}]\ \alpha]\ [\text{not } \beta]]]\!]^{g,c}$ are defined only if $\mathbb{X} \subseteq \mathbb{A}\ \&\ \mathbb{X} \neq \emptyset$; where defined,
$[\![[[[[\text{almost } \chi_i] \text{ no}]\ \alpha]\ \beta]]\!]^{g,c} = [\![[[[[\text{almost } \chi_i] \text{ every}]\ \alpha]\ [\text{not } \beta]]]\!]^{g,c}$
$= 1$ *iff* $(\mathbb{A}-\mathbb{X}) \subseteq \overline{\mathbb{B}}\ \&\ \textsc{small}_c(\mathbb{X})$

A.2 Proofs

A.2.1 Subtraction is Downward Entailing with respect to the subtrahend

(1) Let $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{A}$ be arbitrary sets — PREMISE

SHOW: $\mathbb{X} \subseteq \mathbb{Y} \rightarrow \mathbb{A}-\mathbb{Y} \subseteq \mathbb{A}-\mathbb{X}$

(2) $\mathbb{X} \subseteq \mathbb{Y}$ — ASSUMPTION

SHOW: $\mathbb{A}-\mathbb{Y} \subseteq \mathbb{A}-\mathbb{X}$

(3) $\mathbb{X} \subset \mathbb{Y} \vee \mathbb{X} = \mathbb{Y}$ — def. of $\subseteq$, (2)

CASE 1: $\mathbb{X} \subset \mathbb{Y}$

(4) $\exists z[z \in \mathbb{Y}\ \&\ z \notin \mathbb{X}]$ — def. of $\subset$

(5) $\neg\exists z[z \in \mathbb{Y}-\mathbb{X}\ \&\ z \in \mathbb{A}] \vee \exists z[z \in \mathbb{Y}-\mathbb{X}\ \&\ z \in \mathbb{A}]$ — TAUTOLOGY

CASE 2.1: $\neg\exists z[z \in \mathbb{Y}-\mathbb{X}\ \&\ z \in \mathbb{A}]$

(6) $\mathbb{A}-\mathbb{Y} = \mathbb{A}-\mathbb{X}$ — def. of $-$, $\mathbb{X} \subset \mathbb{Y}$

(7) $\mathbb{A}-\mathbb{Y} \subseteq \mathbb{A}-\mathbb{X}$ — (6), def. of $\subseteq$

CASE 2.2: $\exists z[z \in \mathbb{Y}-\mathbb{X}\ \&\ z \in \mathbb{A}]$

(8) $\exists z[z \in \mathbb{A}-\mathbb{X}\ \&\ z \notin \mathbb{A}-\mathbb{Y}]$

(9) $\mathbb{A}-\mathbb{Y} \subset \mathbb{A}-\mathbb{X}$ — def. of $-$, $\mathbb{X} \subset \mathbb{Y}$

(10) $\mathbb{A}-\mathbb{Y} \subseteq \mathbb{A}-\mathbb{X}$ (10), def of $\subseteq$

CASE 2: $\mathbb{X} = \mathbb{Y}$

(11) $\mathbb{A}-\mathbb{Y} \subseteq \mathbb{A}-\mathbb{X}$ TRIVIAL

Q.E.D.

A.2.2 Subtractive entailments

We will now make general claims about entailment between an exception χ in a subtractive LF φ and alternative exceptions as a function of the monotonicity of φ with respect to the argument α from which χ is subtracted. These proofs are about Logical Form, and will make use of the S.E.S. theorem, so that we can make claims about all of our subtractives, regardless of the syntax in which they appear.

We omit the size constraint, but it should be obvious what its inclusion will add. For any context c and arbitrary set $\mathbb{X}$, if $\textsc{small}_c(\mathbb{X})$ is true, then any subset of $\mathbb{X}$ is also small in c; non-subsets might not be.

(1) Let Σ be a subtractive operator, Δ be a quantificational determiner, and α and χ be arbitrary expressions denoting functions of type $\langle$e,t$\rangle$ PREMISE

(2) Let φ be subtractive LF that contains particular occurrences of α and χ, call them α^* and χ^*, and a constituent γ such that $\gamma = \left[\Delta \left[\alpha^* \left[\Sigma\ \chi^*\right]\right]\right]$ PREMISE

A.2.3 φ is Upward Entailing with respect to α

SHOW: If φ is Upward Entailing with respect to α^*, then $\forall\chi' \left[[\![\chi']\!] \subseteq [\![\chi^*]\!] \rightarrow [\![\varphi]\!] \Rightarrow [\![\varphi[\chi^*/\chi']]\!]\right]$

(3) φ is Upward Entailing with respect to α^* ASSUMPTION

SHOW: $\forall\chi' \left[[\![\chi']\!] \subseteq [\![\chi^*]\!] \rightarrow [\![\varphi]\!] \Rightarrow [\![\varphi[\chi^*/\chi']]\!]\right]$

(4) Let $\overline{\chi}$ be such that $[\![\overline{\chi}]\!]$ characterizes $\emptyset$; let ψ be the LF such that $\psi = \varphi[\alpha^*/\alpha^*-\chi^*, \chi^*/\overline{\chi}]$

(5) $[\![\alpha^*]\!]-[\![\chi^*]\!] = [\![\alpha^*-\chi^*]\!]-[\![\overline{\chi}]\!]$ def. of $-$

(6) $[\![\varphi]\!] = [\![\psi]\!]$ 4., 5., S.E.S

(7) ψ is Upward Entailing with respect to $\alpha^*-\chi^*$ 3., 6.

SHOW: $\forall\chi' \left[[\![\chi']\!] \subseteq [\![\chi^*]\!] \rightarrow [\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi']]\!]\right]$

(8) Let χ^{**} be an arbitrary expression such that $[\![\chi^{**}]\!] \subseteq [\![\chi^*]\!]$ ASSUMPTION

SHOW: $[\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi^{**}]]\!]$

(9) $[\![\alpha^*-\chi^*]\!] \subseteq [\![\alpha^*-\chi^{**}]\!]$ A.1.2

(10) $[\![\psi]\!] \Rightarrow [\![\psi[\alpha^*{-}\chi^*/\alpha^*{-}\chi^{**}]]\!]$ 7., 9., def. of $\Rightarrow$

(11) $[\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi^{**}]]\!]$ 10.

(12) $\forall\chi'\big[[\![\chi']\!] \subseteq [\![\chi^*]\!] \rightarrow [\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi']]\!]\big]$ 8., 11.

(13) $\forall\chi'\big[[\![\chi']\!] \subseteq [\![\chi^*]\!] \rightarrow [\![\varphi]\!] \Rightarrow [\![\varphi[\chi^*/\chi']]\!]\big]$ 4., 6., 12.

Q.E.D.

A.2.4 φ is Downward Entailing with respect to α

SHOW: If φ is Downward Entailing with respect to α^*, then $\forall\chi'\big[[\![\chi^*]\!] \subseteq [\![\chi']\!] \rightarrow [\![\varphi]\!] \Rightarrow [\![\varphi[\chi^*/\chi']]\!]\big]$

(14) φ is Downward Entailing with respect to α^* ASSUMPTION

SHOW: $\forall\chi'\big[[\![\chi^*]\!] \subseteq [\![\chi']\!] \rightarrow [\![\varphi]\!] \Rightarrow [\![\varphi[\chi^*/\chi']]\!]\big]$

(15) Let $\overline{\chi}$ be such that $[\![\overline{\chi}]\!]$ characterizes $\emptyset$; let ψ be the LF such that $\psi = \varphi[\alpha^*/\alpha^*{-}\chi^*, \chi^*/\overline{\chi}]$

(16) $[\![\alpha^*]\!] - [\![\chi^*]\!] = [\![\alpha^*{-}\chi^*]\!] - [\![\overline{\chi}]\!]$ def. of $-$

(17) $[\![\varphi]\!] = [\![\psi]\!]$ 15., 16., S.E.S

(18) ψ is Downward Entailing with respect to $\alpha^*{-}\chi^*$ 14., 17.

SHOW: $\forall\chi'\big[[\![\chi^*]\!] \subseteq [\![\chi']\!] \rightarrow [\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi']]\!]\big]$

(19) Let χ^{**} be an arbitrary expression such that $[\![\chi^*]\!] \subseteq [\![\chi^{**}]\!]$ ASSUMPTION

SHOW: $[\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi^{**}]]\!]$

(20) $[\![\alpha^*{-}\chi^{**}]\!] \subseteq [\![\alpha^*{-}\chi^*]\!]$ A.2

(21) $[\![\psi]\!] \Rightarrow [\![\psi[\alpha^*{-}\chi^*/\alpha^*{-}\chi^{**}]]\!]$ 18., 20., def. of $\Rightarrow$

(22) $[\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi^{**}]]\!]$ 21.

(23) $\forall\chi'\big[[\![\chi^*]\!] \subseteq [\![\chi']\!] \rightarrow [\![\psi]\!] \Rightarrow [\![\psi[\chi^*/\chi']]\!]\big]$ 18., 22.

(24) $\forall\chi'\big[[\![\chi^*]\!] \subseteq [\![\chi']\!] \rightarrow [\![\varphi]\!] \Rightarrow [\![\varphi[\chi^*/\chi']]\!]\big]$ 15., 17., 23.

Q.E.D.

A.3 Subtractive non-entailment

In these proofs, we use specific subtractive-determiner combinations to show a general recipe for constructing proofs of Strawson non-entailment. The subtractive equivalences underwrite their extension to the relevant other combinations.

For any g, c, w, $[\![\textbf{EXH}\ \varphi]\!]^{g,c}(w)$ yields truth just in case $[\![\varphi]\!]^{g,c}(w)$ is true, and all ψ which are alternatives to φ and not Strawson-entailed by φ are not true (i.e. false or undefined) in w. More formally:

(25) For any φ, and any admissible g, c:
$[\![\textsc{exh}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\psi\big[\psi \in \textbf{ALT}(\varphi)\ \&\ [\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\psi]\!]^{g,c} \rightarrow [\![\psi]\!]^{g,c}(w) \neq 1\big]$

We're restricting ourselves to subtractive LFs; their alternatives are given below. These facts allow us to rewrite the contribution of **EXH**.

(26) ***SUBTRACTIVE ALTERNATIVES***
For a subtractive LF φ, with an exception χ,
$\textbf{ALT}(\varphi) := \{\varphi[\chi/\chi'] \mid \mathbb{X}' \subseteq D_e\ \&\ \varphi[\chi/\chi']$ is grammatical$\}$

(27) For a subtractive LF φ, with an exception χ,
$[\![\textsc{exh}\ \varphi]\!]^{g,c}(w) = 1$ *iff* $[\![\varphi]\!]^{g,c}(w) = 1\ \&\ \forall\chi'\big[[\![\varphi]\!]^{g,c} \not\Rightarrow_{\text{ST}} [\![\varphi[\chi/\chi']]\!]^{g,c} \rightarrow [\![\varphi[\chi/\chi']]\!]^{g,c}(w) \neq 1\big]$

We know that those χ' which characterize sets $\mathbb{X}'$ that are not small cannot be Strawson entailed, but they will satisfy the consequent above. Given the presuppositions of subtractives, those χ' which characterize sets $\mathbb{X}'$ that are empty will not falsify **EXH**, either. We're interested, then, only in those χ' that are small and and non-empty. The proofs reflect this.

A.3.1 *Almost every*

(1) Let α, β, χ be arbitrary expressions, α and β of type ⟨s,et⟩, and χ of type ⟨e,t⟩ — PREMISE

(2) $\varphi = [11_s\ [[[[\text{almost}\ \chi]\ \text{every}]\ [\alpha\ pro_{11,s}]]\ [pro_{11,s}\ \beta]]]$ — PREMISE

(3) $\forall w \forall c\Big[[\![\varphi]\!]^{g,c}(w)$ is defined only if $\mathbb{X} \subseteq \mathbb{A}_w\ \&\ \mathbb{X} \neq \emptyset$; where defined:
$[\![\varphi]\!]^{g,c}(w) = 1$ *iff* $\mathbb{A}_w - \mathbb{X} \subseteq \mathbb{B}_w\ \&\ \textsc{small}_c(\mathbb{X})\Big]$ — meaning of *almost*; 2.

(4) $\neg\forall w'\big[\mathbb{A}_{w'} \subseteq \mathbb{B}_{w'}\big]$ — PREMISE

SHOW: $\forall c \forall \mathbb{X}'\Big[\big[\mathbb{X} \not\subseteq \mathbb{X}'\ \&\ \textsc{small}_c(\mathbb{X})\ \&\ \textsc{small}_c(\mathbb{X}')\ \&\ \mathbb{X} \neq \emptyset\ \&\ \mathbb{X}' \neq \emptyset\big]$
$\rightarrow \exists w'\big[[\![\varphi]\!]^{g,c}(w') = 1\ \&\ [\![\varphi[\chi/\chi']]\!]^{g,c}(w') = 0\big]\Big]$

(5) a. Let χ^{**} be an arbitrary expression such that $\mathbb{X} \not\subseteq \mathbb{X}^{**}$
b. Let $\mathbb{X}$ and $\mathbb{X}^{**}$ be non-empty
c. Let c^* be such that $\textsc{small}_{c*}(\mathbb{X})$ and $\textsc{small}_{c*}(\mathbb{X}^{**})$ — ASSUMPTIONS

SHOW: $\exists w'\big[[\![\varphi]\!]^{g,c*}(w') = 1\ \&\ [\![\varphi[\chi/\chi^{**}]]\!]^{g,c*}(w') = 0\big]$

(6) $\forall w \forall c\Big[[\![\varphi[\chi/\chi^{**}]]\!]^{g,c}(w)$ is defined only if $\mathbb{X}^{**} \subseteq \mathbb{A}_w\ \&\ \mathbb{X}^{**} \neq \emptyset$; where defined:
$[\![\varphi[\chi/\chi^{**}]]\!]^{g,c*}(w) = 1$ *iff* $\mathbb{A}_w - \mathbb{X}^{**} \subseteq \mathbb{B}_w\ \&\ \textsc{small}_c(\mathbb{X}^{**})\Big]$ — meaning of *almost*, 2.

(7) Let w^* s.t. $\mathbb{X} = \mathbb{A}_{w*} \cap \overline{\mathbb{B}}_{w*}\ \&\ \mathbb{X}^{**} \subseteq \mathbb{A}_{w*}$

(8) $\mathbb{A}_{w*} - \mathbb{X} \subseteq \mathbb{B}_{w*}$ — 7., def. of −

(9) $[\![\varphi]\!]^{g,c*}(w^*) = 1$ 3., 5., 7., 10.

(10) $\exists x\big[x \in \mathbb{X}\ \&\ x \notin \mathbb{X}^{**}\big]$, call it **a** 5., def. of $\not\subseteq$

(11) $a \in \mathbb{A}_{w*} \cap \overline{\mathbb{B}}_{w*}\ \&\ a \notin \mathbb{X}^{**}$ 7., 10.

(12) $a \in \mathbb{A}_{w*} - \mathbb{X}^{**}\ \&\ a \notin \mathbb{B}_{w*}$ 11., def of $-$

(13) $\mathbb{A}_{w*} - \mathbb{X}^{**} \not\subseteq \mathbb{B}_{w*}$ 12.

(14) $[\![\varphi[\chi/\chi^{**}]]\!]^{g,c*}(w^*) = 0$ 5., 6., 7., 13.

(15) $\exists w'\big[[\![\varphi]\!]^{g,c*}(w') = 1\ \&\ [\![\varphi[\chi/\chi^{**}]]\!]^{g,c*}(w') = 0\big]$ 9., 14.

Q.E.D.

A.3.2 *Almost some*

(1) Let α, β, χ be arbitrary expressions, α and β of type $\langle s,et\rangle$, and χ of type $\langle e,t\rangle$ PREMISE

(2) $\varphi = [11_s\ [[[[\text{almost}\ \chi]\ \text{some}]\ [\alpha\ \textit{pro}_{11,s}]]\ [\textit{pro}_{11,s}\ \beta]]]$ PREMISE

(3) $\forall w \forall c\Big[[\![\varphi]\!]^{g,c}(w)$ is defined only if $\mathbb{X} \subseteq \mathbb{A}_w\ \&\ \mathbb{X} \neq \emptyset$; where defined:
$[\![\varphi]\!]^{g,c}(w) = 1$ *iff* $\mathbb{A}_w - \mathbb{X} \cap \mathbb{B}_w \neq \emptyset\ \&\ \textsc{small}_c(\mathbb{X})\Big]$ meaning of *almost*, 2.

(4) $\neg\forall w'\big[\mathbb{A}_{w'} \subseteq \mathbb{B}_{w'}\big]$ PREMISE

SHOW: $\forall c \forall \mathbb{X}'\Big[\big[\mathbb{X}' \not\subseteq \mathbb{X}\ \&\ \textsc{small}_c(\mathbb{X})\ \&\ \textsc{small}_c(\mathbb{X}')\ \&\ \mathbb{X} \neq \emptyset\ \&\ \mathbb{X}' \neq \emptyset\big]$
$\rightarrow \exists w'\big[[\![\varphi]\!]^{g,c}(w') = 1\ \&\ [\![\varphi[\chi/\chi']]\!]^{g,c}(w') = 0\big]\Big]$

(5) a. Let χ^{**} be an arbitrary expression such that $\mathbb{X}^{**} \not\subseteq \mathbb{X}$

b. Let $\mathbb{X}$ and $\mathbb{X}^{**}$ be non-empty

c. Let c^* be such that $\textsc{small}_{c*}(\mathbb{X})$ and $\textsc{small}_{c*}(\mathbb{X}^{**})$ ASSUMPTIONS

SHOW: $\exists w'\big[[\![\varphi]\!]^{g,c*}(w') = 1\ \&\ [\![\varphi[\chi/\chi^{**}]]\!]^{g,c*}(w') = 0\big]$

(6) $\forall w \forall c\Big[[\![\varphi[\chi/\chi^{**}]]\!]^{g,c*}(w)$ is defined only if $\mathbb{X}^{**} \subseteq \mathbb{A}_w\ \&\ \mathbb{X}^{**} \neq \emptyset$; where defined:
$[\![\varphi[\chi/\chi^{**}]]\!]^{g,c}(w) = 1$ *iff* $\mathbb{A}_w - \mathbb{X}^{**} \cap \mathbb{B}_w \neq \emptyset\ \&\ \textsc{small}_c(\mathbb{X}^{**})\Big]$ meaning of *almost*, 2.

(7) Let w^* s.t. $\mathbb{X}^{**} = \mathbb{A}_{w*} \cap \mathbb{B}_{w*}\ \&\ \mathbb{X} \subseteq \mathbb{A}_{w*}$

(8) $\exists y\big[y \in \mathbb{X}^{**}\ \&\ y \notin \mathbb{X}\big]$; call it **a** 5., def. of $\not\subseteq$

(9) $a \in \mathbb{A}_{w*} \cap \mathbb{B}_{w*}\ \&\ a \in \mathbb{A}_{w*} - \mathbb{X}$ 7., 8., def. of $-$

(10) $\big(\mathbb{A}_{w*} - \mathbb{X}\big) \cap \mathbb{B}_{w*} \neq \emptyset$ 9.

(11) $[\![\varphi]\!]^{g,c}(w^*) = 1$ 3., 5., 7., 10.

(12) $\big(\mathbb{A}_{w*} - \mathbb{X}^{**}\big) \cap \mathbb{B}_{w*} = \emptyset$ 7.

(13) $[\![\varphi[\chi/\chi^{**}]]\!]^{g,c}(w^*) = 0$ 5., 6., 12.

(14) $\exists w' \left[[\![\varphi]\!]^{g,c*}(w') = 1 \ \& \ [\![\varphi[\chi/\chi^{**}]]\!]^{g,c*}(w') = 0 \right]$ 11., 13.

Q.E.D.

Bibliography

Amaral, Patrícia. 2007. *The meaning of approximative adverbs: Evidence from European Portuguese*: The Ohio State University dissertation.

Amaral, Patrícia & Fabio Del Prete. 2010. Approximating the limit: The interaction between quasi 'almost' and some temporal connectives in Italian. *Linguistics and Philosophy* 33(2). 51–115. doi:10.1007/s10988-010-9075-0.

Anderson, Curt. 2015. Numerical approximation using some. In *Proceedings of Sinn und Bedeutung*, vol. 19, 54–70.

Anscombre, Jean-Claude & Oswald Ducrot. 1983. *L'argumentation dans la langue*. Editions Mardaga.

Atlas, Jay David. 1984. Comparative adjectives and adverbials of degree: An introduction to radically radical pragmatics. *Linguistics and Philosophy* 7(4). 347–377. doi:10.1007/BF00631072.

Atlas, Jay David. 1997. Negative adverbials, prototypical negation and the De Morgan taxonomy. *Journal of Semantics* 14(4). 349–367.

Bale, Alan Clinton. 2008. A universal scale of comparison. *Linguistics and Philosophy* 31(1). 1.

Bale, Alan Clinton. 2011. Scales and comparison classes. *Natural Language Semantics* 19(2). 169–190. doi:10.1007/s11050-010-9068-0.

Bar-Lev, Moshe E. 2018. *Free Choice, Homogeneity, and Innocent Inclusion*: Hebrew University of Jerusalem dissertation.

Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence*, 241–301. Springer.

Bassi, Itai, Guillermo Del Pinal & Uli Sauerland. 2021. Presuppositional exhaustification. *Semantics and Pragmatics* 14. 11.

Beaver, David & Emiel Krahmer. 2001. A partial account of presupposition projection. *Journal of logic, language and information* 10(2). 147–182.

Błaszczak, Joanna. 2000. *Investigation into the Interaction between Indefinites and Negation in Polish*: Humboldt-Universität zu Berlin dissertation.

Brasoveanu, Adrian. 2008. Measure noun polysemy and monotonicity: Evidence from Romanian pseudopartitives. In *Proceedings of the 38th meeting of the north east linguistic society*, 139–150. GSLA-UMASS MA.

Bresnan, Joan. 1973. Syntax of the comparative clause construction in English. *Linguistic inquiry* 4(3). 275–343.

Bresnan, Joan. 1975. Comparative deletion and constraints on transformations. *University of Massachusetts Occasional Papers in Linguistics* 1(1). 4.

Brisson, Christine. 1998. *Distributivity, Maximality, and Floating Quantifiers*: Rutgers University dissertation.

Brisson, Christine. 2003. Plurals, 'all', and the Nonuniformity of Collective Predication. *Linguistics and Philosophy* 26(2). 129–184.

Buccola, Brian & Benjamin Spector. 2016. Modified numerals and maximality. *Linguistics and Philosophy* 39(3). 151–199. doi:10.1007/s10988-016-9187-2.

Büring, Daniel. 2007a. Cross-polar nomalies. In *Proceedings of SALT 17*, CLC Publications, Cornell University.

Büring, Daniel. 2007b. More or less. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, vol. 43, 3–17. Chicago Linguistic Society.

Bylinina, Lisa. 2013. Degree infinitival clauses. In *Semantics and Linguistic Theory*, vol. 23, 394–411.

Bylinina, Lisa & Rick Nouwen. 2018. On "zero" and semantic plurality. *Glossa: a journal of general linguistics* 3(1).

Bylinina, Lisa & Rick Nouwen. 2020. Numeral semantics. *Language and Linguistics Compass* 14(8). e12390.

Carlson, Gregory. 1981. Distribution of free-choice *any*. In *Papers from the Regional Meeting Chicago Linguistic Society 17* 17, 8–23.

Chen, Sherry Yong. 2018. Zero degrees: Numerosity, intensification, and negative polarity. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, vol. 54, 53–67. Chicago Linguistic Society.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. *Semantics: An international handbook of natural language meaning* 3. 2297–2332.

Cresswell, Max J. 1976. The semantics of degree. In *Montague grammar*, 261–292. Elsevier.

Crnič, Luka. 2018. A note on connected exceptives and approximatives. *Journal of Semantics* 35(4). 741–756. doi:10.1093/jos/ffy011.

Cruse, David Alan. 1986. *Lexical semantics.* Cambridge University Press.

Dayal, Veneeta. 1998. Any as inherently modal. *Linguistics and philosophy* 21(5). 433–476.

Dowty, David R. 1979. *Word Meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*, vol. 7. Dordrecht: Holland: Springer Science & Business Media.

Ducrot, Oswald. 1973. *La preuve et le dire.* Paris: Maison Mame.

Eckardt, Regine. 2007. *Almost*—A theory. Ms., University of Göttingen.

Eisner, Jason. 1994. …-less in Wonderland?-Revisiting Any. In *In ESCOL*, Citeseer.

Fiengo, Robert & Robert May. 1994. *Indices and identity.* MIT Press.

von Fintel, Kai. 1993. Exceptive constructions. *Natural Language Semantics* 1(2). 123–148.

von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148.

von Fintel, Kai. 2011. Conditionals. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of meaning*, vol. 2, 1515–1538. Berlin/Boston: De Gruyter Mouton.

von Fintel, Kai & Sabine Iatridou. 2020. Prolegomena to a theory of X-marking. *Ms. under review for Linguistics and Philosophy* .

Fleisher, Nicholas. 2016. Comparing theories of quantifiers in than clauses: Lessons from downward-entailing differentials. *Semantics and Pragmatics* 9. 4–1.

Fox, Danny. 2000. *Economy and semantic interpretation*, vol. 35. MIT press.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, 71–120. Springer.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586.

Fox, Danny & Benjamin Spector. 2009. Economy and embedded exhaustification. Talk, Cornell University.

Fox, Danny & Benjamin Spector. 2018. Economy and embedded exhaustification. *Natural Language Semantics* 26(1). 1–50.

Gajewski, Jon. 2002. L-analyticity and natural language. Ms., MIT.

Gajewski, Jon. 2008. NPI any and connected exceptive phrases. *Natural Language Semantics* 16(1). 69–110.

Gajewski, Jon. 2013. An analogy between a connected exceptive phrase and polarity items. In E. Csipak, R. Eckardt, M. Liu & M. Sailer (eds.), *Beyond 'any' and 'ever': New Explorations in Negative Polarity Sensitivity*, vol. 262, 183–212. Boston: De Gruyter.

Gotzner, Nicole & Jacopo Romoli. 2022. Meaning and Alternatives. *Annual Review of Linguistics* 8(1). 213–234.

Grice, Herbert P. 1975. Logic and conversation. In *Speech acts*, 41–58. Brill.

Hackl, Martin. 2000. *Comparative quantifiers*: Massachusetts Institute of Technology dissertation.

Heim, Irene. 1983. On the projection problem for presuppositions. *Formal semantics–the essential readings* 249–260.

Heim, Irene. 1996. Lectures on ellipsis. Lecture notes, MIT.

Heim, Irene. 2000. Degree operators and scope. In *Semantics and Linguistic Theory*, vol. 10, 40–64.

Heim, Irene. 2008. Decomposing antonyms? In *Proceedings of Sinn und Bedeutung*, vol. 12, 212–225.

Hirsch, Aron. 2016. An Unexceptional Semantics for Expressions of Exception. *University of Pennsylvania Working Papers in Linguistics* 22(1).

Hitzeman, Janet. 1992. The selectional properties and entailments of *almost*. In *Papers from the 28th Regional Meeting of the Chicago Linguistic Society*, 225–238. Chicago Linguistic Society: Chicago.

Hoeksema, Jacob. 1987. The logic of exception. In *Proceedings of ESCOL*, vol. 4, 100–113. The Ohio State University Columbus.

Hofstetter, Stefan. 2016. Measure Phrase Constructions in English, German, and French: The (Non-) Occurrence of Antonyms and Effects of Evaluativity. *Quantitative Approaches to Grammar and Grammatical Change: Perspectives from Germanic/ed. by S. Featherston, Y. Versley. Berlin: Walter de Gruyter GmbH & Co KG* 85–112.

Horn, Laurence. 1972. On the semantic properties of logical operators in English: UCLA dissertation. *Distributed by IULC* .

Horn, Laurence. 1989. *A natural history of negation.* Chicago, IL: University of Chicago Press.

Horn, Laurence R. 1996. Exclusive company: Only and the dynamics of vertical inference. *Journal of semantics* 13(1). 1–40.

Horn, Laurence R. 2002. Assertoric inertia and NPI licensing. In *Chicago Linguistic Society*, vol. 38, 55–82.

Horn, Laurence R. 2005. Airport'86 revisited: Toward a unified indefinite any. *The Partee Effect* 179–205.

Husband, Edward Matthew. 2011. Severing scale structure from the adjective. In *LSA Annual Meeting Extended Abstracts*, vol. 2, 31–1.

Ionin, Tania & Ora Matushansky. 2006. The Composition of Complex Cardinals. *Journal of Semantics* 23(4). 315–360. doi:10.1093/jos/ffl006.

Jayez, Jacques & Lucia Tovena. 2008. *Presque* and *almost*: How argumentation derives from comparative meaning. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical Issues in Syntax and Semantics (CSSP)* 7, 23.

Kadmon, Nirit & Fred Landman. 1993. *Any*. *Linguistics and Philosophy* 16(4). 353–422.

Kagan, Olga & Sascha Alexeyenko. 2011. Degree Modification in Russian Morphology: The Case of the Suffix *-ovat*. In *Proceedings of Sinn und Bedeutung*, vol. 15, 321–336.

Kamp, Hans & Antje Rossdeutscher. 1994. DRS-construction and lexically driven inference. *Theoretical Linguistics* 20(2-3). 165–236.

Karttunen, Lauri & Stanley Peters. 1979. Conventional lmplicature. In *Presupposition*, 1–56. Brill.

Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30(6). 669–690.

Keenan, Edward L. & Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9(3). 253–326.

Kennedy, Christopher. 2001. Polar opposition and the ontology of 'degrees'. *Linguistics and Philosophy* 24(1). 33–70.

Kennedy, Christopher. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1). 1–45.

Kennedy, Christopher. 2013. A scalar semantics for scalar readings of number words. *From grammar to meaning: The spontaneous logicality of language* 172. 200.

Kennedy, Christopher. 2015. A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8. 10–1.

Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 345–381.

Kilbourn-Ceron, Oriana. 2016. Embedded Exhaustification: Evidence from *Almost*. *Journal of Semantics* ffw002. doi:10.1093/jos/ffw002.

Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics and philosophy* 4(1). 1–45.

Klein, Wolfgang. 1994. *Time in language.* London: Routledge.

Kratzer, Angelika. 1986. Conditionals. Chicago Linguistics Society 22 (2). 1–15. *TWO TYPES OF DONKEY SENTENCES* 163.

Krifka, Manfred. 2007. Approximate interpretation of number words.

Križ, Manuel. 2015. *Aspects of homogeneity in the semantics of natural language*: Universiteit Wien dissertation.

Križ, Manuel. 2016. Homogeneity, Non-Maximality, and *all*. *Journal of Semantics* 33(3). 493–539. doi:10.1093/jos/ffv006.

Križ, Manuel & Benjamin Spector. 2021. Interpreting plural predication: Homogeneity and non-maximality. *Linguistics and Philosophy* 44(5). 1131–1178.

Ladusaw, William. 1979. *Polarity sensitivity as inherent scope relations*: University of Texas at Austin dissertation.

Lassiter, Daniel & Noah D. Goodman. 2013. Context, scale structure, and statistics in the interpretation of positive-form adjectives. *Semantics and Linguistic Theory* 23(0). 587–610. doi:10.3765/salt.v23i0.2658.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use and interpretation of language*, 302–323. Berlin; New York: De Gruyter.

Löbner, Sebastian. 2002. *Understanding semantics.* Routledge.

Magri, Giorgio. 2014. An Account for the Homogeneity Effect Triggered by Plural Definites and Conjunction Based on Double Strengthening. In Salvatore Pistoia Reda (ed.), *Pragmatics, Semantics and the Case of Scalar Implicatures*, 99–145. London: Palgrave Macmillan UK. doi:10.1057/9781137333285_5.

Marty, Paul P. 2017. *Implicatures in the DP domain*: Massachusetts Institute of Technology dissertation.

McCawley, James D. 1972. *Syntactic and Logical Arguments for Semantic Structures.* Indiana University Linguistics Club.

McKenzie, Andrew & Lydia Newkirk. 2019. *Almost* at-a-distance. *Linguistics and Philosophy* doi:10.1007/s10988-019-09275-6.

Mendia, Jon Ander. 2018. Epistemic numbers. *Semantics and Linguistic Theory* 28. 493. doi:10.3765/salt.v28i0.4433.

Moltmann, Friederike. 1995. Exception sentences and polyadic quantification. *Linguistics and Philosophy* 18(3). 223–280.

Morgan, Jerry. 1969. On Arguing About Semantics. *Papers in Linguistics* 1. 49–70.

Morzycki, Marcin. 2001. *Almost* and Its Kin, Across Categories. In *Semantics and Linguistic Theory*, vol. 11, 306–325.

Nouwen, Rick. 2006. Remarks on the polar orientation of almost. *Linguistics in the Netherlands* 23(1). 162–173.

Ogihara, Toshiyuki & Kiyomi Kusumoto. 2020. The Interpretation of Tense: "I Didn't Turn Off the Stove". *The Wiley Blackwell Companion to Semantics* 1–24.

Partee, Barbara. 1986. *Any*, *almost*, and superlatives (the airport squib). Unpublished Ms., reprinted in *Compositionality in formal semantics: selected papers by Barbara H. Partee*, Blackwell, 2004, 31–40.

Partee, Barbara Hall. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 70(18). 601–609.

Penka, Doris. 2006. Almost There: The Meaning of *almost*. In *Proceedings of Sinn und Bedeutung*, vol. 10, 275–286. doi:10.18148/sub/2006.v10i2.690.

Percus, Orin. 2000. Constraints on some other variables in syntax. *Natural language semantics* 8(3). 173–229.

Potts, Christopher. 2005. *The logic of conventional implicatures* 7. Oxford University Press on Demand.

Potts, Christopher. 2007. Into the conventional-implicature dimension. *Philosophy compass* 2(4). 665–679.

Prior, Arthur N. 1962. Tense-logic and the continuity of time. *Studia Logica: An International Journal for Symbolic Logic* 13. 133–151.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English language*. London; New York: Longman.

Rapp, Irene & Arnim von Stechow. 1999. *Fast* 'almost' and the visibility parameter for functional adverbs. *Journal of Semantics* 16(2). 149–204. doi:10.1093/jos/16.2.149.

Rett, Jessica. 2008. *Degree modification in natural language*: Rutgers University-Graduate School-New Brunswick dissertation.

Rett, Jessica. 2015. *The semantics of evaluativity*. OUP Oxford.

van Rooij, Robert. 2011. Vagueness and linguistics. In *Vagueness: A guide*, 123–170. Springer.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116. doi:10.1007/BF02342617.

Rotstein, Carmen & Yoad Winter. 2004. Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12(3). 259–288.

Rullmann, Hotze. 1995. *Maximality in the Semantics of* Wh-*constructions*: University of Massachusetts Amherst dissertation.

Rusiecki, Jan. 1985. *Adjectives and comparison in English: A semantic study.* Longman Publishing Group.

Sadock, Jerrold. 1981. Almost. In Peter Cole (ed.), *Radical Pragmatics*, 257–272. New York: Academic Press.

Schwarz, Bernhard. 2010. A note on *for*-phrases and derived scales. Handout, Sinn und Bedeutung 15.

Schwarzschild, Roger. 2005. Measure phrases as modifiers of adjectives. *Recherches linguistiques de Vincennes* 34. 207–228.

Schwarzschild, Roger. 2013. Degrees and segments. In *Semantics and Linguistic Theory*, vol. 23, 212–238.

Schwarzschild, Roger & Karina Wilkinson. 2002. Quantifiers in comparatives: A semantics of degree based on intervals. *Natural Language Semantics* 10(1). 1–41.

Seuren, Pieter AM. 1978. The structure and selection of positive and negative gradable adjectives. In *14th Regional Meeting, Chicago Linguistic Society*, 336–346. CLS.

Sevi, Aldo. 1998. *A Semantics for* almost *and* barely: Tel-Aviv University MA thesis.

Solt, Stephanie. 2012. Comparison to arbitrary standards. In *Proceedings of Sinn und Bedeutung*, vol. 16, 557–570.

Solt, Stephanie. 2015a. Q-Adjectives and the Semantics of Quantity. *Journal of Semantics* 32(2). 221–273. doi:10.1093/jos/fft018.

Solt, Stephanie. 2015b. Vagueness and imprecision: Empirical foundations. *Annu. Rev. Linguist.* 1(1). 107–127.

Sorensen, Roy. 2018. Vagueness. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University summer 2018 edn.

Spector, Benjamin. 2013. Bare numerals and scalar implicatures. *Language and Linguistics Compass* 7(5). 273–294.

Spector, Benjamin. 2014. Global positive polarity items and obligatory exhaustivity. *Semantics and Pragmatics* 7. doi:10.3765/sp.7.11.

von Stechow, Arnim. 1984a. Comparing semantic theories of comparison. *Journal of Semantics* 3(1-2). 1–77.

von Stechow, Arnim. 1984b. My reaction to Cresswell's, Hellan's, Hoeksema's and Seuren's comments. *Journal of Semantics* 3(1-2). 183–199.

von Stechow, Arnim. 2009. The temporal degree adjectives *früher/später* 'early(er)'/'late(r)'and the semantics of the positive. In Anastasia Giannakidou & Monika Rathert (eds.), *Quantification, Definiteness and Nominalization* Oxford Studies in Theoretical Linguistics, 214–233. Oxford: Oxford University Press.

Unger, Peter. 1978. *Ignorance: A case for scepticism.* OUP Oxford.

Van der Sandt, Rob A. 1992. Presupposition projection as anaphora resolution. *Journal of semantics* 9(4). 333–377.

Vendler, Zeno. 1967. *Linguistics in Philosophy.* Cornell University Press.

Wellwood, Alexis. 2014. *Measuring predicates*: University of Maryland, College Park dissertation.

Yoon, Youngeun. 1996. Total and partial predicates and the weak and strong interpretations. *Natural Language Semantics* 4(3). 217–236. doi:10.1007/BF00372820.

Zanuttini, Raffaella. 1991. *Syntactic properties of sentential negation: A comparative study of Romance languages*: University of Pennsylvania dissertation.

Ziegeler, Debra. 2000. What *almost* can reveal about counterfactual inferences. *Journal of Pragmatics* 32(12). 1743–1776. doi:10.1016/S0378-2166(99)00119-8.