MEDIATED MODIFICATION: FUNCTIONAL STRUCTURE AND THE INTERPRETATION OF MODIFIER POSITION

A Dissertation Presented

by

MARCIN MORZYCKI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2005

Department of Linguistics

MEDIATED MODIFICATION: FUNCTIONAL STRUCTURE AND THE INTERPRETATION OF MODIFIER POSITION

A Dissertation Presented

by

MARCIN MORZYCKI

Approved as to style and content by:

_____
Angelika Kratzer, Chair

_____
Kyle B. Johnson, Member

_____
Lisa Matthewson, Member

_____
Barbara H. Partee, Member

_____
Robert A. Rothstein, Member

_____
Elisabeth O. Selkirk, Department Head
Department of Linguistics

ACKNOWLEDGMENTS

‘Sell your ideas. They are totally acceptable.’

—Fortune cookie, received at
Teapot in Northampton, 2001

The biggest difficulty in writing the acknowledgments in a dissertation is, of course, that it is impossible. Dissertations cannot in principle be infinitely long, apparent counterexamples notwithstanding. So whatever these remarks may be, they are certainly not adequate acknowledgment of the debts I owe.

That cliché about writing a dissertation being like having a baby is false—for one thing, a dissertation involves more pain-killers, nausea, and mood swings. I’m grateful to my committee for getting me through. Angelika Kratzer, my chair, is awe-inspiring, like a benign deity or hilltop guru or mysterious force of nature, and I have no idea how to go about thanking her. (Maybe some sort of sacrifice is in order.) She knew everything about everything, understood everything, and saw everything long before I saw it. Sometimes, she said things that were like little koans—one insight wrapped tightly in another inside another still—the full meaning of which would reveal itself to me incrementally as I thought a problem through. She was kinder and more understanding and encouraging than I deserved, yet unfailingly skeptical and unflinchingly direct. And she is owed a humanitarian medal for putting up with me. Amazingly, Barbara Partee turns out

to actually be *the* Barbara Partee—she somehow lives up to the larger-than-life image. I've benefited from her perspective, encouragement, and advice. If she hasn't fully purged from me the inclination to disparage my own work, she has certainly ensured that I will always feel a little guilty doing it. If it weren't for Lisa Matthewson, I would have left grad school halfway through. When every thought I had seemed self-evidently hopeless, she sat with me for hours, patiently searching for anything to salvage. She always made it seem possible to actually *do* linguistics—and not only that, she made it seem possible to do *semantics* in particular, which is truly a feat. Kyle Johnson listens more intelligently than most people talk. He's helped me maneuver through both linguistic and professional problems, and thereby prevented me from stumbling over all sorts of obstacles and breaking the, uh, metaphorical china. And if for some bizarre reason I wanted to make sense of this rambling, ill-conceived metaphor—or of any convoluted conceptual jumble—he would be a good person to sit down and hash it over with.[1] Bob Rothstein has been an ideal outside committee member, and I wish I had taken better advantage of his Slavic expertise.

Sandy Chung is primarily responsible for originally inflicting me on linguistics, and vice versa. Her pedagogical skills border on the occult, and I strongly suspect her of being telepathic. Had she not been such a provocative and engaging teacher, I would almost surely be squandering my life in an entirely different fashion, and in some respects I'd be a different person. Somehow, she managed to be simultaneously skeptical, enthusiastic, ironically detached, and completely serious; and she's the polar opposite of dogmatic, whatever that is. She's cat-matic. My other undergrad professors—including Jim McCloskey, Donka Farkas, Bill Ladusaw, Judith Aissen, and Jaye Padgett—made discovering linguistics like a regular series of religious epiphanies.

---

[1] This sentence is grammatical despite violating the Coordinate Structure Constraint. That's for Kyle. (Alternatively, it might be Right Node Raising, but it has the wrong intonation for that.)

Since a long-ago undergrad syntax class, Meredith Landman and I have shared a common academic history, and we have similar methodological and pedagogical predispositions. She has been an indispensable release valve whenever we discovered that everyone else in the world had evidently been replaced by a completely insane but otherwise perfect duplicate. Anne-Michelle Tessier is not a phonologist. She is under the impression that she is a phonologist, but she is mistaken. Her contributions to me, intellectual and otherwise—and to other syntax-semantics people in the department—are far too great for someone who claims not to work on this stuff. Therefore, she is not a phonologist. QED. Lyn Frazier makes trying to figure out the connection between mind and grammar exactly as exciting as it should be, and I never talked to her about anything in linguistics, psycho- or otherwise, without afterwards seeing things in a different way.

I've benefited from talking to Kylito Rawlins about adverbs and other things, but really I owe him an apology—it must be endlessly annoying that I take such obvious satisfaction in a former student of mine having turned into this brilliant rock-star linguist. In my defense, it's objectively clear that having failed to botch his first encounter linguistics is and may well remain the most valuable thing I've ever done in the field. He's kind to put up with me. I would have afflicted Matt Smith similarly, had he lived, though I'm grateful to have seen his sentence generator achieve a beautiful fluency and, if I'm remembering right, the ability to control the weather. And I owe a collective debt to many of my other former students for serving as both a constant challenge and a constant reminder of how much fun linguistics actually is.

Nancy Hall, Ania Łubowicz, and Ana Arregui are particularly good at entertaining and tolerating me—neither of which I can do myself, so I'm very grateful. Sources of semantics- or syntax-flavored amusement, commiseration, liquor-consumption, despair-prevention, conversation, and intellectual stimulation also include Maria Gouskova, Ji-Yung Kim, Uri Strauss, Helen Majewski, Paula Menéndez-Benito, Shai Cohen, Luis

Alonso-Ovalle, Min-Joo Kim, Peggy Speas, Ellen Woolford, Mike Terry, Jan Anderssen, Mariko Sugahara, and Youri Zabbal, and semantics reading group people generally. Among the people responsible for creating a friendly, lively, and entertaining social and intellectual atmosphere in which to work and live have been Adam Werle, Andre Isaak, Andries Coetzee, Beata Gyuris, Elliott Moreton, Eva Juarros-Daussa, Helen Stickney, Jen Smith, Joe Pater, John Kingston, John McCarthy, Lisa Selkirk, Mako Hirotani, and others I am accidentally omitting. I'm grateful to Steve Weisler, Lee Spector, and Leni Bowen at Hampshire College for the opportunity to teach there, which was immensely fun and satisfying. For heroically rescuing me countless times, thanks to Kathy Adamczyk and Lynne Ballard. I was ridiculously lucky to have Tom Ernst—who knows more about adverbs than god—around the department in my last few years. People outside the department whose comments and suggestions I have benefited from are too numerous to mention, but, in radically abbreviated form, include Chris Kennedy, Richard Larson, Roger Schwarzschild, and audiences at SALTs XI and XII, WCCFL XX, the LSA, and elsewhere. Roselyn Tantraphol was entitled to cure me of my repellent personality with a shotgun at many points over the last few years but inexplicably chose not to. I couldn't begin to thank her adequately here.

A few days or weeks after having written this, I will slap my forehead and cringe, mortified at the realization that I have forgotten to include someone very important to whom I owe a large intellectual or personal debt. If you are such a person, consider yourself retroactively thanked—all I can say in my defense is that the depth of my gratitude far exceeds the breadth of my addled sleepless memory.

ABSTRACT

MEDIATED MODIFICATION: FUNCTIONAL STRUCTURE AND

THE INTERPRETATION OF MODIFIER POSITION

FEBRUARY 2005

MARCIN MORZYCKI, B.A., UNIVERSITY OF CALIFORNIA, SANTA CRUZ

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Angelika Kratzer

For all the progress that has been made in understanding the syntax-semantics interface, many long-familiar connections between the meaning and position of modifiers remain largely mysterious. Why should color adjectives, for example, occupy different positions than size adjectives and evaluative adjectives? Why should subject-oriented adverbs occupy different positions than manner adverbs? More generally, why should the lexical semantics of modifiers often seemingly determine their position in this intricate way?

The analytical thread unifying the independent case studies in this thesis is an attempt to address this longstanding question by attributing part of the apparent lexical semantics of certain modifiers instead to their position itself. In Chapter 1, I discern in previous research a common analytical impulse to posit grammatical devices that mediate between a particular modifier and the expression it modifies. I then propose a generalized understanding of how such mediation might work, building broadly on McConnell-Ginet (1982) and Cinque (1999), in which certain modifiers are treated as arguments of semantically-interpreted features in the functional structure of the modified expression.

This permits a kind of decomposition, in which a modifier's lexical semantics is, in the spirit of Ernst (2002), reduced to a single simple core meaning that remains constant irrespective of syntactic position, and the semantic peculiarities that vary from one position to another arise instead directly form the denotations of features. After sketching how this might provide some analytical traction on some otherwise vexing puzzles, I move on in other chapters to examine some particular not-well-understood modifiers.

Chapter 2 focuses on a class of evaluative adverbs such as *remarkably* that occur inside the adjectival projection, arguing that they are interpreted as arguments of degree morphology that imputes to them a roughly exclamative-like domain-widening interpretation. Chapter 3 develops an account of the distinguishing properties of weak DP adverbials—low scope, fixed position, and a durativity presupposition—in which they are linked to Aktionsart information in the VP and mirror the syntax of other types of measure phrase. Chapter 4 examines modifiers such as *almost* and *virtually* cross-categorially, assigning them an intensional semantics. Chapter 4 analyzes *whole* and *entire* as fundamentally non-quantificational expressions that reduce exception tolerance.

TABLE OF CONTENTS

CHAPTER 1

MEDIATED MODIFICATION

1.1 Introduction

It is an old observation and a longstanding puzzle that the distribution of modifiers, in particular their distribution relative to each other, is not arbitrary. Adverbs aren't merely sprinkled into a sentence any which way, or even in a way that reflects only obvious and straightforward scope relations. Nor are adjectives sprinkled into nominals haphazardly, in no particular order. Rather, the distribution of adverbs, adjectives, and certain other modifiers seems to track their semantics in some intricate and largely mysterious way.

Close relations between syntax and semantics are of course commonplace in the grammar. But despite much progress and theoretical development elsewhere in both syntax and semantics over many years, this area has remained uncomfortably foggy, and has not received nearly the attention one might have expected. Very few widely-held general assumptions have developed; no single overall theory of the relative order of modifiers has become standard. Even many fairly simple, concrete, and relatively specific questions (e.g., why do size adjectives occupy different syntactic positions than color adjectives do?) have no well-established satisfactory answers.

So there is a broad and difficult question here that's pressing. In this domain, lexical semantics seemingly determines syntactic position quite directly. Why should that be? Why should it be that modifiers belonging to certain semantic classes seem to occur in certain corresponding positions?

The unifying, overarching aim of this thesis is to develop some very general assumptions about the semantics and syntax of modification that might provide a consistent framework for explaining such an intimate and fine-grained connection, and to adduce evidence for these assumptions from a number of distinct, empirically diverse areas in which they find independent support. Wedding principally certain guiding intuitions underlying the approach to adverb semantics of McConnell-Ginet (1982) and the approach to modifier syntax of Cinque (1994, 1999), the core proposal around which the story centers is to attribute some of what might otherwise be understood to be the lexical semantics of various modifiers instead directly to the positions they occupy. More precisely, it will be argued that many syntactically 'small' modifiers that occupy relatively fixed positions acquire their distinguishing semantic characteristics as a consequence of having been introduced into semantic composition as arguments of licensing features with non-trivial denotations in the functional structure of the modified expression.

This can be understood as a reformulation of the observation above in reverse—on this view, it is not so much that certain modifiers occur in particular positions because they mean certain sorts of things, but rather that these modifiers mean certain sorts of things because they occur in particular positions.

Naturally, this all requires substantial elaboration and illustration to be made precise. This chapter will take some steps in that direction. What will follow in subsequent chapters is a sequence of several case studies—though 'case study' may not be quite an adequate word, since none of them will take as its sole purpose advancing this very general idea. Rather, the methodology will be to examine a few distinct phenomena in some detail, separately, and each on its own terms. Each chapter is intended to be able to stand alone as a self-contained argument (and can be read that way as well, I think). The phenomena they deal with are all, to varying extents and in varying respects, off the beaten

track of previous research in a way that I hope might provide a novel perspective on these larger longstanding questions. From each of these investigations—all focused primarily on an empirical puzzle rather than on a larger agenda—an argument for the the unifying idea should emerge independently.

As for this chapter, section 1.2 very briefly surveys some of the empirical terrain in this domain, chiefly to situate the case studies to follow with respect to the much broader empirical puzzle. Section 1.3 argues that there is a recurring analytical impulse in the literature—manifest most clearly though not exclusively in work that tries to analyze many kinds of modifiers in a uniform fashion—to suppose that the grammar of many modifiers crucially involves semantic and sometimes structural 'enrichment' operations. Section 1.4 explores ways this intuition might be distilled into a principle of the grammar. Section 1.5 concludes.

1.2 Empirical Background

On one simple view of the semantics of modification, setting aside many important complications, adverbs prototypically denote properties of events (Davidson 1967, Parsons 1990) and adjectives prototypically denote properties of individuals. This, along with a principle of composition such as predicate conjunction (the Predicate Modification rule of Heim and Kratzer 1997, say) and usual assumptions about the syntax of adjunction, yields a surprisingly general and elegant understanding of the grammar of these modifiers. But in a number of respects, it is on its own inadequate.[2]

---

[2] Framed in this radically simplified way, this view is of course essentially a caricature. But it provides a useful starting point nonetheless.

Without further elaboration, such a view predicts that all adjectives and adverbs should be able to occur in the same positions as all others and in any order relative to others with no semantic consequences, since both of these categories would be semantically and syntactically homogeneous. But this is emphatically not the case. Semantically-based restrictions on adverb and adjective position are numerous—sufficiently so that only a sampling will be mentioned here.

1.2.1 *Some Adverb Facts*

Among adverbs, one of the more essential distinctions is the one between speaker-oriented adverbs, subject-oriented adverbs, and manner adverbs.[3] Each of these types of adverbs—or more precisely, as I will argue, each type of adverb semantics—is associated with a particular position (an observation that, in the generative tradition, goes back to Jackendoff 1972). Speaker-oriented adverbs are associated with the extreme left periphery of the clause:

(1) a. Frankly, most people find Clyde irritating.
b. #Most people frankly find Clyde irritating.
c. #Most people find Clyde irritating frankly.

The relative oddness of (1b) and (1c) can in both cases be remedied with parenthetical intonation. To the extent these sentences are good without such intonation, one feels an inclination to interpret them in ways that happen to be inconsistent with what *frankly*

---

[3] The taxonomy of adverbs and the terminology surrounding it is complicated and sometimes inconsistent among authors. I will generally try to favor more familiar terms over less familiar ones.

permits—that is, one is inclined to take (1b) to attribute frankness to most people, and (1c) to involve a 'frank manner' of finding Clyde irritating. More will be said about this momentarily.

Subject-oriented adverbs (Jackendoff 1972; Bellert 1977; McConnell-Ginet 1982; Wyner 1994, 1998; Geuder 2000; Ernst 1984, 2002, among others) are associated in English with, pretheoretically, certain sentence-medial positions:

(2) a. #Clyde said that carelessly, Floyd forgot his mother's birthday.
b. Clyde said that Floyd carelessly forgot his mother's birthday.
c. #?Clyde said that Floyd forgot his mother's birthday carelessly.

Again, (2a) and (2c) are not straightforwardly ungrammatical, but certainly marked relative to (2b). Both (2a) and (2c) require either special intonation, or else a special kind of interpretation. In (2c), for example, imagining a particularly careless manner of forgetting improves the sentence significantly.

Manner adverbs are associated with relatively lower, in English more right-peripheral positions (examples from McConnell-Ginet 1982):

(3) a. #Lavishly, Josie has furnished the house.
b. #Josie lavishly has furnished the house.
c. Josie has furnished the house lavishly.

As before, other positions are marked. Again, the oddness can be alleviated by assigning the adverb an interpretation would not normally have had—for example, by taking (3b) to suggest that Josie was lavish to have furnished her house (and should presumably have been more modest and lived in an unfurnished one).

Still more fine-grained versions of these observations could in principle be made, and this is one of the major descriptive tasks undertaken by Cinque (1999) and Ernst (1984, 2002), from quite different perspectives.

1.2.2 *Some Adjective Facts*

Analogous observations could be made for adjective position (though the data require a slightly different kind of coaxing). Again, it will suffice to merely hint at the full range of facts. (Discussions of adjective position include Bolinger 1967, Sproat and Shih 1988, Valois 1991, Bernstein 1993, Cinque 1994, Laenzlinger 2000 among many others, as well as work in the typological tradition, such as Hetzron 1978 and Dixon 1982.) One essential observation in this domain, alluded to above, is that adjectives of evaluation, color, and size occur in a quite consistent relative order, in English exactly this one (evaluation $\prec$ size $\prec$ color):

(4) a. the awful big red ball
b. #the awful red big ball
c. #the big awful red ball
d. #the big red awful ball
e. #the red awful big ball
f. #the red big awful ball

In the Romance DP, where the noun raises overtly into a medial position in this order (as Valois 1991, Bernstein 1993, Cinque 1994 and Laenzlinger 2000 among others argue), the positions can be even more clearly distinguished:

(5) a. un gros ballon rouge (French;
a big balloon red
Laenzlinger 2000)

'a big red balloon'

b. *un rouge ballon gros

Even in English, though, certain possible adjective positions can be distinguished by the position of the noun and by the semantic consequences they give rise to (Bolinger 1967, Larson 1999, and references therein):

(6) a. the river navigable (Bolinger 1967)

b. the navigable river

(7) a. the stars visible (Bolinger 1967)

b. the visible stars

In these cases, there is a distinctly perceptible semantic reflex of adjective position, as Bolinger famously observed. The stars visible, for example, are at any given moment the stars not obscured by clouds or washed out by the glare of city lights. On the other hand, the visible stars could just as easily be the stars which can be seen in principle. As an initial approximation, one might take this to indicate that the postnominal adjectives have an obligatorily stage-level interpretation which the prenominal adjectives are not obligatorily subject to.

1.2.3 *Commonalities, and the Unity of the Problem*

Importantly, for current purposes, it is not the case that adverbs and adjectives simply fall into entirely disjoint classes each with a fixed syntactic position. Rather, the

situation is a bit more fluid—when misplaced, these modifiers do not trigger straightforward ungrammaticality so much as semantic consequences of one sort or another. These come in two varieties. One possible consequence, reflected in the natural response to the degraded examples in (4) and in some of the degraded examples in (1–3) with appropriate intonation on the adverb, is a kind of markedness one might associate with pragmatically unmotivated scrambling. Another and perhaps more interesting consequence, reflected in the degraded examples in (1–3) with neutral intonation, is a kind of coercion. One can easily enough imagine, for example, what it might mean to utter #*Most people find Clyde irritating frankly*—namely, that the manner in which most people find Clyde irritating is frank. But it is difficult to imagine how irritation might be frank, so this very natural interpretation is pragmatically a dead end.[4] It is all the more striking, then, that when *frankly* occurs here, we insist on a manner interpretation, even though it is counterpragmatic. Not all adverbs lead down this road, however:[5]

(8) Happily, Clyde would happily play the tuba happily.

Here, each instance of *happily* makes a distinct semantic contribution associated with its position. The leftmost instance ascribes to the speaker an attitude of happiness about the proposition expressed by the sentence; the medial instance ascribes to Clyde a cheerful willingness to play the tuba; and the rightmost instance contributes that the manner of tuba-playing involved is happy. The relationship among these semantic effects is certainly not an accident. Here, in each instance it is the position *happily* occupies that determines what interpretation it will receive.

---

[4] Manner readings of *frankly* are perfectly possible in more sensible contexts, as in *He spoke frankly*.

[5] This is built around an example due to Jackendoff (1972).

This effect—like the 'unmotivated scrambling' effect—occurs with adjectives as well. In (9), *intellectual*, for example, receives distinct readings obligatorily, as does *social* in (10):[6]

(9) a. a Texan intellectual midget
b. an intellectual Texan midget

(10) a. a religious social masochist
b. a social religious masochist

A Texan political figure of limited intellectual resources can truthfully be described with (9a), but certainly not with (9b); a particularly short and bookish Texan, on the other hand, could be described with (9b), but not with (9a). Similarly, (10a) can truthfully be predicated of a devout person who enjoys social situations she finds painful, while (10b) can truthfully be predicated of a particularly sociable person who enjoys pain imposed by one's religion. As in the *happily* example in (8), two instances of the same modifier in different positions receive different interpretations:

(11) a. an (unnervingly) intellectual intellectual fraud
b. an athletic athletic mascot

---

[6] Much more could be said about contrasts like (7) and (8)—a bit more is, in fact, in 1.4.4—and they have to my knowledge received particularly little attention from the perspective of formal semantics. One important observation about the lower (pretheoretically, immediately-prenominal) readings of e.g. *intellectual* here is that they bear a close relationship to domain adverbs, both semantically and syntactically (domain adverbs include *politically* in e.g. *Politically, this was a serious mistake*; Ernst 2002, Rawlins 2003, and, with slightly different terminology, Bartsch 1976, Moltmann 1997). Some aspects of this connection, along with some relevant cross-linguistic observations, are explored in Rawlins (2003). In the adjective literature, adjectives of this general sort, or ones very much like them, have sometimes been termed simply 'relational' or 'classificatory' (e.g., *musical* in *musical comedy*; Bosque and Picallo 1996, McNally and Boleda Torrent 2003, Cinque 2003). To what extent these involve the same phenomenon is, I think, an interesting and apparently largely neglected question, as is the clear connection these adjectives manifest syntactically and semantically to domain adverbs.

Each instance of the adjective makes a distinct contribution—in (11a), for example, the leftmost instance of *intellectual* characterizes an individual as cerebral, while the rightmost instance indicates that it is with respect to intellectual pursuits that an individual is a fraud.

Collectively, then, these facts show that each position is not so much the natural home for a particular kind of *modifier* as for a particular kind of *interpretation*. None of this would be expected if all adverbs and adjectives in all positions were interpreted uniformly, as the simple view sketched above would suggest.

Importantly, adverbs and adjectives seem to parallel each other with respect to these sorts of characteristics—both are sensitive to similarly semantically-based ordering restrictions in a similar way, and for both of them, the results of re-ordering are similar. This is not to suggest that there aren't significant differences in the syntax and semantics of these categories, of course—there are—but the similarities in their sensitivity to syntactic position are great. Given the many other profound similarities between adjectives and adverbs (many languages don't even distinguish them, after all), it seems that there is a good case for regarding the adjectival and adverbial manifestations of these ordering effects as different aspects of the same larger puzzle, and that the burden of proof should rest with any analysis that treats them as unrelated. This is not a trivial point, since exactly such a fundamental distinction is often made implicitly without argument, and despite the similarities, the two are discussed in unmotivated isolation from each other.

These examples, of course, represent just a diverse sampling of such effects, but they will suffice to frame the subsequent discussion. None of the case studies in the chapters to follow focus directly on any of these examples. The hope, however, is that the phenomena that are examined there will—precisely because they are (even) less well-studied and not normally linked to data like this—provide a new window onto these issues.

1.3 Modification with 'Enrichments': A Recurring Impulse

Around 20 years ago, a kind of promissory note seems to have been issued regarding the syntax-semantics of modifiers. With the elimination of particularized phrase structure rules (as in Stowell 1981 and Chomsky 1981), the distribution of elements of phrase structure was to be assured instead by independent principles. For arguments, these principles were made fairly explicit. For modifiers, though, there was no such clarity—only the understanding that an explanation would follow from semantic principles of some kind. This was in this respect an odd kind of promissory note, inasmuch as it was issued by syntacticians and payable by semanticists.

For the most part, it remains unpaid. To varying extents, most of the puzzles noted in the previous section—and others that could have been mentioned in a similar vein—have been the focus of investigation on their own, but attempts to construe these together, with an eye toward the bigger picture, are relatively few.

In one respect, this is perhaps unsurprising. Clearly, a diverse range of effects are involved here, and it seems plausible—indeed, very likely—that some of them might genuinely be fundamentally different in kind from the others, and in need of an entirely separate account. It seems rather less likely, though, that all of them are completely unrelated to each other, and that nothing general needs to be said at all.

In this section, after some further prefatory remarks, I'll consider some of the few fairly big-picture attempts to grapple with such phenomena. In all of these, I will suggest, there is a recurring analytical strategy of supposing that the grammar of such modifiers involves either supplementing their semantics beyond what a simple intersective interpretation would provide (or even a straightforward operator interpretation,[7] for that matter),

---

[7] For the sake of (somewhat greater) explicitness: by 'operator' I mean a function that applies to (the

or supplementing their syntax beyond simple adjunction, or both. In each account, what this will amount to is a claim that the relationship between a modifier and the expression it modifies is not direct, but rather mediated by some additional grammatical mechanism.[8]

1.3.1 *Modifier Position and Ontological Distinctions*

The guiding intuition behind the notion that modifier order should 'follow from the semantics' seems to be, in large part, that different kinds of modifiers occur in different positions because they modify different semantic objects. One prototypical example of this is the difference between sentence adverbs like *allegedly*, which modify propositions, and a VP-adverbs like *softly*, which modify properties of events.[9] Purely because sentences denote propositions and VPs denote properties of events, *allegedly* will attach to sentences and hence higher than *softly*, which will attach to VPs. This is an ontological approach to explaining modifier order, in the sense that for every distinct kind of modifier position, this sort of explanation requires finding a corresponding ontological distinction in what the modified expression denotes. Ernst (2002) calls this the 'Scope Theory' of adverb licensing.

This kind of reasoning is of course very elegant and (largely) uncontroversial, motivated as it is by virtually unavoidable semantic considerations. But despite its elegance,

---

denotation of) a modified expression and yields something of the same semantic type as the modified expression itself (i.e., an operator modifier denotation is of type $\langle\alpha,\alpha\rangle$, where α is the type of the modified expression). In most cases, this will amount to meaning exactly what is normally meant by 'predicate modifier'—I'm not using this otherwise preferable term only because it can easily be taken to exclude any modifiers that apply to propositions.

[8] Since I am not undertaking a review of the literature for its own sake, but rather constructing an argument for a particular view of it with respect to certain empirical questions, I will make no attempt to be comprehensive in discussing potentially relevant work. Some specific, more narrowly focused discussion of this sort takes place in subsequent chapters, and some older foundational work (e.g. Davidson 1967, Thomason and Stalnaker 1973, Jackendoff 1972, 1977) is in various ways a presence throughout.

[9] Of course this reasoning works similarly irrespective of what one supposes VPs denote.

and despite the apparently widely-held feeling that it is the most natural or best way to account for modifier position, there have been surprisingly few attempts to apply it systematically to a wide range of modifier data and evaluate or defend it as a theory. Rather, it is usually a tacit background assumption, appealed to by semanticists to distinguish various kinds of semantic objects using modifiers in much the same way that syntacticians tacitly assume that adverbs don't move and use them to distinguish landing sites for verb movement (Emonds 1978b and its descendants, such as Pollock 1989).

This is despite some clear shortcomings. One of these is that it is not obvious that it has been particularly fruitful in explaining phenomena like those discussed in section 1.2. They have not in general proved to simply fall out as predictions from a sufficiently articulated understanding of the semantics of the modified constituents. At some point, they may, of course; but simply delaying any explanation of these problems until that moment does not seem an especially satisfying or promising methodological strategy.

Another significant shortcoming of such an approach is that that with respect to many regularities in modifier position, it seems virtually a dead end. For example, it seems very hard indeed to imagine what independent motivation there might be for drawing an ontological distinction between nominal expressions modified by shape adjectives and ones modified by size adjectives—yet on this approach, exactly such a distinction would be required to explain why size and shape adjectives occupy distinct syntactic positions.

More generally, the spirit of such an approach is to require proliferating ontological distinctions in exact proportion to distinctions in syntactic distribution.[10] This is a great and rather precarious commitment. Surely, such proliferation should come at some significant theoretical cost. And absent independent motivation for such ontological distinctions,

---

[10] Of course, purely type-theoretical distinctions (such as $\langle e,t\rangle$ vs. $\langle e,\langle e,t\rangle\rangle$, say) exist independently, so in that sense they can't in principle be said to be 'proliferated'. But types *with some linguistic importance* can be proliferated.

this cost should presumably be higher than the cost of introducing purely grammatical theoretical distinctions instead.

It seems reasonable, then, to conclude that while this sort of explanation has a very important role to play, it cannot be the *only* means available for explaining these sorts of distinctions. Something else is needed. Interestingly, the most sustained, fully-developed, and analytically ambitious attempt to apply such a strategy to a wide range of facts—Ernst (2002)—implicitly settles on this conclusion too, though in a sense that requires some elaboration.

1.3.2 *Ernst (2002)*

Ernst (2002), building on some of his previous work, examines an extremely wide range of adverbs, and proposes a unified theory of their distribution. One of the major theses of the book is that adverb distribution should in general not be handled directly through syntactic means, but rather should follow principally from semantic considerations. Ernst is pragmatic about this—he doesn't claim that no other explanation could ever be necessary, and indeed explicitly proposes certain syntactic principles, like linearization rules ('Directionality Principles' and 'Weight Theory'). Still, the core aim is to develop and defend a theory of adverb position firmly rooted in an approach based on ontological distinctions. Interestingly, though, what is actually developed is a hybrid theory in which a crucial role is played by mostly unacknowledged semantic enrichment operations that intervene between the denotations of modifiers and those of the expressions they modify.

There are two respects in which this is so. One of them has to do with the assumptions about semantic composition Ernst adopts. His semantic representations are cast in a broadly DRT-inspired style (Kamp 1981, Kamp and Reyle 1993), and he inherits from DRT

the notion of 'construction rules' that relate syntactic structures to semantic representations. In DRT, construction rules are usually devoted to governing how to organize the pieces of a discourse representation structure to reflect the meaning of a particular linguistic expression. This fundamentally compositional role would seem to differ in a fundamental way from the principal role of a content word like a noun, which contributes a predicate which a construction rule can manipulate in building a DRS, but does not include (idiosyncratic) compositional instructions.

Ernst, however, proposes a rule for interpreting manner adverbs that folds together both sorts of tasks. His 'Manner Rule' does several things apart from building up a semantic representation out of a syntactic structure. In Ernst's notation, it looks like this:[11]

(12) MANNER RULE:

A predicational adverb within PredP, selecting an Event [F(x, . . .) . . .] denoted by its sister, may yield

$[_{E'}\ [_{E}\ F(e)\ \&\ \theta\ (e, x), \ldots]\ \&\ P_{ADJ}([_{E}\ F(e)\ \&\ \theta\ (e, x), \ldots], \ldots], x)]$,

where the designated relation in P_{ADJ} is $[_{REL}$ manifests], and (if P_{ADJ} maps Fact-Event Objects to a scale) the comparison class for P_{ADJ} is all events of x F-ing.

I won't spell out in full what all this means,[12] focusing instead on only the portions of this rule whose role is not essentially compositional: this rule bleaches the adverb of a part of its lexical semantics Ernst calls the 'designated relation'; it supplies a new relation

---

11 This the version of the rule from his chapter two; he makes some small revisions subsequently.

12 Though more on this notation follows below.

which Ernst calls *manifest* to take its place; and, if the adverb is scalar, it determines the comparison class with respect to which the adverb is evaluated. This, then, involves significant extrinsic additions to the semantics of the expression. In order to interpret a manner adverb, Ernst is suggesting, something is necessary beyond a purely compositional rule. What is needed is a rule that contributes a new predicate to the semantics, a predicate that does not correspond to any syntactic constituent (or even any part of any syntactic constituent) of the expression being interpreted. In this respect, he is proposing a kind of constructional meaning, in the sense of Kay and Fillmore (1999), Goldberg (1995) and other work in Construction Grammar and its relatives—for him, what one might call the 'manner modification construction' contributes a meaning of its own independent of any of the lexical items that enter into it. It is a syntactic configuration with *inherent* meaning. Importantly too, this additional 'constructional' meaning does not follow from anything about the ontology he assumes. The rule is formulated to manipulate events, and there is nothing in the nature of an event that would by itself lead us to expect any of these semantic effects. Nor is there any necessary relationship among these semantic effects, so if any of them could be derived from something else, it is not at all clear that any of the others could be made to follow. This, then, is is an approach committed to interpreting manner adverbs with special and substantive additional semantic mechanisms—certainly, it couldn't be characterized as involving either an intersective interpretation or a straightforward operator one.

There is another, more subtle respect in which the theory Ernst develops involves 'adding something' to the interpretation of adverbs. It concerns the interaction of the Manner Rule, subject-oriented interpretations, and adverb lexical semantics. One of the differences among adverb classes for Ernst is that their lexical entries reflect different 'interpretation schemata' or 'templates'. For example, for two kinds of subject-oriented

adverbs—agent-oriented (e.g. *cleverly, stupidly, secretly*) and mental-attitude (e.g. *reluctantly, eagerly, frantically*)—he proposes lexical entries that accord with (13) and (14) respectively, and for evidential adverbs (e.g. *clearly, obviously*) he proposes ones that accord with (15):[13]

(13) AGENT-ORIENTED ADVERBS:

$\text{ADV}\ [_{E}\ldots] \rightarrow [_{E'}\ [_{E}\ldots]\ldots \&\ P_{ADJ}([_{E}\ldots], \text{Agent})]$,

where the designated relation in P_{ADJ} between the event and the Agent is $[_{REL}$ warrants positing], and the comparison class for P_{ADJ} is all relevant events in context.

(14) MENTAL-ATTITUDE ADVERBS:

$\text{ADV}\ [_{E}\ldots] \rightarrow [_{E'}\ [_{E}\ldots]\ldots \&\ P_{ADJ}([_{E}\ldots], \text{Experiencer})]$,

where the designated relation between e and Experiencer is

a. $[_{REL}$ is accompanied by] (for state mental-attitude adverbs)

b. $[_{REL}$ is intended with] (for intentional mental-attitude adverbs), and the comparison class for P_{ADJ} is Experiencers,

(15) EVIDENTIAL ADVERBS:

The perception of $\left\{\begin{array}{l}\text{(the true) Proposition } [_{REL}\text{'s truth shows }]\\ \text{the Theme of Event } [_{REL}\text{ manifests }]\end{array}\right\}$ a greater degree of P_{ADJ} than the norm for $\left\{\begin{array}{l}\text{Proposition}\\ \text{Event}\end{array}\right\}$.

---

13 Ernst never spells out the full lexical entry of an evidential adverb in a way that includes the sort of combinatorial information that (13) and (14) include, presumably because in these cases the combinatorics are straightforward. Elsewhere, he indicates that *clearly* can (on a non-manner reading) yield semantic representations like (i):

(i) $[_{p}\ldots] \&\ \text{CLEAR}(p)$

In all of these, P_{ADJ} represents 'the property designated by the adjective' which is the counterpart of the adverb,[14] and $[_{E}\ \ldots\ ]$ and $[_{E'}\ \ldots\ ]$ represent events. The expressions bracketed by '$[_{REL}\ \ldots\ ]$' are the 'designated relations' which the Manner Rule above can remove from the interpretation these adverbs. As (13) mentions, these designated relations are thought of as being embedded somewhere inside the lexical semantics of the adjective denotation P_{ADJ}. What this amounts to for an adjective like *rudely*, Ernst explains, is an interpretation in which an event '$[_{REL}$ warrants positing] more rudeness in Agent than the norm for events' (except on a manner reading, in which the Manner Rule comes into play and yields something else).

It is not altogether clear precisely what status these templates have in the theory, but the most straight-forward interpretation is to suppose that they are just expository devices rather than true rules of the grammar. If that is the case, though, they leave something pretty significant unexplained: why should it be that all these adverb classes should have so much of their lexical entries precisely in common? That is, if such templates can be extracted from the lexical entries of large classes of adverbs, does that not itself reflect a generalization about these adverbs sorely in need of capturing?

To address this, one might alternatively take templates like this to be rules of the grammar after all. But what sort of rules? One option is to take them to be lexical redundancy rules of some kind—perhaps this is actually closer to what Ernst intends. Alternatively, one might take them to be rules of compositional semantics of some kind—this is not what Ernst intends, but certainly a possibility.

Considering the first possibility first, the difficulties with construing these as lexical redundancy rules are severe. The signature trait of a fundamentally lexical generalization

---

[14] Apparently, 'property' here is understood to mean, potentially, a relation between an agent and an event, such as P_{ADJ} in in (13).

is not being completely exceptionless, or at least tolerating a small number of exceptions in principle. But these rules are not like that. It is not, for example, merely an overwhelming *tendency* for certain adverbs to target the agents or experiencers rather than, say, the instrument or goal theta roles. Rather, it appears to be an absolute, systematic regularity. Nor is it the case that merely *most* agent-oriented adverbs are paraphrasable in roughly the 'warrants positing' way Ernst suggests.[15] Rather, it appears to be all of them. If these rules were lexical, we would expect that the kind of alternations these rules are designed to reflect should not be perfectly regular. Just as *rudely* has an agent-oriented reading in medial positions, for example, we should expect to find some fleetingly few adverbs that have a manner reading in the same positions *rudely* does, but that in medial positions have a reading that varies in some virtually arbitrary other way. Nothing, on this view, would rule out an adverb *blargly* that means, say, 'loudly' on its manner reading but on a higher reading means something like 'the event warrants positing *lack* of loudness in the agent' or 'the event might lead a cynical person to suspect the agent of loudness' or even 'the event does not warrant positing loudness in any of the agent's first cousins'.

More problematic still, if these rules were construed as lexical, they would leave unexplained why these generalizations are relatively invariant across many languages. The generalizations that underlie (13–15) are for the most part not facts about English—they are facts about natural language. For a lexical rule to hold exceptionlessly in any one language would be extremely odd, but for a lexical rule to hold exceptionlessly in languages generally would be a spectacularly unlikely coincidence.

None of these difficulties arise if (13–15) are construed instead as rules of semantic composition of some kind, perhaps like the Manner Rule. Rules of semantic composition,

---

[15] That is, setting aside the issue of how good a paraphrase this is and considering only the meaning it intends to reflect.

irrespective of their particular role in a particular theory, differ fundamentally from lexical rules in that they are exceptionless. In this respect, unlike lexical rules, they have exactly the property that is required here. Because of this, the complete regularity of the phenomena that motivate these rules would be expected. Moreover, it is not problematic in the slightest for rules of semantic composition to hold across languages—indeed, it a highly desirable and perhaps expected and necessary (Matthewson 1996, 2001) characteristic of such rules.

Construing (13–15) this way would also relieve a certain conceptual difficulty surrounding the notion of 'designated relation'. If these rules are fundamentally lexical, what the Manner Rule does in bleaching an adverb of its designated relation and replacing it with another one amounts to looking inside of the lexical semantics of an expression, and removing part of its meaning. This runs counter to the idea of what the lexical semantics of a word is. The lexical semantics of an expression, as normally conceived, is invariant; the principal thing that remains invariant in this system is P_{ADJ} minus its designated relation.[16] If a word's basic lexical semantics were instead the sort of thing that could change with its syntactic position and be altered by syntactic rules, we might, for example, expect there to be a word that means 'dog' when it's a passivized subject and 'cat' in all other positions.[17] But this seems an alarming and highly unfamiliar state of affairs. And if there were a word of this sort, one would probably not be inclined to say that either 'dog' or 'cat' is part of its lexical semantics. Fundamentally, the difficulty is that compositional semantics is normally understood to assemble bits of meaning, but not to simply throw certain bits of

---

[16] This is actually particularly vexing in light of the larger agenda Ernst advances of deriving the various readings of an adverb from a single invariant lexical semantics. If the lexical semantics of the adverb can be subtracted from and otherwise altered in the syntax, this very appealing hypothesis is drained of much of its force.

[17] This example actually is perhaps a bit unfair, in that 'dog' and 'cat' involve an intuitively content-word kind of meaning, whereas what the Manner Rule manipulates is an intuitively function-word kind of meaning, something a bit more subtle, elusive, and understood to relate closely to syntactic position.

meaning away.[18] If (13–15) are not lexical, however, this problem wouldn't arise, since on this view the designated relation would either not be part of the lexical semantics of the adverb, or else perhaps the conditions on the designated relation could be understood instead as presuppositions.

Of course, when (13–15) are construed as some kind of rules of compositional semantics, they—like the Manner Rule—are also rules that intervene between adverbs and the expressions they modify, adding some significant additional semantics.

Ernst does not discuss any of adjectival data of the sort briefly mentioned in 1.2, and translating his proposals to the adjectival domain would not be a trivial task. It seems likely, though, that because this data involves additional kinds of semantic relationships between modifier and modified, this would involve making use of further semantic enrichment operations. Thus even though Ernst argues in favor of attributing adverb distribution largely to the distribution of various semantic objects in the clause, his proposals nonetheless implicitly rely also on the strategy of appealing to this sort of semantic enrichment.

1.3.3 *McConnell-Ginet: Adverbs and 'Augmentation' of Verbs*

One early influential discussion of the formal semantics of adverbs that pays particular attention to correlations between adverb position and interpretation is McConnell-Ginet (1982). Hers is also perhaps the most explicit instance of this analytical strategy.

McConnell-Ginet proposes a view of adverbial modification that departs in a fundamental way from the more familiar view that adverbs are interpreted either intersectively

---

[18] A possible counterargument to this may center on the presence or absence of vacuous quantification in natural language (Potts 2002 and references there).

or as higher-type operators. After noting some of the adverb position-interpretation correlations above, she argues against treating manner adverbs as operators (as Thomason and Stalnaker 1973 do) by observing that even apart from any immediate empirical puzzle, there is in one respect something suspiciously aesthetically unsatisfying about such a view. A common intuition about adverbs, she points out, is that they 'add something' or, in Jackendoff (1972)'s phrase, 'attach semantic markers' to a verbal meaning. But on an operator approach, it is not entirely clear in what sense this is the case. (Indeed, an adverb with a operator denotation could just as easily have the effect of 'subtracting' something—that is, of rendering the claim made in the sentence in which it occurs weaker than if it hadn't been present.) This intuition is perhaps satisfied a bit better on a view in which adverbs denote properties of events, but of course on its own this wouldn't naturally explain the sorts of contrasts she is interested in.

McConnell-Ginet devises another means of reconciling the analysis with this intuition: she proposes to turn the situation on its head and treat adverbs as semantic arguments of verbs. Of course, she does not mean by this merely the trivial claim that verbal denotations should be type-lifted to take operators as arguments. Rather, she suggests that adverbs are arguments in essentially the same sense that direct objects are arguments. In *walk quickly*, for example, she suggests that the denotation of *walk* is 'augmented' in a particular way to take as an argument the denotation of *quickly*, which can now have a simpler denotation than it otherwise might (for McConnell-Ginet, a property of a 'rate').

It will suffice here to only briefly sketch out the formal implementation of this. The idea is that for any combination of a verb and an adverb, there is a unique augmentation of the verb's meaning determined by the adverb, and that it must satisfy two requirements. One of them is that the augmented denotation must have exactly one more argument place than the unaugmented denotation did. The other is that the augmented denotation must be

such that, to put things informally, that the additional argument be droppable[19] in the sense that deleting it could not render a sentence false. So, taking great liberties with McConnell-Ginet's formalism for convenience of exposition,[20] the first step in interpreting *run quickly* is to augment the denotation of *run* in the appropriate way, that is, in the unique way *quickly* determines. Since *quickly* characterizes a rate of speed, the augmented denotation $[\![\textit{run}]\!]'$ is just like $[\![\textit{run}]\!]$, except that it takes one additional argument, a rate of speed:

(16) a. $[\![\textit{run}]\!] = \lambda x \,.\, x \text{ runs}$

b. $[\![\textit{run}]\!]' = \lambda r\text{: r is a rate of speed } \lambda x \,.\, x \text{ runs at r}$

The denotation of *quickly*, then, can be merely a property of rates:

(17) $[\![\textit{quickly}]\!] = \lambda r\text{: r is a rate of speed} \,.\, r \text{ is quick}$

These can be composed by a rule that existentially quantifies over the rate argument introduced by augmentation and predicates $[\![\textit{quickly}]\!]$ of it:

(18) $[\![\textit{run quickly}]\!] = \lambda x \, \exists r\text{: r is a rate of speed } [r \text{ is quick} \wedge x \text{ runs at r}]$

Thus *run quickly* means what perhaps its most natural paraphrase does: *run at quick rate of speed*.

---

19 This is roughly the sense of 'droppable' in Wyner (1994). It is normally used of adverb(ial)s, though for obvious reasons it makes sense to extend its use to adverb-related arguments in McConnell-Ginet's framework. (A little more precisely, an adverb is droppable iff a sentence containing it entails an otherwise identical sentence in which it is absent. Thus *badly* in *Clyde cooks badly* is droppable because *Clyde cooks badly* entails *Clyde cooks*.)

20 The chief liberty being taken here is to recast her proposal in the Heim and Kratzer (1997) style, which makes it better accord with the notation and assumptions in other parts of the thesis. The only difference that goes beyond notation is that a presupposition is added to the effect that the additional argument is a rate—this is not present explicitly in McConnell-Ginet, but seems a natural move, makes things perhaps a bit more transparent, and hopefully does her proposal no violence.

This is a very satisfying view, and accords elegantly with most people's pre-theoretical intuitions about how things work. It is also a straightforward instance of taking the relationship between an adverb and the expression it modifies to be indirect, mediated by something else—in this case, a semantic augmentation operation.

This, however, is only the starting point for a theory of adverb position. Because McConnell-Ginet defines augmentation to apply only to Vs (and not, say, VPs) and to add arguments outermost, adverbs introduced by this means will necessarily occur low, in the immediate vicinity of the verb. To account for higher, subject-oriented uses of these adverbs, she proposes that these are interpreted through the same process of augmentation, except that what they augment is an abstract higher verb, which she suggests means something like 'act'. This choice is based her observation that e.g. *Louisa departed rudely* can be paraphrased 'Luisa acted rudely to depart'.[21] To derive this result, she supposes that adverbs that apply to VP will be interpreted in a way that inserts this 'act' predicate in the semantics. In discussing such adverbs in passives, though, she suggests (foreshadowing Wyner 1998) that an agentive form of the passive auxiliary spells out this predicate overtly. Again, taking the same notational liberties as before, *reluctantly was instructed* would be interpreted by first augmenting the agentive passive auxiliary in (19a) with an additional manner argument place, yielding (19b):

(19) a. $[\![was_{passive}]\!] = \lambda P\ \lambda x\ .\ x$ acted to be P-ed

b. $[\![was_{passive}]\!]' = \lambda m$: m is a manner $\lambda P\ \lambda x\ .\ x$ acted in m to be P-ed

The denotation of *reluctantly* will be a property of manners:

---

[21] I actually find this paraphrase a bit awkward, and would prefer something like 'Luisa acted rudely in departing'.

(20) ⟦ *reluctantly* ⟧ = λm: m is a manner . m is reluctant

Combining the pieces by the same principles as before results in the denotation in (21):

(21) ⟦ *reluctantly was instructed* ⟧ = λx $\exists$m: m is a manner [m is reluctant $\wedge$ x acted in m to be instructed]

This is the desired result. Had *reluctantly* been lower (*was reluctantly instructed*) and interpreted by augmenting the *instructed* instead, the result would have been as in (22):

(22) ⟦ *was reluctantly instructed* ⟧ = λx $\exists$m: m is a manner [m is reluctant $\wedge$ x acted to be [instructed in m]]

This seems to reflect the essential contrast.

So, the explanation of the correlation between position and interpretation offered here hinges on which of two distinct verbs—one of them either abstract or the passive auxiliary—is targeted for augmentation. To extend the account to adjectives, or to more fine-grained generalizations about adverbs, the most straight-forward move would be to suppose that other nouns and verbs—potentially abstract or else expressed syntactically by some functional element[22]—can similarly be augmented. Given the great number of subtle, fine-grained correlations that exist, though, one might have reservations about the extent to which this reasoning could be applied more widely. (This concern is discussed further in section 1.3.4.)

---

[22] Of course, 'noun' and 'verb' here have to be taken in a very syntactically loose sense that also includes nominal and verbal functional heads.

A natural alternative is to suppose that at least in some instances, the process of augmentation is itself subject to ordering restrictions. Stating things in old-fashioned rule-ordering terms, perhaps augmentation with respect to some dimensions (maybe manner, for example) obligatorily precedes augmentation with respect to other dimensions (duration, say). But cast in this very rule-ordered, procedural way, this is an idea born many years too late and with no congenial home in most current conceptions of grammar. This is not just an issue of fashion, of course—there genuinely does not seem to be an independent need for this kind of serial rule-ordering in semantics.

Setting these questions of implementation aside for the moment, it seems quite clear that McConnell-Ginet's proposal very directly and explicitly manifests the strategy of appealing to operations that supplement modifier semantics and mediate the relationship with the modified expression.

1.3.4 *The Cinquean Approach: Modifiers and Functional Heads*

Cinque (1994, 1999) offers an intensely syntactic perspective on modifier interpretation-position correlations. He proposes an account of adverb and adjective ordering restrictions rooted in the idea that for each position in which an adjective or adverb can occur, there is a functional head whose specifier position it occupies. In Cinque (1999), he surveys a large range of cross-linguistic data to arrive at independent and corresponding generalizations about the relative order of adverbs and the relative order of overt functional heads. Given the assumption that adverbs occupy unique specifier positions to functional heads, and further assuming a fixed universal hierarchy of such heads, this cross-linguistic result is expected. Such an approach has been pursued, and further articulated in various ways, by others as well (including Alexiadou 1997, which is a particularly sustained effort

of this sort, and Laenzlinger 2000, who further develops this approach for adjectives). Since the mechanism bonding modifiers and functional heads together in this sort of theory is syntactic features, Ernst calls this general approach the 'Feature Theory' of modifier licensing.

Importantly, on the Cinquean approach the explanation of modifier distribution is entirely a syntactic matter. Adverbs are clearly divided into many syntactically distinct classes in the lexicon, each associated uniquely by syntactic principles with a unique functional projection. A central virtue of this approach is that it can in principle account for a large number of ordering restrictions, just as there may be a large number of possible functional heads. (And on Cinque's view, famously, there is.) But there is an downside to using these purely syntactic tools.

It is on this view purely an accident that there should be a systematic correspondence between the lexical semantics of a class of adjectives or adverbs and the functional projection with which it is associated. If, as Cinque suggests, it is syntactic features such as [+HABITUAL] or [–NECESSARY] that encode which functional projection each modifier class is associated with, there is no reason why all these syntactic features should be ones with names that seem at least superficially so semantic. We might expect that there should be classes of adverbs or adjectives distinguishable by syntactic position with nothing at all in common semantically. For example, we might expect that a particular position be restricted to the adjectives *red*, *huge*, and *Spanish*, to the exclusion of *blue*, *enormous*, and *Portuguese*. Put another way, if these features have no truth-conditional semantic import, it ought to be the case that they could just as easily have had names like [+ONE] and [–TWO] and so on. Yet this is not the case. Fundamentally, taking so purely syntactic an approach leaves the connection between syntax and semantics at the heart of the issue unexplained, or at best quite tenuous and uncertain. What it seems to miss is that each position is associated with a particular *interpretation*, not merely with an arbitrary lexical class of modifiers.

Even so, it does provide syntactic equipment powerful enough to be useful in the analysis of a wide range of facts.

Of course, one might also worry about the wisdom of proliferating as many functional heads as would be necessary to reflect the full range of ordering restrictions Cinque notes. On the face of it, positing numerous functional heads undetectable in any given language except by the purely syntactic purpose for which they were proposed may be understood to run afoul of the impulse toward a syntax geared to feeding phonological and semantic interfaces economically. But since there are apparently no other accounts of comparable empirical coverage—especially with respect to adjective order generalizations—it seems reasonable to regard avoidance of aesthetic uneasiness as a less important concern than descriptive adequacy.

Notably, both McConnell-Ginet and Cinque share an intuition that adjectives and adverbs are in some sense argument-like. For Cinque, this is interpreted in a purely syntactic sense, through rejection of an adjunction structure in favor of a more subject-like specifier structure. For McConnell-Ginet, this is interpreted in a purely semantic sense, through rejection of predicate-modifier denotations in favor of denotations in which these modifiers are literally logical arguments.

And in a straightforward way, the Cinquean approach, like McConnell-Ginet's, exemplifies the analytical impulse toward positing additional mechanisms that mediate between modifier and modified to explain these facts. For Cinque, these additional mechanisms are purely syntactic, and consist principally of functional heads (and the features that associate them with particular modifiers).

1.3.5 *Wyner (1994)*

Among the principal concerns in Wyner (1994), as in McConnell-Ginet (1982), is the difference between manner adverbs and subject-oriented adverbs.[23] Wyner develops a semantics for both of these that relies in large part on developing a particular, lattice-theoretic view of the ontology of events. This is structured in such a way that, according to Wyner, the principal semantic difference between manner adverbs and subject-oriented ones can be expressed using a lattice-theoretic minimality operator MIN: subject-oriented adverbs, he suggests, require that events of which they are predicated be minimal (in a sense he makes precise), while manner adverbs do not impose this requirement.[24]

Further, he suggests that subject-oriented adverbs introduce existential quantification over events, while manner adverbs do not. This on its own accounts for why manner adverbs cannot (normally) occur above subject-oriented ones, since a subject-oriented adverb binds off an event argument, thereby making unavailable to any manner adverbs above it.

What is important about this approach for current purposes is only that it too exemplifies the analytical strategy at issue here. In introducing existential closure and a lattice-theoretic minimality operator along with subject-oriented adverbs, Wyner is proposing that the relationship between these adverbs and the expressions they modify is mediated by the introduction of substantive additional semantic mechanisms.

---

[23] With respect to both of these terms, I am purely for convenience using more widely familiar (though somewhat sloppier) terminology than what Wyner uses, since it will suffice here.

[24] Geuder (2000) argues persuasively that this distinction doesn't have quite the necessary effect.

1.3.6 *Wyner (1998)*

A different approach to the semantics of subject-oriented adverbs, and in particular to the problem of passive sensitivity, is taken in Wyner (1998). Like McConnell-Ginet, Wyner proposes that the English passive auxiliary is ambiguous, with one reading more 'volitional' than the other. Wyner's implementation of this differs from McConnell-Ginet's, though. His version is stated in terms of a thematic role 'Volition' assigned to the surface subject. This role this assumption plays in his proposal is roughly similar to the role it plays in hers. However, Wyner understands the semantics of subject-oriented adverbs themselves differently. The example he gives defines *reluctantly* as an operator introducing a state of reluctance:

(23) $[\![\textit{reluctantly}]\!] = \lambda P \lambda e \,.\, P(e) \wedge \exists s[\text{reluctant}(s) \wedge \text{Experiencer}(s)=\text{Volition}(e)]$

The experiencer of the state of reluctance in (23) is also the bearer of the volition role of the modified event. Presumably, (23) should represent the general lexical-semantic architecture of the class of such adverbs.

This proposal, then, is like Ernst's in placing much of the core semantic machinery in the lexical entries of adverbs. As such, much of what was said about Ernst's approach in section 1.3.2 could be said about this proposal too. Just as much as for Ernst, and for the same reasons, if we were to take something like (23) to be a general lexical semantic architecture for subject-oriented adverbs, the generalization that they have this interpretation systematically would be missed. Again, we would have no account of why there aren't just as many adverbs that identify the experiencer of a state with the bearer of the source thematic role of the modified event, or the theme of a state with the instrument of the modified event. Moreover, it is unclear how to extend this to provide a theory of the alternation between subject-oriented and manner readings of e.g. *rudely*—are all adverbs to

be systematically ambiguous, with distinct lexical entries for subject-oriented and manner counterparts?

As before, these problems could be overcome straight-forwardly by taking much of this semantic architecture out of the lexicon, and placing it in the compositional semantics semantics instead (by whatever means). And again, doing so reveals that the underlying analytical move here is to assume that by some means, additional semantics is somehow 'inserted' between the adverb and the expression it modifies.

1.3.7 *Geuder (2000)*

Geuder (2000) examines the lexical semantics of several distinct adverb classes, and a number of issues in adverbial modification more generally. In doing so, he appeals to semantic enrichment operations in a few different respects, of which I'll only mention two here for the sake of brevity. At many points, he discusses various notions of such enrichment quite explicitly himself.

One of these involves resultative adverbs—by which he means adverbs such as those in (24), which resemble adjectival resultatives (see also Parsons 1990):

(24) a. He sliced the bread thinly.

b. They loaded the cart heavily.

In characterizing these, he would like to predict that these adverbs have the flavor of being predicated of an (ordinary, non-event) individual and manifest a clear relation to adjectives that predicate of individuals. In (24b), for example, it is the cart that will as a result of the loading event be heavy. Following Nunberg (1995), he appeals to a mechanism of

'predicate transfer' by which one predicate is shifted to another, related one. Here, the transfer can shift the adjective *heavy*, a predicate of (ordinary) individuals, into an event-predicating denotation for *heavily* like (25):

(25) $[\![\textit{heavily}]\!] = \lambda e\ \exists x[h(x)=e \wedge \text{heavy}(x)]$

where h is a function that assigns a created object the whole event by which it is brought about

Thus *heavily* will be true of any whole event by which a heavy object is created. It is the fact that loading events necessarily involve a 'resultant individual'—something that winds up loaded—that makes this shift possible. Straightforwardly, this sort of shift is a kind of semantic enrichment.

Another example is his analysis of psych adverbs—his term for adverbs derived from psych predicates—such as those in (26):

(26) a. John left the room sadly.

b. He read the review of his book angrily.

The VP *leave sadly*, he suggests, will have a denotation like (27):[25]

(27) $[\![\textit{leave sadly}]\!] = \lambda e\,.\,\text{leave}(e)(x) \wedge \exists s[\text{sad}(s)(x) \wedge f\ \psi\text{-CAUSE}\ s \wedge f \approx \text{leave}(e)(x)]$

This crucially involves a psychological causation relation Geuder terms 'ψ-CAUSE', which in this case relates the 'factive entity' of e being an event of x leaving and a state of x being

[25] I have fixed what I suspect is a typo in this formula. In Geuder's original version, the event variable is bound by an existential quantifier rather than a lambda. Alternatively, of course, I might be overlooking something. In any case, this should be an independent issue.

sad. Very loosely paraphrasing, then, *John left sadly* will mean that there was an event of John leaving, the fact of which caused John to be sad. To arrive at this denotation, he suggests that the last two conjuncts are 'interpolated' by 'a lexical extension of the adjectival meaning that is driven by context'. Earlier in the thesis, he indicates that by 'lexical extension', he means a process proposed by Wunderlich (1997), which Wunderlich himself called 'lexical adjunction'. This is a close cousin of McConnell-Ginet's augmentation—its crucial characteristic is that it applies to a verb, creating an additional argument position in the verb's denotation for a property and supplementing its semantics in a way that makes use of this property. Again, this is a straightforward form of semantic enrichment.

1.3.8 *Kennedy (1997)*

Kennedy (1997) proposes an approach to the interpretation of measure phrases in the adjectival domain (e.g. *six feet tall*, *twelve miles longer*) that embodies this analytical impulse in a very clear-cut way—indeed, it involves both additional semantics and additional structure. This will be discussed at much greater length in subsequent chapters, but for current purposes it is sufficient merely to note that he suggests that such measure phrases are interpreted as arguments of a Degree head (see also Abney 1987, Corver 1990, and Grimshaw 1991) in the extended adjectival projection. Such a head applies to the AP and measure phrase denotations, contributes the semantics that relates them, and yields the appropriate type of denotation for the extended AP. In this way, the head intervenes between modifier and modified and introduces additional semantics.

1.3.9 *Assessment: So What?*

The aim of this section was to establish that it is a very common, perhaps even pervasive, strategy in work on modifier semantics and position to propose additional mechanisms that relate modifiers and expressions they modify, beyond simple intersective or operator interpretations and adjunct syntax. For many phenomena, indeed, it is not clear that there is any alternative. This strategy seems to be adopted universally, in one form or another, in semantically explicit work on subject-orientation, for example, and it may be essential to accounting for many adjective order restrictions.

Indeed, the impulse to approach modifiers this way sometimes manifests itself even in work that is not explicitly concerned with modifier position-interpretation correlations—though this is best left as simply an unjustified announcement here, since demonstrating it would take us far afield.[26]

Of course, it's not at all the case that all these additional mechanisms have much in common. Some, in fact, seem to have very little in common. So this certainly isn't a particular *analysis* that recurs, or a single, clear-cut analytical tool. But neither is it a completely heterogeneous collection of grammatical devices (more on this below). The most essential commonality among them is the very fact that they are 'additional'.

This all supports a strong negative claim: there is widespread though implicit agreement—perhaps even an unrecognized consensus—that the standard assumptions about modifier syntax and semantics are insufficient. This is a bit surprising, since these

---

[26] I have in mind things like Guttiérrez Rexach (1997), who appeals explicitly to McConnell-Ginet-style augmentation in a generalized-quantifier analysis of questions; Bouma (2003), who, working in HPSG, appeals to a rule under which 'adjuncts may be added lexically to the syntactic valency of a verb' in an analysis of adjunct scope in Dutch verb clusters; or Potts (2003), who proposes a semantics for nominal appositives (e.g., *Lance, a cyclist*) that relies crucially on a syntactic feature [COMMA] triggering a particular kind of type shift.

standard assumptions are fundamental and very well-established. Interpreting modifiers either intersectively or else as operators is a basic notion in semantic theory, and analyzing them as adjuncts is equally basic in syntactic theory. Moreover, the mode of explaining modifier position purely through ontological distinctions is deeply entrenched. Nothing here suggests that any of these need to be rejected, of course. Surely, there are *some* modifiers that are adjoined and interpreted in the conventional way, and *some* modifier order facts that follow directly from ontological distinctions—in fact, there are clearly many of both. But perhaps this does suggest that the standard assumptions on their own lead to a systematic gap in empirical coverage. That many of these areas have for so long remained relatively unexamined and at best imperfectly understood may be a testament to this. After all, it seems natural that work would be limited and progress slow where the available analytical tools systematically fail.

Maybe, then, what is necessary is some articulation, as precise and general as possible, of what this covert (quasi-)consensus is—that is, some theory that can embody it in a consistent and integrated way, and ideally, replace much of the menagerie of more local stipulations that are otherwise necessary. In some sense, what may be required is not necessarily a new proposal[27] but some common language or framework for expressing old ones. Developing one is a difficult task. The only other options, though, seem to be methodologically undesirable: either neglecting a wide range of interesting and cross-linguistic phenomena or else sticking one's head in the analytical sand (to put it perhaps too pejoratively) with respect to the recurring need to reach beyond the existing tools.[28]

---

[27] Though this wouldn't hurt either.

[28] Of course, there is a certain personal judgment to be made on this point. Despite the tendentious wording here, certainly it's entirely possible that the right approach is simply to wait for a new insight to

1.4 From Analytical Intuition to Grammatical Principle

This section will explore ways to distill this recurring analytical intuition into a principle of the grammar, or, perhaps more accurately, into a consistent framework in which this intuition can naturally be expressed.

Sections 1.4.1 and 1.4.2 elaborate some theoretical goals a bit further and sketch some general ideas toward an understanding that might achieve them. Sections 1.4.3 and 1.4.4 present a more explicit implementation of the these ideas and a demonstration of how they work and of some analytical benefits they offer. An alternative implementation is developed starting in 1.4.5 in which a particular notion of feature bundling provides some additional analytical flexibility and phrase-structural advantages. Section 1.4.9 considers the significance of modifier 'size'; 1.4.10, an argument for the proposed approach from learnability; and 1.4.11, some possibilities for restricting the theory.

1.4.1 *Some Desiderata*

The challenge in undertaking this task lies in part in determining as precisely as possible where conventional assumptions about modifier interpretation are systematically insufficient, and in identifying what the additional mechanisms that have been proposed

---

emerge serendipitously from work on other, more tractable problems.

It is worth mentioning here that there is a promising novel approach toward certain modifier-order facts developing that is fundamentally different from any discussed above. In recent work, Øystein Nilsen (Nilsen 2001, 2003) suggests that adverb order might better be understood on a par with negative-polarity licensing phenomena or quantifier scope. This seems a natural idea, and its intuitive appeal is great. In Nilsen (2001), however, he frames his discussion in terms of a Type Logical Grammar (Barker 2003 is a linguistically-oriented introduction; Nilsen cites Morrill 1994 and Moortgat 1996), and in that respect it is couched in a framework architecturally different from the one assumed here. As a consequence, a meaningful comparison of this proposal to those considered here would require undertaking the non-trivial task of translating things into more directly comparable terms. More recently, in Nilsen (2003), this stumbling block has apparently been eliminated.

have in common. This is of course in some sense a goal of the entire enterprise, but the picture that has emerged so far provides some general ideas in this direction.

First, it seems clear that the modifiers that participate in the effects at issue, and the ones least amenable to analysis in conventional terms, are all 'small'. That is, they are modifiers that project less syntactic structure than some others normally do. The examples considered so far are almost exclusively adverbs and adjectives. These, of course, do involve syntactic complexity and do project further structure—but the extended adverbial and adjectival projections are in a meaningful sense smaller than, for example, PP or CP modifiers. As mentioned very briefly above, certain DPs—primarily various kinds of measure phrases—seem to count as small in this sense. Much more discussion of these follows in subsequent chapters, but one point relevant here is that in a number of languages with overt case, these are marked with accusative (see chapter 3). This may itself be a kind of smallness, since accusative case may involve fewer nominal projections than lexical cases do (as in e.g. Hopf, Bayer, Bader, and Meng 1998; cf. Bittner and Hale 1996).

The observation that smaller modifiers pattern differently from larger ones has been made before for adverbs. Cinque (1999) specifically distinguishes between adverbs closely tied to particular syntactic positions and various larger adverbials, such as locative and temporal PPs, whose distribution seems to be relatively less restricted. He calls these larger, more syntactically free modifiers 'circumstantial' (following e.g. Fillmore 1994), though this semantic-sounding term is in some respects misleading, at least with respect to the primarily syntactic distinction he has unearthed. Modifiers with (some version of) this circumstantial distribution are a highly heterogeneous class semantically, including among other things rationale clauses, instrumental PPs, and even certain manner PPs (e.g., *in a clumsy way*). Moreover, they maintain their semantic character irrespective of their position—no matter how one dislocates a locative PP like *in Cleveland*, it will make essen-

tially the same semantic contribution.[29]

So, one property that it would be desirable to explain is the rough correlation between a modifier being phrase-structurally small and it participating in the kind of meaning-position relations discussed here. It would, of course, also help to have a clearer notion of what counts as small.

Also important, a theory should make it possible to understand at least certain modifier order facts as genuinely syntactic, and not necessarily as a consequence of any deep property of the semantics—after all, one of the difficulties faced in this domain seems to be a paucity of essentially syntactic tools, so some should be provided.

On the other hand, it is equally important not to *cut off* the possibility for certain effects to be explained in truly semantic terms, on the basis of ontological distinctions. While there may be evidence that this mode of explanation doesn't on its own suffice, there is no reason to get rid of it completely or even to systematically disfavor it wherever it is available. The optimal situation would be for these sort of explanations to remain an option, but not for them to be required by the absence of alternatives.

It would also be desirable to reflect something of the connection between modifier classes and functional heads Cinque notes. A related desideratum is that the fundamental cross-categoriality of these effects not be unexpected, as it would be on an explanation that hinges on something particular to adverbs to the exclusion of adjectives, for example. These cross-categorial facts suggest the need for a similarly cross-categorial explanation.

Since what is necessary here is apparently the addition of some general theoretical tools—ones flexible enough to provide a consistent framework in which existing proposals could be expressed—the theoretical machinery will unavoidably have to be made more

---

[29] Modulo purely scopal facts and information-structure aspects of interpretation like focus.

powerful. That being the case, though, one has to be careful not to make things so powerful that the notion of a 'consistent' framework starts to become meaningless. The additional power must not be so great that it is not clear how it might later in principle be constrained.

And of course, the core goal has to be some way to reflect a notion of 'enrichment' or 'augmentation', some way of providing a regular means by which the relationship between modifier and modified can be mediated, as well as the broader observation that particular syntactic positions are associated with particular meanings. To avoid some of the pitfalls of placing too much of this work in the lexicon, most of the moving parts of the machine will have to be in the syntax and compositional semantics.

1.4.2 *Toward Integration, and the Outlines of an Approach*

The outlines of a framework with these general properties may begin to emerge in one form from simply synthesizing, articulating, and reformulating certain elements of existing approaches. This yields a fairly direct distillation of existing insights, and doesn't add much that is extrinsic. A slightly different, possibly preferable but in some respects more complicated conception of essentially the same idea is developed in 1.4.7.

The first ingredient in putting things together will be McConnell-Ginet's suggestion that it's natural to think of adverbs semantically as arguments of a verb, in some extended sense of 'argument'. Since it is of course crucial that these modifiers-qua-arguments be distinguished in some way from arguments proper, this entails adopting some notion of augmentation as well. The alternative—supposing that modifiers are routinely[30] *true*

---

[30] I'm setting aside here famously tricky cases in which something that is prototypically a modifier appears to be an argument, such as *Floyd behaved *(badly)*.

arguments of heads—is highly undesirable, for reasons already discussed (among many others). If modifiers-qua-arguments and arguments proper are all arguments of lexical heads, some further explanation is required of their differences. But more problematically, this would confound idiosyncratic characteristics of particular heads, such as that *wiggle* can take a direct object but that *writhe* cannot, with characteristics that are general to various lexical categories, such as that verbs can be modified by adverbs but nouns cannot.

So conceptually, the notion of augmentation must be placed squarely in the compositional semantics. One might imagine doing this quite directly in the way McConnell-Ginet does, by simply defining an augmentation rule and understanding it to be part of the basic mechanics underlying semantic composition, like the rule of functional application or intersective interpretation of modifiers. There might be some worry at this point about how an adequate augmentation rule would square with these existing rules. At least as she formulates it (working in Montague Grammar), it seems quite different in character and complexity from rules like these. In Montague Grammar, this would be less of a concern than it is in a framework that eschews such specific interpretation rules in favor of very general ones, such as that of Heim and Kratzer (1997).

Moreover, to achieve what we'd like it to, we would need multiple augmentation operations, not just one general one—or else at least several particular variants of a general augmentation relation. If a wide variety of interpretive effects are to be attributed to this notion in some form, the other work above makes clear that many modifiers that we might want to assume are interpreted by some such operation would not fit into the neat mold of having a semantics of the form of, to put it schematically, 'V in/at an Adj_{Adv} N'—that is, of involving only predicating something of a particular aspect or dimension of an event (or of a verb, to characterize her idea less anachronistically). This frame certainly gets us very far, and accounts for much of the intuitive appeal of her idea, but for the purposes I want to

put it to, a bit more flexibility is required. And, to serve as the beginning of a more general theory of adverb order, different sorts of augmentation will likely have to be ordered with respect to each other.

Coping with this sort of situation would be relatively straightforward in a framework in which rules of semantic composition come fairly cheap—that is, one in which there may be many such rules, and they can be quite construction-specific and refer directly to the lexical content of predicates with respect to notions like ‘manner’ and ‘speed’. Ernst’s flavor of DRT, for example, might provide this sort of power.

But in a framework in which rules of semantic composition are costly, like the one adopted here, the situation looks a bit different. The price of stipulating an inventory of such rules would be too high, and the price of imposing an extrinsic ordering on them higher still, so the only analytical course available is to look to the syntax—that is, effectively, to complicate the syntax in favor of a simpler semantics. To put it another way, the price of a stricter notion of compositionality is a more complicated view of syntactic representations.

Here, then, is where the second principal ingredient comes in. As it stands, we are driven to look for elements of syntactic structure—morphemes, really—that will do work semantic rules might otherwise have to do. In his language of modifier licensing features, Cinque provides something like this. If adverbs and adjectives co-occur with particular syntactic features in the functional structure of the modified expression, as he proposes, these features could provide the additional ‘morphemes’ we are looking for. Importantly, this does not yet require making syntactic commitments about whether this needs to be done in a spec-head configuration and whether distinct functional projections are necessary for every feature—more on this in 1.4.7. Supposing further with Cinque that these features occur in a fixed hierarchical order, this provides a means of talking about structural positions.

The next step is to take these licensing features, precisely contrary to Cinque's intentions, very seriously semantically. Every such feature can be assigned a denotation and given some semantic work to do—specifically, exactly the semantic work associated with the modifier position corresponding to feature. Such features, once assigned a semantics, are no longer purely syntactic devices and have detectable consequences apart from word order, so they are perhaps less likely to rouse any aesthetic uneasiness than they otherwise might be. In this way, it is possible to encode into the grammar the notion that modifier structural positions in themselves have meanings—various structural positions are associated with various licensing features, and these features themselves make semantic contributions.

The natural way to spell this out is to suppose that licensing features perform operations in some respect like the 'augmentation' McConnell-Ginet envisions—they increase the valence of the modified expression in a systematic way so that the modifier can be taken as an argument. This preserves McConnell-Ginet's analytical intuition that adverbs should be treated as argument-like. In syntactifying this idea in this way, it also solves the problem of how one might impose a fixed order on the rules that augment a predicate. On this view, the fixed order of licensing features imposes a fixed order on these augmentation operations.

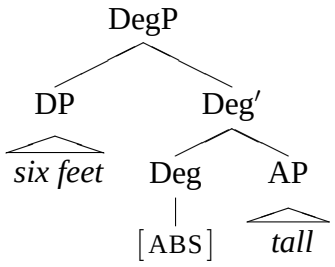
This, then, provides a starting point toward expressing the general intuition of modifiers enriching representations from the previous section in a sufficiently flexible way. The essential notion is that certain modifiers may make their semantic contribution not directly, but rather through an indirect process mediated by semantically-interpreted syntactic licensing features that establish the relation between the modifier and the modified.

For convenience, I'll refer to this means of introducing modifiers as 'mediated modification', for the obvious reason that it involves mediating bits of syntax and seman-

tics. There is of course something a little presumptuous about assigning names to things, in the sense that it may create the impression that one is proposing a Grand Idea. But this general idea, at least at the level of abstraction assumed at the moment, is both something typical assumptions lead us to expect as a possibility, and something that has actually been proposed for particular modifiers.

For example, in the analysis of nominal measure phrases in the adjectival projection mentioned in the previous section and discussed in more detail in subsequent chapters, Kennedy (1997) adopts a structure headed by a degree head corresponding to e.g. *very* and *quite* but, with absolute[31] adjectives—that is, neither comparatives nor superlatives—occupied by a feature [ABS] (Abney 1987, Corver 1990, and Grimshaw 1991 among others propose similar structures):

(28)



Degree heads in general for Kennedy have independently-motivated semantic work to do. The role of [ABS] in particular is to take both the AP and the measure phrase as arguments and to relate them, contributing the additional meaning that the maximal degree to which an individual is mapped on the scale designated by the adjective is at least that designated by the measure phrase. Formulating this to echo McConnell-Ginet's idea, *Floyd is six feet tall* on this view means very roughly something like 'Floyd is tall to a six-foot degree'.

---

[31] Kennedy now uses the alternative term 'positive' for such adjectives, and reserves 'absolute' for use in another sense. I follow his original terminology here.

This is exactly an instance of what I intend by 'mediated modification', and arrived at through reasoning that had nothing to do with McConnell-Ginet's or Cinque's concerns. Implemented in the very direct way Kennedy does—with a modifier in a specifier position and a phonetically empty head occupied by a syntactic feature—it resembles Cinque's syntax very closely. This, I think, reflects the sense in which this particular arrangement is a natural way, perhaps even *the* natural way, of expressing the analytical intuitions in section 1.3 without making new assumptions about the shape of syntactic structures and the semantic combinatorics that interpret them. Though independently motivated, it couples Cinque's general conception of things syntactically and McConnell-Ginet's semantically.

This, then, if understood as a more general means by which certain modifiers are interpreted, provides the essentials of the theoretical language we're looking for. But what would it mean, really, to apply this more generally? And how would this provide the theoretical resources that are necessary? And just how much additional explanatory weight must the syntax bear?

1.4.3 *Mediated Modification, Implementation I: The Christmas Tree Version*

To begin to address these questions, it will be necessary to implement this core idea through some more precisely spelled out theoretical assumptions. I'll develop one picture of how to do this in this section, and consider an alternative implementation in 1.4.7.

The first implementation of this approach is to simply take it as a way of reconceptualizing very loosely Cinquean structural assumptions by assigning them a semantics—indeed, by finding semantic motivation for these structural assumptions.

On this kind of implementation, the features whose interpretations embody various 'augmentations' in McConnell-Ginet's sense each correspond to a distinct functional head,

and are therefore hierarchically ordered with respect to each other by the same principles that order functional heads more generally. And, following Cinque, the modifier each feature introduces can occupy the specifier position of the head that hosts it.[32]

Importantly, unlike for Cinque, these features need not play any (further) syntactic role. That is, no additional mechanism of feature checking is required. For Cinque, such a mechanism is needed principally to assure that modifiers occur sufficiently close to the heads that introduce them. But here, each modifier (of the relevant sort) is a semantic argument of its corresponding feature, and so must occur in a particular position with respect to it for this purely type-theoretical reason. Equally important, though, there is nothing here that would *prevent* these features from playing a further syntactic role, either. For similar reasons, the Cinquean assumption that the modifier occupies a specifier position loses one of its principal purposes: its role in ensuring that a modifier occur near its corresponding head is no longer necessary. This assumption is maintained here, though, largely because it also ensures that the modifier occupies a left branch, which of course the semantics cannot do.

Thus a subject-oriented adverb such as *rudely*[33] might, in this sort of framework, be introduced in a specifier position occupied by a feature [+AGENTIVE], as in (29):[34]

---

[32] As Ernst (p.c.) points out, it is not always obvious exactly what the content is of a claim that something occupies a specifier position without defining precisely what notion of 'specifier' is intended. Here, at least for current purposes, it will suffice to suppose simply that a specifier of a head X is the unique phrasal constituent that is the daughter of (the lowest segment of) the maximal projection of X, does not dominate X, and occurs on a left branch. Certainly, though, a number of more fully-developed understandings are compatible with everything I will say here. Indeed, it is not really crucial for my purposes that the modifier occupy a specifier position in the first place—for reasons discussed further below, an adjunct syntax will do just as well, provided some independent means is available for ruling out right-adjunction of the modifier (at least in English).

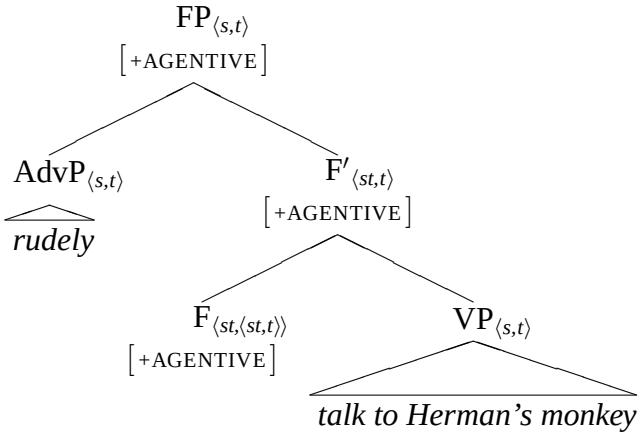
[33] I'm using *rudely* in these examples because it renders the ambiguity between manner and subject-oriented readings quite starkly, more so than the examples of subject-oriented adverbs encountered already, *carelessly* and *reluctantly*.

[34] At this point, I'm using features with plus-or-minus values here mostly to echo Cinque's use.

(29) Floyd [$_{FP}$ rudely [$_{F'}$ [+AGENTIVE] [talked to Herman's monkey]]].

Spelling things out a bit more explicitly, this might have a structure like (30), in which semantic types have been indicated with subscripts, and be interpreted as in (31) (assuming Wyner 1998's semantics for subject-orientation, purely for the sake of the example):[35]

(30)



(31) a. $[\![\,[\text{+AGENTIVE}]\,]\!] =$

$$\lambda P_{\langle s,t\rangle}\ \lambda A_{\langle s,t\rangle}\ \lambda e_s\,.\,P(e) \wedge \exists s[A(s) \wedge \text{Experiencer}(s)=\text{Volition}(e)]$$

b. $[\![$ *rudely* $]\!] = \lambda e_s\,.\,\text{rude}(e)$

c. $[\![$ *talk to Herman's monkey* $]\!] = \lambda e_s\,.\,\text{talk}(\text{Herman's monkey})(e)$

d. $[\![\,[\text{+AGENTIVE}]$ *talk to Herman's monkey* $]\!]$

$= [\![\,[\text{+AGENTIVE}]\,]\!]([\![$ *talk to Herman's monkey* $]\!])$

$= \lambda A_{\langle s,t\rangle}\ \lambda e_s\,.\,\text{talk}(\text{Herman's monkey})(e) \wedge \exists s[A(s) \wedge$

$\text{Experiencer}(s)=\text{Volition}(e)]$

---

35 Following Kratzer (1996, 2002a), I'm assuming in this structure that the agent is introduced in a neo-Davidsonian way above VP. This is not crucial. I am using *s* to indicate the type of eventualities and e and s as variables over events and states respectively.

e. ⟦ *rudely* [+AGENTIVE] *talk to Herman's monkey* ⟧

= ⟦ [+AGENTIVE] ⟧(⟦ *talk to Herman's monkey* ⟧)(⟦ *rudely* ⟧)

= λe_s . talk(Herman's monkey)(e) ∧ ∃s[rude(s) ∧

Experiencer(s)=Volition(e)]

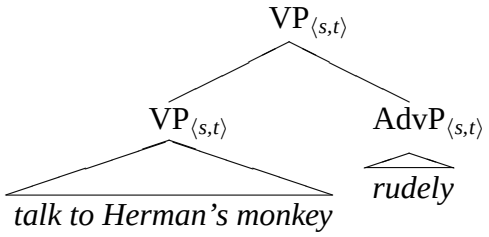
So [+AGENTIVE] takes as arguments the VP and *rudely*, which can now simply denote a property, and relates them in the appropriate way. Its semantic contribution is precisely that which distinguishes the subject-oriented reading from other readings.

In the absence of a subject-oriented adverb, of course, one could either assume, paralleling Cinque's framework, that a negatively-valued (and semantically vacuous) counterpart of the [+AGENTIVE] feature is present, or else that this feature is absent altogether along with the functional head that hosts it. Either of these options are viable given what has been said so far.

An important virtue of this approach—and a fundamental departure from Cinque's view that different uses of an adverb are lexically distinct and essentially syntactically unrelated—is that it allows *rudely* to have exactly the same denotation, the one in (31b), on its manner reading. This is illustrated in (32). I assume here that manner adverbs are introduced in the conventional way—that is, as adjuncts and without any kind of mediation—mostly for the sake of demonstrating that the assumptions developed so far play well with others. Though this treatment of manner adverbs is certainly not a necessary one, it does seem natural to suppose that manner uses are in some sense basic. Thus (32), the manner counterpart of (29), will have a structure like (33) and be interpreted intersectively as in (34):

(32) Floyd [[talked to Herman's monkey] rudely].

(33)



(34) ⟦ *talk to Herman's monkey rudely* ⟧

$= \lambda e_s$. talk(Herman's monkey)(e) $\wedge$ rude(e)

This, of course, is the usual Davidsonian way of interpreting a manner adverb, and yields the usual Davidsonian denotation.

More examples of how things might work—particularly with multiple modifiers—follow in 1.4.4. Before proceeding further, though, it may be useful to dwell momentarily on an aesthetic point. Unavoidably, this implementation of the idea (though not the alternative implementation to follow) requires proliferating functional heads. It inherits this property from the Cinquean framework, which famously involves positing well over a dozen functional heads in the verbal extended projection alone. This strikes some people as something like a reductio ad absurdum, so unappealing it must somehow be wrong. These sort of structures are, in a memorable pejorative characterization due to Roger Higgins, Christmas trees, seemingly existing independent of any lexical material, on whose branches linguistic expressions can be hung here and there like ornaments. While it is possible to alleviate some of this aesthetic uneasiness with respect to this proposal by pursuing instead the alternative implementation below, I think it may be valuable to consider simply accepting this charge—embracing it, really—and treating this sacrifice of structural elegance as a price willingly and justifiably paid. What paying this aesthetic price buys is increased descriptive adequacy, which in this empirical domain is sorely needed, for

reasons already discussed—there is no single widely accepted understanding of the sorts of facts considered here, and progress has long been relatively slow. It also buys an integration of distinct local theories of particular modifier position effects, each of which would otherwise itself require paying an independent price. Perhaps, then, a certain sacrifice in theoretical elegance is warranted, at least provisionally and in the short term, if progress is to be made, especially if the alternative is to simply ignore a wide range of interesting and often cross-linguistically robust phenomena in hopes that an elegant account of them will ultimately emerge of its own accord. To put it another way, perhaps at least in some cases it may be methodologically preferable for a theory to be insufficiently elegant and even insufficiently restrictive than for it to be prematurely elegant and restrictive.

There is a related, more empirical objection to this sort of proliferation of heads: one might expect each such modifier-introducing head to be spelled out in some language. This is exactly the challenge Cinque (1999) takes up, and he provides a large amount of typological evidence in support of his conception of things. His is, of course, a different hierarchy of heads and features than what would be required here, so it cannot be taken to directly address this concern. But it is sufficiently similar as to suggest that these are two probably reconcilable attempts to model essentially the same linguistic phenomenon, and that the differences arise from the different methodologies being pursued. Thus it seems reasonable to tentatively take each of these to be independent support for the other, with the assumption that here, as elsewhere, tensions between where syntax guides us and where semantics does can be resolved (and indeed, may turn out to be revealing). It is, after all, desirable that a theory lead us to look for new facts and make new connections.

Another perspective on this issue is available, though. For entirely different reasons, Koopman (1996) proposes a ‘generalized doubly-filled Comp filter’, which requires that in

a single projection, either the head or the specifier can be spelled out, but not both.[36] If this is right, we would on independent grounds *expect* it to be the case that when modification proceeds in this way, heads would not be spelled out.

This approach has a number of advantages. Of course, it provides a unified language in which a certain kind of generalization can be stated. It does this without requiring any special combinatorics or ad-hoc rules—no new mode of syntactic or semantic composition is proposed. Importantly, the denotations of (the relevant) modifiers themselves can now be very simple, since the features that introduce them absorb much of the complexity. This is desirable on its own, and may have important learnability advantages. Perhaps most important, though, is that this approach allows a single modifier to have the same denotation across its uses. In this regard, this is a fundamentally Ernstian model, in that it commits to this single-denotation idea very strongly rather than taking the Cinquean path of treating this kind of semantic alternation as a lexical ambiguity. More precisely, at least on this implementation it uses more or less Cinquean tools to construct an Ernstian theory.

This implementation of the idea is the one that generally surfaces in subsequent chapters. In many respects, I actually regard the more flexible and phrase-structurally conservative alternative in 1.4.7 as preferable, but will not explicitly make use of it because it requires making additional assumptions that are not immediately relevant to the particular phenomenon each chapter focuses on.

1.4.4 *Taking the Idea for a Test Drive (on Some Rocky Terrain)*

It may be worthwhile to illustrate in some more detail how things would work in the case of multiple modifiers introduced in this way, in part for the sake of explicit-

---

[36] More precisely, it requires that no projection have an overt head of a chain in both its head and its specifier.

ness, in part to briefly consider some ad-nominal modifiers, and in part to demonstrate that these assumptions provide some traction on some neglected problems that might otherwise appear too daunting or mysterious not to sweep under the rug. The examples will involve color adjectives, attributive nouns that designate a material of which an object is (materially) composed, and classificatory adjectives, all of which occur in a rigidly fixed order with respect to each other.

CLASSIFICATORY ADJECTIVES Among the adjectives restricted to especially low positions are so-called 'classificatory' or 'relational' adjectives such as *musical* in *musical comedy* (Bosque and Picallo 1996, McNally and Boleda Torrent 2003, Cinque 2003), which bear a close resemblance to the kind of adjectives that occur in the *intellectual midget* examples in (9–11). Obligatorily, these occur closer to the noun than adjectives of a number of other classes, including evaluative adjectives as in (35), size adjectives as in (36), and, most relevant at the moment, color adjectives as in (37–38):

(35) a. an awful *pulmonary* disease
b. *a *pulmonary* awful disease

(36) a. a huge *political* problem
b. *a *political* huge problem

(37) a. a beige *dental* instrument
b. *a *dental* beige instrument

(38) These horrible invitations are bad enough, but for god's sake, did they really have to print them on...
a. ...pink *nuptial* stationery?
b. *...*nuptial* pink stationery?

This fixed syntactic position correlates with a particular interpretation—it is exactly this fact that permits the use of (quasi-)semantic characterizations such as 'classificatory' to identify this class. Before proceeding further with this puzzle, it will help to introduce the others.

COMPOSITION NOUNS There is a distinct class of attributive nouns that designate the material of which something is materially composed. These include *leather* in *leather shoes* and *plastic* in *plastic box*. They have been referred to simply as adjectives (in e.g. Kamp and Partee 1995, who call *stone* in *stone lion* an adjective), but this can suffice only as a pretheoretical characterization—these expressions fail standard tests of adjectivehood, including ability to occur in comparatives or the complement position of *seem*:

(39) a. That hat is $\left\{\begin{matrix}\text{more}\\\text{very}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{*steel}\\\text{metallic}\\\text{*leather}\\\text{leathery}\end{matrix}\right\}$.

b. That hat seemed $\left\{\begin{matrix}\text{*steel}\\\text{metallic}\\\text{*leather}\\\text{leathery}\end{matrix}\right\}$.

True adjectives such as *leathery* and *metallic* are perfectly grammatical in these contexts, but their nominal counterparts are clearly not, despite their semantic similarity.

Importantly, though, these are not simply members of noun-noun compounds, either.[37] Unlike nouns in compounds, composition nouns can be excluded in *one* pronomi-

[37] This is not, of course, to say that composition nouns can't be lexicalized in particular cases, such as *paper bag*. They can, but of course so can adjectives (e.g. *old maid*) and members of virtually any other syntactic category (at least in idioms), so this is not evidence that composition nouns are *always* members of a noun-noun compound.

nalization:

(40) a. *Floyd bought the coffee maker, and I bought the bread one.

b. Floyd bought the steel coffee maker, and I bought the plastic one.

(41) a. *Greta is wearing gym pants, and Herman is wearing sweat ones.

b. Greta is wearing spandex pants, and Herman is wearing leather ones.

So (41a) is ungrammatical because *sweat pants* (like *gym pants*) is a noun-noun compound, and *ones* cannot stand in for something that's a proper part of a compound. In (41b), on the other hand, *leather ones* is fine because *leather pants* is not a noun-noun compound. That these are not parts of compounds is also reflected in their position. Composition nouns can occur only outside of a noun-noun compound. Thus a coffee maker made of steel can be characterized as in (42a), but not as in (42b); and similarly for the *pants* example in (43):

(42) a. steel coffee maker

b. *coffee steel maker

(43) a. leather gym pants

b. *gym leather pants

Maybe the most important evidence for not treating composition nouns as part of noun-noun compounds is that their semantic contribution is non-idiosyncratic, as the contribution of nouns in noun-noun compounds may be. To my knowledge, *mouse hat* is not an existing noun-noun compound in English. Accordingly, it has no fixed meaning—perhaps a *mouse hat* is a hat designed to be worn by mice, or perhaps it's a hat with big mouse ears on it (like

a *Mickey Mouse hat*), or perhaps its a hat one wears to repel mice (cf. *mosquito net*). But in *steel mouse hat*, we can be quite sure what the contribution of *steel* is. Irrespective of what *mouse hat* means, *steel mouse hat* can be used to characterize an instance of something in the extension of *mouse hat* that also happens to be made of steel. It is of course also perfectly possible to lexicalize *steel mouse hat* (particularly on the parse not relevant here, *[steel mouse] hat*), but this would not eliminate the possibility of interpreting *steel* as a composition noun. And certainly, one might call a metallic hat worn by mice a *steel mouse hat* and not a **mouse steel hat*. So, to indulge in the inevitable pun, composition nouns are compositional.

One striking bit of evidence that composition nouns are not simply members of noun-noun compounds is relevant to the broader point here: they obligatorily occur above classificatory adjectives:

(44) a. a steel dental instrument
b. *a dental steel instrument

(45) a. a leather bridal gown
b. *a bridal leather gown

(46) a. a stone religious relic
b. *a religious stone relic

So again, since composition nouns can occur outside of classificatory adjectives, they cannot be simply members of compounds.[38] It also reflects that these composition nouns

[38] Unless, of course, one were to treat all classificatory adjectives as themselves part of a single complex lexical item. This, though, would be profoundly unsatisfying in several respects, chief among them that the semantics of classificatory adjectives is completely compositional.

occupy a relatively fixed position. This position is below the one most naturally occupied by color adjectives:[39]

(47) a. a blue cotton shirt
b. *?a cotton blue shirt

(48) a. a yellowish metal shelf
b. *?a metal yellowish shelf

(49) a. a grey stone lion
b. *?a stone grey lion

The generalization, then, is that composition nouns are independent modifiers with their own syntactic distribution, one which places them systematically above classificatory adjectives and below color adjectives.

COLOR ADJECTIVES The final part of the picture to be drawn is color adjectives. As we have already seen, color adjectives occupy a relatively fixed position in the DP. In addition to occurring obligatorily before composition nouns and classificatory adjectives, they are also restricted to positions below evaluative and size adjectives:

(50) a. the big red ball
b. *the red big ball

---

[39] Some of these judgments are not perfectly clear-cut. One likely reason for this is that it's possible in principle (though not easy) to assign color adjectives classificatory adjective interpretations, so if a *yellowish shelf* can be construed as a particular variety of shelf, *metal yellowish shelf* will be fine for precisely the same reason e.g. *steel dental instrument* in (44) is.

(51) a. the beautiful red ball

b. *the red beautiful ball

This is a more familiar puzzle than the previous two, but it remains one which for which a satisfying explanation is elusive. The only one available that I am aware of is in the (broadly) Cinquean tradition, in which color adjectives occupy a particular position in the hierarchy of functional heads. Cinque (1994) places them in the specifier position of a ColorP, whose position is fixed with respect to corresponding projections for evaluative and size adjectives.[40] This, though, leaves the semantics curiously out of the picture in the same way the Cinquean approach to adverbial modification does. It treats 'color adjective' as merely a lexical class, whose membership could just as well have been utterly unpredictable and idiosyncratic—but of course, it is not. Any adjective that has the appropriate meaning acts in precisely this way, exceptionlessly. So some explanation is required for why this particular position should correlate with color modification.

ASSEMBLING THE PIECES All these puzzles fit quite naturally into the bigger picture here—all accord with what supposing modifiers can be introduced in this mediated way leads us to expect, and all can be given an explanation in these terms. In each case, there is a systematic correlation between position and interpretation that does not seem to follow from anything else.

A general picture of how this will proceed may have become clear by now. The strategy will be, as before, to simply factor out the semantics associated with a particular position, thereby deriving the interpretation from the position and not vice versa. For

---

[40] It is worth stressing that this apparently remains the only explanation of color adjective position in the mainline generative tradition, even though the facts in this area are quite well-known and systematic, and in this respect these facts must be taken as evidence for a broadly Cinquean approach until some alternative explanation is provided.

classificatory adjectives, this crucial feature—[+CLASS], say—can be assigned a semantics roughly along the lines proposed by McNally and Boleda Torrent (2003). They suggest that classificatory adjectives are introduced by a special rule of semantic composition that predicates the adjective of a kind, and yields a property of realizations of that kind as the denotation of the modified NP.[41] Thus [+CLASS] can do the work this rule would otherwise have to:[42]

(52) $[\![[\text{+CLASS}]]\!] = \lambda P_{\langle e,t\rangle}\ \lambda A_{\langle e,t\rangle}\ \lambda x_e\ .\ \exists k_e[A(k) \wedge P(x) \wedge x \text{ realizes } k]$

This will take as arguments an NP denotation and a classificatory adjective denotation (here construed following McNally and Boleda Torrent 2003 as a property of kinds), and yield a property of individuals that realize a kind that satisfies the denotations of both the adjective and the head noun.

For composition nouns, for current purposes it will suffice to assume a feature [+COMPOSITION] with a semantics that contributes a 'composed of' relation that won't be examined further here:

(53) $[\![[\text{+COMPOSITION}]]\!] = \lambda P_{\langle e,t\rangle}\ \lambda C_{\langle e,t\rangle}\ \lambda x_e\ .\ P(x) \wedge \exists y[C(y) \wedge x \text{ is composed of } y]$

This will take as arguments the denotations of the NP and the composition noun, and yield a

---

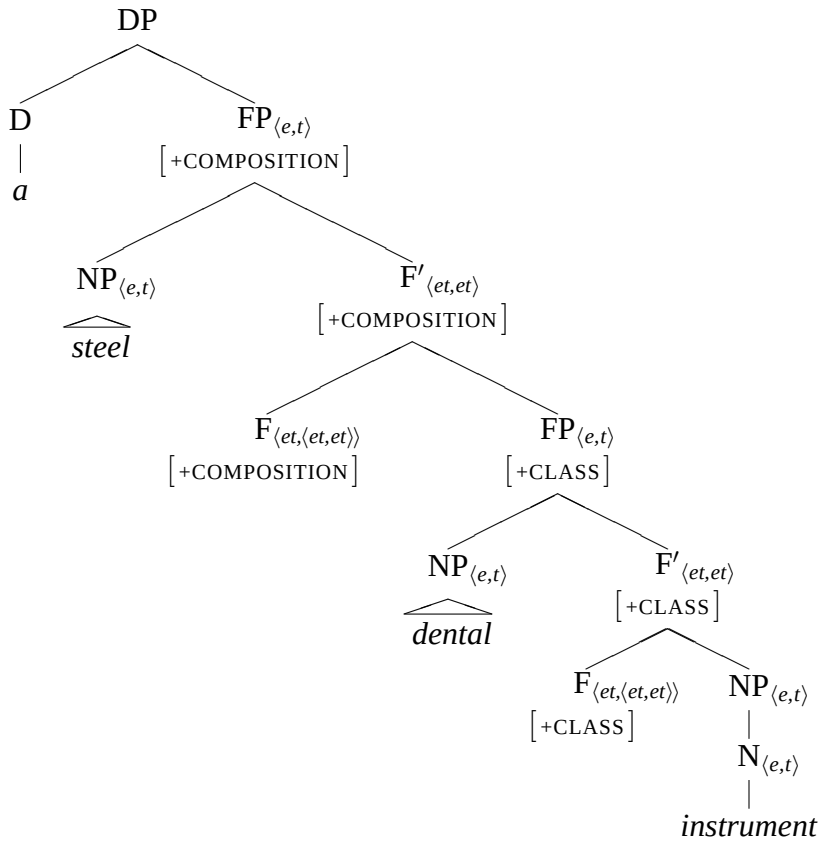
[41] More precisely, the realization relation itself is not provided by the rule, but independently by the semantics of every noun that occurs with a classificatory adjective. It doesn't seem to me to do any harm to the semantics they propose or to the analysis of these structures more generally to extract the realization relation from the noun's semantics in this way. They also propose that the kind is contextually-provided rather than bound by an existential quantifier as it is in (52)—I take this route here for simplicity of exposition, but it does not strike me as obviously wrong.

[42] I'll reflect the sortal distinction between kinds and ordinary individuals by using *k* as a variable over kinds.)

property of individuals which are composed from something that satisfies the composition noun denotation.

With this in place, *a steel dental instrument* can be assigned the structure in (54), and the denotation in (55):

(54)



(55) a. ⟦ *instrument* ⟧ = λx . instrument(x)

b. ⟦ *dental* ⟧ = λk_e . dental(k)

c. ⟦ *dental* [+CLASS] *instrument* ⟧

= ⟦ [+CLASS] ⟧(⟦ *instrument* ⟧)(⟦ *dental* ⟧)

= λx . $\exists k_e$[dental(k) $\wedge$ instrument(x) $\wedge$ x realizes k]

d. ⟦ *steel* ⟧ = λy . steel(y)

e. ⟦ *steel* [+COMPOSITION] *dental* [+CLASS] *instrument* ⟧

= ⟦ [+COMPOSITION] ⟧(⟦ *dental* [+CLASS] *instrument* ⟧)(⟦ *steel* ⟧)

= λx . $\exists k_e$[dental(k) $\wedge$ instrument(x) $\wedge$ x realizes k] $\wedge$ $\exists$y[steel(y) $\wedge$ x is composed of y]

At the cost of the stipulation of the relative order of these features, then, this accounts for the (otherwise unexplained) facts about why composition noun and classificatory adjective interpretations should be restricted to particular positions in this way—and it explains how these interpretations arise in the first place.

But what about color adjectives? The fundamental hypothesis underlying the approach pursued here is that structural positions themselves have meanings. Apparent associations between a position and a particular kind of lexical semantics are therefore treated by distilling some component of the apparent lexical semantics and attributing it instead directly to the position the modifier occupies. So for color adjectives, the natural thing to suppose is that the position they occupy is simply associated with the semantics of color. But the next step this would lead to is a bit awkward—it's not clear how one could factor out the meaning associated with color from *red* or *blue* without losing the very heart of what these adjectives mean. One could, perhaps, suppose that the lexical semantics of *red* doesn't really (in some tortured sense of 'really') directly say anything about having a color, but *red* happens to be true only of things that have a red color (perhaps for some reason having to do with nonlinguistic facts about the world). This, though, seems to be slicing things a bit too thin.

An alternative would be to suppose that *red* means exactly what one might think it does, and that the (prenominal) position it occupies has a semantics associated with a 'x is of the color y' relation. This, in other words, doesn't factor out the color meaning so

much as mirror it. This would do the trick—it would enforce the association between that position and color semantics as a kind of selectional restriction, essentially by ensuring that only color adjectives could be interpreted there because only they are compatible with the color semantics required of an adjective that occurs there. But while this is all perfectly compatible with what has been considered so far, it isn't very satisfying, either. If the essential idea here is that positions can have meanings, evidence for this idea will have to come from instances where positions have meanings that are independently detectable, not ones that simply mirror the independent lexical semantics of a modifier. To find real evidence for this approach from color adjectives, then, it will be necessary to find some independent meaning associated with the prenominal color adjective position that can be disentangled from the lexical semantics of color adjectives themselves.

Led to look for it in this way, we indeed find it. There are aspects of the interpretation of color adjectives that are particular to their prenominal position—that is, that are found there systematically and without exception but not elsewhere. The contrast between attributive and predicative uses of *green* in (56) provides an illustration. The scenario here is that Floyd and Clyde have met on St. Patrick's Day, and that they observe the custom that anyone having failed to wear green in observance of St. Patrick's Day is implicitly inviting an act of minor violence against them.[43] Floyd is sadistic, and joyously observes that Clyde has failed to wear green and has thereby consented to have a minor act of violence performed against him. Clyde objects, however, pointing to the lining on the inside of his jacket, which is green, and uttering (56):

(56) *Clyde (displaying the green lining of his jacket):*
But this jacket is green!

---

[43] There appears to be some variation among those who have attended American grade schools in the act of minor violence expected, ranging from pinching, the most typical, to punching. This an observation the relevance of which is not linguistic.

This is a likely defense for Clyde to adopt, and he has said something true. But not so if Clyde had instead uttered (57):

(57) But this is a green jacket!

This is false because it is not sufficient for something to be a *green jacket* that only its lining be green. There is some additional special requirement that is imposed by combining *green* with *jacket* via attributive adjectival modification that is not present when these are combined in other ways.

Another example of this contrast, again involving *green*, might arise upon opening a friend's refrigerator and discovering some particularly moldy bread:

(58) Holy crap—your refrigerator is absolutely revolting! Rotting vegetables! Discolored cheese! Disintegrating fruit!...

a. ...And, ugh, look at this bread! It's green!

b. #...And, ugh, look at this green bread!

The judgment is perhaps a bit less clear-cut here than in the St. Patrick's Day scenario, but things seem to work similarly: bread that has spots of green mold might be said to be *green*, but it cannot truthful be characterized as *green bread*.

The common thread, then, is that the prenominal position introduces an additional, stronger requirement in addition to the lexical semantics of the color adjective, which has to do with the 'way' in which the color adjective is predicated, that is, with how or in what respect the jacket or the bread is green—internally or externally, visibly or not, partly or wholly, etc. It is, of course, well beyond the scope of this discussion to try to isolate this difference, but it might, very loosely, be said to involve what the *principal* color of an

object is. Thus green may be the color of the jacket internally or on the inside, but it is not the principal color of the jacket. What counts as the principal color would likely depend on a number of subtle, probably contextual factors and of course also on what the particular object is.[44] This, then, could be said to be the special meaning of the position occupied by color adjectives, and expressed as the denotation of a licensing feature [+COLOR]:

(59) $[\![[+\text{COLOR}]]\!] = \lambda P_{\langle e,t\rangle}\ \lambda A_{\langle e,t\rangle}\ \lambda x_e\ .\ P(x) \wedge A(\text{PRINCIPAL-COLOR}(x))$

The PRINCIPAL-COLOR function maps an individual onto its principal color.[45] Thus [+COLOR] takes the denotation of an NP and that of a color adjective as arguments and yields a property of individuals whose principal color satisfies the color adjective.

There is, then, some independently observable meaning associated with the position of color adjectives, namely that in (59). But in expressing this meaning, the restriction that only color adjectives can occur as arguments to [+COLOR] was captured as well. The mirroring of color lexical semantics considered above instead falls out as a straightforward selectional restriction of the 'principal color' relation—blue or red or green might be the principal colors of an object, but ugly or dental can never be.

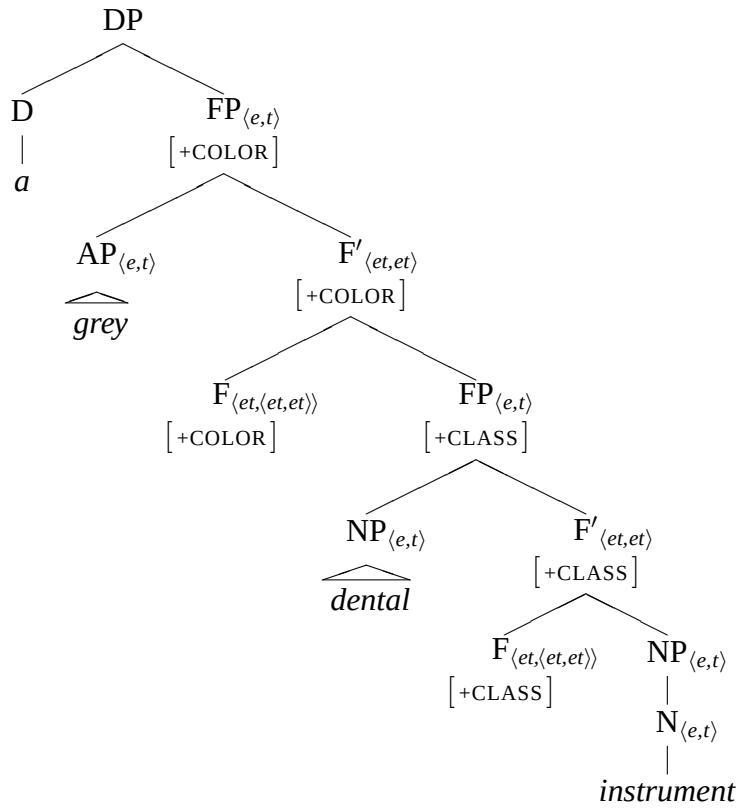
And this, of course, allows the semantics to be distributed across the tree in a way that has the desired result, again with only the stipulation of the relative order of the independently motivated features. So the structure of *grey dental instrument* will be as in (60), and this will be interpreted as in (61):

---

[44] This story is not, of course, sufficient as it stands. One difficulty is that while the jacket in (57) is not a *green jacket*, it is in fact an *internally green jacket*—yet of course 'internally green' cannot be said to be the principal color of the jacket. Pragmatic factors governing the process by which the color of an object is determined are discussed in Blutner (1998) (and, approached with a somewhat different set of concerns, in the philosophical literature; see Blutner for citations).

[45] Whether this should in fact be a function is not clear to me. Certainly, one might have a *green and black jacket*, and hence perhaps one with two principal colors.

(60)



(61) a. ⟦ *dental* [+CLASS] *instrument* ⟧

= λx . $\exists k_e$[dental(k) $\wedge$ instrument(x) $\wedge$ x realizes k]

b. ⟦ *grey* ⟧ = λx . grey(x)

c. ⟦ *grey* [+COLOR] *dental* [+CLASS] *instrument* ⟧

= ⟦ [+COLOR] ⟧(⟦ *dental* [+CLASS] *instrument* ⟧)(⟦ *grey* ⟧)

= λx_e . $\exists k_e$[dental(k) $\wedge$ instrument(x) $\wedge$ x realizes k] $\wedge$

grey(PRINCIPAL-COLOR(x))

This approach, then, seems to provide a general means by which the pairing of modifier interpretations and positions can be accounted for, one that is equally natural in both the nominal and the verbal projection and and that can be scaled up to accommodate multiple different co-occurring modifiers. More than that, though, and probably more important, it

seems to provide some traction on difficult problems of a sort that are often neglected (or not observed in the first place), and raises expectations that lead to some novel questions and novel observations. In this way, perhaps it earns its theoretical keep—though it comes at a theoretical price, the price buys broader empirical coverage in places where it is sorely needed.

1.4.5 *A Step Toward an Alternative Implementation: Feature Bundling*

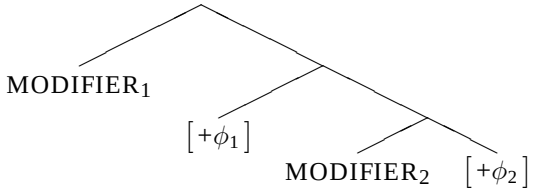
However justified it might be, it is worth considering a bit further whether the theoretical cost of this sort of approach can be, to stretch the metaphor a bit, bartered down. Might there be a way to execute the same larger idea, and get the same or nearly the same benefits, in a more theoretically conservative way? In this section, I'll develop a different implementation of a notion of mediated modification. In a nutshell, this will amount to a version of the implementation already considered supplemented with a theory of feature bundling. The result will be a theory that is more conservative phrase structurally—that is, it will involve smaller, less abstract trees with fewer or no phonetically unrealized functional projections—but whether this actually amounts to a lower theoretical cost hinges on a number of largely aesthetic judgments, chiefly because the savings may be offset by the costs it imposes in other areas.

The core difference in this alternative implementation is that the features that serve the mediating role between modifiers and what they modify will not need to be distributed among distinct heads. Indeed, they could in principle all be on the same head. Schematically, the result will be that trees like (62) can instead look like (63):[46]

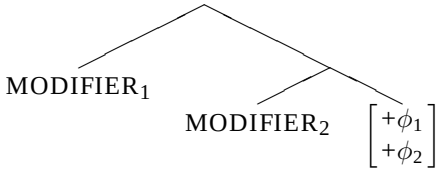
---

[46] My use of ϕ here does not reflect any special connection to ϕ-features—indeed, the features of interest here aren't ϕ-features.

(62)



(63)



Phrase-structurally, (63) is of course simpler then (62) (if only in the sense of having fewer branches). In that respect, it would be in general be preferable to (62), all else being equal. This would be particularly the case in the absence of independent syntactic motivation for the intermediate head position $[+\phi_1]$ occupies in (62). The pivot on which this hinges, then, is the 'all else being equal'. To make all else equal, it would be necessary at a minimum to extract from (63) the same advantages in the grammar of modifiers that trees like (62) proved to have.

But how? In order to make progress toward this question, it necessary to first consider a more basic one: How should feature bundles like the one in (63) be interpreted? This is, of course, is a general question that arises independent of anything proposed here.

Clearly, if the denotations of these features are to be the same as of their unbundled counterparts, they can't simply be applied to each other, for example. That is, given that (62) is interpreted as in (64a), one couldn't for straightforward type-theoretic reasons simply hold the denotations constant and interpret (63) as in (64b):

(64) a. $[\![[+\phi_1]]\!]([\![[+\phi_2]]\!]([\![\text{MODIFIER}_2]\!]))([\![\text{MODIFIER}_1]\!])$

b. $[\![[+\phi_1]]\!]([\![[+\phi_2]]\!])([\![\text{MODIFIER}_2]\!])([\![\text{MODIFIER}_1]\!])$

This could, of course, be taken as an indication that the denotations *shouldn't* be held constant—perhaps the $[+\phi_1]$ that occurs in (62) and (64a) can't be the same $[+\phi_1]$ that occurs in (63) and (64b). But merely adjusting the types in this way, and consequently distinguishing bundled and unbundled instances of what otherwise would seem to be the same feature, has some highly undesirable consequences. If we would like $[+\phi_1]$ to occur in a single language in both bundled and unbundled forms, there would have to be two distinct interpretations for this single feature, a kind of featural lexical ambiguity that would have to be propagated throughout the lexicon for every such feature. If we would like $[+\phi_1]$ to occur unbundled in some languages and bundled in others, this would require this sort of type-theoretic distinction across lexicons. Yet it seems likely that the meaning of functional (as opposed to lexical) linguistic material is subject to severe cross-linguistic constraints, particularly type-theoretically. Famously, for example, determiners seem to be conservative across languages, and generalized quantifier theory has at its heart the idea that determiners as a class denote functions that yield generalized quantifiers; Matthewson (1996, and elsewhere) suggests that cross-linguistic semantic variation is highly constricted in at least this way; and aspectual morphemes may consistently relate properties of events and properties of intervals across languages (Klein 1994, Kratzer 1998).

This difficulty cannot simply be taken as evidence against feature bundling. In one form or another, under one name or another, it's a pervasive notion in the grammar. Indeed, if one were to hew strictly to the notion that distinct syntactic features always occupy distinct nodes, there would be no real distinction between features and null morphemes more generally.[47] In some sense, the whole point of features is that they can be bundled—that is, they exist to provide a way of talking about a proper subset of the properties of a lexical item without structurally decomposing it.

---

[47] An alternative not so easily set aside is to suppose that distinct features always occupy distinct nodes *at LF*—indeed, this seems plausible in a number of respects, and would avoid, via syntactic means, many

The real difficulty confronted here, then, is with functional application itself. It just doesn't seem to be the right operation for interpreting the elements of a feature bundle. Certainly, however fundamental an operation functional application is, it isn't grammatically holy, and alternatives should be considered. If the denotations of features are to stay the same in bundled and unbundled forms, what appears to be necessary instead (something like) function composition. In its general form, this is not a new idea. That function composition play a role in the grammar is a fundamental assumption of Variable Free Semantics (Jacobson 1999 and elsewhere), and von Fintel (1995) suggests that is an engine of semantic historical change as well. And taking this kind of approach specifically below the word level (i.e., below X^0) has antecedents including Di Sciullo and Williams (1987) and Kratzer (2000). In light of this, and particularly in light of this independent evidence for its necessity below X^0 as a kind of morphological rule, it seems natural to take this strategy toward the interpretation of features (which are, after all, sub-lexical entities). To execute this, a distinct semantic rule will be necessary, but of course one would be necessary independently (irrespective of mode of composition) to encode in the grammar how feature bundles are interpreted, if only because they don't meet the structural description of a functional application rule (at least as it is formulated in e.g. Heim and Kratzer 1997).

1.4.6 *A Brief Interlude: Function Composition*

Before proceeding further, it may be useful to briefly illustrate function composition itself. (If it is not useful, one can safely proceed to section 1.4.7.) Informally, function

---

of the difficulties encountered here (cf. the 'feature scattering' of Giorgi and Pianesi 1997 and the treatment of number features in this spirit of Kratzer 2002a). This would, of course, require a dramatic accordion-like unpacking of trees at LF into structures far larger than they would otherwise be, which might provoke exactly the same aesthetic objections bundling of features may be able to address. More generally, this approach in any form would wind up resembling quite closely the first, bundle-free implementation considered in previous sections, and so in this respect it's is the less interesting course to take in the context of the current discussion.

composition involves applying one function to the result of another. It is the operation represented by ∘ in (65):

(65) $A \circ B = \lambda c \,.\, A(B(c))$

So, to take an unrelated linguistic example, one could in principle function-compose the denotations of *help* and *build* in (66):[48]

(66) Clyde will help build our mechanical ferret.

That is, instead of first interpreting *build our mechanical ferret*, as one normally would, function composition would make it possible to interpret *help build* first while maintaining exactly the same denotations for all the terminal nodes in the sentence. The complex denotation of *help build* could than be applied to the denotation of *our mechanical ferret*, effectively 'postponing' interpretation of this argument. The order in which things are interpreted, then, would reflect a constituent structure like (67b) rather than (67a):

(67) a. Clyde will help [build our mechanical ferret].
b. Clyde will [help build] our mechanical ferret.

Assuming the denotations in (68), the result of function-composing *help* and *build* would be (69):

---

[48] For the sake of the example, I will simply skip over the interpretation of all sorts of functional structure that may be present in the extended projection of *build*. More generally, none of this is intended as an analysis of this construction, or to suggest that there is any particular merit of using function composition in interpreting it. This is all only to illustrate function composition in a way that bears some resemblance to what will be proposed below.

(68) a. $[\![\textit{help}]\!] = \lambda P_{\langle s,t \rangle}\ \lambda e\ .\ \exists e'[P(e') \wedge \text{help-to-happen}(e')(e)]$

b. $[\![\textit{build}]\!] = \lambda x\ \lambda e'\ .\ \text{build}(x)(e')$

c. $[\![\textit{help}]\!] \circ [\![\textit{build}]\!]$

$= \lambda x\ .\ [\![\textit{help}]\!]([\![\textit{build}]\!](x))$

$= \lambda x\ \lambda e\ .\ \exists e'[\text{build}(x)(e') \wedge \text{help-to-happen}(e')(e)]$

The result can be applied to *our mechanical ferret*:

(69) $[\![\textit{help build}]\!]$(our-mechanical-ferret)

$= \lambda e\ .\ \exists e'[\text{build}(\text{our-mechanical-ferret})(e') \wedge \text{help-to-happen}(e')(e)]$

This yields exactly the same denotation for *help build our mechanical ferret* as the normal procedure would have:

(70) $[\![\textit{help build our mechanical ferret}]\!]$

$= [\![\textit{help}]\!]([\![\textit{build}]\!]([\![\textit{our mechanical ferret}]\!]))$

$= \lambda e\ .\ \exists e'[\text{build}(\text{our-mechanical-ferret})(e') \wedge \text{help-to-happen}(e')(e)]$

The effect of function composition, then, is to 'postpone' the interpretation of an argument.

1.4.7 *Mediated Modification, Implementation II: The Feature-Bundling Version*

There is, however, a slight complication that should be overcome before we can make use of this—function composition has the effect of 'postponing' the interpretation of exactly one argument. But, in light of other assumptions about how arguments are introduced, this will not be enough. Sometimes, multiple arguments will have to be postponed

in this sense. If a feature and a transitive verb are interpreted by function composition, for example, exactly one argument would be postponed, namely, the object; but if that same feature and a *di*transitive verb were to be interpreted this way, it would be necessary to postpone *two* arguments, namely, both objects of the ditransitive verb. So things have to be a little more flexible.

One way to achieve this is to borrow the 'Geach rule' from Variable Free Semantics and categorial grammar, though in a much more limited way. This rule is, in Jacobson (2000)'s words, 'just a unary (Curry'd) version of function composition', and is expressed as the function g (the formulation here is Jacobson's):

(71) $g(\alpha) = \lambda V\ \lambda C\ .\ \alpha(V(C))$

for V of type $\langle c,a \rangle$ and C of type c

Applying g to A and then applying the result to B yields the same result as having function composed A and B in the more ordinary way:

(72) $g(A)(B) = [\lambda C\ .\ A(B(C))] = A \circ B$

This can provide a means of extending or generalizing function composition in a way that will provide the necessary flexibility. Applying g once has the effect of 'skipping over' one argument; applying it more than once, though, makes it possible to 'skip over' more than one argument. So in (73), one individual argument of B is skipped over; in (74), two individual arguments are:

(73) A is of type $\langle et,et \rangle$ and B is of type $\langle e,et \rangle$

a. $g(A) = \lambda P\ \lambda x\ .\ A(P(x))$

b. $g(A)(B) = \lambda x\ .\ A(B(x))$

(74) A is of type $\langle et,et\rangle$ and B is of type $\langle e,\langle e,et\rangle\rangle$

a. $g(A) = \lambda P\ \lambda x\ .\ A(P(x))$

b. $g(g(A)) = g(\lambda P\ \lambda x\ .\ A(P(x)))$

$= \lambda Q\ \lambda y\ .\ [\lambda P\ \lambda x\ .\ A(P(x))](Q(y))$

$= \lambda Q\ \lambda y\ \lambda x\ .\ A(Q(y)(x))$

c. $g(g(A))(B) = \lambda y\ \lambda x\ .\ A(B(y)(x))$

This effect is exactly what's needed here. But this particular way of thinking about it, in which g is a kind of type-shift that can be applied repeatedly, gives it the flavor of a grammatical operation of some sort, and particularly when interpreting the object language directly, one that might be expressed in the syntax and perhaps even in some language with overt morphology. This might, of course, be precisely the case, but in the current context it is for this reason at least expositorily awkward. To express the same idea of a more flexible kind of function composition via repeated geaching, the operation $\circ\circ$ in (75) will be used as a kind of shorthand to represent this more promiscuous cousin of function composition:

(75) $A \circ\circ B = g^n(A)(B)$

where g^n is n instances of applying g and n is the smallest integer ≥ 0 such that B is in the domain of $g^n(A)$

So $\circ\circ$ is unlike function composition in that it skips over as many arguments as it needs to. Intuitively, the effect is to peel off as many lambdas from B as necessary to have something of which A can be predicated, then prefix these lambdas to the result.

With this in place, feature bundles can now be interpreted by combining the denotations of the features with the $\circ\circ$ operation. More precisely, the rule for interpreting feature

bundles will be as in (76):[49]

(76) FEATURE RULE

If α is a terminal node bearing the features $[\phi_1]$ through $[\phi_n]$,

$$[\![\begin{bmatrix} \alpha \\ \begin{bmatrix} \phi_1 \\ \vdots \\ \phi n \end{bmatrix} \end{bmatrix}]\!] = [\![[\phi_1]]\!] \circ\circ \cdots \circ\circ [\![[\phi_n]]\!] \circ\circ [\![\alpha]\!]$$

This, then, requires the denotations of features to be combined with each other by the $\circ\circ$ operation starting at the top of the feature bundle and working toward the bottom, then combining the result by $\circ\circ$ with the denotation of the node bearing the feature bundle. This loosening up of the interpretive procedure is implemented very narrowly here, targeting only feature structures.[50] Above the word level, no new principles of interpretation will be available that weren't available before.

---

[49] It might be possible to simplify this rule. Because $\circ\circ$ is associative, it would be sufficient to assume that any immediately adjacent features in a feature structure can be interpreted via the $\circ\circ$ operation. Thus the rule might be something like:

(i) FEATURE RULE (ALTERNATIVE FORMULATION)
If $[\phi_1]$ and $[\phi_2]$ are features or collections of adjacent features,
$[\![\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}]\!] = [\![[\phi_1]]\!] \circ\circ [\![[\phi_2]]\!]$

An independent rule would be necessary to combine the features' denotation with that of the head that bears them:

(ii) FEATURE RULE (ALTERNATIVE FORMULATION): PART II
If α is a terminal node bearing the feature structure ϕ,
$[\![\begin{bmatrix} \alpha \\ [\phi] \end{bmatrix}]\!] = [\![[\phi]]\!] \circ\circ [\![\alpha]\!]$

I don't follow this course mostly because it's a bit more slippery conceptually, and because the simplification in (i) comes at the price of having to state (ii) separately.

[50] Indeed, this formulation is probably *too* restrictive. It may be useful to assume that something like this is available below X^0 more generally, as a specifically morphosemantic means of assembling meanings. A bit of relatively direct independent motivation for this comes from the suggestion of von Fintel (1995) that diachronically, the meaning of *many* developed by function-composing the meaning of a pre-existing

A crucial but unusual assumption underlying (76) is that the relative order of features in a feature structure is significant. This warrants some elaboration, and I will return to this in 1.4.8 after a brief illustration of how this machinery works.

The principal appeal of this implementation is that it can achieve exactly the same interpretation as the implementation without feature bundling, but with more minimal trees. Considering again the sentence that first illustrated the bundle-free approach, repeated in (77), a simpler syntactic structure can now be assigned:

(77) Floyd rudely talked to Herman's monkey.

For example, we could now place the [+AGENTIVE] feature that introduces the adverb directly on the verb rather than according it an independent phrase-structural location.[51] No further syntactic projections will be required (to achieve the ends sought here). As far as the semantics is concerned, the position the adverb occupies need only be at the right hierarchical level.

This can be achieved by adjunction—the course I'll adopt here—or alternatively, by supposing the adverb occupies a specifier position as before. The major disadvantage of this is that an independent explanation is required for why the adverb seems to necessarily occupy a left branch, and not merely a particular hierarchical position. Here one could appeal to e.g. Ernst (2002)'s linearization principles, or the Linear Correspondence Axiom (Kayne 1994). Of course, this would not be a disadvantage if it were desirable to suppose

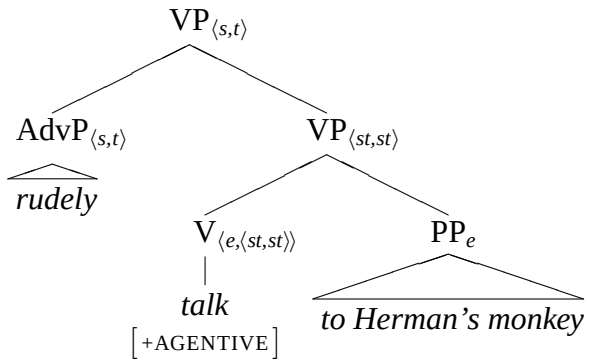
---

adjective *many* and a null determiner. If, as he proposes, this is a regular avenue of historical change, it would be somewhat odd in that it would be possible only when one of the morphemes it targets has precisely one argument to 'skip over', which doesn't seem to be a grammatically relevant property. But supposing that the relevant operation is ○○ instead would eliminate this restriction.

[51] In this sense, the Feature Rule in (76) actually introduces analytical possibilities not only by permitting feature bundles to be interpreted, but also by permitting features borne by semantically interpreted heads to be interpreted, too.

that certain modifiers introduced this way occupy left branches, and certain others don't—if distinct principles govern linearization, linear position could vary among these modifiers independent of hierarchical position. Whether this would in fact be desirable is, I think, not an easily resolved empirical question, though certainly one that merits continued inquiry. On the other hand, the advantage of taking the specifier approach would be that it ensures entirely on its own that the adverb will occupy a left branch. But in order to provide the adverb with a specifier position to occupy, it would be necessary to assume multiple specifiers, a move that, whatever its other merits, isn't consistent with the more conventional X-bar Theoretic assumptions that will be observed here.[52] Pursuing the adjunction analysis here, then, (77) can be assigned the structure in (78):

(78)



The crucial step in interpreting this structure is in (79). The Feature Rule combines the denotation of *talk* in (79a) and the denotation of the feature [+AGENTIVE] it bears in (79b)—importantly, both precisely the same denotations assumed earlier—to yield (79c):

---

[52] Of course, the adverb could also be placed in the specifier of a distinct functional projection—but this would eliminate the phrase-structural simplicity that is an important advantage of the feature-bundling approach.

(79) a. $[\![\textit{talk}]\!] = \lambda x_e\ \lambda e_s\ .\ \text{talk}(x)(e)$

b. $[\![[\text{+AGENTIVE}]]\!] = \lambda P_{\langle s,t\rangle}\ \lambda A_{\langle s,t\rangle}\ \lambda e_s\ .\ P(e) \wedge \exists s[A(s) \wedge$
Experiencer(s)=Volition(e)]

c. $\left[\!\!\left[\begin{matrix} \textit{talk} \\ [\text{+AGENTIVE}] \end{matrix} \right]\!\!\right]$
$= [\![[\text{+AGENTIVE}]]\!] \circ\circ [\![\textit{talk}]\!]$ (by the Feature Rule)
$= \lambda y\ .\ [\![[\text{+AGENTIVE}]]\!]([\![\textit{talk}]\!](y))$ (by the definition of $\circ\circ$)
$= \lambda y\ .\ [\![[\text{+AGENTIVE}]]\!](\lambda e_s\ .\ \text{talk}(y)(e))$
$= \lambda y\ .\ [\![[\text{+AGENTIVE}]]\!](\lambda e_s\ .\ \text{talk}(y)(e))$
$= \lambda y\ \lambda A_{\langle s,t\rangle}\ \lambda e_s\ .\ \text{talk}(y)(e) \wedge \exists s[A(s) \wedge \text{Experiencer}(s)=\text{Volition}(e)]$

So $\begin{matrix} \textit{talk} \\ [\text{+AGENTIVE}] \end{matrix}$ is a kind of augmented form of *talk*, in a way that corresponds even more closely to McConnell-Ginet (1982)'s vision. Simple *talk*, which has a straightforward transitive-verb denotation, and [+AGENTIVE], which is a function that applies to VP denotations, are combined by the Feature Rule so that $\begin{matrix} \textit{talk} \\ [\text{+AGENTIVE}] \end{matrix}$ will have a denotation that takes as its arguments first an individual—here, Herman's monkey—then an adverb denotation—here, $[\![\textit{rudely}]\!]$—and yields a VP denotation. In this sense, it is a form of *talk* that has been augmented to take the adverb as an additional argument, added after the internal arguments of *talk* have been saturated.

From this point, things proceed by functional application in the familiar way. Again assuming exactly the same denotations for *rudely* and *to Herman's monkey* as before, the VP is interpreted as in (80):

(80) a. $[\![\textit{to Herman's monkey}]\!]$ = Herman's monkey

b. $[\![\textit{rudely}]\!] = \lambda e_s\ .\ \text{rude}(e)$

c. $[\![\begin{matrix} talk \\ [+\text{AGENTIVE}] \end{matrix} \textit{ rudely to Herman's monkey}]\!]$

$= [\![\begin{matrix} talk \\ [+\text{AGENTIVE}] \end{matrix}]\!] ([\![\textit{to Herman's monkey}]\!])([\![\textit{rudely}]\!])$

$= \lambda e_s$. talk(Herman's monkey)(e) $\wedge$ $\exists$s[rude(s) $\wedge$

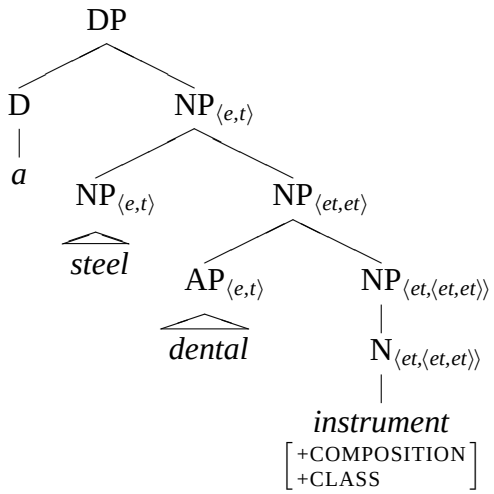
Experiencer(s)=Volition(e)]

This denotation is precisely the same denotation arrived at on the other implementation of the idea. Thus the largely the same mediated theory of modifier interpretation—and the fundamentally the same understanding of the link between modifier position and interpretation—is achieved here without requiring specialized projections.

An important feature of this system is that it establishes precisely the same semantic bond as the original implementation between a modifier and the feature that introduces it—as before, the modifier is effectively an argument of the feature. Thus for precisely the same reasons as before, there is no need for a syntactic mechanism (feature checking, say, as in Cinque 1999) to do this work.

And as before, the hierarchical order in which modifiers occur will mirror the order in which the features occur. Because of the way the Feature Rule and the $\circ\circ$ operation are defined, a modifier introduced by a feature higher in the feature structure will occur higher than one introduced by a feature lower in the feature structure. Returning now to one of the the two-modifier examples considered before, *a steel dental instrument*, it is now possible to assume a structure as in (81), again treating the modifiers here as adjuncts:

(81)



The Feature Rule will ensure that the *instrument* along with the bundled features it bears will be interpreted as in (82), as before maintaining the same denotations for the features and *instrument*:

(82) a. $[\![\textit{instrument}]\!] = \lambda x \,.\, \text{instrument}(x)$

b. $[\![[\text{+COMPOSITION}]]\!] = \lambda P_{\langle e,t \rangle}\ \lambda C_{\langle e,t \rangle}\ \lambda x_e \,.\, P(x) \wedge \exists y[C(y) \wedge x \text{ is composed of } y]$

c. $[\![[\text{+CLASS}]]\!] = \lambda P_{\langle e,t \rangle}\ \lambda A_{\langle e,t \rangle}\ \lambda x_e \,.\, \exists k_e[A(k) \wedge P(x) \wedge x \text{ realizes } k]$

d. $[\![\textit{instrument}_{\left[\begin{smallmatrix}\text{+COMPOSITION}\\ \text{+CLASS}\end{smallmatrix}\right]}]\!]$

$= [\![[\text{+COMPOSITION}]]\!] \circ\circ [\![[\text{+CLASS}]]\!] \circ\circ [\![\textit{instrument}]\!]$

(by the Feature Rule)

$= [\![[\text{+COMPOSITION}]]\!] \circ\circ [\lambda A_{\langle e,t\rangle}\ \lambda x_e\ .\ \exists k_e[A(k) \wedge \text{instrument}(x) \wedge$

x realizes k]] (by the definition of $\circ\circ$)

$= \lambda A_{\langle e,t\rangle}\ .\ [\![[\text{+COMPOSITION}]]\!](\lambda x_e\ .\ \exists k_e[A(k) \wedge \text{instrument}(x) \wedge$

x realizes k]) (by the definition of $\circ\circ$)

$= \lambda A_{\langle e,t\rangle}\ \lambda C_{\langle e,t\rangle}\ \lambda x_e\ .\ \exists k_e[A(k) \wedge \text{instrument}(x) \wedge x \text{ realizes } k] \wedge \exists y[A(y) \wedge$

x is composed of y]

Again, this is essentially an augmented lexical head, with new argument places introduced for modifiers. Crucially, the modifier corresponding to the higher feature is introduced first (the $\lambda A_{\langle e,t\rangle}$ above, which came from $[\text{+CLASS}]$ and will be saturated by the classificatory adjective) and the modifier corresponding to the lower feature is introduced later (the $\lambda C_{\langle e,t\rangle}$ above, which came from $[\text{+COMPOSITION}]$ and will be saturated by the composition noun:

(83) a. $[\![\textit{dental}]\!] = \lambda k_e\ .\ \text{dental}(k)$

b. $[\![\textit{steel}]\!] = \lambda y\ .\ \text{steel}(y)$

c. $[\![\textit{steel dental}\ \textit{instrument}_{\left[\begin{smallmatrix}\text{+COMPOSITION}\\ \text{+CLASS}\end{smallmatrix}\right]}]\!] = [\![\textit{instrument}_{\left[\begin{smallmatrix}\text{+COMPOSITION}\\ \text{+CLASS}\end{smallmatrix}\right]}]\!]([\![\textit{dental}]\!])([\![\textit{steel}]\!])$

$= \lambda x_e\ .\ \exists k_e[\text{dental}(k) \wedge \text{instrument}(x) \wedge x \text{ realizes } k] \wedge \exists y[\text{steel}(y) \wedge x \text{ is}$

composed of y]

As before, this leads to precisely the same denotation that the original implementation did.

1.4.8 *Bundled Feature Order*

As before, things hinge to a large extent on the order in which these features occur. On the original implementation, their ordering could follow from the same principles—or stipulations—as the ordering of functional heads comes from, and nothing special needed to be said. Here, though, what is relevant, in light of the way the Feature Rule is constructed, is the order of these features in a feature structure. This is a distinct notion. Yet, fundamentally, the order of features here is intended to express part of the same thing—it's simply a way of talking about the order in which they are interpreted. The only substantive difference is the shape of the trees. So it would be desirable not only for the order of features in a feature structure to mirror the order of features when distributed across different nodes, but also for this ordering to be derived the same way.

To the extent that the order of functional heads can be explained semantically—say, by facts about semantic selection broadly construed—nothing further would need to be said here with respect to the order of features in a feature structure, since their semantic relations are defined by the Feature Rule.

But a theory of functional head order that doesn't rest wholly on semantic explanation would not carry over so straightforwardly. Because such a theory is very likely to be built on some notion of c-command, though, we can remedy this by elaborating the definition of c-command so that some features can c-command others they are bundled with. More precisely, the precedence relation associated with a feature structure—graphically, the top-to-bottom order to which the Feature Rule appeals—can be understood to be itself a kind of c-command. The definition of c-command, then, could be extended (disjunctively, which isn't ideal) to include something like (84):

(84) If a feature $[\phi_1]$ precedes (graphically, is above) a feature $[\phi_2]$ in a feature structure, $[\phi_1]$ c-commands $[\phi_2]$

This adapts the likely foundation of any syntactic theory of functional head order to feature structures, and so should make it carry over directly. If, for example, such a theory that takes the form of a simple universal list of stipulations about possible c-command relations—say, that α must c-command β, and that β must c-command γ—these stipulations would ensure that if α and β are features, they can only be bundled in such a way that α is above β.

There is a different and more radical way of looking at this issue, though. It seems to be a fact about modification—and one that doesn't seem to be avoidable—that the syntactic category of an expression, and thus also the syntactic category of its head, at least partly determines what modifiers it may have. Thus it is the fact that *instrument* is a noun that permits it to have as modifiers, among many other things, classificatory adjectives and composition attributives and color adjectives. Put in terms of features, it is because *instrument* bears an [N] feature that it can bear the features [+CLASSIFICATORY] and [+COMPOSITION] and [+COLOR]. It might be desirable for this to follow from something. A particularly bold analytical move one might make here is to actually understand the [N] feature as a kind of abbreviation for features like [±CLASSIFICATORY] and [±COMPOSITION] and [±COLOR]—that is, for the features (at least the modifier-licensing features of interest here, but perhaps others as well) that collectively reflect the syntactic properties of a noun:[53]

(85) [N] abbreviates $\begin{bmatrix} \vdots \\ \pm\text{CLASSIFICATORY} \\ \pm\text{COMPOSITION} \\ \pm\text{COLOR} \\ \vdots \end{bmatrix}$

[53] The features are represented with ±values here to reflect that even when negatively valued, they are in some sense present, if only because a noun might have been lexically inserted with a positive value for that feature instead.

And of course, other categories, including both lexical categories like V and functional ones like T and Deg, might be understood in a corresponding way, as abbreviations for the modifiers they accept. This is, in a sense, a very direct way of encoding into the grammar the idea that syntactic categories should be defined by the syntactic environments in which they occur. As bold a move as this seems, it is not clear that it amounts to a rejection of any profound assumptions. The tradition of decomposing lexical categories into more basic features (e.g., understanding the category P to abbreviate $\begin{bmatrix}-\text{N}\\-\text{V}\end{bmatrix}$) has a long history in generative grammar, and this is certainly entirely consistent with that. Beyond that, as Baker (2003) notes, this has for the most part not been taken much further.[54]

This has an unexpected advantage that suggests there might be something right about it. One of the more important open questions currently in research into the syntax-semantics interface is to what extent the notion of syntactic notion of interpretability, which of course plays a central and in many ways driving role in contemporary syntax, can be identified with the semantic notion of interpretability. It would of course be a very interesting and highly desirable discovery for these to turn out to be the same. One roadblock in the program of identifying them, though, is that it often seems to be necessary for the syntax to refer to lexical categories as interpretable features—yet it is not obvious what meaning is contributed by simply being a noun, for example, apart from the meaning of individual nouns. The view (85) reflects, though, suggests an answer. Each of the features in (85) has, when positively valued, an interpretation (and when negatively valued, can be assumed to denote the identity function). It would follow then, that the D feature—that is, the collection of features that are characteristic of a determiner—would be interpretable in

[54] Though of course, Baker proceeds to do so, and arrives at a theory that *is* significantly different from the approach that (85) suggests.

the semantic sense on a determiner and not on a tense morpheme. This, of course, is the basis of the explanation of the EPP in English in which a subject position must be filled to check an uninterpretable D feature on T. It is normally necessary to stipulate that the D feature is uninterpretable on T; here, it would follow independently from the semantic types of the denotations of the features that the D feature abbreviates.

So at this point two largely distinct implementations of the larger idea have been suggested. Each has its advantages, and of course it's ultimately an empirical question which is preferable. Importantly, though, the second implementation is in some respects just a more flexible version of the first, and it can comfortably be 'mixed' with it. It may be the case—and indeed, seems likely—that languages might vary in how they group modifier-licensing features. Some languages may group together in a single node features other languages distribute among several, for example. The important thing is that on the second implementation, this is possible, but nothing that was possible on the first is ruled out. I regard this as an advantage, providing analytical flexibility that is probably necessary. But one could reasonably regard it instead as an empirically undermotivated weakening of the more restrictive first implementation, justified as it is mostly by the desire to prune trees. Either way, it does demonstrate that there is no necessary relation between the essential notion of mediated modification and phrase structural complexity.

1.4.9 *On Bigness and Smallness*

One of the very general observations noted earlier—and a rather vague one—is that there seems to be a loose correlation between the size of a modifier, in some sense of 'size' that needs elaborating, and the extent to which its syntactic position seems fixed. This is the observation that underlies Cinque (1999)'s distinction between adverbs and less rigidly ordered large modifiers like PPs (circumstantials, in his sense of the word).

While it will certainly not be possible to nail down a firm generalization here, some progress in this direction can be made by supposing that some modifiers are introduced via mediated modification (on either implementation) and others are not. This establishes a fundamental theoretical distinction that seems correlate with the apparent empirical distinction. Any modifier introduced via the mediation of a feature in the extended projection of the modified expression will necessarily be bonded to a position where that feature can 'find' it—that is, where a feature can take it as an argument. On the other hand, any modifier introduced through more conventional means will not be similarly restricted because it is not dependent in this way on the expression it modifies.

But why should being a 'small' modifier correlate with being introduced via mediation?

There is no necessary connection. Perhaps this is desirable, in the sense that the empirical correlation is neither clear or robust enough to warrant it. There is, however, a looser connection. Large modifiers—PPs and CPs—bring with them functional elements specialized for building modifiers. A PP like *with great malice*, for example, owes its modificational character to its preposition—while the core semantic content comes from *great malice*, the job of relating this content and a modified expression falls entirely to *with*. One can understand this as fact about types (or sorts); it is *with* that maps *great malice*, which on its own could not compose with a VP, onto a modificational type that can, say a function from VP denotations to VP denotations. In this respect, PPs are self-contained and hence more independent. They have internally the equipment they need to become modifiers.

APs and AdvPs, on the other hand, are more impoverished. To be sure, they (or their extended projections, to speak more strictly) have functional structure of their own—degree and agreement morphemes, for example—but the functional structure they have in

their modifier incarnation is precisely the same as they have as arguments.[55] What they lack is precisely the specialized equipment for creating modifiers that PPs have.

It is, of course, not always clear that they need it—indeed, this is clearer for nominal modifiers—but this ultimately beside the point. APs and AdvPs may well be just as willing to be introduced in a non-mediated way as PPs are. What is crucial is that larger modifiers like PPs be less willing to be introduced in a mediated way. And, in light of the fact that PPs are self-contained modifiers, this is likely to be so, for the simple reason that they don't *need* mediation. Relying on a mediating feature would, for a PP, generally be redundant. These mediating features are, after all, in a sense reverse prepositions—rather than helping an expression be a modifier, as prepositions do, they help an expression accept a modifier. And, if one respect in which language is economical is that it abhors overkill, it seems reasonable to expect that PPs and CPs be less likely to be introduced via unnecessary mediation. To put it metaphorically, PPs can pry their way into semantic composition on their own with the brute force of prepositions, so they don't need to resort to slipping in through cracks afforded them by the modified expression itself.

This is actually a version of the familiar intuitive quasi-functionalist story about why languages with rich system of overt morphological case are less likely to insist on rigid word order and permit more scrambling—just as on this story, arguments in such languages don't need to rely on strict structural positions to convey their thematic roles because they express that information internally, so too in this domain, PPs don't need to rely on strict structural positions to convey the sort of modification they achieve because they express that information internally.

---

[55] Of course, it's open to question whether AdvPs are really *ever* arguments; but then everything that's said here could apply equally to modifier DPs, which are discussed at length in subsequent chapters.

This, though, is probably at best a part of the story. It explains why larger modifiers may be able to be interpreted (in the same way) in more positions than smaller modifiers, but it doesn't address why syntactic movement should be less likely to target smaller modifiers than larger ones. This is an interesting question, but one beyond the scope of the current discussion. It's worth noting, though, that there is a longstanding but probably insufficiently well-examined assumption in generative grammar that adverbs in particular don't move. Early on, there was no such conventional wisdom, and indeed Keyser (1968) proposed rules governing what he called adverb 'transportability'. So while I'm in good company in setting this issue aside, it's one that warrants further exploration. There are, though, some reasons for the relative immobility of particular smaller modifiers that emerge more straightforwardly—as subsequent chapters attest, DP modifiers, for example, are not mobile, but this is to be expected for both type-theoretic and case reasons.

1.4.10 *A Word About Learnability*

There is an independent argument from learnability for supposing that some modifiers are introduced in a mediated way.

An important part of the traditional distinction between functional and lexical categories is the notion that functional categories are a kind of glue that holds together the syntactic and semantic building blocks provided by members of lexical categories. The semantic part of idea is naturally expressed in a corresponding type-theoretical distinction—generally, lexical categories have simple, first-order denotations, and functional categories have more complicated, higher-order ones.

This can't be coincidence. There are relatively few functional morphemes in any language—and hence relatively few to be learned. And what functional morphemes there

are in a language usually manifest clear correspondences to those in others—and hence more about them is predictable and plausibly innate, and less must be learned. In these respects, there are independent facts that ease the burden of learning the inventory of functional morphemes in a language. Lexical categories, on the other hand, provide neither of these benefits. They are open classes, and their meaning can vary from language to language in profound and utterly unpredictable ways. So, in light of the fact that they are indeed learned despite their productivity and unpredictability, they must be *inherently* easier to learn than functional categories can be. This is precisely what restricting them to lower-type denotations accomplishes. A learner will be able to profit by this restriction by eliminating from consideration any possible meanings that violate it. Functional categories, though, don't need to be easy to learn in this inherent way, and thus can bear the price of not observing this restriction.

This presents a profound challenge to any analysis of a modifier as an expression with a lexically higher-order type, such as any predicate-modifier (i.e., operator, non-first-order) account of adverb or adjective semantics. If adjectives and adverbs as a class are to be learnable, it would be of profound help for them to be severely constrained type-theoretically. If all adjectives and adverbs—or at least the vast majority of them—could be restricted to first-order types, this would pose less of a learnability puzzle. But non-first order denotations for adjectives and adverbs seem unavoidable, as classic examples like *alleged* reflect for adjectives and e.g. quantificational adverbs show.

A mediated approach, though, may at least mitigate these problems. Because it involves a kind of decomposition of modifier meaning into two parts—the part contributed by the modifier proper and the part contributed by its position, that is, by the feature that licenses it—the higher type meaning may be able to be factored out from the denotation of the modifier proper and attributed instead to the licensing feature. This would accomplish

the goal, since it would shift the type-theoretic heavy lifting from a lexical category onto a functional one. Some of this will be illustrated in passing in subsequent chapters, but—to take a cursory glance at a different place this might turn out to be possible—the adverbs of quantification in (86), all (quantificationally) weak, raise this sort of problem:

(86) a. On Fridays, Chester is very *often* drunk.

b. Floyd quite *frequently* talks to Herman's monkey.

c. George rather *rarely* reads a newspaper.

All of these occur here with degree words—*very, quite* and *rather*—of precisely the same sort as occur in non-quantificational AdvPs and APs. This in itself suggests that the these adverbs aren't of a different type than non-quantificational adverbs, since the alternative would be to suppose that all degree words are systematically ambiguous between versions that apply to adverbs of quantification and ones that do not. This would suggest that these adverbs simply denote properties, which is natural enough in light of them being weak. But how then to compose them with the verbal projection they modify in a way that adequately captures their quantificational effect? Simple intersection wouldn't seem to work, at least not without something more being said. But if these AdvPs were instead understood to be introduced by a particular licensing feature that does the additional semantic work, this problem could more easily be sidestepped. This probably couldn't be done for strong quantificational adverbs, but then those are more plausibly analyzed as functional heads themselves. Indeed, this may be the case for many unavoidably higher-type adverbs and adjectives; cf. Travis (1988) with respect to adverbs and Abney (1987) with respect to adjectives.

Even so, it doesn't seem likely that this could ultimately eliminate *any* need for non-first-order types for certain particularly intransigent (kinds of) adjectives and adverbs

in a straightforward sense. But if the problem could be beaten back sufficiently in other respects, independent explanations may be able to overcome what remains. For example, intensional adjectives like *alleged* or *apparent* indeed do not seem to submit readily to the decomposition into lower- and higher-type parts envisaged here. But neither are these staples of most young children's vocabularies. To be less glib, these may be 'peripheral' in a meaningful sense and not part of the core of the language children acquire.[56] They seem like good candidates for what Sobin (1997) has termed 'grammatical viruses': apparent elements of the grammar that are in fact generally learned very late and consciously, and about which most speakers do not have consistent, effortless, and fully confident intuitions. One clear case of such a virus Sobin points to is the 'rule' of English that the overt morphological case of conjuncts must correspond to the case assigned to the whole conjoined DP. Some speakers do indeed claim a preference for (87a) over (87b), for example, and certainly there is a register difference (with (87a) more formal than (87b)):

(87) a. Floyd and I get indignant about inane prescriptivism.
b. Me and Floyd get indignant about inane prescriptivism.

So while this 'rule' is a quasi-systematic feature of English, and one which it may be socially advantageous to be aware of, it is not really part of the grammar. Something similar might be said for the form *whom*, as Lasnik and Sobin (2000) propose. It might be the case that intensional adjectives fall into this class too—that they are semantic viruses, burdened with meanings we conventionally ascribe to them by a kind of collective conscious decision that it takes continual effort to uphold. That these words are acquired late and associated

---

[56] This isn't true of privative adjectives like *fake*, but, as Partee (2001) suggests, these may not be truly intensional (or truly 'privative', really).

with fairly formal registers, and usually written or careful speech, supports this view, as do our sometimes unnervingly unsteady intuitions about their precise meaning and felicitous use (if I, profoundly confused, claim that Castro is king of New York, can Castro then be properly characterized as *the alleged king of New York*?). So, since these adjectives may require swimming upstream against the grammar and may not be strictly speaking part of it, it may be possible to discount them as counterexamples to the claim that all adjectives—in the theoretical sense, that is, As—are first-order.

Thus learnability considerations lead us to expect lower-type denotations for (lexical) modifiers generally, and assuming this kind of mediated approach in many cases advances us toward this goal.

1.4.11 *Possible Constraints*

As it stands, though, this mode of explanation loosens the constraints on the grammar—opens up more analytical possibilities, to put it more positively—but does not offer new constraints to reign things in. It's probably the case that, if any of this is at all on the right track, not enough is known about the sorts of modifiers and modifier interpretation effects of concern here to venture many guesses or to propose anything too confidently. But some initial ideas of some meaningful possible constraints to present themselves.

TYPES One kind of constraint is simply type-theoretic, and flows naturally both from the learnability discussion in the previous section and from the logic of the proposal itself. First, since modifier licensing features would only seem to be licensing modifiers (rather than doing other, independent work) if they ultimately yield objects of the same semantic type as they initially apply to, a constraint like (88) might be adopted:

(88) Modifier features that are positively valued must be of type $\langle\alpha,\langle\beta,\alpha\rangle\rangle$, where α is the type of the modified expression and β is the type of the modifier

This may, in fact, be able to serve as a definition of a modifier-licensing feature—though interpreted that way, it could hardly also be viewed as a constraint on them. Second, as already suggested, this approach provides a way of dividing the labor of modification by separating the type-theoretic heavy lifting from the lexical semantics of the modifier. Elevating this to a principle, one might adopt a constraint like (89) as a supplement to (88):

(89) A modifier introduced by a modifier feature must have a first-order denotation.

This, then, would impose on (88) the additional requirement that β be first order. It would be even more appealing—but also harder to sustain—to suppose that β must actually be a property.

STRUCTURAL LOCATION All the features suggested so far, even tentatively, have been borne by the modified expression, not the modifier, and in this way linked modifiers with particular positions. It might have been otherwise—modifiers might have borne their own modifier features. Again, this may be more a definitional point than a substantive constraint, but stating it officially even so:

(90) Modifier features are interpretable on some head in the extended projection of the modified expression.

NEGATIVELY-VALUED FEATURES If these features are to be implemented in the plus-or-minus fashion, as the discussion in 1.4.8 suggests it should be and which accords with

Cinque's approach, it will be a consistent property of the system—and thus potentially a constraint on it, a principle—that negatively valued features do *not* introduce modifiers, and in fact contribute no semantics of their own at all:

(91) Negatively valued features denote the identity function.

FEATURES CONTRIBUTE PRESUPPOSITIONS Maybe the most interesting of these possibilities is a constraint that accords with something that has independently been proposed about the interpretations of various features. As Schlenker (2002b) notes, numerous features have been treated as contributing presuppositions, including ϕ-features such as gender (Cooper 1983), tense features (Heim 1994), and mood (von Fintel 1997, Schlenker 2002a). Abusch (1997) also suggests this kind of presuppositional view of tense, and Sauerland (2002, and elsewhere) has elaborated and generalized this understanding. It is therefore a satisfying result that all of the modifier feature denotations entertained here—and, indeed, in subsequent chapters—seem to contribute, at heart, only presuppositional meaning.

They do, of course, perform operations that map one type to another, and this aspect of meaning probably shouldn't be thought of as presuppositional.[57] But it is also an essentially combinatoric kind of meaning, and in this respect not quite the same as the more contentful kind of meaning one normally intends when using 'meaning' in its ordinary sense. The [+CLASS] feature was earlier given the denotation in (92a), but this might more accurately be rendered as (92b):

(92) a. $[\![[\text{+CLASS}]]\!] = \lambda P_{\langle e,t\rangle}\ \lambda A_{\langle e,t\rangle}\ \lambda x_e\ .\ \exists k_e [A(k) \wedge P(x) \wedge x \text{ realizes } k]$

b. $[\![[\text{+CLASS}]]\!] = \lambda P_{\langle e,t\rangle}\ \lambda A_{\langle e,t\rangle}\ \lambda x_e\ .\ \exists k_e\text{: } x \text{ realizes } k\ [A(k) \wedge P(x)]$

---

[57] Though of course the restrictions on what types their arguments may have are presuppositions.

That is, the only contentful non-combinatoric meaning meaning [+CLASS] has is that there is a kind of which its individual argument is a realization—and this is naturally understood as a presupposition. The other features encountered in this section similarly seem to contribute only presupposed non-combinatoric meaning, and their denotations could be similarly reformulated to reflect that:[58]

(93) a. $[\![[\text{+COMPOSITION}]]\!] = \lambda P_{\langle e,t\rangle}\ \lambda C_{\langle e,t\rangle}\ \lambda x_e\ .\ P(x) \wedge \exists y[C(y) \wedge x \text{ is composed of } y]$

b. $[\![[\text{+COMPOSITION}]]\!] = \lambda P_{\langle e,t\rangle}\ \lambda C_{\langle e,t\rangle}\ \lambda x_e\ .\ P(x) \wedge \exists y\text{: } x \text{ is composed of } y\ [C(y)]$

(94) a. $[\![[\text{+AGENTIVE}]]\!] =$

$\lambda P_{\langle s,t\rangle}\ \lambda A_{\langle s,t\rangle}\ \lambda e_s\ .\ P(e) \wedge \exists s[A(s) \wedge \text{Experiencer}(s)=\text{Volition}(e)]$

b. $[\![[\text{+AGENTIVE}]]\!] =$

$\lambda P_{\langle s,t\rangle}\ \lambda A_{\langle s,t\rangle}\ \lambda e_s\ .\ P(e) \wedge \exists s\text{: Experiencer}(s)=\text{Volition}(e)\ [A(s)]$

So it might be fruitful to view contributing only presupposed non-combinatoric meaning as a constraint on what (modifier) features can mean:

(95) Any meaning a modifier feature contributes that is not purely combinatoric is a presupposition.

There is, of course, a certain potential slipperiness in the notion of 'purely combinatoric' meaning, but the essence is clear enough: meaning of the sort that one might expect a type-shift to have, for example, and of the sort that members of lexical categories do not.

---

[58] I exclude the [+COLOR] feature here only because, as formulated in (59), it already reflects (if obliquely) that it is a presupposition that its individual argument has a unique principal color through its use of a PRINCIPAL-COLOR function that would, of course, not be defined in the absence of (a unique) one.

The flip side of these observations, though, is that all of the features considered here even in passing *do* have a meaning that is not purely combinatoric. None of them are simply recipes for assembling denotations, essentially idiosyncratic composition rules or type shifts encoded directly into the object language and associated with particular positions. Each contributes something else, something beyond this sort of recipe, some general but non-trivial element of meaning of the same flavor as the meanings of other functional morphemes. So (95) could perhaps be supplemented with the requirement in (96):

(96) Every modifier feature contributes a presupposition.

1.4.12 *Summary*

The aim of this section was to articulate a broad, consistent understanding in which the recurring analytical intuition discussed in 1.3 can be expressed. The core of this understanding is the notion that the interpretation of certain modifiers is to be understood as a two-part affair, uniting the lexical semantics of the modifier and a distinct semantics contributed by the position itself. This distinct semantics is attributed to semantically-interpreted features in the extended projection of the modified expression that introduce modifiers as their arguments, creating various 'slots' in fixed positions into which modifiers can be plugged. This view embodies many of the theoretical desiderata. It provides a way of associating modifiers, or really modifier interpretations, with positions in a semantically-based way without requiring that this effect be derived only from the types of the expressions involved—but, importantly, neither does it preclude these sorts of essentially ontological explanations, where they are possible. In this respect, it permits some facts about

modifier order to be truly ‘syntactic’ rather than purely semantic. It permits simplifying the types of modifiers by shifting higher-type work onto functional elements, which has learnability advantages. It sheds some light on the rough connection between being a small modifier and manifesting a particularly close relationship between position and interpretation. It is powerful enough to supply the kind of consistent understanding needed, but not so powerful as to be unconstrainable—and indeed, potential constraints suggest themselves quite readily. And of course, it captures the notion of semantic enrichment or augmentation at the core of the varied proposals considered in 1.3.

1.5 A Final Word and a Glance Ahead

This chapter adopted a kind of top-down mode of argumentation, taking a big-picture view of a very broad empirical area. It approached a heterogeneous but related set of puzzles by what unites them, working from the general puzzle to the specific ones rather than vice versa. If there is anything right about this, it should be possible to make the same argument in the other direction as well. So in the remaining chapters, the methodology is precisely the opposite. Each chapter considers a particular specific semantic puzzle involving modifier position, and each focuses first on developing an analysis of the puzzle on its own terms, and only secondarily on exploring this broader theoretical approach. In this respect, subsequent chapters proceed bottom-up.

They also proceed bottom-up—or perhaps better, from the outside in—in another sense. They each examine phenomena that aren’t among the most obvious or straightforward examples of the sorts of effects discussed here. Part of the merit of taking this course with respect to the argument made in this chapter is that any case that can be made for it from these less obvious examples would constitute more nearly independent support. And

to the extent that this approach can meet with some success in addressing puzzles that are relatively mysterious, it may be an indication of more general usefulness.

Another reason for examining these particular puzzles, though, is simply that they have received very little attention, and in this respect warrant examination irrespective of whether they support any particular larger conclusions about how modification works. And indeed, this is how things will proceed in each chapter. Each follows an analytical path more or less in its own way and justified only by the particular facts of the phenomenon being examined, and each develops in isolation of the others, without explicitly relying on any of them—or for that matter, on anything said in this chapter, either.

Because each chapter is standalone, there is a certain nontrivial amount of redundancy from chapter to chapter. For the same reason, certain background analytical assumptions not directly relevant to the argument at hand shift slightly from chapter to chapter as convenient.

Chapter 2 examines how evaluative adverbs such as *remarkably* or *disturbingly* work in the adjectival projection, in e.g. *remarkably tall*, and how these uses relate to other uses of these adverbs. The semantic account proposed hinges on domain widening with respect to degrees. Chapter 3 turns to measure DP adverbials such as *twenty minutes* in e.g. *Floyd sang twenty minutes*. Among the main theoretical puzzles with respect to this is how to account for the obligatory low scope of these expressions, and for how they differ from their PP paraphrases. Chapter 4 is concerned with modifiers such as *almost*, *virtually*, and *nearly*, which have a cross-categorial distribution that presents a puzzle for any account of their semantics. Assumptions like those advanced in this chapter, it will turn out, can be helpful in accounting for this cross-categoriality. Despite the important role *almost* has played and continues to play in linguistic theory, neither its semantics nor the question of how it relates to other members of its class have received much attention, so these questions

are further probed as well. Chapter 5 focuses on the grammar of adjectives such as *whole* and *entire*, and chiefly on providing an account of what their semantic contribution is that does not impute to them universal quantificational force they don't seem to have. These adjectives occupy a relatively fixed position inside the DP, and their semantics is very similar and perhaps actually identical, so they might seem like natural candidates for the kind of analysis advanced here. But this turns out to be harder to motivate for these than for the modifiers considered in other chapters.

Though these topics are all examined ultimately for their own sake and not for what larger theoretical goals they might serve, the essential idea presented in this chapter is a unifying theme that ties these distinct studies together. It's worth noting that this idea reflects a certain methodological gamble. What it presupposes—with some justification, but even so—is that modifier position-interpretation correlations of the relevant sort do indeed require an explanation in similar terms as a kind of unified phenomenon, and that an explanation will not ultimately emerge on its own for all such correlations from existing assumptions about modification along with a sufficiently articulated understanding of the semantics of the various constituents modified. Perhaps pursuing each of these empirical puzzles independently might support this presupposition, at least to the extent that they all converge on some similar points.

By elevating the notion of mediated modification to a grammatical principle—by committing to a particular understanding in which expressions must be semantically 'primed' to support certain kinds of modification using truth-conditionally meaningful features just as walls must be primed to support certain kinds of ornamentation using hooks—this is where this approach places its bets. This may, of course, be entirely wrong. But either way, there is no reason a priori why assumptions about modification must remain constant, so experimenting with changing them—in any way, really, even if not in this one—should prove fruitful or at least revealing.

CHAPTER 2

EVALUATIVE ADVERBIAL MODIFICATION IN THE ADJECTIVAL EXTENDED PROJECTION

2.1 Introduction

One of the principal analytical challenges of adverbial modification is how to account for the intricate and often subtle correlation between an adverb's syntactic position and its interpretation. Why, to consider one familiar class of examples, should subject-oriented readings be associated with an intermediate position in the clause, as in (1a)? Why should manner readings be associated with a relatively lower position, as in (1b)? Why should speaker-oriented readings be associated with a higher position, as in (1c)?[59]

(1) a. Clyde would happily play his tuba.
rough paraphrase: 'Clyde would be happy to play his tuba.'

b. Clyde would play his tuba happily.
rough paraphrase: 'Clyde would play his tuba in a happy way.'

c. Happily, Clyde would play his tuba.
rough paraphrase: 'I'm happy Clyde would play his tuba.'

---

[59] These sentences are patterned after some examples of Jackendoff (1972). The paraphrases here reflect the most natural readings of these sentences, but these are not, of course, the only ones possible.

Attempts to grapple with these issues—from Jackendoff (1972) and McConnell-Ginet (1982) to Cinque (1999) and Ernst (2002), among others—have typically focused on adverbial modification in the verbal and sentential domain, from which the paradigm in (1) is drawn. This is of course no accident. These positions are, after all, the prototypical ones for adverbs. Even so, adverbial modification can be found elsewhere as well—in English and many other languages, it can also occur in the extended adjectival projection. Importantly, the interpretation adverbs receive in these less understood 'ad-adjectival' positions varies predictably from the one they receive elsewhere. Because of this, adverbial modification in the extended adjectival projection may offer an avenue not often taken for the exploration of this larger problem.

This chapter examines one large natural class of such AP-modifying adverbs, which have a kind of evaluative interpretation and include *remarkably*, *surprisingly*, and *breathtakingly*, among many others. The central analytical proposal will be that these adverbs are interpreted as arguments of unrealized degree morphology in the functional structure of the APs they modify, in much the same way as measure phrases have been proposed to be. This approach turns out to extend naturally to uses of these adverbs in other positions.

Section 2.2 identifies the class of adverbs of interest here and explores its distinguishing characteristics. Section 2.3 develops an analysis of the semantics of sentences containing *remarkably* adverbs based in part on a notion of domain widening in the degree domain, assimilating them to certain exclamatives. Section 2.4 confronts problems of compositionality these adverbs pose, and arrives at a kind of decomposition in which part of the interpretation of a *remarkably* adverb is contributed by its lexical semantics and part is contributed directly by its place in the architecture of the extended adjectival projection. Section 2.5 sketches how these syntactic and semantic assumptions can be the foundation of a more general theory of how the meaning of these adverbs is related to the meaning they have in other structural positions. Section 2.6 concludes.

2.2 *Remarkably* Adverbs

2.2.1 *The Cast of Characters*

If, in encountering Clyde, I was struck by his height, I can report this impression in a number of ways. I might simply say that he is tall; alternatively, I might make a slightly stronger claim and say that he is very tall; or I might instead be more precise and say that he is six and a half feet tall. All of these strategies have in common that they convey this information by indicating the relation between, pre-theoretically, points aligned vertically—either the maximal point of Clyde's height and some minimal height one must attain to count as tall,[60] or else the maximal point of Clyde's height and some zero point at which his height begins. These are all extensional characterizations of Clyde's height.

In contrast, the adverbs of interest here—henceforth *remarkably* adverbs—provide a fundamentally different, intensional means by which to comment on Clyde's height. They characterize Clyde's height not in terms of the relation between points in the actual world, but rather in terms of my attitude toward Clyde's height:[61]

(2) a. Clyde is *remarkably* tall.

b. Floyd is *surprisingly* ugly.

c. Many voters are *horribly* conservative.

d. Floyd's SUV is *unacceptably* inefficient.

e. Tranquility is *heart-breakingly* difficult to attain.

f. Self-referential example sentences are often *unpleasantly* distracting.

---

[60] More precisely, some minimal height one such as Clyde must attain in the relevant context.

[61] If the adverb receives parenthetical intonation, these sentences can have a reading other than the one at issue here. This reading, discussed a bit more in the next section, is the same reading as the one these adverbs receive in higher, clausal positions, as in (5).

Very roughly, *remarkably* adverbs might be said to have a semantics that gives rise to a judgment about having a property to a particular degree—that it is, say, remarkable or surprising or horrible.

This class of adverbs is quite large—indeed, it is an open class. Among its many other members are *amazingly, astoundingly, arousingly, calmingly, disappointingly, earth-shatteringly, excitingly, extraordinarily, frighteningly, grotesquely, heart-breakingly, horribly, illegally,*[62] *impressively, inappropriately, inconceivably, infuriatingly, interestingly, irritatingly, laughably, mind-numbingly, nauseatingly, provocatively, revoltingly, ridiculously, satisfyingly, shockingly, stimulatingly, stunningly, sufficiently, terribly, terrifyingly, typically, (un)acceptably, unbelievably, unexpectedly, unnervingly, (un)pleasantly, (un)remarkably, unusually, upsettingly, uselessly,* and *wonderfully*.

New adverbs of this class can be coined quite easily. It is perfectly natural, for example, to characterize particularly uncomfortable shoes with a neologism like *foot-shatteringly*:

(3) How can you wear those things? They look *foot-shatteringly* uncomfortable.

This seems to be the case even in coinages without relatively transparent internal structure. If we accept a novel adjective *blarg*, it's quite natural to coin a corresponding *remarkably* adverb *blargly*:

(4) Those things look *blargly* uncomfortable.

It seems important, though, that *remarkably* adverbs seem to be dependent on corresponding adjectives in this way, so much so that coining a novel *remarkably* adverb seems

---

[62] *Illegally* in particular gives rise to interesting and especially clear semantic differences in various positions it occupies, as Rawlins (2003) shows.

to entail having coined a corresponding adjective. This is the case even when that adjective hasn't been explicitly uttered. If uttered out of the blue, (3) seems to be a simultaneous coinage not only of the new *remarkably* adverb *foot-shatteringly* but also of a new adjective, *foot-shattering*; to the extent that one can imagine making sense of (4) out of the blue, it seems to have the same property. Indeed, there does not seem to be any *remarkably* adverb without a corresponding adjective.

This connection holds semantically as well. The meaning of an *remarkably* adverb and that of its corresponding adjective stand in a fixed relation, and one is always predictable from the other. If we know what *foot-shattering* means, we also know what *foot-shattteringly* means in (3), and vice versa.

2.2.2 *Contrast with Clause-Modifying Uses*

These adverbs can occur high in a clause-modifying position as well, where they receive a different reading entirely:[63]

---

[63] In this clausal position these are speaker-oriented evaluative adverbs, adopting the adverb taxonomy of Ernst (2002) (Cinque 1999 refers to these as simply 'evaluative adverbs'). There do not seem to be any *remarkably* adverbs that occur in clause-modifying positions as speaker-oriented speech-act adverbs such as *frankly* or *honestly*.

Of course, there is a sense in which the reading which *remarkably* adverbs receive can be characterized as speaker-oriented, in that it can reflect a judgment made by the speaker. Like speaker-oriented adverbs (proper), it's not normally possible to use a *remarkably* adverb to indicate a judgment made by the addressee with which the speaker disagrees—this is certainly the case for all the sentences in (1) and (5). But unlike true speaker-oriented adverbs, *remarkably* adverbs can in intensional contexts receive an interpretation in which the judgment they reflect is made by the holder of an attitude:

(i) a. Greta thinks that Clyde is remarkably tall.
b. Harriet suspects that Floyd is surprisingly ugly.

If what Greta thinks is that Clyde is six feet tall but does not regard being six feet tall as remarkable, (ia) could be true if the speaker regards being six feet tall as remarkable. It could also be true irrespective of what the speaker regards as remarkable if Greta's only thought about how tall Clyde is is that however tall he is, being that tall is remarkable. In contrast, it is not clear that (ii) is even grammatical:

(5) a. *Remarkably*, Clyde is tall.

b. *Surprisingly*, Floyd is ugly.

c. *Horribly*, many voters are conservative.

d. *Unacceptably*, Floyd's SUV is inefficient.

e. *Heart-breakingly*, tranquility is difficult to attain.

f. *Unpleasantly*, self-referential example sentences are often distracting.

With the adverb in this position, no judgment is being rendered specifically about having a property to any particular degree. Rather, to characterize things very crudely for the moment, the judgment in these sentences is about the proposition expressed by the sentence as a whole. These readings are truth-conditionally distinct—if Clyde is a professional basketball player and therefore expected to be very tall, *Clyde is remarkably tall* could be true while (5a) could be false.[64] It is not possible to construe any of the sentences in (5) as having the interpretations of their counterparts in (1).

Just as the meaning of a *remarkably* adverb is predictable from its adjective counterpart, the meaning of sentences like those in (1) is predictable from their counterparts in (5). And, as with adjectives, this fixed semantic relation is apparent in neologisms. Returning

---

(ii) a. *?Greta thinks that remarkably, Clyde is tall.
b. *?Harriet suspects that surprisingly, Floyd is ugly.

To the extent that these can be made sense of, the judgment of remarkability or surprisingness is attributed to the speaker (unless the embedded clause is taken to be in some way quotative). In a nutshell, then, *remarkably* adverbs can have *de dicto* readings but their clausal counterparts cannot. This difference is significant, but for current purposes it is sufficient to note that the meaning of *remarkably* adverbs should not be regarded as speaker-oriented in the same sense as the meaning of their clausal counterparts is.

[64] Perhaps (5a) would better be characterized as infelicitous rather than false. This is because clausal uses of these adverbs contribute something other than run-of-the-mill straightforwardly truth-conditional meaning; rather, their contribution seems to be a variety of conventional implicature (Grice 1975, Potts 2003) or expressive meaning (Kratzer 1999b and references therein). While this is certainly an important difference between *remarkably* adverbs and their clausal counterparts, it will generally be safe to disregard it here for convenience.

to *foot-shatteringly*, once we accept it as a *remarkably* adverb as in (3), we expect (6) to have a particular interpretation paralleling (5) in the relevant respect:

(6) ?Foot-shatteringly, they look uncomfortable.

In this case, assigning this interpretation results in (mild) pragmatic anomaly, yet this anomalous interpretation is the only one available (6).[65] In particular, it cannot mean what (3) means, even though interpreting it this way would yield a non-anomalous interpretation. It seems clear, then, that these semantic patterns reflect robust, apparently exceptionless grammatical regularities.

2.2.3 *Restricted Distribution*

The distribution of *remarkably* adverbs is quite narrowly restricted. They can occur only at the left edge of (the extended projection of) AP. In particular, they cannot occur in right-peripheral positions:

(7) a. *Clyde is tall remarkably.
b. *Floyd is ugly surprisingly.
c. *Many voters are conservative horribly.

To the extent that one might be able to salvage sentences like these, it is necessary to provide comma interpretation to the adverb. Doing so, though, results in interpretations

---

[65] Actually, it might be better to say that there is actually a *family* of interpretations available for (6), all anomalous. One might imagine interpreting *foot-shatteringly* as a speech-act adverb like *frankly*, for example; but this wouldn't help.

like those in (5). By assigning this comma intonation, one is apparently rendering the adverb a kind of parenthetical. In this respect, it is not surprising that in these cases any available interpretation parallels the reading available in higher positions, since irrespective of their position parentheticals more generally receive (something like) wide-scope interpretations (Potts 2003, others). Accordingly, this strategy for salvaging uses like those in (7) is unavailable in sentences in which this flavor of speaker-oriented adverb cannot occur:

(8) a. *Remarkably, how tall is Clyde?

b. *How tall remarkably is Clyde?

Speaker-oriented adverbs cannot occur in questions, as (8a) reflects, so (8b) cannot be rescued by interpreting *remarkably* this way.

Remarkably adverbs must precede certain other AP-modifying adverbs:[66]

(9) a. George seems remarkably intellectually inadequate.

b. *George seems intellectually remarkably inadequate.

(10) a. How surprisingly socially awkward is Herman?

b. *How socially surprisingly awkward is Herman?

---

[66] Specifically, *remarkably* adverbs must precede AP-modifying adverbs with a domain adverb (Ernst 2002, Rawlins 2003; under different names, also Bartsch 1976, Moltmann 1997) interpretation:

(i) a. How inadequate intellectually is George?
b. How awkward socially are many semanticists?

Again, there exists a parenthetical rescue strategy involving comma intonation that is possible in (9); and again, that route is cut off in (10).[67]

2.2.4 *Not Degree Words*

One natural analytical impulse is to suppose that *remarkably* adverbs are in fact a species of degree word (that is, of Degree head; I will use these interchangeably), like *too*, *very*, *pretty*, or comparative morphology. But does not seem to be the right approach, for several reasons.

Perhaps the clearest of these is that, unlike degree words, *remarkably* adverbs support degree words of their own:[68]

---

[67] This time for different reasons—the clausal counterpart of *socially* is possible in questions, but parentheticals are apparently not possible immediately following *how*.

[68] These may be marginally possible for some speakers on a reading in which a property is ascribed to the proposition expressed by the whole sentence, paralleling the interpretation of clausal uses of these adverbs in (5). This appears to be the consequence of interpreting the adverb as a parenthetical.

Another potential complication for some speakers with respect to all the judgments in this section is metalinguistic comparatives. As with most other syntactic categories, these are possible with *remarkably* adverbs:

(i) a. ?Clyde is more remarkably tall than somewhat lanky.
b. ?Floyd is less surprisingly ugly than he is a minor annoyance.

But these uses of comparatives are clearly special and licensed by a distinct set of principles having at least as much to do with phenomena like metalinguistic negation as with true comparatives. Among the ways in which this is manifested are that metalinguistic comparatives can be used with DPs, as in (iia); that they are not in complementary distribution with true comparatives, as in (iib); and that they cannot be expressed using the *-er* morpheme, as in (iic):

(ii) a. ?Clyde is more goofy than he is a fool.
b. A: I think Donald is even worse than George.
B: Well, no, I wouldn't say that. ?It seems to me that he's less WORSE than he is just more frequently INTERVIEWED.
c. *Clyde is goofier than he is a fool.

(11) a. Clyde is [more remarkably] tall.

b. *Clyde is [more quite] tall.

(12) a. Floyd is [quite surprisingly] ugly.

b. *Floyd is [quite too] ugly.

(13) a. Many voters are [pretty horribly] conservative.

b. *Many voters are [pretty more] conservative.

One might object at this point that there is a conceivable alternative parse of the sentences in (11–13) in which the degree word is actually associated with the adjective rather than the adverb, as indicated in (14):

(14) Clyde is [more [remarkably tall]]. (parse to be rejected)

If this were the structure of (14), a puzzle arises immediately—*more tall* is not the comparative form of *tall*; *taller* is. Yet what we find in (14) on this structure is comparative morphology applying to an AP headed by *tall*, so we would expect *taller* to occur here. Assuming that the way comparative morphology and adjectives combine morphologically is by head movement of the adjective to a higher position where it finds the comparative morpheme, we would expect that the adjective would move over *remarkably*, as in (15):[69]

(15) *Clyde is [tall-*er* [remarkably *t*]]

This, as indicated, results in an ungrammatical sentence. Nor is there evidence for a struc-

[69] This sets aside the possibility of a parse in which *remarkably* is itself a head.

ture like (14) from interpretation. Certainly, it's true that if Clyde is said to be very remarkably tall, he must also be very tall. But this is not evidence for construing *very* as applying to *remarkably tall*, because of the way being tall is related to being remarkably tall. The only way Clyde's height can be more remarkable (in the way relevant to *remarkably* adverbs) is to be greater; the only way for it to be less remarkable is for it to be smaller. Consequently, increasing or decreasing the extent to which Clyde's height is remarkable also increases or decreasing his height correspondingly. The effect of a degree word, then, will be in this respect the same irrespective of which structure is adopted.

There are broader considerations that militate against treating *remarkably* adverbs as degree words. Degree words do not share the principal properties of *remarkably* adverbs noted in the previous sections.

While new *remarkably* adverbs can be coined with ease and essentially on-the-fly, coining degree words is comparatively harder. Though new degree words do of course occasionally arise,[70] they cannot be coined on-the-fly in the course of a conversation, or readily accommodated by one's interlocutor. No doubt related to this is the relative scarcity of degree words—it does not seem at all out of the question that one might be able to compile an exhaustive list. Compiling an exhaustive list of *remarkably* adverbs, on the other hand, would be an enormous undertaking at best, and perhaps nearly as futile as attempting to compile an exhaustive list of nouns might be.

Among the signature characteristics of *remarkably* adverbs noted in the previous sections are their relationship to their adjective counterparts and to their corresponding

---

[70] Presumably, the degree words *hella* (in some dialects of English) and *wicked* (in New England dialects) are relatively recent coinages:

(i) Clyde is $\left\{\begin{matrix}\text{hella}\\\text{wicked}\end{matrix}\right\}$ tall.

uses in higher, clausal positions. Degree words manifest neither of these characteristics. In general, degree words do not have adjective counterparts. There is a small handful of degree words that might be said to, but in these cases, unlike with *remarkably* adverbs, the meaning of the degree word is not predictable from its corresponding adjective or vice versa. Among the potential suspects in this regard are *real*, *pretty*, *mighty*, and *wicked*. It is not clear what relates the meaning of these degree words to their homophonous adjectives. And it is highly unlikely that it is any single semantic relation applying systematically. To the extent that one might claim a consistent morphological relationship between these degree words and adjectives, it is different from the one that holds between *remarkably* adverbs and adjectives—these degree words are derived by zero affixation, while *remarkably* adverbs are derived by suffixing *-ly*.

Finally, unlike *remarkably* adverbs, degree words cannot occur in higher, clause-modifying positions at all, so the question of this relationship does not arise.

2.3 Developing an Interpretation

2.3.1 *Some Paraphrases*

The essential semantic contribution of *remarkably* adverbs seems to be systematically paraphrasable in terms of the corresponding adjectives. Since this is unlikely to be an accident, given the close relationship between *remarkably* adverbs and adjectives, it seems appropriate to construct the denotations of *remarkably* adverbs in terms of their adjective counterparts, and to take these paraphrases as a starting point in identifying what *remarkably* adverbs mean.

There are several varieties of such paraphrases that get relatively close to what *remarkably* adverbs mean:

(16) Clyde is remarkably tall.

a. It is remarkable that Clyde is as tall as he is.

b. It is remarkable to be as tall as Clyde is.

c. It is remarkable how tall Clyde is.

(17) Floyd is surprisingly ugly.

a. It is surprising that Floyd is as ugly as he is.

b. It is surprising to be as ugly as Floyd is.

c. It is surprising how ugly Floyd is.

(18) Floyd's SUV is unacceptably inefficient.

a. It is unacceptable that Floyd's SUV is as inefficient as it is.

b. It is unacceptable to be as inefficient as Floyd's SUV is.

c. It is unacceptable how inefficient Floyd's SUV is.

(19) Many voters are horribly conservative.

a. For many voters x, it is horrible that x is as conservative as x is.

b. For many voters x, it is horrible to be as conservative as x is.

c. It's unacceptable how conservative many voters are.

Not all of these paraphrases are equally good. The (a) and (b) paraphrases all suffer from a problem of ambiguity, though it is remedied easily enough. For (16a), for example, there is a reading in which what is remarkable is the fact that Clyde is as tall as Clyde. Similarly, in (17a), what is surprising could be the fact that Floyd is as ugly as Floyd.[71] The *remarkably* adverb sentences do not have this reading. But this problem could be avoided easily

---

[71] This is essentially the same ambiguity as in Russell (1905)'s *Your yacht is larger than I thought it is.*

enough—one could imagine pursuing paraphrases of the form *Floyd is tall to some degree, and it's remarkable that he's that tall*, or, in linguist quasi-English, *Floyd is d-tall and it's remarkable to be d-tall*. There is, however, a deeper problem.

An inkling of this problem is reflected in (16a) and (16b). If what is remarkable about Clyde's height is that he is very short, both of these paraphrases would be true; but of course, the *remarkably* adverb sentence cannot mean this. This is still only an inkling of the problem, in that it too could be solved relatively straight-forwardly, in this case by adding to the denotation a requirement that, in this instance, Clyde be tall.

The full measure of the problem emerges more clearly in a situation in which we know Clyde to be the victim of a particular creepy numerological accident. We know that he was born at precisely 5:09 in the morning, on the fifth day of the ninth month of 1959. We further know that he currently lives at 59 Fifty-ninth Street. Discussing this strange happenstance, I inform you that Clyde's height is precisely five feet and nine inches. So Clyde is not very tall, but he is not very short either. It would be quite natural for you to say, upon having heard this news, that it is remarkable that Clyde is five feet nine inches tall, or to utter (16a). But it would not be natural at all to say that Clyde is *remarkably tall*—indeed, given typical contemporary expectations about adult male height, it would be false.[72]

In this situation, the problem cannot simply be simply that Clyde is not tall. If we increment all the numbers that seem to haunt Clyde to the point that he might qualify as just barely tall but not very tall, the result stays the same—it is still remarkable that he is as tall as he is, in light of the numeric coincidences in his life, but he is certainly not *remarkably tall*.

---

[72] It could, of course, be true if the context provides a sufficiently unusual comparison class, as it might if we also know that Clyde is a race horse jockey or president of the International Federation of Unusually Short Taxidermists.

What this demonstrates is that to qualify Clyde as *remarkably tall*, it is not sufficient that he be tall and that there be something remarkable about his height. It must also be the case that what is remarkable about his height is how great it is. Similar facts hold for other *remarkably* adverbs—for (17), for example, what is surprising must be how great Floyd's ugliness is, not simply that he is ugly.

This suggests strongly that there is something fundamentally inadequate about the (a) and (b) paraphrases above, and more generally about paraphrases that involve predicating an adjective of a proposition in a straight-forward way. But all this also strongly suggest that the (c) paraphrases above, which involve embedding *wh*-clauses, are in some important way on the right track. They face none of these difficulties. They don't give rise to the undesirable ambiguity discussed above—they have only the interpretation that *remarkably* adverbs have. Nor do they fail to reflect that *remarkably* adverbs always seem to require that the degree in question be high, and that it must be the highness of the degree that leads to the judgment expressed by the *remarkably* adverb. No further stipulations or additions are required to achieve this, and the paraphrase does not have to be altered in any way.

The *wh*-paraphrases also have the advantage that they, like sentences with *remarkably* adverbs, inherently give rise to a kind of factivity entailment of the form in (20):

(20) a. Clyde is remarkably tall.

entails: Clyde is tall.

b. It is remarkable how tall Clyde is.

entails: Clyde is tall.

(21) a. Floyd is surprisingly ugly.

entails: Floyd is ugly.

b. It is surprising how ugly Floyd is.

entails: Floyd is ugly.

(22) a. Floyd's SUV is unacceptably inefficient.

entails: Floyd's SUV is inefficient.

b. It is unacceptable how inefficient Floyd's SUV is.

entails: Floyd's SUV is inefficient.

In some respects, these entailment seem obvious and almost unavoidable, so it is worth pointing out that it is not a priori necessary that such entailments should have been valid. Measure phrases, for example, do not give rise to this effect:

(23) Clyde is five feet tall.

does not entail: Clyde is tall.

Nor do the alternative paraphrases in (16–19) just considered reflect the entailment pattern in (20–22).

In light of the close parallel between these paraphrases and *remarkably* adverbs, then, taking the semantics of these paraphrases as a guide in sorting out the semantics of *remarkably* adverbs seems to be an approach with some empirical support—these really are very close paraphrases, close enough to suggest that that the semantic connection between them and *remarkably* adverbs is genuine.

2.3.2 *Embedded Exclamatives*

There is, however, a complication in taking the semantics of these *wh*-paraphrases as a guide: it is less than clear what the semantics of these paraphrases themselves is. The *wh*-clause in these paraphrases is not, as it might initially seem, an indirect question. Rather, it is an embedded exclamative of the sort discussed in Grimshaw (1979)—a less-studied construction.

Perhaps the clearest evidence for this involves *very*. As Grimshaw observed, *very* is impossible with *wh*-words in questions, as in (24), but possible in exclamatives, as in (25):

(24) a. *How very tall is Clyde?
b. *How very ugly is Floyd?
c. *How very inefficient is Floyd's SUV?

(25) a. How very tall Clyde is!
b. How very ugly Floyd is!
c. How very inefficient Floyd's SUV is!

This contrast holds under embedding as well. Embedded clauses that are relatively clearly indirect questions do not admit *very*:[73]

(26) a. *I wonder how very tall Clyde is.
b. *Someone asked how very ugly Floyd is.
c. *Mildred wondered how very inefficient Floyd's SUV is.

---

[73] As before, there is a certain interpretation here that seems to involve metalinguistic comparison that should be set aside. On this interpretation, *I wonder how very tall Clyde is* is more or less possible, but it reports that what the speaker is wondering about is how appropriate the phrase *very tall* is as a characterization of Clyde.

But embedded exclamatives do:

(27) a. It is remarkable how very tall Clyde is.
b. It is surprising how very ugly Floyd is.
c. It is unacceptable how very inefficient Floyd's SUV is.

Although it is not directly relevant to the application of this diagnostic, *remarkably* adverbs seem to have the same distribution in these sentences as *very* does.

Another diagnostic for exclamatives, due to Zanuttini and Portner (2003), is based on the observation that (alternative) questions license structures like those in (28), while exclamatives do not, as (29) shows:

(28) a. How tall is Clyde—average height or less than five feet?
b. How ugly is Floyd—just slightly or enough to frighten children?

(29) a. *How tall Clyde is—average height or less than five feet!
b. *How ugly Floyd is—just slightly or enough to frighten children!

Zanuttini and Portner use this to diagnose matrix exclamatives only, but it seems to work (somewhat less cleanly) with embedded exclamatives as well:

(30) a. I wonder how tall Clyde is—average height or less than five feet.
b. Someone asked how ugly Floyd is—just slightly or enough to frighten children.

(31) a. *It is remarkable how tall Clyde is—average height or less than five feet.
b. *It is surprising how ugly Floyd is—just slightly or enough to frighten children.

Again, the paraphrases under consideration pattern with embedded exclamatives rather

than with the embedded questions in (30).

A third diagnostic is based on the observation, due to Elliott (1974) and noted by Zanuttini and Portner, that exclamatives do not seem to occur comfortably under negation in declaratives:

(32) a. I don't (particularly) wonder how tall Clyde is.

b. No one asked how ugly Floyd is.

(33) a. *?It isn't remarkable how very tall Clyde is.

b. *?It isn't surprising how ugly Floyd is.

Zanuttini and Portner observe that curiously, in questions the situation is reversed—exclamatives can occur with negation, as in (34), but not without it, as in (35):

(34) a. Isn't it remarkable how tall Clyde is?

b. Isn't it surprising how ugly Floyd is?

(35) a. *?Is it surprising how ugly Floyd is?

b. *?Is it remarkable how very tall Clyde is?

So in this respect too, these paraphrases pattern with embedded exclamatives.

Building on the foundation these paraphrases provide, then, leads to a semantics for *remarkably* adverbs framed in terms of their corresponding adjectives and embedded exclamatives.

2.3.3 *The Interpretation of Exclamatives*

The first challenge in relating the semantics of *remarkably* adverbs to that of exclamatives is that the semantics of exclamatives is itself not entirely clear, at least from a formal perspective.[74] Still less clear is the semantics of exclamatives under embedding. Zanuttini and Portner (2003), who develop an approach to these issues, will serve here as a guide through this thicket of uncertainty.

Their first move is to observe that exclamatives don't have truth values, and hence should not be analyzed as proposition-denoting. One reflection of this is that it is odd to attempt to affirm or deny an exclamative in discourse:

(36) A: How tall Clyde is!

B: { #No, that's not true. / #No, that's a lie! / #Yes, that's right; good point. }

To object to the use of an exclamative, it is necessary to do relatively roundabout things—one might, for example, deny being in a position to utter it (e.g., *I wouldn't say that*).

Zanuttini and Portner suggest that instead, exclamatives have denotations of the same type as questions do—sets of propositions (following, for questions, Hamblin 1973, Karttunen 1977, Groenendijk and Stokhof 1984, and others). This, of course, reflects quite clearly the deep syntactic parallel between questions and exclamatives. It also sets aside the difference between the two in illocutionary force, which can be reflected in other ways (as they convincingly argue).

---

[74] Perspectives on the interpretation of exclamatives from outside of formal semantics include McCawley (1973), Elliott (1974), and Michealis and Lambrecht (1996).

Adopting the Karttunen (1977) view that a question denotes the set of its *true* answers, they they treat exclamatives as likewise denoting a set that includes only true propositions. So an exclamative such as (37a) will denote a set of propositions that might, under the appropriate circumstances involving discussion of chili pepper consumption, look like (37b):[75]

(37) a. What surprising things he eats!

b. {'he eats poblanos', 'he eats serranos', 'he eats jalapeños'}

More generally, then, (37a) will denote the set of true propositions of the form 'he eats x' for some (surprising) value of x:

(38) ⟦ *What surprising things he eats!* ⟧ = {p: p is true and there is a surprising thing x such that p is the proposition that he eats x}

Exclamatives of the sort most relevant here, such as the one in (39), will have similar denotations:

(39) ⟦ *How tall he is!* ⟧ = {p: p is true and there is a degree of height d such that p is the proposition that he is d-tall}

It might be the case in some circumstances, for example, that *How tall he is!* will denote a set among whose members are 'he is four feet tall' and 'he is five feet tall' and 'he is six feet tall'.

---

[75] This example is a variation on an example of theirs in Paduan. Poblanos, serranos, and jalapeños are all chili peppers.

Zanuttini and Portner identify two principal ingredients in the semantics of exclamatives. One of them is *factivity*—exclamatives systematically presuppose the truth of a corresponding declarative, as already partly exemplified in (20–22). While *remarkably* adverbs have a similar property, as these examples show, this will not be a central focus at the moment.

The other ingredient, which will figure prominently in the analysis of *remarkably* adverbs proposed here, is *widening* of the domain of quantification of the displaced *wh*-expression. To illustrate how this works, consider a context in which we are discussing what Herman eats. If I say *Herman eats everything*, the domain of quantification of the universal is of course constrained by a contextual domain restriction, so you probably wouldn't conclude from my utterance that Herman eats light bulbs, his relatives, or presidential elections. It is very probable that what we might expect Herman to eat would be even more constrained than this—assuming the appropriate cultural background, we might also fail to conclude from my statement that Herman eats serrano chilies. Zanuttini and Portner propose that exclamatives affect essentially this sort of domain restriction, widening it to include things we otherwise would not have considered. So if what I had uttered instead was the exclamative *What surprising things he eats!*, its effect would be to cause you to entertain some possibility you previously hadn't—say, that Herman eats serranos. The denotation of the exclamative, then, will because of this widening include more propositional alternatives than it otherwise would have.[76] As Zanuttini and Portner observe, this bears a close family resemblance to Kadmon and Landman (1993)'s analysis of what *any* does.[77]

---

[76] This discussion does not include one important aspect of widening in their sense, which is that the widening must be in accord with some contextually-provided scale. It is possible to set this aside here because for adjectives, this scale is provided lexically, as subsequent sections illustrate.

[77] They are careful to point out, however, that their 'use of the concept [of domain widening] is quite different'.

This idea elegantly gathers together several otherwise slippery and elusive intuitions about what exclamatives mean. Among these are the intuition that exclamatives somehow involve an 'extreme' value for something, and that exclamatives convey that something is unexpected in a particular way.

2.3.4 *Interpreting Exclamatives Embedded*

The next question relevant to understanding exclamative paraphrases of *remarkably* adverb sentences is what happens when an exclamative is embedded. This presents one slight additional complication, but it also eliminates another one.

The additional complication is that some assumptions have to be made about the semantics of the embedding predicate—hardly a minor point here, since this embedding predicate is what corresponds to *remarkably* adverbs. Here too, Zanuttini and Portner lead the way. They suggest that *amazing*, which embeds both exclamatives and finite indicatives, can be understood as having two forms, one for each type of complement. The garden-variety form applies to propositions and hence embeds finite indicatives. Its semantics is relatively straight-forward—it predicates of a proposition that it is amazing:[78]

(40) $[\![\textit{amazing}_{\textit{garden-variety}}]\!] = \lambda p_{\langle s,t \rangle} . \text{amazing}(p)$

The other form of *amazing* applies to sets of propositions and hence embeds exclamatives. It is interpreted as requiring that some proposition in this set be amazing:

(41) $[\![\textit{amazing}_{\textit{exclamative-embedding}}]\!] = \lambda E_{\langle \langle s,t \rangle ,t \rangle} . \exists p[E(p) \wedge \text{amazing}(p)]$

---

[78] This isn't precisely their formalism, but the content is (intended to be) the same. I haven't reflected here in any independent way that *amazing* is factive.

For an exclamative denotation to be amazing, then, it must include a proposition which is amazing. So, supposing that Clyde is 6 feet 4 inches tall, one might utter (42a), and the exclamative will have a denotation something like the one indicated schematically in (42b):[79]

(42) a. It is amazing how tall Clyde is.

b. It is amazing {'Clyde is 6 feet 1 inch tall′, . . . , 'Clyde is 6 feet 2 inches tall′, . . . , 'Clyde is 6 feet 3 inches tall′, . . . , 'Clyde is 6 feet 4 inches tall′}

c. ∃p[p∈{'Clyde is 6 feet 1 inch tall′, . . . , 'Clyde is 6 feet 2 inches tall′, . . . , 'Clyde is 6 feet 3 inches tall′, . . . , 'Clyde is 6 feet 4 inches tall′} ∧ amazing(p)]

In light of (41), (42a) can be interpreted as requiring that one of the propositions in the set in (42b) be amazing, as (42c) reflects. If it's the case that it's amazing to be 6 foot 4, then, this will be true. More generally, we might assume that embedded exclamatives (at least ones embedded under the relevant sort of predicate) are interpreted in a way that parallels (42).

While in some respect complicating things slightly, this simplifies the situation in another respect. In light of the denotation arrived at for these sorts of structures, for current purposes, it will be possible to do away with making reference in these denotations to sets of propositions, replacing them with sets of degrees.[80] The reason is that asserting (42)

---

[79] For reasons of exposition, I gloss over here what is actually an important point, which is that for Zanuttini and Portner an exclamative denotation includes only those propositions that are outside the domain of quantification as it was before widening. This is not crucial to the point being made in this section.

[80] It will, of course, likely still be required in putting together the semantics of embedded exclamatives compositionally; the concern here, though, is only to arrive at an interpretation of *remarkably* adverb paraphrases.

actually amounts to claiming that it's amazing that there's a degree (in a particular set of degrees with the relevant properties) to which Clyde is tall:

(43) amazing(^∃d[d∈{6 feet 1 inch, . . . , 6 feet 2 inches, . . . , 6 feet 3 inches, . . . , 6 feet 4 inches} ∧ Clyde is d-tall])

All embedded-exclamative paraphrases of *remarkably* adverbs involve adjectives, so in all of them it will be possible to make this simplifying move, quantifying over degrees rather than over propositions.

To capture the meaning of embedded exclamatives, and by extension of sentences containing *remarkably* adverbs, it will also be necessary to say something about what the set of degrees being quantified over is—specifically, it will be necessary to capture the effect of domain widening.

2.3.5 *Brief Interlude: Some Assumptions About Adjectives*

Before proceeding further, though, it may be helpful to briefly lay out some background assumptions about the interpretation of adjectives, mostly drawn from Kennedy (1997).

The first of these is that a degree is an interval on a scale abstractly representing measurement (Kennedy 1997, Schwarzschild and Wilkinson 2002). A scale in this sense is a dense, linearly ordered set of points. Different adjectives are in general associated with different scales, though some pairs of adjectives—antonymous pairs like *tall* and *short*—may share the same scale.

Second, I will assume that a gradable adjective denotes a relation between an individual and a degree—a relatively standard assumption (Seuren 1973, Cresswell 1976,

von Stechow 1984, Bierwisch 1989, Klein 1991, Rullman 1995, Kennedy and McNally 2004).[81] In a sentence like (44), then, *tall* relates Clyde to some degree of height, here one measuring six feet:

(44) a. $[\![\textit{tall}]\!] = \lambda x \lambda d\,.\,\text{tall}(x)(d)$

b. $[\![\textit{Clyde is six feet tall}]\!] = \exists d[\text{tall}(\text{Clyde})(d) \wedge \text{the measure in feet of d is 6}]$

If no overt measure phrase is present, the adjective will be interpreted with respect to a contextually-supplied standard degree of tallness. In (45), for example, *tall* relates Clyde and the standard for tallness s_{tall} provided by the context of utterance:

(45) $[\![\textit{Clyde is tall}]\!] = \exists d[\text{tall}(\text{Clyde})(d) \wedge d \geq s_{tall}]$

What (45) requires is that Clyde be tall to some degree and that it meet or exceed the standard s_{tall}.

2.3.6 *The Interpretation of* Remarkably *Adverb Sentences*

Returning to the main thread of the discussion, it will now be possible to propose an interpretation for exclamative paraphrases of *remarkably* adverbs in the spirit of Zanuttini and Portner, and thereby one for the corresponding *remarkably* adverb sentences as well.

Given what has already been said, a sentence such as the now-familiar (46a), along with its exclamative paraphrase (46b), might (in a particular circumstance) receive an interpretation such as (46c):

---

[81] This divergence from Kennedy (1997) is not in any way crucial to the analysis.

(46) a. Clyde is remarkably tall.

b. It is remarkable how tall Clyde is.

c. remarkable(^∃d[d∈{6 feet 1 inch, ..., 6 feet 2 inches, ..., 6 feet 3 inches, ..., 6 feet 4 inches} ∧ Clyde is d-tall])

So, as before supposing that Clyde is 6 foot 4, (46a) might assert that it is remarkable that Clyde is tall to a degree in the set indicated schematically in (46c).

To spell things out a bit more precisely—and in particular, to make explicit the domain widening that is a signature of both *remarkably* adverbs and exclamatives—a means of representing domain restrictions will be needed. One way of doing this, though not the path taken by Zanuttini and Portner, is to make use of resource domain variables (von Fintel 1994, Westerståhl 1985). Thus just as a resource domain variable can be used to reflect contextual domain restrictions on determiner and adverbial quantification, it can also be used to reflect contextual domain restrictions on quantification inside the extended projection of AP. Spelling this out, the denotation of *Clyde is tall* in (45) can be elaborated with the addition of a resource domain variable C, which will restrict an existential quantifier over degrees as in (47):

(47) $[\![\textit{Clyde is tall}_C]\!] = \exists d[d \in C \wedge \text{tall(Clyde)}(d) \wedge d \geq s_{tall}]$

The resource domain variable C has as its value a contextually-salient set of degrees; (47) requires that the degree quantified over be in this set.

It is actually a fairly significant step, and one that will be crucial here, to suppose that quantification over degrees is contextually restricted in the way that quantification over individuals or events (or situations) is. While it is not usual to think of quantification over

degrees in this way,[82] it seems quite natural. Having domain restrictions seems to be a general property of quantification in natural language, so it ought to be surprising to find that degree quantification *didn't* work this way. Indeed, making this assumption is actually simpler than the alternative, since it would otherwise be necessary to stipulate that only quantifiers over individuals and events are subject to contextual domain restrictions.

It is not especially clear, though, that such contextual domain restrictions should be detectable in a relatively simple example like (47). Still, it does seem sensible to suppose that in uttering a sentence like (47), we have some idea of a possible range of heights to which we restrict our consideration. Thus just as there is something surprising about (48a), which involves quantification over individuals, so too with quantification over degrees, there is something surprising about (48b):

(48) a. Someone showed up during office hours.
It was {Queen Victoria / Gadzork the Martian}.

b. Clyde is tall. He's {about 6 or 7 kilometers / the same height as his apartment building}.

One way of understanding the sense of surprise these sentences give rise to is to suppose that the addressee has initially taken the speaker to have intended to quantify existentially over a more narrowly restricted domain than the one the speaker ultimately turns out to have intended.[83]

---

[82] Though something like this does seem to be what Zanuttini and Portner have in mind.

[83] This sort of explanation, of course, does not rule out a pragmatic explanation in which the surprise is attributed to the bizarreness of what has been asserted; indeed, this is a way of formulating such an explanation a bit more precisely.

There is also a certain common flavor between (48b) and the effects of varying the comparison class

An independent worry about introducing contextual domain restrictions in the way suggested in (47) is that this results in a kind of double context-sensitivity that might seem suspiciously redundant. As it stands, (47) is context sensitive both via the contextually supplied resource domain variable C and via the contextually-supplied standard for tallness s_{tall}. There might be ways of eliminating this difficulty (Morzycki in progress), though it will not be remedied here. It is sufficient to note that to the extent that this double-context sensitivity is a concern—and it is an aesthetic concern, or in any case one of parsimony—it would cast doubt on this means of representing contextually-supplied standards just as much as it would on this approach toward introducing contextual domain restrictions into the adjectival projection.

With this in place, the widening effect of *remarkably* adverbs can now be represented fairly straightforwardly. As a first step, without yet reflecting the effect of widening in the denotation, we can take (49a) to have the denotation in (49b):

(49) a. Clyde is remarkably tall.

b. $[\![\textit{Clyde is remarkably tall}_C]\!]$ = (not final)

$$\text{remarkable}(\hat{}\exists d[d \in C \wedge \text{tall}(\text{Clyde})(d) \wedge d \geq s_{tall}])$$

This merely predicates remarkable-ness of the proposition expressed by *Clyde is tall*, yielding a meaning that might be paraphrased ‘it is remarkable that Clyde is tall’ (which is an inadequate paraphrase for reasons discussed in section 2.3.1). To introduce the effect of domain widening, we might merely modify (49b) by existentially quantifying over a

---

in the interpretation of an adjective (e.g. *Clyde is tall; he’s about 4 foot two, which is tall for a five-year-old.*). This apparent similarity between comparison classes and domain restrictions might not be accidental, and perhaps suggests that domain restrictions might be put to other analytical uses in the semantics of adjectives (Morzycki in progress).

domain larger than the contextually-supplied domain of quantification provided by the resource domain variable C:

(50) $[\![\textit{Clyde is remarkably tall}_C]\!]$ = (not final)

remarkable($^\wedge\exists d\exists C'$ [$C' \supset C \wedge d \in C \wedge$ tall(Clyde)(d) $\wedge$ $d \geq s_{tall}$])

This amounts to loosening the requirement that a degree of Clyde's tallness be among the contextually salient degrees, permitting it instead to be either among these degrees or in some larger domain C′ that includes these degrees.

Still, this is not yet quite adequate, because *remarkably* adverbs, like exclamatives, contribute domain widening in a particular sense that (50) does not reflect. Unlike the kind of widening that Kadmon and Landman (1993) argue *any* involves, exclamatives and *remarkably* adverbs actually impose the further requirement that the degree quantified over *not* be in the unwidened portion of the domain. For Clyde to be remarkably tall, it is not sufficient that he be tall to a degree that's either among the contextually salient ones or in some proper superset of these.[84] Rather, Clyde actually has to be tall to some degree that's not among the degrees already contextually salient—he must be tall to a degree that has been added to the domain by widening, as (51) reflects:

(51) $[\![\textit{Clyde is remarkably tall}_C]\!]$ =

remarkable($^\wedge\exists d\exists C'$ [$C' \supset C \wedge d \in C' - C \wedge$ tall(Clyde)(d) $\wedge$ $d \geq s_{tall}$])

This requires that there be a degree to which Clyde is tall which exceeds the standard and

---

[84] In fact, unless something more is said about what the value of a resource domain variable may be, this wouldn't seem to impose any additional requirement at all, since *any* degree is either in the contextually supplied domain or in a proper superset of it.

that it is in the portion of the widened domain C′ that excludes the original domain C.[85]

This denotation seems to be an adequate representation of the meaning of *Clyde is remarkably tall*. It reflects that this sentence involves a claim that something is remarkable, and that what is remarkable isn't merely that Clyde is tall or even that there is some particular degree such that it's remarkable that he's tall to that degree. Rather, what is claimed to be remarkable is that Clyde's height is so great that it exceeds all the heights one would otherwise have entertained. In this way, this denotation reflects the same sort of domain-widening that an embedded exclamative would contribute, thereby explaining the semantic correlation with the embedded exclamative paraphrase. The factivity entailment that is also characteristic of both *remarkably* adverbs and exclamatives is predicted here, too, because this denotation requires that there be a degree to which Clyde is tall that exceeds the standard for tallness. Maintaining this requirement of exceeding the standard is crucial to capturing the factivity entailment—the requirement of widening the domain on its own would not suffice, since it would not rule out the possibility the Clyde is tall to a degree *smaller* than any in the domain, and that what is remarkable about his height is how small it is.[86]

Other *remarkably* adverb sentences can, of course, be given interpretations analogous to this one.

---

[85] The denotation in (51) places no bounds on how big the widened domain may be, so it amounts to requiring only that the degree quantified over not be in the unwidened domain. I'll represent things in these terms, though, because it corresponds better to the intuition about what's happening here; because it makes the connection to Zanuttini and Portner's account of exclamatives (perhaps) a bit clearer; and because constraints on what a possible domain restriction is may constrain what (52) can mean, too.

It seems reasonable to wonder whether the term 'widening' is fully descriptive of the operation involved here. I will stick to it here because it's the term Zanuttini and Portner use.

[86] This presupposes that the standard will always be in the domain of quantification—a assumption natural at least, and perhaps unavoidable. (Rejecting this assumption, though, would be of no help in deriving the factivity entailment.)

Zanuttini and Portner speculate that there might be a way to rule out widening the domain downward on the basis of some fact about the ontology of degrees. This seems like it would be a very desirable result.

2.3.7 *Summary*

This section developed a semantics for sentences with *remarkably* adverbs by pursuing a parallel between them and paraphrases involving embedded exclamatives. The analysis arrived at makes crucial use of the notion of widening a domain of quantification, applying this notion to quantification over degrees.

2.4 Assembling the Pieces

The previous section arrived at a model of the interpretation of *remarkably* adverb sentences, but nothing has so far been said about how this interpretation is assembled compositionality. It will emerge in this section that the familiar means of semantically combing an adverb and an expression it modifies are not adequate for the task that needs to be performed here—and that a further examination of the syntax suggests another path to take.

2.4.1 *The Trouble with an Intersective Interpretation*

The most basic means of interpreting a modifier is intersectively, by a rule like Heim and Kratzer (1997)'s Predicate Modification. There is no straightforward way of doing this for *remarkably* adverbs. The principal difficulty is that for two expressions to be interpreted intersectively they must be of the same semantic type. In order to implement an intersective interpretation for *remarkably* adverbs and the adjectival projections they modify, it will thus be necessary to find a single type for the denotations of both the *remarkably* adverb and its sister. But what could this type be?

One possibility that seems initially appealing is that both the *remarkably* adverb and its sister denote properties of degrees. This though, problematic, and at a minimum requires complicating the ontology of degrees significantly.

To begin with, it would be necessary to find a way to construe the *remarkably* adverb itself as a property of degrees. Given denotations like the one arrived at above, it is at best highly unclear how this might be done. Of course, one might conclude from this that there is something severely wrong with these denotation. It could in principle be that *remarkably* adverbs are interpreted simply by predicating them directly of degrees. This has the appeal of simplicity, but, among other difficulties, such an approach would have to be spelled out quite a bit before it could be made sense of. Certainly, if a degree is simply an interval on a scale as assumed here (following Kennedy 1997 and Schwarzschild and Wilkinson 2002), predicating of this interval that it is remarkable or surprisingly or disappointing or strange would at a minimum fail to make obvious predictions, and at worst might be irredeemably incoherent.

If we view degrees purely as very abstract representations of measurement, asserting that a particular degree is remarkable would be like asserting that (the integer) 16 is remarkable. If we view degrees as perhaps not quite so thoroughly abstract as that, things don't improve. Assuming, as indicated in section 2.3.5, that a degree is an interval on a particular scale, and that scales are distinguished from each other, a degree might perhaps be better conceptualized as something like '16 feet' or '16 pounds' or '16 dollars' rather than simply '16'. But asserting that 16 feet is remarkable or 16 pounds is surprising is odd and probably not quite coherent as well. We can only make sense of such things by taking them to be in some respect more than meets the eye. Thus to make sense of the claim that 16 feet is surprising, we might think of it as really the claim that 16 feet is a surprising height, say—but of course this is would not amount to predicating surprising-ness of 16 feet. Yet clearly, *remarkably* adverbs like *remarkably* can be used with adjectives

whose denotations involve the scales relevant here, like height or length or weight or cost (e.g., *remarkably tall, remarkably heavy, remarkably expensive*). So if *remarkably* adverbs simply denoted properties of degrees, it would really be quite unclear what the extension of, say, *remarkably* is—and if *remarkably* is anything like *remarkable*, it doesn't seem it could include degrees.[87]

Another, perhaps less serious but non-trivial difficulty is what one might do with the type that would result when a *remarkably* adverb and its sister are interpreted—if this type is itself a property of degrees, as would result from an intersective interpretation, an account would have to be provided of how this can ultimately be predicated of individuals. Certainly, there are ways in which this can be done, both by altering syntactic assumptions or semantic ones. One especially interesting semantic approach toward this problem may be available if degrees are formalized, as Faller (2000) proposes, as vectors in a Vector Space Semantics (Zwarts 1997, Zwarts and Winter 2000, Winter 2001). In this sort of framework, there are independently necessary type shifts that map properties of vectors (qua degrees) to properties of individuals. At a minimum, though, whether by adopting a Vector Space Semantics or by other means, some further explanation here is required—and in either case despite the additional complications would be unlikely to help solve the first problem discussed immediately above.

Any approach in which a *remarkably* adverb is predicated directly of a degree also faces the problem of explaining what the relationship is between predicating a *remarkably* adverb of a degree and predicating its adverbial or adjectival cognates of individuals and propositions (and perhaps eventualities). To illustrate, we might imagine that the denotation of *remarkably* in its *remarkably* adverb incarnation is simply (52)—setting aside all

---

[87] Importantly, *remarkably* adverbs are in this respect different from ad-adjectival adverbs of a different class that includes *enormously*, *minimally*, and *slightly*, all of which can much more easily be thought of as properties of degrees.

the other problems discussed above with such an approach—and that its cognates have denotations as in (53):

(52) $[\![\textit{remarkably}_{\textit{ad-adjectival use}}]\!] = \lambda d\,.\,\text{remarkable}(d)$ (to be rejected)

(53) a. $[\![\textit{remarkably}_{\textit{clause-modifying use}}]\!] = \lambda p\,.\,\text{remarkable}(p)$ (to be rejected)

b. $[\![\textit{remarkable}_{\textit{predicated of a proposition}}]\!] = \lambda p\,.\,\text{remarkable}(p)$ (to be rejected)

c. $[\![\textit{remarkable}_{\textit{predicated of an individual}}]\!] = \lambda x\,.\,\text{remarkable}(x)$ (to be rejected)

While certainly in one sense simple, this sort of approach on its own does nothing to explain the relationship between the uses of these cognate expressions, such as the relationship between *remarkably* as an ad-adjectival modifier and *remarkable* in embedded-exclamative paraphrases. Much of the interest of the problem here is concealed inside the metalanguage-predicate 'remarkable'. On the other hand, the puzzle is not concealed frames things in terms of a single metalanguage predicate whose meaning is relatively clear—say, by holding the denotation in (53c) constant, and defining the others in terms of it. But it is exactly this that reveals the difficulty in defining (53) as a property of degrees.

2.4.2 *The Trouble with a Predicate Modifier Interpretation*

When an intersective denotation for a modifier is not possible, one usually simply adopts a higher, predicate-modifier type denotation—construing it as a function that applies directly to the modified expression. But for *remarkably* adverbs, this road too has some dangerous pitfalls.

If *remarkably* adverbs were predicate modifiers, they would presumably denote

functions from AP denotations to AP denotations[88]—given the assumptions here, expressions of type $\langle\langle e,dt\rangle,\langle e,dt\rangle\rangle$. This would certainly help with the problems noted in the previous section, since the *remarkably* adverb could now 'have access' to the adjectival denotation in a way that would make it possible to build up a denotation like the one arrived at in section 2.3.

This, though, is inconsistent with the syntactic behavior of these expressions. As we have already seen (in (11–13)), *remarkably* adverbs project further structure:

(54) a. Clyde is [[quite remarkably] tall].
b. Floyd is [[rather surprisingly] ugly].
c. Many voters are [[pretty horribly] conservative].

In light of this, it is not the *remarkably* adverb itself but rather the extended AdvP in which it occurs which must have the higher-type denotation. But to achieve this, barring some kind of complicated, previously unattested type shift, it would be necessary to assume that other elements of the adverbial extended projection—including comparative morphology, *very*, and all other Degs—are systematically ambiguous between their regular denotations and ones that yield this very high AP-modifying type. This would be an exceptionally implausible and costly assumption at best.

2.4.3 *Building Up More Syntax: Analogy to Measure Phrases*

If, as the previous section argued, *remarkably* adverbs can't be interpret intersectively or as predicate modifiers, how *should* they be interpreted? A closer examination of

---

[88] More accurately, functions from the denotation of some constituent of the extended adjectival projection to denotations of the same type—it is not crucial that these be AP denotations.

the syntax suggests an answer.

One especially clear aspect of the syntax of these expressions is that they resemble nominal measure phrases in at least two ways. First, they occur in the same linear position—both of them can occur only at the left periphery of AP, never at the right:

(55) Floyd is {six feet / remarkably} tall {*six feet / *remarkably}.

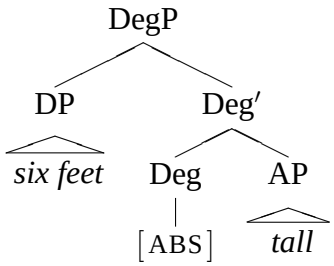
And both kinds of expressions are in complementary distribution with overt degree words in the AP in which they occur:

(56) Floyd is {*six feet / *remarkably} very tall.

It seems reasonable, then, to pursue a parallel syntactic analysis.

APs with absolute adjectives and measure phrases have a structure like the one reflected in (57), in which the measure phrase occupies the specifier position of a Deg(ree) head (Abney 1987, Corver 1990, Grimshaw 1991, Kennedy 1997):

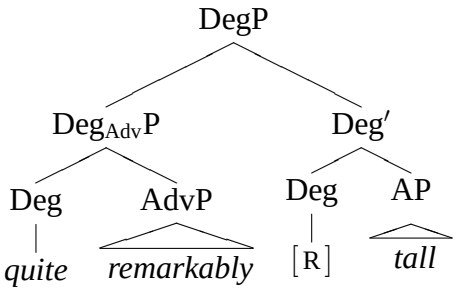
(57)



Under other circumstances, the Deg head can be spelled out overtly as a comparative morpheme (or other degree morpheme) or as a degree word. With absolute adjectives, it is not spelled out overtly. Rather, I'll assume that in these cases the Deg head is instead a null degree morpheme [ABS], following Kennedy (1997).

In light of the similarities, it is natural to assign *remarkably* adverbs, a similar structure, in which their phrasal projections likewise occupy the specifier position of DegP:[89]

(58)



Proposals of this general form for degree adverbs generally—by which is typically meant any true adverbs in AP—have been made before. Abney (1987) suggests a structure similar to (58), with adverbs in a specifier position, and the structures Jackendoff (1977) has in mind would have something like (58) as one more contemporary analogue.

Kennedy's [ABS] has in (58) been replaced with a similar feature [R]. This is intended to be in many respects similar to [ABS]. Although a stronger reason to distinguish these will emerge shortly, there are at least two other, purely syntactic reasons this distinction may be useful. One of these is that [ABS] licenses a DP in its specifier, so it is Case-licensing. *Remarkably* adverbs, on the other hand, have no need to check Case. Another consideration here is a small difference in distribution. Measure phrases, unlike *remarkably* adverbs, are possible in comparatives:

(59) Clyde is $\left\{\begin{array}{c}\text{two feet}\\ \text{*remarkably}\\ \text{*surprisingly}\end{array}\right\}$ taller than Floyd.

It will be necessary, then, to distinguish the ability to license measure phrases and *remark-*

---

[89] I use Deg$_{Adv}$P here to distinguish the degree projection of the adjective and that of the adverb.

ably adverbs in order to reflect that certain Degs may license one but not the other. In light of this independent necessity, there would not be any advantage to uniting the ability to license measure phrases and *remarkably* adverbs in one Deg, [ABS].

This sort of structure has a number of syntactic advantages. It can account for the complementary distribution of measure phrases and *remarkably* adverbs, since these both occupy the same structural position. It can account for why *remarkably* adverbs are obligatorily left of the adjective. And it can account for why they are in complementary distribution with overt Degs, since they require a particular (null) Deg to license them.[90]

2.4.4 *Putting the Syntax and Semantics Together*

With these syntactic structures in place, it is now possible to look on the semantic compositionality puzzle with a fresh eye.

For measure-phrase structures like (57), Kennedy suggests that the semantics is assembled as in (60):

(60) ⟦ *Clyde is six feet* [ABS] *tall* ⟧ = ⟦ [ABS] ⟧(⟦ *tall* ⟧)(⟦ *six feet* ⟧)(⟦ *Clyde* ⟧)

The Deg [ABS] applies first to the AP denotation and then to the measure phrase. It yields a property of individuals as the denotation of the DegP. Here, the Deg does the semantic work of relating the AP and the measure phrase.

Given the parallels between the measure phrase structure in (57) and the *remarkably* adverb structure in (58), it is natural to suppose that semantic composition works similarly

---

[90] This structure also predicts that it should not be possible to stack *remarkably* adverbs, but that it should be possible to introduce them recursively. That is, while exactly one *remarkably* adverb phrase can occur for each AP, a *remarkably* adverb phrase can itself contain a *remarkably* adverb (e.g. *[[surprisingly [terrifyingly]] ugly]*).

in (58). The [R] feature can be taken be interpretable, and paralleling [ABS], to be what relates the AP and the *remarkably* adverb semantically:

(61) ⟦ *Clyde is remarkably* [R] *tall* ⟧ = ⟦ [R] ⟧(⟦ *tall* ⟧)(⟦ *remarkably* ⟧)(⟦ *Clyde* ⟧)

This means of putting the pieces together, via the mediation of [R], will be the key to solving the compositionality problem and arriving at the desired interpretation.

It will now be possible to suppose that the denotation of *remarkably* is actually exactly identical to that of the adjective *remarkable*. To illustrate this, though, it will be useful to make two simplifying assumptions purely for exposition. First, I will omit the degree argument in the denotation of both *remarkably* adverb and their corresponding adjectives. Second, I will for the moment suppose that these corresponding adjectives denote properties of propositions rather than, say, ordinary individuals (more on this point will follow). Adopting these, both *remarkably* and *remarkable* can be taken to have the denotation in (62):

(62) ⟦ *remarkable* ⟧ = ⟦ *remarkably* ⟧ = λp . remarkable(p)

This is, of course, a very simple denotation, and reflects only the barest, most minimal lexical core of the meaning of these expressions. It is a long way from the making the semantic contribution that was attributed to *remarkably* adverbs in section 2.3.

But the challenge of getting from one to the other can now be met straightforwardly—the additional semantic work that needs to be done can be attributed not to the adverb itself, but rather to the [R] feature in Deg that licenses it. Just as the adverb itself can now have as its denotation only the irreducible essence of its lexical semantics, so too the [R] can now have as its denotation only those aspects of meaning that

characterize class of *remarkably* adverbs generally, independent of the particular choice of adverb:[91]

(63) $[\![\,[\,\mathrm{R}\,]\,]\!] = \lambda \mathrm{A}_{\langle e,\langle d,t\rangle\rangle}\ \lambda \mathrm{R}_{\langle st,t\rangle}\ \lambda \mathrm{x}\,.\,\mathrm{R}(^{\wedge}\exists \mathrm{d}\exists \mathrm{C}'\ [\mathrm{C}' \supset \mathrm{C} \wedge \mathrm{d} \in \mathrm{C}' - \mathrm{C} \wedge \mathrm{A}(\mathrm{x})(\mathrm{d}) \wedge \mathrm{d} \geq \mathrm{s}_{\mathrm{A}}])$

This denotation reflects exactly the semantic properties identified in 2.3 as characteristic of *remarkably* adverbs—among the more prominent ones, domain widening. It also serves as a kind of semantic glue, helping hold together type-theoretically the adjective and adverb denotations.

These pieces fit together in a way that yields the desired result:

(64) $[\![$ *Clyde is remarkably* $[\,\mathrm{R}\,]$ *tall* $]\!]$

$= [\![\,[\,\mathrm{R}\,]\,]\!]([\![$ *tall* $]\!])([\![$ *remarkably* $]\!])([\![$ *Clyde* $]\!])$

$= \mathrm{remarkable}(^{\wedge}\exists \mathrm{d}\exists \mathrm{C}'\ [\mathrm{C}' \supset \mathrm{C} \wedge \mathrm{d} \in \mathrm{C}' - \mathrm{C} \wedge \mathrm{tall}(\mathrm{Clyde})(\mathrm{d}) \wedge \mathrm{d} \geq \mathrm{s}_{\mathit{tall}}])$

This is exactly the denotation ultimately arrived at in section 2.3 in (51).

2.4.5 *Problems Averted*

This division of labor avoids the problems raised by the alternative approaches to introducing *remarkably* adverbs into semantic composition.

The difficulties considered raised by supposing that *remarkably* adverbs are simply interpreted intersectively do not arise here because this approach does not impose the

---

[91] For simplicity, I omit here the syntactic representation in the object language of the (index corresponding to the) contextual domain resource variable C. It might be worth noting, though, that since the quantificational force will now come from not the AP denotation but from Deg, it will now be Deg, not the adjective, that must be subscripted with C.

requirement that the *remarkably* adverb and its sister be of the same semantic type. Consequently, we are not forced into any uncomfortable further assumptions to sustain these types. In particular, there is no analytical pressure on this view to treat *remarkably* adverbs as properties of degrees. Rather, the denotation of a *remarkably* adverb is ultimately predicated of a proposition, as seems most natural.

The difficulties raised by supposing that *remarkably* adverbs phrases have predicate modifier denotations and apply directly to their sister are avoided as well. On the current account, the type of the *remarkably* adverb and its projections remains very simple, and more important, the same as the corresponding adverb.[92] Consequently, it is no surprise—and indeed, expected—that *remarkably* adverbs should support their own degree words and project the full adverbial extended projection. It will not be necessary to assume either massive systematic ambiguity of Degs or any novel otherwise unmotivated type-shifts, because the types of all elements of the adverbial projection will be exactly the same as they would otherwise be—and the same as in the adjectival projection.

2.4.6 *Summary*

Neither of the usual ways of interpreting a modifier provides a satisfactory explanation of how *remarkably* adverbs make their semantic contribution. The syntax of these expressions seems to place them in the specifier position of a Deg. By supposing that they are arguments of this Deg, the compositionality problems can be avoided. As an additional benefit, the denotation of *remarkably* adverbs can be dramatically simplified and the semantic characteristics that distinguish the class can instead be associated directly with the Deg that licenses them, and in that sense with the position they occupy.

---

[92] And, as will be suggested in the following section, as other uses of the same adverb.

2.5 Relation to Clausal Counterparts

2.5.1 *A Simple Theory*

What has now been introduced is a kind of decomposition—the apparent meaning of *remarkably* adverbs has been split into two parts, one associated with the *remarkably* adverb itself and one associated with its position. Among the chief advantages of having done things this way is that it quite straightforwardly offers a theory of how (ad-adjectival) *remarkably* adverbs relate to their clause-modifying counterparts and to their adjective counterparts.

In what has already been said, a (possibly maximally) simple theory of the relation to adjectives: the denotation of a *remarkably* adverb and its corresponding adjective are identical.

It turns out that the same can be said of the relation to clause-modifying uses. Given exactly the same denotation for *remarkably* proposed in (62) and repeated in (66a), the right interpretation for its clause-modifying use in (65) will follow:

(65) Remarkably, Clyde is tall.

(66) a. $[\![\textit{remarkably}]\!] = \lambda p \,.\, \text{remarkable}(p)$

b. $[\![\textit{Clyde is tall}_C]\!] = \exists d[d \in C \wedge \text{tall}(\text{Clyde})(d) \wedge d \geq s_{tall}]$

Assuming as before that the denotation of *Clyde is tall* is as in (66b), which repeats (47),[93] *remarkably* can apply directly to (66b) to yield (67):

---

[93] The contextual domain restriction is made explicit in (66b), though it will not be relevant here; it is included for consistency only.

(67) $[\![\textit{Remarkably, Clyde is tall}_C]\!] = \text{remarkable}(^{\wedge}\exists d[d \in C \wedge \text{tall}(\text{Clyde})(d) \wedge d \geq s_{tall}])$

The denotation in (67) requires only that it be remarkable that Clyde is tall. This seems to reflect what the clause-modifying use of *remarkably* means.

2.5.2 *A Slightly Less Simple Theory*

As noted previously, I have been systematically indulging in an expository shortcut with respect to the denotation of *remarkably* and of its adjectival counterpart, in two respects. The first is that their degree argument has been systematically suppressed throughout. This I believe to be genuinely a a kind of purely notational abbreviation.[94] The second simplification is they have been taken here to apply directly to propositions. This may conceal something substantive.

One reason it is probably undesirable to take *remarkably* adverbs and their adjectival counterparts to actually apply to propositions rather than to individuals is that this

---

[94] Thus, undoing only this simplification the denotation of *remarkably* and *remarkable* would be as in (ia), and the denotation of the full extended projection as in (ib):

(i) a. $[\![\textit{remarkable}]\!] = [\![\textit{remarkably}]\!] = \lambda p\, \lambda d\, .\, \text{remarkable}(p)(d)$
b. $[\![[_{DegP}\ \textit{remarkable}\ ]]\!] = [\![[_{Deg_{Adv}P}\ \textit{remarkably}\ ]]\!] =$
$\lambda p\, \exists d[d \in C \wedge \text{remarkable}(p)(d) \wedge d \geq s_{remarkable}]$

In (ib), both of these expressions are taken to be true of a proposition iff it is remarkable to a degree that exceeds the standard for remarkability (and is in the contextually-supplied domain). The means by which (ib) is built up from (ia) is the same means by which the denotation of *Clyde is tall* is built up—following Kennedy (1997), by the application of the intervening Deg head, which (among other things) introduces the standard. Given that $[\,_{R}\,]$ has the denotation in (63), this will yield sentence denotations like (ii):

(ii) $[\![\textit{Clyde is } [_{DegP}\ [_{Deg_{Adv}P}\ \textit{remarkably}\ ]\ [\,_{R}\,]\ [_{AP}\ \textit{tall}\ ]\ ]]\!]$
$= [\![[\,_{R}\,]]\!]([\![[_{AP}\ \textit{tall}\ ]]\!])([\![[_{Deg_{Adv}P}\ \textit{remarkably}\ ]]\!])([\![\textit{Clyde}]\!])$
$= \exists d'\, [d' \in C \wedge \text{remarkable}(^{\wedge}\exists d \exists C'\, [C' \supset C \wedge d \in C' - C \wedge \text{tall}(\text{Clyde})(d) \wedge d \geq s_{tall}])(d') \wedge$
$d' \geq s_{remarkable}]$

This will be true iff it is remarkable to a degree (in the contextually-supplied domain) that exceeds the standard for remarkability that Clyde is tall to a degree which exceeds the standard for tallness and which is not in the contextually-supplied domain but is in the widened domain.

would make them a different type from run-of-the-mill adjectives. This is not a purely aesthetic concern. If it were possible for adjectives and adverbs to vary type-theoretically in this way, a version of one of the compositionality problems encountered in section 2.4 would arise. In order to account for why *remarkably* adverbs and their adjectival counterparts co-occur the same range of Degs that ordinary adverbs and adjectives do, it would be necessary either to assume all Degs are systematically ambiguous in type or that some sort of type shift takes place here. While the first option is a non-starter, assuming a type shift here is plausible. The type shift required would be one that maps back and forth between functions that apply to a proposition and ones which apply to the individual correlate of that proposition. This would fix the problem—a *remarkably* adverb would be type-shifted into an ordinary adjective denotation in this way before a Deg applies to it, and the entire *remarkably* adverb phrase would be type-shifted back to feed $[\mathrm{R}]$ the right type of argument.

There is, however, another alternative.

2.5.3 *A Theory Slightly Less Simple Still*

With respect to the analysis of *remarkably* adverbs themselves, it is not actually necessary to assume this type shift. Instead, one can suppose that the lexical denotation of *remarkably* adverbs is already 'shifted' in the right way—that is, that *remarkably* and *remarkable* inherently apply to the individual correlates of propositions, not to propositions themselves. This would eliminate the need for the first type shift, the one that was necessary to provide a Deg with the right type of argument. The second type shift, the one that was necessary to provide $[\mathrm{R}]$ with the right type of argument, can similarly be eliminated by simply encoding the effect of this type shift into the meaning of $[\mathrm{R}]$ itself—it too could

apply to expressions with ordinary AP/AdvP denotations, and retrieve in its own denotation the propositional correlate. This solves the problem for *remarkably* adverbs, allowing them and their adjectival counterparts to be of an ordinary adjective type without appealing to type-shifting.

But what about the clause-modifying uses? If *remarkably* adverbs apply to directly individuals, a type-shift would be necessary to permit them to apply to propositions, as is necessary to achieve the right clause-modifying reading (as in (67)). Without resort to this type-shift, these adverbs would be uninterpretable.

One could conclude from this that, in order to maintain a unified account that includes these uses, the type-shifting approach is necessary after all. A more intriguing possibility, though, is to take the solution for this problem in the adjectival extended projection to be more general—perhaps this mapping from individuals to propositions in the clause-modifying uses is not a type-shift, but rather the semantic contribution of an element of clausal functional structure in this respect analogous to the Deg [R].

While it is a larger project than can be undertaken here to fully motivate such an explanation, it does seem to have a number of things to recommend it. One of these is that it accords nicely with recent work that strives to further articulate the functional structure in this high region of the clause (Rizzi 1997 and others), as well as to approaches to adverbial modification in these positions that provide exactly the necessary architecture (Cinque 1999). Another is that it would in several ways be independently useful to suppose that some element of clausal functional structure can perform this sort of mapping between propositions and individuals. Thus, it would be useful in accounting for the ability of certain adjectives—the adjectival counterparts of *remarkably* adverbs among them, non-accidentally—to take both propositional and individual-denoting arguments:

(68) a. Esmeralda is amazing.

b. It is amazing that Esmeralda is an astronaut.

(69) a. George is quite shocking.

b. It is quite shocking that George would say that sort of thing publicly.

The (a) sentences here are unproblematic, but if adjectives always apply to individuals, the (b) sentences will require this proposition-individual mapping to take place. A natural view of how this works, from this perspective, is to suppose that this mapping is performed high in the functional structure of the embedded clause—indeed, in the same place where the clause-modifying adverbs relevant here are licensed. Thus one might suppose that two interpretable features, [EMBEDDED] and [SOE-ADV][95], occur in the same structural position, and both perform different sorts of proposition-individual mapping. [EMBEDDED] takes only a proposition as an argument, and yields its individual correlate[96]; this is what is involved in the (b) sentences here. On the other hand, [SOE-ADV] takes as arguments a proposition and the denotation of a clausal use of a *remarkably* adverb, and yields the proposition that results from applying the adverb denotation to the individual correlate of the proposition. This would account for the pattern in (70):[97]

(70) a. Remarkably, Clyde is tall.

b. It is doubtful that Clyde is tall.

c. *It is doubtful that remarkably, Clyde is tall.

In (70a), *remarkably* is interpreted with the aid of [SOE-ADV]; in (70b), the clause is embedded with the aid of [EMBEDDED]. Neither [SOE-ADV] nor [EMBEDDED] could

---

95 [SOE-ADV] abbreviates 'speaker-oriented evaluative adverb'.

96 Kratzer (1999a) suggests that this is the denotation of the complementizer *that*.

97 There is an interpretation on which (70c) is grammatical, in which the adverb is interpreted as though it were not embedded, suggesting that it is an appositive.

perform the mapping necessary in (70c), however—[EMBEDDED] does not accept an adverb argument, and [SOE-ADV] does not yield an individual, which is what *doubtful* requires.

At this level of detail, this is of course all rather speculative, but it does seem to suggest at least that the treatment of *remarkably* adverbs proposed here presents some interesting further analytical possibilities, especially with respect to the interrelations between adverb position, meaning, and functional structure.

2.6 Final Remark

In a nutshell, the central goal here has been to account for the syntax and semantics of one natural class of ad-adjectival adverbs, namely *remarkably* adverbs. The interpretation proposed for these adverbs makes crucial reference to domain-widening for degrees, thereby likening these adverbs to certain exclamatives. This interpretation is built up by treating these adverbs as the semantic arguments of a Deg(ree) head in a way that parallels existing analyses of measure phrases, thereby avoiding a number of compositional difficulties.

An essential element of this proposal is that it executes a kind of 'factoring out' of the positional meaning of the adverb, leaving the adverb's meaning relatively simple and attributing the additional factored-out meaning to an element in the structure of the adjective projection, the degree head. Importantly, this factored-out meaning involves more than combinatorial manipulation; it is not merely a type-shift. Nor is it of a sort that could be derived (in any straightforward way) only from differences in the nature of the semantic object being modified in different positions.

The structure arrived at here, motivated primarily by compositional concerns and independent assumptions about the architecture of AP, involves an adverb occupying the specifier position of a functional head, which accords with the view of Cinque (1994, 1999) and others that adverbs in general are specifiers to functional heads. This is in some respects a startling result, since it was arrived at for different reasons on the basis of a fundamentally different sort of argumentation and in the context of a different set of broader theoretical assumptions.

This suggests too that it may be reasonable to think that the semantic relation between (certain) other modifiers and the expressions they modify might similarly be mediated by elements of the functional structure of the modified expression. As suggested in chapter one, to the extent that this mode of modification can be more generally motivated, it may open up some possibilities in the analysis of various other nettlesome modifiers more broadly.

More narrowly, this represents an attempt to begin to grapple with some of the essential facts of adverbial modification inside AP, an empirical domain largely (and oddly) neglected in the comparatively extensive literature on adverbs more generally. There is far more richness in this domain than it has been possible to touch on here—other apparent natural classes include what might be called adverbs of degree measure (*slightly*, *vastly*, *enormously*), which could more sensibly be predicated directly of a degree, and ad-adjectival analogues of domain and subject-oriented adverbs. Perhaps, then, the conclusion here of which one can be most confident is that adverbial modification inside AP might merit considering in research into the larger question of adverb syntax and semantics.

CHAPTER 3

MEASURE DP ADVERBIALS: MEASURE-PHRASE MODIFICATION IN VP

3.1 Introduction

If DP adverbials such as those in (1) did not occur quite regularly in English and other familiar languages, one might for several reasons be inclined to consider them a fairly odd and exotic construction:

(1) a. It had been raining an hour.
b. Clyde played the ukulele every day.
c. Floyd slept the wrong way again.

These expressions, after all, manifest two characteristics that seem at odds with each other: they are DPs and they are modifiers. Prototypically DPs are not modifiers, and modifiers

---

The discussion in this chapter does not take adequate note of Pereltsvaig (2000), which had come to my attention after the paper that this chapter is based on (Morzycki 2001b) had been completed. The relevance of that work is very direct, in that Perelstvaig reaches broadly the same general syntactic conclusion, but reasoning both from different data and from different observations about that data. She concerns herself with what are called 'accusative adverbials' in the Slavic literature. This is a class of expressions defined chiefly, as the name suggests, by case considerations, but, interestingly, it is apparently either identical to the class of measure DP adverbials or nearly so. The empirical foundation that drives Perelstvaig's analysis includes a wide range of observations about case in Russian and in Finnish which have no analogue in this chapter, and does not include the scope observations that figure prominently here—yet despite addressing a different set of puzzles, she too concludes that these expressions are introduced in the specifier position of a verbal projection associated with aspect. More detailed discussion of how her work relates to what is said here and attempts at integration (particularly of the much more sophisticated account of case her analysis offers with the semantic proposals made here) will sadly have to be left to for future work. There are also interesting connections and parallels to Kratzer (2002b) that will similarly have to wait.

are not DPs. In part because of this tension, perhaps, DP adverbials have long posed an analytical challenge.

The focus of this chapter will be on a proper subset of these—quantificationally weak DP adverbials like the one in (1a), which I'll call 'measure DP adverbials'.[1] These constitute a natural class distinguished not only by quantificational strength but also by restrictions on their scope, distribution, and the Aktionsart of the VP they modify. From these characteristics an argument will be woven that these expressions are in a sense argument-like, licensed by elements of verbal functional structure that directly encode some Aktionsart information. Section 3.2 examines the essential properties of measure DP adverbials, identifying several particular puzzles they present. Section 3.3 strives to address these puzzles and arrives at a treatment of measure DP adverbials in which they are simply the verbal exponent of the more general cross-categorial phenomenon of measure-phrase modification. Section 3.4 concludes.

3.2 Distinguishing Characteristics of Measure DP Adverbials: Some Puzzles

3.2.1 *The Scope Puzzle*

Each of the distinguishing characteristics of measure DP adverbials is a kind of puzzle, and of course part of the large puzzle of why they should be correlated. One of these is that measure DP adverbials obligatorily take narrow scope with respect to sentential

---

[1] I choose this term because 'measure DP adverbial' is a pretheoretical, intuitive characterization in a way in which 'weak DP adverbial' is not—in fact, 'weak DP adverbial' is in itself a step, though tiny, toward an analysis. This term also draws more explicitly the parallel to measure phrases in other categories. Indeed, measure phrases in AP and PP also seem to be obligatorily weak (setting aside certain complications involving expressions that probably do not involve true measure phrases, like *this tall*).

negation, quantified DPs in argument positions, aspectual morphology (i.e., AspP), and—perhaps particularly striking—embedding verbs. Neither the corresponding PPs nor DP adverbials that are quantificationally strong manifest this restriction.

NEGATION Taking these in turn, negation must out-scope *an hour* and *several days* in (2), for example:[2]

(2) a. Clyde didn't sleep an hour. ($\neg \prec$ *an hour*; **an hour* $\prec \neg$)

b. Greta couldn't stay a day. ($\neg \prec$ *a day*; **a day* $\prec \neg$)

If, for example, a car alarm kept Clyde awake all night, (2a) is true. If, however, Clyde went to bed promptly at midnight, was awakened by a car alarm at 4:00 in the morning, and fell asleep again an hour later, (2a) is quite clearly false; it does not have a reading in which it asserts that there was an hour during which Clyde was awake. Likewise, (2b) can only be understood as claiming that the duration of Greta's stay had to be shorter than a day; a reading in which there is a particular day during which Greta can't stay is impossible. This contrasts with PP paraphrases of these sentences, which permit both scope possibilities:

(3) a. Clyde didn't sleep for an hour. ($\neg \prec$ *an hour*; *an hour* $\prec \neg$)

b. Greta couldn't stay for a day. ($\neg \prec$ *a day*; *a day* $\prec \neg$)

In the situation above in which a car alarm interrupts Clyde's otherwise undisturbed and lengthy sleep, (3a) is true, unlike (2a). And if there was a particular day during which Greta

---

[2] In these examples, most of the measure DP adverbials will be indefinites headed by *a(n)*. This is not a coincidence. It is crucial here to keep weak and strong DP adverbials distinct. While indefinites in general may have strong readings, for some reason, this is not possible for adverbial DPs headed by *a(n)*—they have only the weak reading. In order to express the strong reading, it is necessary to resort to *one*, even in contexts that otherwise don't seem to involve counting (as on the true reading of *I expect to be dead one day*; cf. *I expect to be dead a day*, which is false). I take advantage of this convenient property here.

had to be elsewhere, but her stay could otherwise in principle be of indefinite duration, (3b) is true, unlike (2b). Strong DP adverbials also allow both scope possibilities with respect to negation:

(4) Clyde didn't sleep the whole day. ($\neg \prec$ *the whole day*; *the whole day* $\prec \neg$)

This can mean either that during the whole day, Clyde went without sleep, or that Clyde's sleep did not last the whole day.

QUANTIFIED ARGUMENTS The narrow-scope requirement measure DP adverbials manifest also holds with respect to argument DPs. In (5), *no one* and *few chiropractors* must scope above the measure DP adverbial:

(5) a. No one slept an hour. (*no one* $\prec$ *an hour*; **an hour* $\prec$ *no one*)

b. Few chiropractors waltzed ten minutes. (*few chiropractors* $\prec$ *ten minutes*; **ten minutes* $\prec$ *few chiropractors*)

It is not possible to take (5a) to assert that there is a particular hour during which everyone was awake—rather, it asserts that no one experienced an hours' sleep. And (5b) cannot be taken to assert that there was a particular ten-minute interval during which few chiropractors waltzed. As before, this restriction is not shared by the PP paraphrases of these sentences:

(6) a. No one slept for an hour. (*no one* $\prec$ *an hour*; *an hour* $\prec$ *no one*)

b. Few chiropractors waltzed for ten minutes. (*few chiropractors* $\prec$ *ten minutes*; *ten minutes* $\prec$ *few chiropractors*)

Both scope possibilities are available here—these can report interruptions in sleeping or in mass waltzing, respectively, of a particular length. Strong DP adverbials permit both scope possibilities as well:

(7) No one slept all afternoon. (*no one* $\prec$ *all afternoon*; *all afternoon* $\prec$ *no one*)

This may mean that all afternoon, no one slept, or that no one slept through the afternoon.

ASPECT Measure DP adverbials are obligatorily interpreted below aspect. In (8), for example, *an hour* cannot scope above generic (habitual) aspect:[3]

(8) Clyde swam a year. (GEN $\prec$ *a year*; **a year* $\prec$ GEN)

The only reading possible here seems to be a pragmatically implausible one that involves a year-long swimming event by Clyde. Is it not possible to interpret this more plausibly as reporting that for a year, Clyde was a habitual swimmer. The PP counterpart of this sentence in (9), though, is not similarly restricted:

(9) Clyde swam for a year. (GEN $\prec$ *a year*; *a year* $\prec$ GEN)

This *does* have precisely the pragmatically more plausible reading (8) lacks—and the implausible reading is of course possible as well. Strong DP adverbials again pattern with PPs in lacking this narrow-scope restriction:

(10) Clyde swam all year. (GEN $\prec$ *all year*; *all year* $\prec$ GEN)

---

[3] GEN here is represents the generic quantifier over intervals.

This may mean that for all of a year, Clyde habitually or regularly swam, or that Clyde's swimming lasted an entire year.

EMBEDDING VERBS Perhaps most surprisingly, measure DP adverbials must also take narrow scope with respect to embedding verbs. That is, they are necessarily construed with the lower VP in an embedded clause. Thus in (11), *a few minutes* must be construed to be in the scope of *wanted*, and *at least two hours* must be construed in to be in the scope of *feared*:

(11) a. Greta wanted to talk a few minutes. (*wanted* $\prec$ *a few minutes*; **a few minutes* $\prec$ *wanted*)

b. She feared that Clyde would yodel at least two hours. (*feared* $\prec$ *at least two hours*; **at least two hours* $\prec$ *feared*)

If what Greta desires is a few minutes of talking, (11a) is true. If she has been wanting for a few minutes to talk for some other length of time, (11a) is not true. Similarly, (11b) may report a fear of two-hour yodeling, but not a two-hour fear of (even momentary) yodeling. Again, this contrasts with the corresponding PPs:

(12) a. Greta wanted to talk for a few minutes. (*wanted* $\prec$ *a few minutes*; *a few minutes* $\prec$ *wanted*)

b. She feared that Clyde would yodel for at least two hours. (*feared* $\prec$ *at least two hours*; *at least two hours* $\prec$ *feared*)

These both manifest straightforward and familiar structural ambiguities—in (12a), what has a few-minute duration could be either Greta's desire or the talking she desires, and in

(12b) what has a two-or-more-hour duration could be either her fear or the yodeling she fears. Strong DP adverbials also manifest this ambiguity:

(13) Greta wanted to talk every morning. (*wanted* $\prec$ *every morning*; *every morning* $\prec$ *wanted*)

This may mean that Greta's desire was daily, or that her desire was for daily talking.

3.2.2 *The Distribution Puzzle*

Another defining characteristic of measure DP adverbials is their relatively restricted distribution; they are limited to positions fairly low in the tree. They can't front in standard varieties of English:

(14) a. *An hour, Clyde slept.

b. *Several minutes, it had been raining.

The corresponding PPs and strong DP adverbials can:

(15) a. For an hour, Clyde slept.

b. For several minutes, it had been raining.

(16) a. All afternoon, Clyde slept.

b. The whole day, Floyd complained.

Nor can measure DP adverbials occur as modifiers of NP:

(17) a. *Clyde's nap an hour was restful.

b. *His vacation a week was largely wasted.

Again, this contrasts with PPs and strong DP adverbials:

(18) a. Clyde's nap for an hour was restful.

b. His vacation for a week was largely wasted.

(19) a. Clyde's nap every afternoon was restful.

b. His vacation last March was largely wasted.

Measure DP adverbials cannot occur above—that is, in English, right of—PPs and strong DP adverbials:

(20) a. Clyde slept an hour every day.

b. *Clyde slept every day an hour.

(21) a. Clyde usually slept less than six hours for a year.

b. *Clyde usually slept for less than six hours a year.

And again, PPs and strong DP adverbials manifest no analogous ordering restriction:

(22) a. Clyde slept for an hour every day.

b. Clyde slept every day for an hour.

So, measure DP adverbials are distinguished from PPs and strong DP adverbials by both narrow scope and low structural position.

3.2.3 *The Aktionsart Puzzle*

Measure DP adverbials are also distinguished by an Aktionsart restriction: they impose the requirement that the eventuality they measure be durative—or, to put it in a way that doesn't have this necessarily temporal flavor, aspectually homogeneous (Moltmann 1991).[4] Thus they can modify states and activities, as in (23), but are not naturally compatible with achievements and accomplishments, as in (24):

(23) a. He slept an hour.
b. He danced an hour.

(24) a. #He died an hour.
b. #He walked an hour to the corner.

To the extent that the sentences in (24) are good at all, they must be coerced into activity interpretations.

This sort of restriction is, of course, not unique to measure DP adverbials. It is typical of numerous adverbials (as Moltmann 1991 shows), and indeed lies at the heart of the classic distinction between *for* and *in* PPs. But with PPs, the source of this restriction, if clearly identified, is often said to be the preposition (Dowty 1979, Krifka 1989). Here, though, no overt preposition is present. In the face of such a problem, a natural inclination is to posit a null preposition, and in fact null or deleted prepositions have been proposed quite

---

[4] I'll prefer the term 'homogeneous' both because it's consonant with Moltmann's discussion of other adverbials that impose this restriction and because measure DP adverbials are not exclusively temporal modifiers, but also members of the larger class of measure modifiers, which can measure along other dimensions than time. That said, I will not have much to say about (apparent) non-temporal measure DP adverbials such as *slide several inches*, other than to suggest that at least some of them may actually be temporal after all. (See sections 3.2.5 and 3.3.7).

a number of times in the analysis of DP adverbials (Emonds 1978a, 1987, Bresnan and Grimshaw 1978, McCawley 1988, Roeper et al. 2000). This assimilates measure DP adverbials to PPs, which are better-behaved, better-understood, more prototypical modifiers, so it's an appealing impulse—and it may actually be appropriate for *strong* DP adverbials, which pattern with PPs scopally and in distribution. But as the previous sections have observed, measure DP adverbials do not pattern with PPs in a variety of ways, and in assimilating them, such approaches leave no clear means by which to account for their different scope and distribution properties.

An alternative course one might pursue—indeed, one in some respects similar to what will ultimately be advocated here—is to posit not a null preposition, but rather, following Larson (1985), a generalized DP-adverbial-licensing feature that percolates from the head noun of a DP adverbial, which can perhaps be manipulated to impose the necessary restrictions. This, though, fails to make a crucial distinction in another respect—it would fail to distinguish measure DP adverbials from strong DP adverbials.

It's worth pausing to note, though, that it's certainly not the case that strong DP adverbials never impose an apparent homogeneity requirement or in some intuitive sense *feel* durative. There are, in fact, some strong DP adverbials with pretheoretically durative or 'measuring' meanings that do seem to have this effect:

(25) a. {All my life / My entire life / My whole life}, I've been indifferent to rutabagas.

b. {#All my life / #My entire life / #My whole life}, I've written a book about rutabagas.

But such cases are misleading—the internal semantics of the DP itself seems to be responsible for imposing the homogeneity requirement here. The DPs in (25) all involve either

universal quantification or an operation sufficiently similar to universal quantification for current purposes (see chapter 4). To take the clearest example, though, *all my life* quantifies over subintervals of my life. For an eventuality to have taken place all my life, it must be the case that it took place at every subinterval of my life. Only a homogeneous eventuality could satisfy this requirement, virtually by definition. Tellingly, and a bit surprisingly, sentences like (25) but without the universal quantifier(-like element) not only do not impose any such requirement, but are in fact ungrammatical:

(26) *My life, I've been indifferent to rutabagas.

The internal semantics of the DP, then, is crucial to the trick these adverbials perform. Strong DP adverbials that do not have this misleading characteristic do not have pretheoretically durative semantics or impose the homogeneity requirement:

(27) a. Clyde wrote a book every year.
b. Monday, Floyd ate a rutabaga.

This Aktionsart restriction, then, is not a characteristic of strong DP adverbials generally. Measure DP adverbials, on the other hand, systematically impose this restriction, even though it is not readily apparent from independently necessary assumptions about their internal semantics why they should.

3.2.4 *A Puzzle Measure DP Adverbials Don't Raise: The Head Noun Restriction*

DP adverbials are generally subject to idiosyncratic restrictions on which nouns can grammatically head them (Larson 1985):

(28) Floyd played the ukulele $\left\{ \begin{array}{l} \text{the wrong way} \\ \text{*the wrong manner} \end{array} \right\}$.

Measure DP adverbials, though, are not subject to such a restriction. Any noun that can denote a property of intervals (or of spatial regions) seems to yield a grammatical measure DP adverbial.

3.2.5 *Spatial DP Adverbials*

There are are apparent measure DP adverbials that involve spatial nouns:

(29) a. The soap slipped several inches.

b. On his way to Philadelphia, Clyde slept several miles.

The precise nature of these expressions, though, is a bit less clear than for their straightforwardly temporal counterparts. But they do appear to have the distinguishing properties noted above.

One methodological complication is that such measure DP adverbials must not be confused with argument spatial DPs, like the objects of *run* or even *fly*, which can passivize (cf. Kural 1996):

(30) a. Many miles were run by Clyde.

b. Many miles have been flown by this plane

(31) *Seven inches were slipped by the soap.

Argument spatial DPs need not even denote distances, unlike spatial measure DP adverbials:

(32) a. Clyde ran the race.

b. This plane has flown this route.

(33) *The soap slipped the counter.

Among the other complications that make this somewhat treacherous terrain is that it's not always apparent that the Aktionsart restriction imposed is quite the same—*Floyd drove 20 miles to Greenfield* seems quite natural and can be interpreted as an accomplishment—and that it's not entirely clear to what extent some of these are genuinely spatial (see section 3.3.7). So these at least warrant a certain amount of additional caution.

3.2.6 *The Facts in a Nutshell*

Measure DP adverbials are quantificationally weak, cannot take wide scope, occur in a restricted set of positions low in the tree, and impose an aspectual homogeneity requirement.

3.3 Measure DP Adverbials and Functional Structure in the Extended Verbal Projection

3.3.1 *The Essential Idea*

The challenge measure DP adverbials present is twofold (at least). First, an explanation is necessary of the Aktionsart requirement. This does does not obviously arise from properties of the DP itself—that is, there is nothing about the internal semantics of the DP that might be expected to contribute this requirement. Second, an explanation is necessary of the particular pairing of scope and distribution measure DP adverbials manifest.

Both problems can be addressed simultaneously. If the Aktionsart information associated with a measure DP adverbial does not come from within, the natural alternative is to suppose that its source lies outside the adverbial. What might this source be?

An answer to this question may arise from a solution to the other problem, that of scope and distribution. Measure DP adverbials are restricted to a fixed syntactic position, a restriction one can naturally understand as a need to occupy precisely one fixed 'slot' in verbal functional structure. But since they are not actually adverbs—that is, members of the category Adv—but merely adverbial modifiers, this does not follow from the common assumption that adverbs occur in fixed positions around which heads move (Emonds 1978b and many others). That it turns out, unexpectedly, to be true therefore seems important. There is a fairly straightforward means by which one might express this common property, made available by two distinct analytical currents in previous research that converge here on a common structure: the notion that measure phrases occupy (fixed) specifier positions (Abney 1987, Corver 1990, Grimshaw 1991, Kennedy 1997) in the adjectival projection, and the notion that adverbs occupy (fixed) specifier positions in the verbal projection (Cinque 1999, Alexiadou 1997, Laenzlinger 2000, and many others). Much more needs to be said, of course. But, supposing for the moment that measure DP adverbials therefore likewise occupy a specifier position at some appropriate fixed point (in this case, low in the extended VP), a need immediately arises for some independent structural element that they can be the specifiers of.

At this stage, then, the answer to one part of the puzzle would be a component of meaning—the homogeneity requirement—in need of a bit of structure to call home, and the answer to another part of the puzzle would be a bit of structure—the head to which measure DP adverbials may be specifiers—with no semantic work to do. Thus it seems natural to attribute the semantics of the homogeneity requirement to this bit of structure. So, given all

this, measure DP adverbials can be said to be introduced as specifiers to some element in verbal functional structure that imposes the homogeneity requirement and occupies a fixed position relatively low in the tree.

3.3.2 *Phrase-Structural Position*

One might imagine approaching the syntax of these expressions from (at least) two perspectives. Measure DP adverbials do in many ways resemble measure phrases in other categories—like other measure phrases, they occupy a fixed position, are obligatorily weak,[5] may bear morphological accusative case in languages that have it (more on this in a moment), and have a measuring semantics. Perhaps, then, measure DP adverbials and measure phrases in other categories share structural similarities as well.

Following Kennedy (1997) (who is himself following in this respect Abney 1987, Corver 1990, Grimshaw 1991, Kennedy 1997), I'll assume the structure in (34) for measure phrases in the extended AP:[6]

---

[5] That measure phrases more generally are obligatorily weak may not actually be an entirely uncontroversial claim, or in any case, it is certainly not an observation that is universally emphasized. There are apparent counterexamples, but I believe them to be only apparent. Among these are expressions like *this tall*, which seem to involve a measure phrase consisting of only a demonstrative. But *this* is not in fact a measure phrase here, but rather a degree head like *very* or *more* that happens (presumably for the obvious historical reasons) to be homophonous with a demonstrative. This is reflected in its incompatibility with comparatives, as (i) shows, and in its inability to occur with an NP complement, as the determiner *this* can, as (ii) shows:

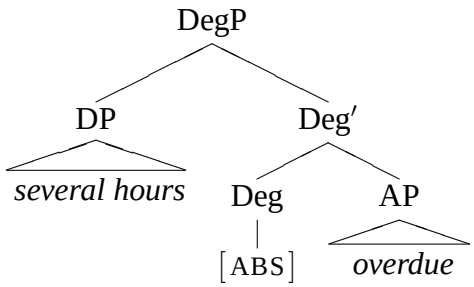
(i) Clyde is $\left\{ \begin{array}{l} \text{three feet} \\ \text{*this} \end{array} \right\}$ taller that his filing cabinet.

(ii) a. George is dishonest to this level.
b. George manifests this level of dishonesty.
c. *George is this level dishonest.

This many is of course a possible measure phrase, but it is also clearly weak (as its ability to occur in the existential construction reflects: *There were this many books on the table*; Milsark 1976)—it is merely an instance of degree-word *this* occurring with a comparative quantifier (Hackl (2000)), structurally similar to *very many*.

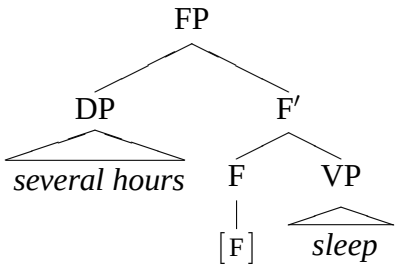
[6] Measure phrases in PP have received relatively less attention (in the generative tradition), though see

(34)



Here, the measure phrase *several hour* occupies the specifier position of a degree head. The particular degree head here happens to be a phonologically null one, [ABS], which Kennedy associates with measure phrases in absolute[7] (i.e., non-comparative non-superlative) APs, though it could just as easily have been *more*.[8] Blindly translating this structure to the verbal domain, (35) results:

(35)



Since DegP lacks an obvious analogue in the verbal domain, the corresponding projection has for the moment been labeled simply FP; for the same reason, the counterpart of [ABS] is rendered as [F]. In (35), the order in which *several hours* and *sleep* occur is of course different from their surface order. This will be corrected by the effect of verb movement. Assuming verb movement in English is present though short (Johnson 1991, Runner 1995

Zwarts (1997), Zwarts and Winter (2000) and Winter (2001); measure phrases inside DP present numerous idiosyncratic complications not relevant here.

7 In more recent work he uses the term ‘positive’.

8 For speakers for whom *overdue* is gradable.

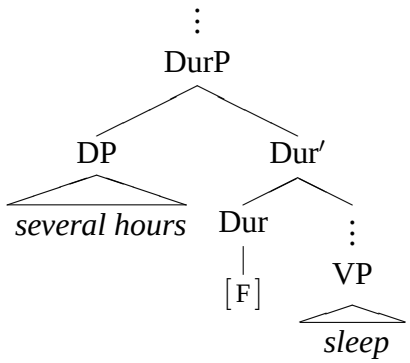
a.o.), it will independently derive the surface order from (35) without anything further being said:

(36) sleep$_i$ [$_{FP}$ several hours [F] [$_{VP}$ t_i]]

This sort of structure, then, maintains a parallel with measure phrases cross-categorially, and through its interaction with independent facts about VP yields the English surface order.

Alternatively, though, one might follow another analytical course. Since measure DP adverbials are after all adverbials, perhaps it may be more appropriate to look instead to adverbs for structural parallels—that is to say, true adverbs, not simply adverbial expressions of some sort, since measure DP adverbials manifest classic characteristics of adverbs proper like syntactic immobility and fixed scope. The adverbs most similar to measure DP adverbials are durative ones like *briefly*, which Cinque (1999) addresses specifically, proposing a functional projection Dur(ative)P whose specifier position these such adverbs occupy. Supposing, as seems natural, that measure DP adverbials occupy precisely the same position, a structure like (37) results:

(37)



This elides the many other verbal functional projections that Cinque (famously) proposes,

since none of them will be directly relevant here individually—all that is important is that DurP be in the appropriate hierarchal position relative to them, i.e., very low in the tree. Cinque proposes a feature-checking mechanism by which adverbs are bonded to the projections whose specifiers they must occupy. Thus the [F] feature[9] occupies the head of DurP, and, for Cinque, adverbs would be forced into its specifier to check this feature. But since on the current account only the feature will be necessary, and not the full feature-checking mechanics Cinque appeals to, the [F] feature is not instantiated on *several hours* in (37). (I will, though, assume this projection is Cinque's DurP, though nothing hinges on this.) Again, verb movement would independently be expected to bring about the English surface order.

Thus, as already suggested, both these analytical roads lead to essentially the same syntactic destination. This is unlikely to be an accident, particularly because the structures on which (35) and (37) are based emerged separately in entirely independent threads of research. So a structure like (37) will be adopted here. Having done this, an obvious suspect emerges with respect to the aspectual homogeneity requirement—namely, the feature that, in both versions of this structure, heads the projection whose specifier the measure DP adverbial occupies. To reflect this semantic role, it will be renamed [+HOMOGENEOUS] here.

Importantly, though, although these structures reflect distinct functional projections associated with the [+HOMOGENEOUS] feature, this is not really what's at the heart of the proposal, and one might imagine other ways of implementing the core idea here. For example, [+HOMOGENEOUS] may be bundled with other features in some other functional projection, or for that matter it may be on the lexical head directly. In these cases, it could

---

9 Cinque doesn't call his feature this; I use [F] again only for consistency.

have denotations of the same type as it would in a structure such as (37), but be interpreted by function composition.[10] Thus the question of whether this feature occupies a distinct functional projection is in principle independent from the basic idea here that a semantically contentful Aktionsart-related feature in the functional structure of VP licenses measure DP adverbials.

That said, (37) in its exactly current form does reflect the intuition behind this approach quite naturally, and has the advantage of placing measure DP adverbials on a left branch—which, it will emerge, is independently necessary. And it accords with Cinque's broader vision, in which adverbs generally are specifiers of functional projections.[11]

3.3.3 *Case*

Since measure DP adverbials are after all DPs, some account is necessary of how their case is licensed—a particularly pressing question given that they are also adverbial, and hence don't enter naturally into the usual system of case-checking mechanisms that license argument DPs.

This issue cannot merely be swatted away by supposing that, precisely because measure DP adverbials are adverbial, they are somehow exempted from the demands case places on other DPs. (Perhaps, for example, one might suppose they are excused in this way by some version of Chomsky 1986's Visibility Criterion, which links the case needs of a DP to its thematic role.) The difficulty is that, measure DP adverbials do not seem to

---

[10] For a more explicit elaboration of this idea, see the discussion of feature bundling in chapter 1.

[11] In addition, this entire approach is an instance of the mediated means of introducing modifiers advocated in other chapters as well and most explicitly in chapter 1. In accord with the goal of motivating the idea from scratch in each empirical area, I relegate this observation to this footnote, and will not rely on other chapters in the argumentation here.

be part of some secondary non-thematic system of oblique cases. Rather, in languages that show case overtly, they are systematically accusative:

(38) POLISH

Dwa lata pracowaliśmy nad tym projektem.
two.ACC years.ACC worked.2PL on this project

'We worked on this project for two years.'

(39) FINNISH (Mitchell 1991, via Pereltsvaig 2000)

Hän asui siellä yhden vuoden.
s/he.NOM lived there one.ACC year.ACC

'S/he lived there one year.'

(40) GERMAN (Paxton 1986)

Er war einen Monat in London.
He was one.ACC month in London.

'He was in London for a month.'

(41) KOREAN (Wechsler and Lee 1996)

Tom-un twu sikan-tongan-ul
Tom-TOPIC two hour-period-ACC

tali-ess-ta.
run-PAST-DECLARATIVE

'Tom ran for two hours.'

(42) LATIN (Wechsler and Lee 1996)

Decem annos regnavit.
ten years-ACC he.reigned

'He reigned for ten years.'

These facts don't seem to be the consequence of some fairly superficial coincidence. In fact, more prototypical measure phrases—ones other than measure DP adverbials—demonstrate this fondness for the accusative as well:

(43) POLISH

Pracowaliśmy dwa lata dłużej.
worked.2PL two.ACC years.ACC longer

'We worked two years longer.'

(44) GERMAN (Paxton 1986)

Ich bleibe nicht einen Tag länger.
I stay NEG one.ACC day longer

'I won't stay a single day longer.'

So at least in these languages, there is a systematic correspondence between measure DP adverbials and other measure phrases, and between measure phrases generally and accusative case.[12]

A structure like the one adopted here provides a straight-forward means of understanding these case properties. Since measure DP adverbials on this view occupy the specifier of Dur, they can check their case features there against it. This immediately frees them

[12] This correspondence also, incidentally, provides support for the Kennedy (1997)-style treatment of measure phrases in AP.

from the need to weasel out of meeting the obligations case imposes. And it seems natural, on this view, that the case they bear should be a structural one, because it is checked in the specifier of a verbal functional projection, just like other structural cases. That this case should be fairly consistent cross-linguistically is also unsurprising from this perspective—after all, functional elements tend to be fairly consistent cross-linguistically. As a kind of bonus, in connecting measure DP adverbials and other kinds of measure phrase by assigning them similar structures, this approach helps shed light on the case properties of measure phrases more generally.

Case, and particularly its connections to semantics, is a rich and complicated domain of inquiry and demands more attention than it could be given here without risking losing sight of the more local empirical goals, so I will not go much further down this road than this. (For more on these issues, see Wechsler and Lee 1996, Pereltsvaig 2000, Kratzer 2002b, and references there.)

3.3.4 *The Denotation of* [+HOMOGENEOUS]

Semantically, as already suggested, the [+HOMOGENEOUS] feature takes as arguments a VP denotation (a property of events) and a measure DP adverbial (a property of intervals), and imposes the homogeneity requirement:[13]

(45) $[\![[\text{+HOMOGENEOUS}]]\!] = \lambda P_{\langle s,t\rangle}\ \lambda Q_{\langle i,t\rangle}\ \lambda e\,.\,P(e) \wedge Q(\tau(e)) \wedge \forall t[t \sqsubseteq \tau(e) \rightarrow \exists e'[P(e') \wedge \tau(e')=t]]$

[13] τ maps an event to its running time, and $\sqsubseteq$ is a part relation (perhaps, for the sake of explicitness, the one of Moltmann 1991, 1997). Though (45) is sufficient for current purposes, a more precise formulation would reflect that the homogeneity requirement is a presupposition. This reflects a quantificational approach to representing the homogeneity requirement (Dowty 1979), but I am not aware of any compelling evidence from measure DP adverbials against having taken a measure function approach (Krifka 1989) instead.

This requires that every part of an event satisfy the predicate expressed by the VP. This formulation is roughly patterned after the denotations for measure adverbials proposed by Moltmann (1991). Essentially, this feature contributes what *for* might be taken (on a Dowty 1979-style approach) to contribute in a PP, with the order of arguments reversed. In this respect, (45) reflects the intuition that the semantic work that might otherwise have been done by *for* here must be done by the position of the modifier itself (relative to verbal functional structure).

That measure DP adverbials are always weak and do not QR—and hence have fixed in situ scope—follows in light of (45) from their property interpretation.[14]

A structure such as (37) would thus be interpreted as in (46):

(46) *an hour* [$[$+HOMOGENEOUS$]$ *sleep*]

$[\![\,[\text{+HOMOGENEOUS}]_{Dur}\,]\!]([\![\,\textit{sleep}\,]\!])([\![\,\textit{an hour}\,]\!])$

$= \lambda e\,.\,\text{sleep}(e) \wedge \text{an-hour}(\tau(e)) \wedge \forall t[t \sqsubseteq \tau(e) \rightarrow \exists e'[\text{sleep}(e') \wedge \tau(e')=t]]$

3.3.5 *The Role of Feature Checking: None*

A desirable consequence of what has been said so far is that there is no need for any specialized feature-checking to account for the distribution of measure DP adverbials. Since the $[$+HOMOGENEOUS$]$ feature on which they depend occupies a fixed position in the verbal projection, they too will occupy a fixed position. But why, one might reasonably ask, can't these DPs occur in other positions?

---

[14] I am being a bit sloppy here, in that if measure DP adverbials denote properties, they're not scope-bearing in a direct sense—to be a bit more precise, the existential quantifier over events has scope, and the measure DP adverbial, being necessarily inside its scope, can in an extended sense to itself have the scope of this quantifier as its scope. I will continue to indulge in this simplification.

The answer comes from two considerations, each rooted in one of the two preceding sections: case and interpretability. Because measure DP adverbials are DPs, they need to check case. So they need to occupy—at least at some point in the derivation—a case position. And because they denote properties of intervals, they will only be interpretable in the relatively few positions where properties of intervals are interpretable. Thus these DPs can't simply adjoin to some higher verbal projection—AspP, say, which might plausibly be semantically appropriate—because they could not their check case features there, and they could not occur in an arbitrary case position unless it meets their semantic requirements.

3.3.6 *Scope*

Since measure DP adverbials are interpreted as arguments of a feature in a fixed position, their scopal characteristics will follow from where this position is located. Thus, to account for the scope facts, it will have to be below the lowest landing site for QR, below negation, and below Aspect (where, I'll assume, the generic quantifier in habituals is located):

(47) a. Clyde didn't sleep an hour.

n't [*Clyde* [*an hour* [[+HOMOGENEOUS] *sleep*]]]

b. No one slept an hour.

no one [*an hour* [[+HOMOGENEOUS] *sleep*]]

c. Clyde swam a year.

[$_{\text{AspP}}$ GEN [*Clyde* [*a year* [[+HOMOGENEOUS] *swam*]]]]

If Aspect is where existential quantification over events is introduced (Kratzer 1998, elsewhere), this position of [+HOMOGENEOUS] will follow from its type and the more

general prohibition on nonpersistent quantifiers inside the scope of existential quantification over events.

If the DP adverbial is on a left branch, as suggested here, low scope relative to embedding verbs will follow as well:

(48) Greta wanted to talk an hour.

Greta wanted to [*an hour* [[+HOMOGENEOUS] *talk*]]

Since the feature in (48) is in the lower clause, it will necessarily scope below *wanted*. If *an hour* and its licensing feature had been introduced in the higher clause, the resulting sentence (after verb movement) would not have been the one in (48), but rather (49):

(49) Greta wanted an hour to talk.

Greta [*an hour* [[+HOMOGENEOUS] *wanted to talk*]

Because, of course, *wanted to talk* cannot head-move like *wanted*, there is no way to derive (48) from (49).

3.3.7 *A Note on Spatial Measure DP Adverbials*

The [+HOMOGENEOUS] feature is defined in (45) to apply to properties of intervals, and it involves a temporal trace function. To extend the account to spatial measure DPs, one might formulate a denotation neutral between these.[15] Alternatively, one might suppose

---

[15] Perhaps what is required is something like (i), again in the spirit of Moltmann (1991), where Q is a property of intervals or regions, a is an interval or region, and AT is a relation between an event and an interval or region that coincides with it (i.e., is either the temporal or spatial trace of e):

(i) $[\![[_{+\text{HOMOGENEOUS}}]]\!] = \lambda P_{\langle s,t\rangle}\ \lambda Q\ \lambda e\ .\ \exists a[P(e) \wedge \text{AT}(e,a) \wedge Q(a) \wedge \forall b[b \sqsubseteq a \rightarrow \exists e'[P(e') \wedge \text{AT}(e', b)]]]$

that there are two distinct features that impose temporal and spatial homogeneity. Despite appearances, though, it may be that spatial measure DPs are only apparently spatial and actually temporal. If Clyde is traveling as a passenger in a car, (50) may be true:

(50) Clyde slept several miles.

If Clyde is sleeping fitfully in bed, however, rolling back and forth, and the bed is several feet across, (51) is not true:

(51) Clyde slept several feet.

Yet it seems clear that the spatial trace of Clyde's sleeping extends several feet. Even in a pragmatically somewhat less plausible circumstance in which Clyde rolls across his bed exactly once during the course of the night, without ever retreating, (51) would not be true. (Thus even what Krifka 1998 calls 'strict movements' may not be sufficient.)

What this may be taken to suggest, then, is that at least certain spatial measure DPs are in fact instances of spatial nouns coerced into temporal interpretations in particular circumstances that naturally support this effect.[16]

3.3.8 *True Adverbs*

True adverbs may provide independent support for this approach. On syntactic grounds, Cinque associates his DurP projection with durative adverbs such as *briefly* and

---

[16] Strong DP adverbials manifest what may be a similar effect. One may claim of an assembly-line worker in a chair factory that she *took a break every third chair*.

Italian *lungamente* ‘long’. It seems reasonable, then, to suppose that such adverbs are also interpreted as arguments of the [+HOMOGENEOUS] feature.

This would predict that all adverbs of this class should impose a homogeneity requirement. This seems to be the case:

(52) a. Clyde slept briefly.

b. Greta ran briefly.

(53) a. #Clyde noticed the difficulty briefly.

b. #Greta died briefly.

Similar facts obtain for *momentarily*[17] and the rather marginal *lengthily*:

(54) a. Clyde slept {momentarily / lengthily}.

b. Greta ran {momentarily / lengthily}.

(55) a. #Clyde noticed the difficulty {momentarily / lengthily}.

b. #Greta died {momentarily / lengthily}.

Without the feature proposed above, it would be necessary to encode this homogeneity requirement independently in the denotation of each of these adverbs:

---

[17] When it occurs on the right, *momentarily* also has a reading paraphrasable as ‘a moment from now’. This is not the reading at issue here.

(56) a. $[\![\textit{briefly}]\!] = \lambda P_{\langle s,t\rangle} . P(e) \wedge \text{brief}(\tau(e)) \wedge \forall t[t \sqsubseteq \tau(e) \rightarrow \exists e'[P(e') \wedge \tau(e')=t]]$

b. $[\![\textit{momentarily}]\!] = \lambda P_{\langle s,t\rangle} . P(e) \wedge \text{momentary}(\tau(e)) \wedge \forall t[t \sqsubseteq \tau(e) \rightarrow \exists e'[P(e') \wedge \tau(e')=t]]$

c. $[\![\textit{lengthily}]\!] = \lambda P_{\langle s,t\rangle} . P(e) \wedge \text{lengthy}(\tau(e)) \wedge \forall t[t \sqsubseteq \tau(e) \rightarrow \exists e'[P(e') \wedge \tau(e')=t]]$

A means of accounting for the distribution of these adverbs would still independently be required.

If they were interpreted instead as arguments of the feature proposed above, though, they could simply denote properties of intervals, as in (57):

(57) a. $[\![\textit{briefly}]\!] = \lambda t . \text{brief}(t)$

b. $[\![\textit{momentarily}]\!] = \lambda t . \text{momentary}(t)$

c. $[\![\textit{lengthily}]\!] = \lambda t . \text{lengthy}(t)$

The homogeneity requirement, then, would follow from the interpretation of the feature of which these adverbs are an argument, so it need not be encoded redundantly in the lexical entry of every such adverb. This would simultaneously capture the generalizations that both durative adverbs and measure DP adverbials occur in a particular structural position and that they both systematically introduce this presupposition.

A secondary prediction of this is that these adverbs should obligatorily scope low. This appears to generally be the case:

(58) a. Clyde didn't sleep briefly. ($\neg \prec$ *briefly*; **briefly* $\prec \neg$)

b. No one slept briefly. (*no one* $\prec$ *briefly*; **briefly* $\prec$ *no one*)

c. Greta wanted to talk briefly. (*wanted* $\prec$ *briefly*; *?*briefly* $\prec$ *wanted*)

In (58a) and (58b), *briefly* does indeed seem to require narrow scope. In the clausal embedding example in (58c), though, the judgment is rather precarious, significantly more so than one might wish.

So, from this perspective, it's not any one particular expression that's associated with requiring homogeneity, but rather a particular position.

3.3.9 *Overt Morphology*

In languages in which Cinque's Dur head is spelled out overtly, it appears to have roughly the kind of durative, Aktionsart-related semantics suggested for the [+HOMOGENEOUS] feature here. One of the potential overt occupants of Dur Cinque suggests is the Guyanese Creole particle *de*, which, according to Gibson (1992) 'occurs as a durative aspect marker':[18]

(59) GUYANESE CREOLE (Gibson 1992)

a. Yu get fren wid hii, yuu an hii de nais.
   'You become friendly with him, you and him will get along well'

b. Fu faiv yeerz ii de woking in di bush.
   'For five years he has been working in the bush'

Something similar could be said for the Tauya formative *tei* 'for a long time' (MacDonald 1990), and for Central Alaskan Yupik *uma* 'for long periods' (Mithun and Ali 1996).

---

[18] Gibson doesn't provide word-by-word glosses.

3.4 A Final Word

The analysis of measure DP adverbials requires an account of their narrow scope and low structural position, their obligatory weakness, and the durativity/homogeneity requirement they impose. In the approach suggested here, part of the apparent semantic contribution of a modifier—a measure DP adverbial or a *briefly*-class adverb—is attributed instead to a fixed position in functional structure. This provides a way for accounting for the essential properties of measure DP adverbials, and leads to a simplified view of the interpretation of durative adverbs as well. Importantly, though, it also results in a treatment of these expressions in which they are parallel to measure phrases in other categories, simply the verbal exponent of the more general phenomenon of measure-phrase modification. Correlations of the kind explored here between the position and interpretation of modifiers are quite common, though, and measure DP adverbials (and *briefly*-class adverbs) are not unusual in manifesting them. Other chapters of this thesis consider in the same spirit other modifiers that seem to manifest such correlations, thus perhaps supporting to this approach by demonstrating that it might fruitfully be applied more widely. But on the basis of the facts considered in this chapter alone, it seems fair to say that measure DP adverbials have a cluster of properties consistent with supposing that they are in a sense argument-like—to be sure, not true arguments of the verb, but also not modifiers in the purest sense. Rather, they show signs of a dependence not on a particular verb, as true arguments do, but rather on verbhood itself, that is, on properties general to all verbs. Treating them as arguments of an element of verbal functional structure accords with this intuition.

CHAPTER 4

ALMOST MODIFIERS ACROSS CATEGORIES

4.1 Introduction

In some respects, *almost* appears to be not one modifier but two. One incarnation of *almost* is a verbal modifier:

(1) Clyde almost killed Floyd.

This *almost* gives rise to a variety of readings that have been taken as evidence for lexical decomposition of verbs (Morgan 1969, McCawley 1972, Rapp and von Stechow 1999). In another incarnation, *almost* is a nominal modifier:

(2) Almost every chiropractor plays the ukulele.

This incarnation imposes a restriction on quantifiers that has made it useful as a diagnostic for universal quantification (Carlson 1981, Kadmon and Landman 1993). It is apparently not usual to consider these forms of *almost* together. Indeed, one might suppose that these are two entirely distinct modifiers that happen to be homophonous, simply by chance.

This chapter is essentially identical to Morzycki (2001a).

But this cannot be the case. If it were indeed some accident of the lexicon that *almost* manifests these two apparently divergent sets of syntactic and semantic characteristics, it should be unique, or nearly so, in this respect. Yet in fact, *almost* shares all of its essential characteristics with a natural class of modifiers—henceforth '*almost* modifiers'—of which it is one quite unexceptional member. Among its kin in this class are *virtually, nearly, damn near, pretty much, not quite,* and *just about* (though not bare *about*). All the modifiers of this class occur in the same positions, give rise to the same range of readings, and impose similar restrictions on the expressions they modify. Moreover, the restrictions they impose on different categories are intuitively similar to each other. Thus, because *almost* shares its properties with a significantly larger class of expressions, its behavior in the nominal and verbal domains can be no accident, but rather must reflect a more general puzzle.

The aim here is to elucidate that puzzle and suggest an approach to solving it. The essential distinguishing properties of the *almost* class are examined in section 3.1. Several possible approaches to the problem are considered in section 3.2. Then, in section 3.3, an analysis of *almost* modifiers in terms of intensional similarity is pursued that may be extended across categories. The remaining steps toward an account of this cross-categoriality are taken in section 3.4, which proposes a division of semantic labor between *almost* modifiers themselves and a syntactic feature in functional structure that licenses them. Section 3.5 concludes.

4.2 Some Data: Distribution and Selectional Restrictions

4.2.1 *Distribution: Cross-Categorial and Left Peripheral*

Almost modifiers occur in (the extended projections of) DP, VP, AP, PP, and AdvP, always on the extreme left periphery:[1]

(3) a. {Almost / Nearly / Practically} every chiropractor plays the ukulele.

b. Tilden {almost / nearly / practically} won.

c. The soup is {almost / nearly / practically} free of insects.

d. Floyd sailed {almost / nearly / practically} around the world.

e. The candidate behaved {almost / nearly / practically} inexcusably.

For each of these categories, the left edge is the only possible location an *almost* modifier may occupy.

---

[1] Three *almost* modifiers are taken to be representative here and subsequently, but similar facts hold for the whole class.

4.2.2 *The Nominal Restriction: Universals and Numerals*

In DP, *almost* modifiers normally require either a universal quantifier (Carlson 1981, Kadmon and Landman 1993) or a numeral (Partee 1986):

(4) a. {Almost / Nearly / Practically} every chiropractor plays the ukulele.

b. {Almost / Nearly / Practically} all chiropractors play the ukulele.

c. {Almost / Nearly / Practically} twenty chiropractors play the ukulele.

(5) a. #{Almost / Nearly / Practically} {a / some / the} chiropractor plays the ukulele.

b. #{Almost / Nearly / Practically} {some / several / many} chiropractors play the ukulele.

Negative determiners are of course possible as well, presumably a reflection of the same fact:

(6) a. {Almost / Nearly / Practically} no chiropractor plays the ukulele.

b. {Almost / Nearly / Practically} none of the chiropractors plays the ukulele.

With measure nouns, though, existential quantifiers are possible: *almost a pound of cheese*. Likewise, determiners otherwise incompatible with *almost* modifiers are legal in the presence of the adjectives entire and whole, which apparently suffice to satisfy the requirement of universal quantification: *almost the entire book*.

4.2.3 *The Adjectival Restriction: Non-Gradable Adjectives*

In AP, *almost* modifiers are odd with gradable adjectives (Hitzeman 1992):[2]

(7) a. Herbert is { almost / nearly / practically } dead.

b. a(n) { almost / nearly / practically } defunct restaurant

(8) a. #Herbert is { almost / nearly / practically } tall.

b. #a(n) { almost / nearly / practically } popular restaurant

The gradable adjectives in (8) may occur with these modifiers only to the extent that they can be coerced into non-gradable interpretations. For example, (8a) is well-formed only to

---

[2] Sevi (1998) argues that this is not quite the case—he finds *almost happy* relatively normal. This may, however, be a difference in how easily various normally gradable adjectives can be coerced into nongradable interpretations.

the extent that one can conceive of tall as a clear-cut, discrete predicate that partitions its domain into the clearly tall and the clearly non-tall.

As might be expected, though, an AP whose degree argument is saturated through comparative morphology or an overt measure phrase may occur with an *almost* modifier:

(9) a. Herbert is {almost / nearly / practically} taller than Clyde.

b. Herbert is {almost / nearly / practically} six feet tall.

4.2.4 *The Verbal Restriction: A Telicity Effect*

In VP, there is a contrast between the effect of *almost* modifiers on telic and atelic eventualities:

(10) a. Clyde {almost / nearly / practically} ran into a shopping cart.

b. Clyde {almost / nearly / practically} died.

c. Clyde {almost / nearly / practically} reached the top.

(11) a. (#)Clyde { almost / nearly / practically } ran around.

b. (#)Clyde { almost / nearly / practically } slept.

c. (#)Clyde { almost / nearly / practically } knew the answer.

An interpretation[3] in which the event is claimed not to have culminated is possible only with VPs denoting telic eventualities, as in (10). With atelic eventualities, as in (11), no such interpretation is normally possible. It may perhaps be coerced, but only to the extent that the eventuality in question can be coerced into a telic reading. (None of this is to deny, of course, that there are other interpretations possible for (11).)

4.2.5 *The Adverbial Restriction: Non-Gradability Again, Universality Again*

In adverbial categories, restrictions similar to those already noted reemerge. *Almost* modifiers may occur in AdvP under roughly the same circumstances as in AP—gradable de-adjectival adverbs are odd:

(12) a. The candidate behaved { almost / nearly / practically } inexcusably.

b. #The candidate behaved { almost / nearly / practically } strangely.

---

[3] I use 'interpretation' as a term neutral with respect to whether this is ambiguity or vagueness.

As with adjectives, a non-gradable interpretation must be coerced to salvage a sentence like (12b).

The similarity is not to AP alone, however. With quantificational adverbs a universality requirement analogous to that imposed in DPs manifests itself:

(13) a. Chiropractors can {almost / nearly / practically} always play the ukulele.

b. #Chiropractors can {almost / nearly / practically} {often / usually / sometimes} play the ukulele.

So, as with their determiner counterparts, *almost* modifiers require universality of quantificational adverbs.

Somewhat more mysteriously, *almost* modifiers cannot modify adverbs of certain classes, including speaker-oriented and domain adverbs:

(14) a. *{Almost / Nearly / Practically} {frankly / honestly / truthfully}, Greta can play the ukulele.

b. *{Almost / Nearly / Practically} {legally / officially / anatomically}, Clyde is a chimpanzee.

Indeed, *almost* modifiers seem best compatible with manner and quantificational adverbs.

The now-familiar restrictions arise in PP as well. One way to understand the contrast in(15) is as a kind of universality restriction:

(15) Floyd learned Portuguese $\left\{\begin{matrix}\text{almost}\\\text{nearly}\\\text{practically}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{without}\\\text{*with}\end{matrix}\right\}$ assistance.

Without may involve universal quantification over its object.[4] *With*, however, likely does not involve universal quantification.

4.2.6 *A Glance Outside English*

It does not appear to be a quirk of English that words meaning 'almost' have a cross-categorial distribution and occupy a left-peripheral position. This is the case in a number of languages:[5]

(16) GERMAN (Rapp and von Stechow 1999)

a. Fast alle Pflanzen waren vertrocknet
almost all plants were dry

b. Alle Pflanzen waren fast vertrocknet.
all plants were almost dry

c. ... weil David seinen Hasen fast erwürgte.
because David his rabbit almost strangled

(17) POLISH (vetted by Ania Łubowicz, p.c.)

a. $\left\{\begin{matrix}\text{Prawie}\\\text{Niemal}\end{matrix}\right\}$ każda roślina jest sucha.
almost every plant is dry

---

[4] Say, in this case taking the form of quantification over pieces or instances of assistance: ∀x[assistance(x) → ¬Floyd learned to speak Portuguese with x].

[5] Bob Rothstein (p.c.) points out that the difficulty with *niemal* in (17c) is not present with the variant form *omal nie*.

b. Każda roślina jest $\left\{\begin{matrix}\text{prawie}\\ \text{niemal}\end{matrix}\right\}$ sucha.
every plant is almost dry

c. David $\left\{\begin{matrix}\text{prawie}\\ \text{?niemal}\end{matrix}\right\}$ zadusił jego królika.
David almost strangled his rabbit

(18) HEBREW

a. kim?at kol cemax hu yaveš (Uri Strauss, p.c.)
almost every plant HU dry

'Almost every plant is dry.'

b. ha-ramzor kim?at yarok (Sevi 1998)
the-traffic.light almost green

'The traffic light is almost green.'

c. dani kim?at lo tiken et ha-mexonit ha-zot (Sevi 1998)

'Danny almost didn't fix this car.'

(19) ST'ÁT'IMCETS (LILLOET SALISH) (Lisa Matthewson, p.c.)

a. tqilh t'u7 tákem i ts'í7a qi-7-cw
almost PART all PL.DET deer-DET scared.away(inch)

'Almost all the deer left.'

b. tqilh t'u7 áolsem k Mary
almost PART sick DET Mary

'Mary almost got sick'

c. tqilh t'u7 kúkwpi7 k Mary
almost PART chief DET Mary

'Mary almost became a chief.'

Naturally, none of this is evidence that the semantics of the various expressions glossed here as 'almost' is invariant across these languages—in fact, it is not.[6] But it is evidence that the cross-categoriality and left-peripherality of English *almost* modifiers reflects a more general phenomenon.

4.2.7 *Summary*

Almost modifiers constitute a natural class distinguished by position and interpretation, whose members may occur in left peripheral positions across a range of syntactic categories. In all these categories, they impose intuitively similar semantic restrictions.

4.3 Some Analytical Possibilities

4.3.1 *Lexical Decomposition*

The various interpretations *almost* gives rise to have been a central argument for lexical decomposition of verbs (McCawley 1972, recently Rapp and von Stechow 1999). Although this debate is of course directly relevant to the analysis of *almost* generally, it appears to be largely independent of the questions immediately at hand here, and no new evidence will be brought to bear on it, so it will be noted only briefly.

Both McCawley and Rapp and Stechow take (20) to be three-ways ambiguous:

(20) John almost killed Harry.

---

[6] Rapp and von Stechow (1999), for example, discuss some differences between German *fast* and English *almost*.

Each of the three interpretations is slightly more gruesome than the next. On one interpretation, John is claimed to have almost done something. This would be how (20) is interpreted if it reports, for example, a circumstance in which John got extraordinarily angry at Harry, but resisted acting on his rage. On another interpretation, John is claimed to have actually done something which almost caused Harry's death. This would be the interpretation involved in reporting a circumstance in which John shot at Harry, missing him narrowly. On the third interpretation, John is claimed to have actually caused Harry to be almost dead. This would be the interpretation involved in reporting an instance in which John attacked Harry and wounded him seriously.

Notably, the entire class of *almost* modifiers has this range of interpretations:

(21) John $\left\{\begin{array}{l}\text{nearly}\\\text{practically}\\\text{virtually}\end{array}\right\}$ killed Harry.

Thus irrespective of whether these interpretations involve ambiguity or vagueness, they, like the distribution and selectional restriction facts noted above, are not a peculiarity of *almost*.

In a McCawley-style lexical decomposition theory, these effects can be understood as the consequence of whether *almost* takes scope over various elements of the decomposed verb. Rapp and Stechow approach *almost* (and its German counterpart *fast*) from this perspective, implementing lexical decomposition as in (22):

(22) John almost killed Harry.

a. John almost did something:

almost(w)(λw$\exists$e[AGENT$_{ew}$(John) $\wedge$

BECOME$_{ew}$(λwλs . DEAD$_{sw}$(Harry)))]

b. John did something, and it almost caused Harry's death:

$\exists$e[AGENT$_{ew}$(John) $\wedge$

almost(w)(λw BECOME$_{ew($(λwλs . DEAD$_{sw}$(Harry)))]

c. John did something, and it caused Harry to be almost dead:

$\exists$e[AGENT$_{ew}$(John) $\wedge$

BECOME$_{ew($(λwλs . almost(w)(λwDEAD$_{sw}$(Harry))])

These different interpretations are derived in this account on the basis of the relative scope of *almost*, the existential quantifier, and the BECOME predicate. Since for Rapp and Stechow, this semantic decomposition mirrors a decomposition in the syntax, these scope possibilities correlate with syntactic positions occupied by *almost* (or *fast*). This permits an account of correlations between the position of fast and the interpretations available.

A core virtue of such an approach is that there is a single *almost* with a single denotation—for Rapp and Stechow, a modifier of propositions. Variation in the interpretation arises entirely from its position.

Moreover, the telicity effect can be explained in these terms. If the interpretations of *almost* depend on its scope relative to various predicates in the decomposition, the absence of one of these predicates would reduce the number of possible scope possibilities and hence the number of interpretations. Since the BECOME predicate is absent in the representation of atelic eventualities, one interpretation of *almost* should be missing in these cases.

This amounts to exactly the selectional restriction noted above—the 'almost-culminated' reading is possible only with telic eventualities.

4.3.2 *Intensional Similarity*

Decomposition on its own, however, is not intended as a theory of what *almost* means, so much as a theory of why its meaning may vary. A theory of the essential meaning of *almost* is required independently.

The decomposition account of Rapp and Stechow builds on the denotation for *almost* in (22), though this is similar to denotations proposed in non-decompositional accounts as well (Sadock 1981, Sevi 1998):[7]

(23) almost(w)(p) = 1 iff there is a world w′ which is almost not different from w and p(w′) = 1 and p(w) = 0

This is appropriate in a lexical decomposition theory for VP *almost* (which is what Rapp and Stechow chiefly concern themselves with), but it does not extend straight-forwardly to DP *almost*.

First, (23) would have to be adjusted—through a type shift, say—so that *almost* can apply to generalized quantifiers, to yield (24):[8]

(24) $\text{almost}_2 = \lambda w \lambda Q_{\langle\langle e,st\rangle,st\rangle}\ \lambda P_{\langle e,st\rangle}$. almost(w)(Q(P)(w))

---

[7] This is altered from Rapp and Stechow's original in some purely notational ways.

[8] This assumes that *almost* applies to a DP rather than a D, but this is not crucial—a similar argument could be made from the alternative assumption.

This leads to a problem, however. It predicts that (25a) should be assigned an interpretation like (25b):

(25) a. Almost every plant is dry.

b. $\text{almost}_2(w)(\llbracket \textit{every plant} \rrbracket)(\llbracket \textit{is dry} \rrbracket)$

$= \text{almost}(w)(\llbracket \textit{every plant} \rrbracket)(\llbracket \textit{is dry} \rrbracket(w))$

$= \text{almost}(w)(\forall x[\text{plant}(x)(w) \rightarrow \text{dry}(x)(w)])$

That is, (25a) is predicted to mean roughly ‘it is almost the case that every plant is dry’. In a situation in which every plant is only minimally moist but no plant is dry, (25a) would be predicted to be true. This is not a possible reading.

Sevi (1998) proposes a denotation for *almost* like that in (23), but in which *almost* may, depending on context, operate on not just worlds but also intervals or standards of precision. This additional elaboration would not on its own help here, since what is required is a reduction rather than an increase in the multiplicity of possible interpretations.

4.3.3 *Exceptives*

If the problem in (25) is due to attempting to extend an account of VP *almost* to other uses, perhaps it could be avoided by taking an account of DP *almost* as a starting point instead.

There are apparently no worked-out accounts of DP *almost*, but von Fintel (1992–1993, 1994) mentions in an analysis of exceptive constructions that it may be possible to understand *almost* in these terms. In other words, perhaps (26a) is analogous to (26b):

(26) a. Almost every plant is dry.

b. Every plant except for a few is dry.

Von Fintel notes that certain exceptive constructions—English but phrases—actually impose a universality restriction on quantifiers rather like the one imposed by *almost*:

(27) $\left\{\begin{array}{c}\text{Every}\\ \text{*Some}\end{array}\right\}$ student but John attended the meeting.

Maybe, then, if DP *almost* can be understood as a kind of exceptive, other incarnations of *almost* may as well. Thus, perhaps (28a) is analogous to (28b):

(28) a. The plant is almost dry.

b. The plant is dry, $\left\{\begin{array}{l}\text{except for a few drops}\\ \text{except for some residual moisture}\\ \text{except for the topmost leaves}\end{array}\right\}$

In other words, one might suppose that *almost* allows a kind of fudging of the requirements a predicate would normally impose analogous to the kind of exceptions it permits in DP. A notion of exceptions may be developed that extends to properties: an exception to *dry*, for example, would be any property that an individual must have to be dry but in fact lacks.[9] *Almost* would require that there be few such exceptions.

There is something appealingly simple and intuitively satisfying about such an approach, and (26–28) can be taken as evidence that there is something to the analogy that lies at its core. And certainly, some general notion of exceptions or counterexamples plays an important role in everyday reasoning. But this approach presents some substantial difficulties that are not easily resolved, and that might be avoided while preserving some of its advantages by taking a slightly different route.

---

[9] More generally, and a bit more precisely: C is an exception to A(B) iff $\Box[A(B) \rightarrow C(B)] \wedge C(B)$.

The central problem involves the requirement that these exceptions be few. If exceptions may be properties, this would require counting properties. Counting properties, though, is trouble: Is having three wet leaves a single exception, three exceptions, or both?

One natural solution is to suppose the context of utterance provides a way of structuring the requirements a predicate imposes—for example, in a particular context a particular set of properties may count as the components of being dry. This may permit talk about properties being (proportionally) few to be sensible, but it requires making very deep commitments about what context does.

4.3.4 *Slack Regulation*

Building on the intuition that *almost* modifiers involve relaxation of truth conditions, one might imagine assimilating them to other expressions that may be said to have this effect. Lasersohn (1999) identifies a class of expressions he calls 'slack regulators', which includes exactly, more or less, perfectly, and all. These 'serve as signals of the intended degree of approximation to the truth'. But *almost* modifiers do not appear to be of this class, for several reasons.

First, slack regulators are not typically cross-categorial in the way *almost* modifiers are. Although there are slack regulators for various different categories, different slack regulators are specialized for different categories. Exactly and perfectly are apparently restricted to AP, all is restricted to DP (setting aside its floated form), and more or less is apparently relatively unrestricted. This contrasts sharply with *almost* modifiers, which are quite homogeneous with respect to the range of categories they may modify.

Second, slack regulators do not systematically impose consistent selectional restrictions in the way *almost* modifiers do. Every *almost* modifier imposes the same restrictions,

but the restrictions imposed by slack regulators vary. Exactly and perfectly, for example, differ in their compatibility with numerals: *exactly thirty chiropractors* is possible, but not **perfectly thirty chiropractors*.

Third, and probably most significant, slack regulators do not in general give rise to the inference that the sentence they occur in would be false if they were absent:

(29) a. More or less thirty chiropractors play the ukulele.
does not entail
b. It is not the case that thirty chiropractors play the ukulele.

(30) a. Clyde's head is perfectly spherical.
does not entail
b. Clyde's head is not spherical.

Yet this is an important property of *almost* modifiers:

(31) a. Almost thirty chiropractors play the ukulele.
does entail
b. It's not the case that thirty chiropractors play the ukulele.

There is some controversy about whether this is in fact an entailment (Sevi 1998, Rapp and von Stechow 1999) or an implicature (Sadock 1981, Atlas 1984). But whatever its status, this inference is a signature characteristic of *almost* modifiers. It would be hard to understand in terms of slack regulation.[10]

---

[10] Despite all these empirical differences, though, the conceptual machinery Lasersohn suggests for

4.4 Intensional Similarity, and Building Toward Cross-Categoriality

Each of the analytical possibilities considered so far presents certain problems. This section will attempt to resolve the principal problem confronting the most straight-forward extension of the intensional similarity approach, and consider some ways in which this may lead to an understanding of the restrictions *almost* modifiers impose.

4.4.1 *An Initial Word about Intensional Similarity*

The intensional approaches to *almost* involve asserting that there exists a world in which a proposition is true that is close or similar to the evaluation world (in the case of Rapp and von Stechow 1999, 'almost not different' from it). It is, of course, ultimately important exactly what is meant by these closeness relations. For current purposes, though, this issue will be sidestepped as in (32), an initial stab at the denotation of the AP incarnation of *almost*:

(32) AP *almost*: (not final)

$$[\![\, \textit{almost}_{AP} \,]\!] = \lambda P \lambda x \lambda w \,.\, \neg P(x)(w) \wedge \exists w'[P(x)(w') \wedge \textsc{close}(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge P(x)(w'')] \rightarrow w'' =_w w']]$$

What this requires is that a proposition be false in the evaluation and true in some close world, and that there be no closer world in which it is true. The closeness relation between

---

these expressions—'pragmatic halos'—bears a certain intuitive resemblance to what may be required for *almost* modifiers. Moreover, at least some apparent slack-regulators—absolutely and positively—bear a greater resemblance than others to *almost* modifiers. Absolutely and positively occur apparently quite cross-categorially and always left-peripherally, like *almost* modifiers, and may even impose some similar selectional restrictions. Horn (1972) notes that absolutely in DP may be restricted to universal quantifiers, a property which positively seems to share.

worlds here is CLOSE. A notion of comparative closeness between worlds is also put to use here: $w' \leq_w w''$(w' is at least as close to w as w'' is) iff all the propositions true both in w and in w'' are also true in w' (cf. Lewis 1981). Presumably, CLOSE and comparative closeness are not unrelated. The requirement that only maximally close worlds of the right sort be considered is not crucial at this point, but will prove useful later in defining *almost* modifiers like nowhere near, so it is included here for consistency of exposition.

4.4.2 *Trying to Fix the Problem with* Almost Every Plant

The problem noted above with simply extending a propositional analysis of *almost* to DP is that it would falsely predict that (33a) could mean (33b):

(33) a. Almost every plant is dry.
b. Every plant is almost dry.

A denotation like (34) for DP *almost* would yield sentence denotations like (35):

(34) $[\![\textit{almost}_{DP}]\!] = \lambda Q_{\langle\langle e,st\rangle,st\rangle} \lambda P_{\langle e,st\rangle} \lambda w \,.\, \neg Q(P)(w) \wedge \exists w'[Q(P)(w') \wedge$ (not final)
$\text{CLOSE}(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge Q(P)(w'')] \rightarrow w'' =_w w']]$

(35) $[\![\textit{Almost every plant is dry}]\!] = \lambda w \,.\, \neg\forall x[\text{plant}(x)(w) \rightarrow \text{dry}(x)(w)] \wedge$
$\exists w'[\forall x[\text{plant}(x)(w') \rightarrow \text{dry}(x)(w')] \wedge \text{CLOSE}(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge$
$\forall x[\text{plant}(x)(w') \rightarrow \text{dry}(x)(w'')]] \rightarrow w'' =_w w']]$

What (35) requires is that not every plant be dry in the evaluation world, and there be a world among the closest in which every plant is dry that is close to the evaluation world.

The source of the difficulty reflected in (33), then, is that the evaluation world and the close world to which it is being compared may vary not only in what proportion of plants are dry but also in which plants are dry, or indeed in whether any are. In other words, the worlds may vary not in how closely the evaluation world approximates universality of dryness, but also in how closely the evaluation world approximates total or actual dryness.

This can be fixed by simply imposing a requirement that the worlds not vary with respect to the extension of dry. Thus in (36), a requirement is added that only the DP denotation may vary between the compared worlds:

(36) DP *almost*:

$$[\![\textit{almost}_{DP}]\!] = \lambda Q_{\langle\langle e,st\rangle,st\rangle} \lambda P_{\langle e,st\rangle}\ \lambda w\,.\,\neg Q(P)(w) \wedge \exists w'[Q(P)(w') \wedge [\lambda x\,.\,P(x)(w)] = [\lambda x\,.\,P(x)(w')] \wedge \textsc{close}(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge Q(P)(w'')] \rightarrow w'' =_w w']]$$

Since this enforces identity of dry things between worlds, the problem in (33) will not arise. The desired reading, however, can still be attained by varying the quantity of plants among worlds. With this adjustment, *almost every plant is dry* would receive an interpretation roughly like ‘if there were slightly fewer plants, every plant would be dry’.

4.4.3 *The Universality Restriction*

The denotation in (36) may account for the incompatibility of *almost* with existentially quantifying determiners.

A sentence such as (37a) will receive the interpretation in (37b):

(37) a. #Almost $\left\{\begin{matrix}\text{a}\\ \text{some}\end{matrix}\right\}$ plant is dry.

b. $\lambda w\,.\,\neg\exists x[\text{plant}(x)(w) \wedge \text{dry}(x)(w)] \wedge \exists w'[\exists x[\text{plant}(x)(w') \wedge \text{dry}(x)(w')] \wedge$
$[\lambda x\,.\,\text{dry}(x)(w)] = [\lambda x\,.\,\text{dry}(x)(w')] \wedge \text{CLOSE}(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge$
$\exists x[\text{plant}(x)(w'') \wedge \text{dry}(x)(w'')]] \rightarrow w'' =_w w']]$

What (37b) requires is that there be no dry plants in the evaluation world, and that, in a nearby world in which all the same things are dry, there be some dry plants. This amounts to a requirement that a dry thing that is not a plant in the evaluation world be a plant in a nearby world.

This, however, is an exceptionally strange thing to say. Certainly, it does not amount to saying something like 'it is almost the case that a plant is dry'. To the extent that one can access intuitions about what sentences such as (37a) might mean—which is not great—it seems to me that they can indeed mean this, though it is fiendishly difficult to imagine a context in which one might want to express such a thing. The strangeness of existentially-quantifying determiners with *almost*, then, may follow from this pragmatic strangeness.

DPs with numerals—*almost thirty plants*—will thus be ruled out on the parse in which *almost* modifies the DP. But the alternative structure in which *almost* modifies thirty, a non-gradable adjective, will not be ruled out, just as the *almost* dead plants would not be.

4.4.4 *The Telicity Effect*

This sort of semantics may explain the impossibility of 'almost-culminated' interpretations with atelic eventualities. Since in all worlds, atelic eventualities by definition end but do not culminate, it will never be the case that there exists a world in which a particular atelic eventuality culminated. Yet this is exactly what this interpretation of *almost* would require:

(38) VP *almost*: (not final)

$[\![\textit{almost}_{VP}]\!] = \lambda P\ \lambda e\ \lambda w\ .\ \neg P(e)(w) \wedge \exists w'[P(e)(w') \wedge \textsc{close}(w)(w') \wedge$
$\forall w''[[w'' \leq_w w' \wedge P(e)(w'')] \rightarrow w'' =_w w']]$

(39) $[\![\textit{Clyde}_1 \textit{ almost slept } t_1]\!] = \lambda w\ .\ \neg \text{slept}(\text{Clyde})(e)(w) \wedge \exists w'[\text{slept}(\text{Clyde})(e)(w') \wedge$
$\textsc{close}(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge \text{slept}(\text{Clyde})(e)(w'')] \rightarrow w'' =_w w']]$

It is perfectly possible for (39) to be true, but it can never be true by virtue of Clyde's sleep having culminated in some close world.

4.4.5 *The Non-Gradability Restriction*

At first gloss, denotations such as those proposed above may appear to account for the strangeness of gradable adjectives with *almost*. Given what has been suggested so far, AP *almost* requires that an adjective clearly not hold of an individual in the evaluation world and clearly hold in a maximally close world. To determine which worlds are maximally close, it is necessary to know, for every world, whether in that world the adjective holds of the individual. For a gradable adjective, such a determination would seem impossible. In the case of #*Clyde is almost tall*, for example, as we consider worlds in which Clyde's height gradually increases, we would need to know exactly when we have reached a world in which Clyde counts as tall. To do this, it would be necessary to identify with absolute precision the particular height an individual must attain to be tall and fail to attain to be not tall. Unmodified tall does not permit this.

But, in fact, typical assumptions about gradable adjectives make this explanation unavailable. A usual view is that gradable adjectives relate an individual and a degree

(Seuren 1973, Cresswell 1976, and many others since; discussion and further references in Kennedy 1997). In the absence of overt degree phrases, the degree argument is a standard value provided by context. Once the degree argument position has been saturated, gradable adjectives become indistinguishable from non-gradable adjectives. By the time *almost* encounters an AP denotation, then, the distinction between gradable and non-gradable adjectives has been rendered invisible.

In (40), for example, if the contextually-supplied standard degree of height is six feet, (40a) will be interpreted along the same lines as (40b):

(40) a. Clyde is tall.

b. Clyde is six feet tall.

With respect to *almost* modifiers, though, there is a stark difference:

(41) a. ?Clyde is {almost / nearly / practically} tall.

b. Clyde is {almost / nearly / practically} six feet tall.

If both of these sentences simply involve saturation of a degree argument, this contrast is unexpected. The *almost* modifier in both cases would apply to the property of individuals that are at least six feet tall.

It is possible to conceive of circumstances, however unusual, in which (41a) would be less strange, apparently by virtue of being coerced into an interpretation like (41b). An organization of tall people—say, one devoted to political advocacy for extended legroom in

cars—may require that its members be at least six feet tall. In discussing Clyde's potential membership, one might felicitously utter (41a). But it is exactly this precise identification of a standard degree of height that renders this context unusual. If such a precisely identified standard degree were provided by any context in which a gradable adjective is used without degree modification, it would not be necessary to consider unusual contexts such as this, nor indeed would such contexts seem especially unusual. It seems reasonable, then, to regard the view that all contexts provide precise standards according to which unmodified gradable adjectives may be interpreted as an idealization.

The *almost* facts run afoul of this idealization. Perhaps, then, what is required to explain the contrast in (41) is not a more refined semantics for *almost*, but rather an understanding of the semantics of standard degrees that might better permit distinguishing different means of saturating the degree argument position.[11]

4.5 Final Steps toward Cross-Categoriality: Another Kind of Decomposition

Denotations have now been proposed for *almost* in a number of categories, but these remain distinct denotations, and nothing has been proposed for other *almost* modifiers. This section sets out some syntactic assumptions motivated by the distribution of *almost* modifiers that make possible an account of the cross-categoriality of the class.

---

[11] It does not seem clear to me whether supervaluation approaches to adjective semantics (e.g. Kamp 1975) might be better equipped to avoid this sort of difficulty. Interpreting #*almost tall* in such a framework, given what has been said here, would not seem any more problematic than interpreting *almost dead*—both would simply involve identifying the closest worlds in which Clyde is in the positive extension of the adjective. Another approach entirely to the problem would be to relate the contrast in (41) to the possibility that the *almost* modifier in (48) modifies the measure phrase alone: *[almost six] feet tall*(Karina Wilkinson, p.c.). A difficulty with such an approach, though, is that one would not want to suppose that only measure phrases, and not adjectives themselves, may have *almost* modifiers—this would rule out *almost* dead. Moreover, if six is itself a non-gradable adjective as assumed here, it would not be possible to rule out *almost* modification of adjectives without also ruling out *almost* modification of numerals and hence of this measure phrase.

4.5.1 *Type Shifting?*

Before other possibilities are considered, though, it is sensible to first ask whether the necessary cross-categoriality can be achieved through type shifting (e.g. Partee 1987). This would seem a rather rocky road to travel, though.

One reason for skepticism is conceptual. The necessary type shift would be from modifiers of properties to modifiers of generalized quantifiers (or, alternatively, vice versa). There is no reason to doubt that this type shift should be possible in principle. But it would, at least, be quite unusual. It is not obvious that there is any independent evidence for such a type shift. So before embracing this option too enthusiastically, it would seem prudent to find some other area of the grammar where such a type shift might prove useful.

There is an empirical difficulty as well, which has already been noted in section 3.2 above. Variation in the interpretation of *almost* modifiers across categories is not in type alone. As suggested in 3.2, DP *almost* modifiers impose an additional requirement that property-modifying ones do not. So the requisite type shift would have to do more than shift types—it would have to impose or suspend this requirement as well. This would be an operation significantly more complex than a mere type shift.

4.5.2 *Distribution*

The chief syntactic property of *almost* modifiers, as noted above, is that for every category they modify, they occur in a high, left-peripheral position (repeating (3)):

(42) a. [{Almost / Nearly / Practically} every chiropractor] plays the ukulele.

b. Tilden [{almost / nearly / practically} won].

c. The soup is [{almost / nearly / practically} free of insects].

d. Floyd sailed [{almost / nearly / practically} around the world].

e. The candidate behaved [{almost / nearly / practically} inexcusably].

To find that modifiers with a particular kind of semantics occur in a particular range of syntactic positions is not surprising, of course.

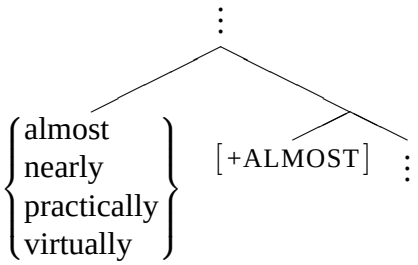
At least in this instance, it may be sensible to take this as a fundamentally syntactic fact in need of a syntactic solution. A semantic explanation of exactly this distribution would be nice, and maybe even more satisfying, but it is not obvious where it would come from. Certainly, there is nothing in the semantics proposed so far that would favor a distribution high in the extended projection, much less on the left periphery. The positions in which *almost* modifiers occur may plausibly be a syntactic natural class, but they do not appear to be a semantic one. It might in principle have been the case, for example, that all property-denoting expressions could license *almost* modification. But not so—nouns and NPs (in the Abney 1987 sense) clearly do not:

(43) a. *a(n) {almost / nearly / practically} gorilla

b. *a(n) {almost / nearly / practically} bottle

A syntactic explanation, however, is eminently available, in the language of modifier-licensing features Cinque (1999) provides.[12] To account for a wide variety of distributional facts involving a large number of adverb classes across a large number of languages, he proposes that different classes of adverbs occur as specifiers to distinct functional projections in fixed syntactic positions, where they check a licensing feature. Perhaps, then, *almost* modifiers are licensed in this general fashion, by checking a feature—here dubbed [+ALMOST]—in a fixed position in functional structure:

(44)



Whether a separate functional projection, as (44) reflects, is really required or this feature can be bundled with others in some other projection will not be explored further here, though I will assume the former here for simplicity.[13]

---

[12] And more particularly in the mediated approach toward modification introduced in chapter 1—but, at the risk of being irritatingly coy, I will observe the principle set out there that this approach should be motivated independently on each instance without relying on previous chapters, and therefore I build here only on the broadly Cinquean foundation rather than on my own.

[13] If evidence that might help decide this question turns out to be lacking, it may be preferable to assume the bundling approach rather than a distinct projection. The semantics that will be proposed momentarily to

The distribution of *almost* modifiers can therefore follow from the fixed position of their licensing feature in the functional structure of the categories in which it occurs. That *almost* modifiers are obligatorily phrase-peripheral would follow from supposing that this fixed position is high in the extended projection of each category. That *almost* modifiers obligatorily occupy leftward positions would follow from assuming, in the Cinque style, that they occupy specifier positions.

4.5.3 *A Division of Labor*

As noted in section 3.2.1 above, *almost* has been witness to a great deal of lexical decomposition. It will now be made the victim of some.

With the introduction of this licensing feature in functional structure, the semantics of *almost* modifiers can now be divided into two parts. Most of the work can be assigned to the licensing features themselves:

(45) $[\![\,[\text{+ALMOST}_{\text{DP}}]\,]\!] = \lambda Q\ \lambda R\ \lambda P\ \lambda w\,.\,\neg Q(P)(w) \wedge \exists w'[Q(P)(w') \wedge [\lambda x\,.\,P(x)(w)] = [\lambda x\ P(x)(w')] \wedge R(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge Q(P)(w'')] \rightarrow w'' =_w w']]$

(46) $[\![\,[\text{+ALMOST}_{\text{DP}}]\,]\!] = \lambda Q\ \lambda R\ \lambda P\ \lambda w\,.\,\neg Q(P)(w) \wedge \exists w'[Q(P)(w') \wedge [\lambda x\,.\,P(x)(w)] = [\lambda x\,.\,P(x)(w')] \wedge R(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge Q(P)(w'')] \rightarrow w'' =_w w']]$

(47) $[\![\,[\text{+ALMOST}_{\text{AP}}]\,]\!] = \lambda P\ \lambda R\ \lambda x\ \lambda w\,.\,\neg P(x)(w) \wedge \exists w'[P(x)(w') \wedge R(w)(w') \wedge \forall w''[[w'' \leq_w w' \wedge P(x)(w'')] \rightarrow w'' =_w w']]$

---

interpret the structure in (44) would extend to this alternative conception as well, provided that function composition is permitted below the X^0 level. On such a view, the licensing feature [+ALMOST] would be interpreted by function-composing with the other features it is bundled with. (Such a theory is developed in chapter 1.)

(48) $[\![\,[\text{+ALMOST}_{\text{VP}}]\,]\!] = \lambda P\ \lambda R\ \lambda e\ \lambda w\,.\,\neg P(e)(w) \wedge \exists w'[P(e)(w') \wedge R(w)(w') \wedge$
$\forall w''[[w'' \leq_w w' \wedge P(e)(w'')] \rightarrow w'' =_w w']]$

Here, licensing features take a modified expression and an *almost* modifier as arguments. All the *almost* modifiers themselves need to do is provide various values for the intensional closeness relation R. For example:

(49) $[\![\,\textit{almost}\,]\!] = \lambda w \lambda w'\,.\,\text{CLOSE}(w)(w')$

This allows *almost* modifiers themselves to have entirely uniform denotations across categories. On this approach, it is not necessary to assume that every *almost* modifier has distinct homophonous forms for each category it can modify. It is still necessary, though, to suppose that there are several incarnations of the [+ALMOST] licensing feature.

But this may not be so shocking. Certainly, there are many independently attested correlations in functional structure across categories. Degree-related projections, for example, occur both in AP and AdvP, and indeed it is not even entirely clear to what extent these categories are distinct. The relation between nominal and verbal projections, moreover, is famously close. So perhaps assuming different flavors of this feature, while less than maximally elegant, may be necessary for a principled reason—and considerably more elegant than reproducing this categorial multiplicity independently for every *almost* modifier.

This approach, then, captures the correlation between the position and interpretation of the *almost* class through a division of labor, wherein part of the semantic burden is borne by the *almost* modifier itself, and part by its structural position through licensing features in fixed positions in functional structure.

4.5.4 *How Do* Almost *Modifiers Vary?*

The major differences among *almost* modifiers can now be understood as variation in the intensional closeness relation provided:

(50) a. $[\![\textit{very nearly}]\!] = \lambda w\ \lambda w' .\ \text{VERY-NEAR}(w)(w')$

b. $[\![\textit{nearly}]\!] = \lambda w\ \lambda w' .\ \text{NEAR}(w)(w')$

c. $[\![\textit{nowhere near}]\!] = \lambda w\ \lambda w' .\ \text{FAR}(w)(w')$

Thus *very near* provides a relation of greater proximity between worlds than *nearly*, and *nowhere near* provides a relation of distance. Since these relations will ultimately be used to compare the evaluation world to the closest worlds in which a proposition holds, a sentence containing *nowhere near* will require the closest worlds in which the proposition holds to be far:

(51) a. Herbert is [nowhere near [$[+\text{ALMOST}_{AP}]$ dead]].

b. $[\![[+\text{ALMOST}_{AP}]]\!]([\![\textit{dead}]\!])([\![\textit{nowhere near}]\!])([\![\textit{Herbert}]\!]) =$

$\lambda w .\ \neg\text{dead}(\text{Herbert})(w) \wedge \exists w'[\text{dead}(\text{Herbert})(w') \wedge \text{FAR}(w)(w') \wedge$

$\forall w''[[w'' \leq_w w' \wedge \text{dead}(\text{Herbert})(w'')] \rightarrow w'' =_w w']]$

Because of the requirement that only maximally close worlds be considered, it is not necessary to independently encode into the semantics that, in addition to there being a distant world in which Herbert is dead, there must also not be any close ones in which he is.

Like a number of other *almost* modifiers, nowhere near appears to consist of two prosodic words, but like other members of the class, it is not internally compositional or syntactically complex. Even *very nearly*, which would otherwise seem likely to be compositional, is frozen—substituting other degree words for *very* is impossible: **rather nearly*

*every chiropractor; *extremely nearly every chiropractor*. Likewise, just about is an *almost* modifier, but not bare about: **(just) about every chiropractor; *(just) about around the world*.[14]

Apart from the intensional closeness relation provided, *almost* modifiers may also vary in register, as between the casual *damn near* and the formal *virtually*.[15]

4.6 Concluding Remarks

One quite intuitive notion that this sort of approach does not reflect, perhaps to its detriment, is that *almost* modifiers involve, in some sense that would have to be made precise, points on a scale (cf. Hitzeman 1992). Naturally, this is not in principle incompatible with the intensional approach taken here—one may imagine constructing scales by ranking worlds in a particular way. Indeed, the closeness relations among worlds may be a step in this direction. Another intriguing possibility is comparing not worlds, but focus alternatives (Maribel Romero, p.c.). These possibilities will have to be left unexplored here.

To summarize, the empirical aim of this paper has been to establish that *almost* modifiers constitute a natural class—they share a cross-categorial distribution, occur in the same position in each category, give rise to the same range of readings, and impose similar restrictions on the categories they modify. Because of this, the behavior of *almost* cannot be understood as a mere idiosyncrasy. Headway toward explaining the properties of

---

[14] An interesting case in this respect is *not quite* and *quite*. These both appear to be *almost* modifiers, but the latter appears to be a negative polarity item: *I *(don't) think quite everyone has left yet*. (The degree word *quite—quite tall, quite upset*—is another beast entirely, and is not a negative polarity item.)

[15] Sadock (1981) suggests another possible difference. *Not quite* and *almost*, he claims, differ in that not quite entails that the proposition it applies to is false, while *almost* merely gives rise to this inference as an implicature. Hitzeman (1992) argues against this.

almost modifiers may be made by combining a semantics rooted in intensional similarity with a model of modification in which semantic labor is divided between modifiers and the syntactic features in functional structure that license them.

CHAPTER 5

MAXIMALITY IN THE INTERPRETATION OF SINGULAR DPs: WHOLES HAVE COVERS THAT FIT

5.1 Introduction

Adjectives such as *whole* and *entire* lend themselves quite readily to an intuitive characterization: What they have in common is, roughly, a requirement that a predicate hold of all the parts of an individual. This is straightforward enough, but underlying this initially clear intuition are puzzles that relate ultimately to the semantics of part structure, distributivity, and quantification and scope mechanisms.

Perhaps perversely, in view of this natural intuition, the argument advanced here is that DPs containing *whole* and *entire*—henceforth '*whole* adjectives'—are obligatorily non-quantificational expressions. Their apparent quantificational force will be understood instead as a restriction on the tolerance normally extended to exceptions to a predication, building on Brisson (1998)'s theory of nonmaximal interpretations of plurals. To implement this, the notion of covers (Schwarzschild 1996) will be imported, with some substantial modifications, from the semantics of plurals into the semantics of singulars, and a Brisson-style theory of nonmaximal interpretations of singulars will be developed on that

---

This chapter is a more elaborated version of Morzycki (2002)—some additional empirical observations are considered, slightly more attention is devoted to the relation between these adjectives and the syntactic position they occupy, and the exposition is a bit more leisurely.

basis. Along the way, the analysis will lead to adopting a particular conception of choice-function indefinites (Winter 1997, Reinhart 1997, Kratzer 1998, Matthewson 1999).

Section 5.2 makes some observations about *whole* adjectives, characterizing their interpretive effect in general terms, noting restrictions these adjectives impose on determiners and the anaphoric and scope properties they give rise to. Section 5.3 briefly notes the more general puzzle of how part structure phenomena involving singular count nouns should be analyzed, in light of similarities and differences between the notion of part structure they require and that required by plural and mass nouns. Section 5.4 argues against analyzing *whole* adjectives as a species of universal quantifier, and then considers the highly intensional treatment proposed by Moltmann (1997). Section 5.5 advances an alternative analysis of *whole* adjectives as 'maximizing modifiers' in the spirit of Brisson, in the process extending the use of covers to the interpretation of singulars and finding reason to permit intermediate existential closure of choice functions (Winter 1997, Reinhart 1997). Section 5.6 concludes and presents some puzzles for further inquiry.

5.2 Characteristics of *Whole Adjectives*

5.2.1 *What Do They Do?*

The essential role of *whole* adjectives is to impose a requirement that no part of an individual be exempted from having to satisfy a predicate. The sentences in (1) differ in just this respect:

(1) a. The ferret is submerged.

b. The $\left\{\begin{matrix}\text{entire}\\\text{whole}\end{matrix}\right\}$ ferret is submerged.

For (1a) to be true, it is under most circumstances quite sufficient for a significant proportion of the ferret to be submerged. If some stray fur or a tail protrudes above the surface of the water, one is not inclined to take this as evidence that (1a) is false. On the other hand, (1b) makes a stronger claim, one of total ferret submersion. Any unsubmerged ferret-matter, even if only a tail, renders (1b) false. So, to put this in slightly less atheoretical terms, one may conceive of the contribution of a *whole* adjective here as universal quantification over parts of the ferret or, alternatively, as a requirement that the tolerance that would otherwise be extended for a few exceptional unsubmerged ferret parts be withheld.

These properties mirror those of *all*:

(2) All of the ferret is submerged.

This seems to mean something very much like what (1b) means. More will be said about this parallel shortly.

Despite their superficial resemblance, *whole* adjectives differ from adjectives like complete or partial:

(3) a. The $\left\{\begin{matrix}\text{entire}\\\text{whole}\end{matrix}\right\}$ ferret is submerged.

b. The $\left\{\begin{matrix}\text{complete}\\\text{partial}\end{matrix}\right\}$ ferret is submerged.

Unlike the *whole* adjectives in (3a), the *complete* adjectives in (3b) merely provide information about the structural integrity of the ferret. They are intersective, so (3b) could be paraphrased as in (4):

(4) The ferret is submerged, and the ferret is $\left\{\begin{matrix}\text{complete}\\\text{partial.}\end{matrix}\right\}$

Thus (3b) and (4) are true if the ferret is submerged, more or less, and if it is a complete ferret, with all its ferret parts in the appropriate ferret configuration, or if it is a partial ferret, with not all of its ferret parts in the appropriate ferret configuration, respectively.[1] Somewhat confusingly, *whole* actually has a reading as an adjective of this class, too. This is the only reading available for *whole* in predicative position:

(5) The ferret is whole.

Entire, which is exclusively a *whole* adjective and lacks a *complete* adjective reading, cannot very naturally occur predicatively at all:

(6) *The ferret is entire.

Whole adjectives, then, unlike *complete* adjectives, do not appear to have property denotations.

There are nominal and adverbial relatives of *whole* adjectives:

(7) a. The $\left\{\begin{matrix}\text{whole}\\\text{entirety}\end{matrix}\right\}$ of the ferret is submerged.

b. The ferret is $\left\{\begin{matrix}\text{wholly}\\\text{entirely}\end{matrix}\right\}$ submerged.

Though these are of course not irrelevant to the topic, they will for the most part be set aside here.

---

[1] That it is in fact necessary for a ferret to be complete that all the ferret parts not only be present but also configured appropriately can be verified by reflecting on whether a bag containing all the parts necessary for the assembly of a ferret can be described as containing a complete ferret. It is, in my judgment, at best a ferret starter kit.

5.2.2 *Determiner Restrictions*

Whole adjectives may occur only with a restricted set of determiners. Definites, demonstratives, and genitives are fine:

(8) a. The {whole / entire} ferret is submerged.

b. {This / That} {whole / entire} ferret is submerged.

c. {My / Floyd's} {whole / entire} ferret is submerged.

So are DPs with indefinites headed by *a*, numerals, *many*, and less clearly *some*:

(9) a. A {whole / entire} ferret is submerged.

b. Twenty {whole / entire} ferrets are submerged.

c. Many {whole / entire} ferrets are submerged.

d. Some {whole / entire} ferrets are submerged.

Strong, inherently quantificational determiners are systematically incompatible with *whole* adjectives. In these contexts, *entire* results in ungrammaticality, and *whole* receives only the complete adjective reading, which will be set aside here:

(10) a. *{Every / Each} {whole / entire} ferret is submerged.

b. *All the {whole / entire} ferrets are submerged.

c. *Most $\left\{\begin{matrix}\text{whole}\\\text{entire}\end{matrix}\right\}$ ferrets are submerged.

This is similar to the pattern of determiner restrictions Abusch and Rooth (1997) note for modifiers such as *unknown* or *undisclosed*, which will be revisited in section 5.5.6.[2]

5.2.3 *Scope*

Whole adjectives appear to give rise to scope effects with respect to negation that would otherwise be absent:

(11) a. The whole ferret wasn't submerged. (*the whole f.* $\prec \neg$; $\neg \prec$ *the whole f.*)

b. Clyde didn't eat an entire sandwich. ($\neg \prec$ *an entire s.*; *an entire s.* $\prec \neg$)

c. No one likes that whole subject. ($\neg \prec$ *that whole s.*; *that whole s.*$\prec \neg$)

In (11a), there seems to be a wide-scope reading of *the whole ferret* in which all of the ferret is claimed to have been unsubmerged, and a narrow-scope reading in which it is denied that all of the ferret was submerged. Since (11a) may in fact be judged true in a context in which parts of the ferret were submerged, *the whole ferret* appears to scope above negation. Similarly, (11b) can be used to deny that Clyde consumed all of a sandwich or to assert that all of a sandwich was unconsumed. And (11c) can be used to deny that anyone likes all of that subject, or to assert that no one likes any of that subject. These effects are a bit striking, since definites and demonstratives do not normally participate in scope interactions.

---

[2] Though *whole* adjectives can occur with a definite determiner and they can occur with a plural noun, they cannot occur with both: **The entire ferrets were submerged*. Also perhaps not as good as maybe it should be is *?*Some entire ferrets were submerged*.

In this respect also, *whole* adjectives differ from *complete* adjectives, which do not give rise to scope ambiguities:

(12) a. The complete ferret wasn't submerged. (unambiguous)

b. Clyde didn't eat a partial sandwich. (unambiguous)

This, of course, is less surprising, since *complete* adjectives have simple property denotations.

Oddly, *whole* adjectives do not give rise to scope ambiguities with respect to indefinites:

(13) a. The whole ferret was submerged by a child. (unambiguous)

b. Someone ate this entire sandwich. (unambiguous)

c. The entire committee complained about an especially misguided proposal. (unambiguous)

It is very difficult to judge (13a) true in a situation in which no single child is guilty of ferret submersion, but rather different children submerged different parts of the ferret. Likewise, (13b) cannot easily involve an act of cooperative sandwich consumption involving several people. Even a noun that provides a relatively clear and individuated part structure like *committee* in (13c) does not quite mitigate this effect—(13c) would not normally be taken to describe a situation in which there is no single proposal drawing universal criticism. With respect to existential quantifiers, then, *whole* adjectives do not seem to induce additional scope possibilities. (Scope effects with respect to universals are unexpected a priori because *whole* adjectives are themselves either universals or else something rather like them, as the facts in (11) seem to suggest.)

This odd unwillingness to scope like normal, well-behaved universal quantifiers with respect to existentials is further reflected in the inability of *whole* adjectives to license the particular reading of *different* that arises only in the scope of universals (Beck 2000):

(14) a. Each part of the ferret was submerged by a different child.
b. The whole ferret was submerged by a different child.

(15) a. Every member of the committee complained about a different proposal.
b. The entire committee complained about a different proposal.

As (14a) and (15a) demonstrate, in the scope of a universal *different* may receive a reading in which it requires (roughly) that the individuals quantified over by the existential vary with those quantified over by the universal. But *whole* adjectives do not give rise to this reading of *different*—(15b), for example, does not have a reading in which children must vary with ferret parts.

5.2.4 *Discourse Anaphora*

Among the characteristic properties of universals is that they do not introduce discourse referents, and indeed make any introduced inside their scope inaccessible beyond it (Heim 1982, Kamp 1981, others). *Whole* adjectives do not have these properties.

True, *whole* adjectives, like their quantificational-determiner paraphrases, do not introduce portions of individuals as discourse referents:

(16) a. Every member of the committee complained. *He was often cranky before lunch.
b. The whole committee complained. *He was often cranky before lunch.

But unlike their quantificational determiner counterparts, *whole* adjectives do not impede the DP in which they occur from itself introducing a discourse referent:

(17) a. Every ferret was submerged. *It didn't seem pleased.
b. A entire ferret was submerged. It didn't seem pleased.

(18) a. Cockroaches ate each sandwich. *It had been sitting out too long.
b. Cockroaches ate a whole sandwich. It had been sitting out too long.

Although the universally-quantified every ferret and each sandwich do not introduce discourse referents, DPs with *whole* adjectives do.

Whole adjectives do not prevent discourse referents from being introduced inside their scope, either, again contrasting with well-behaved universals:

(19) a. Every building with a faulty $roof_i$ was demolished. *It_i had been leaking.
b. An entire building with a faulty $roof_i$ was demolished. It_i had been leaking.

(20) a. Each recent movie was marred by an irritating $soundtrack_i$. *It_i will probably be released on CD.
b. The whole movie was marred by an irritating $soundtrack_i$. It_i will probably be released on CD.

So, *every* prevents *a faulty roof* from licensing discourse anaphora in (19a); in (19b),*whole* has no such effect. Likewise, in (20a), *each* prevents *an irritating soundtrack* from licensing discourse anaphora; in (20b), *whole* has no such effect.

5.2.5 *Summary*

Whole adjectives require that a predicate hold of all parts of an individual, but are distinct from adjectives such as complete in that they do not provide information about the structural integrity of an individual. They cannot grammatically occur in DPs with inherently quantificational determiners. Although *whole* adjectives seem to give rise to scope ambiguities with respect to negation, they do not do so with respect to existential quantifiers. Nor do they impede discourse anaphora, as universal quantifiers do.

5.3 An Interlude: Singulars, Parts, and Parts of Singulars

5.3.1 *Treacherous Terrain*

Any understanding of the semantics of *whole* adjectives will likely require some notion of part structure for singular individuals. It is hard to imagine expressing the contribution of such adjectives, even in the most pretheoretical impressionistic terms, without appealing to parts. Part structure for singulars is, however, treacherous terrain, and, to the extent possible, this minefield will be circumnavigated here. But since this much more general issue is fundamental to the issues more immediately at hand, the part relation necessary for count singulars will be briefly considered here, then the issue will be put to rest, and the discussion of *whole* adjectives themselves will resume in section 5.4.

A complete discussion of part structure that encompasses singulars is undertaken in Moltmann (1997), who brings to light a paradox concerning the notion of 'part of' in natural language.[3] On the one hand, it seems that the semantics of plural and mass nouns

---

[3] Moltmann does not actually call this a paradox, and perhaps would not regard this situation as especially paradoxical. But it is at least a puzzle. More generally, Moltmann does not formulate the problem in precisely this way.

should be modeled using part structures with certain formal characteristics unsuitable to modeling the semantics of singular count nouns. On the other hand, there are parallels between singular count nouns and plural and mass nouns that strongly suggest that a unified treatment is required.

5.3.2 *Well-Behaved Part Structures: Plurals and Mass Nouns*

There are several formal characteristics associated with both plural and mass noun denotations: cumulativity, divisibility/atomicity, transitivity, and extensionality. Link (1983) illustrates cumulativity—which he calls the 'cumulative reference property', following Quine (1960) and Bunt (1979)—by example:

(21) a. If a is water and b is water then the sum of a and b is water.
b. If the animals in this camp are horses, and the animals in that camp are horses, then the animals in both camps are horses.

This demonstrates, then, that both mass nouns like water and plurals like horses denote predicates that hold of any sum of individuals in their extension. Divisibility/atomicity involves inference from individuals to their parts:

(22) a. If the sum of a and b is water, then a is water.
b. If the animals in this camp and the animals in that camp are horses, then the animals in this camp are horses.

What these examples suggest is that mass nouns like water and plurals like horses both denote predicates that hold of parts of individuals in their extensions.[4] The part structure needed for the semantics of plural and mass nouns must further encode transitivity of the part relation. As before, this may be illustrated by example:

(23) a. If the water in this teaspoon is part of the water in this bowl, and the water in this bowl is part of the water on the table, then the water in this teaspoon is part of the water on the table.

b. If the horses in this camp are part of the horses in the state, and the horses in the state are part of the horses in the country, then the horses in this camp are part of the horses in the country.

Finally, the part relation needed for plural and mass nouns is extensional, in the sense that any two individuals with all the same parts are identical:

(24) a. If the water in this glass is made up of these 50 teaspoons, and the water in front of Clyde is made up of these same 50 teaspoons, the water in this glass is the same as the water in front of Clyde.

b. If the horses in this state include only the horses in this camp, and the horses Clyde owns include only the horses in this camp, then the horses in this state are the horses Clyde owns.

---

[4] Strictly speaking, divisibility and atomicity, though similar, must be distinguished. Certainly, not just any part of an individual in the extension of horses is itself horses. Hooves, for example, are not horses, despite being parts of a horse. Nevertheless, all the atomic individuals (in Link's sense, say) of which the extension of horses is made up are in fact horses. So, more precisely, plural denotations manifest atomicity and mass noun denotations may plausibly manifest divisibility.

Thus plural and mass nouns support a view of part structure that reflects transitivity, extensionality, cumulativity, and divisibility/atomicity. Together, these formal characteristics provide a fairly clear and precise view of how part structure should work.

5.3.3 *Ill-Behaved Part Structures: Count Singulars*

But things are not so tidy. Count singulars have denotations that require a different notion of part structure that lacks all the formal properties noted above. They are not cumulative, so the inference in (25), parallel to (21) above, is not valid:

(25) If Ted is a senator and Strom is a senator, then Ted and Strom are a senator.

Of course, one can infer that the sum of Ted and Strom is in the extension of senators, since plural nouns are cumulative. But no such inference can be made for senator. Nor are count singular denotations divisive. The inference in (26), parallel to (22) above, is not valid:

(26) If Ted, Strom, and Jesse are a subcommittee, then Ted is a subcommittee.

Nor are they necessarily compatible with a transitive part-of relation:

(27) If Ted's nose is part of Ted, and Ted is part of a subcommittee, then Ted's nose is part of a subcommittee.

This, of course, is not a valid inference. Nor are they necessarily compatible with an extensional part-of relation:

(28) If the ethics subcommittee includes only Ted, Strom, and Jesse, and the bribe solicitation subcommittee includes only Ted, Strom, and Jesse, then the ethics subcommittee and the bribe solicitation subcommittee are the same subcommittee.

Again, this is not a valid inference. Singular count nouns, then, lack all these distinguishing characteristics of part structure in plural and mass nouns.

5.3.4 *Not So Different After All*

Despite the stark formal difference between singular count nouns and plural and mass nouns, there is evidence that a unified treatment of all three is required.

Perhaps the clearest parallel among them is in partitives:

(29) a. $\left\{\begin{matrix}\text{All}\\\text{Much}\\\text{Some}\end{matrix}\right\}$ of the water spilled onto the floor.

b. $\left\{\begin{matrix}\text{All}\\\text{Much}\\\text{Some}\end{matrix}\right\}$ of the horses ran into the field.

c. $\left\{\begin{matrix}\text{All}\\\text{Much}\\\text{Some}\end{matrix}\right\}$ of the ferret was submerged.

This is more than merely a syntactic parallel. In all three cases, the same sort of quantification seems to be taking place: over parts of the referent of the embedded DP.

Moltmann observes another parallel involving the possibility of collective predication: (her examples)

(30) a. The men gathered.

b. The wood is heavy.

c. The sum of these numbers is greater than two.

d. The weight of the wood is ten pounds.

In all of these cases, predication is collective in that the VP denotation is not necessarily true of proper parts of the subject denotation. Thus no individual man needs to have, impossibly, gathered by himself in order to render (30a) true. Likewise, (30b) remains true if some particle of the wood is not on its own heavy. And analogously, (30c) does not require that both the numbers being added be greater than two, and (30d) does not require that every particle of the wood weigh ten pounds.

5.3.5 *A Simplifying Assumption*

Moltmann develops a theory of part structure consistent with these facts by relativizing every predicate in the grammar to a reference situation which determines part structure, and by further assuming that every argument position of every predicate is filled with a pair one of whose members is a reference situation.

For current purposes, though, it will be sufficient to suppose that there is some part relation for count singulars (s (s for ‘singular’) which is consistent with the observations above. This is a simplifying assumption, essentially a placeholder for an articulated theory of singular part structure. Although it seems likely that some of the additional theoretical machinery appealed to momentarily—namely, covers (Schwarzschild 1996)—may be able to replicate some of the context-sensitivity Moltmann achieves with reference situations, nothing here should be construed as a proposal about the character of the singular part-structure relation itself.

5.4 Universal Part Quantification and Moltmann's Intensional Alternative

This section considers two approaches to *whole* adjectives that are fundamentally different from the cover-theoretic approach that will ultimately be pursued here. The first of these, the universal part quantification approach, is in a sense logically prior; the other, that of Moltmann (1997), is temporally prior.

5.4.1 *Universal Part Quantification: The Idea*

The initial intuition about *whole* adjectives we began with was that they impose a requirement that a predicate hold of all the parts of an entity without exception. One can express this intuition quite simply and directly by analyzing *whole* adjectives as universal part quantifiers. Thus a sentence such as (1) could be interpreted as in (31):

(31) $\left[\!\left[\textit{the} \left\{ \begin{matrix} \textit{whole} \\ \textit{entire} \end{matrix} \right\} \textit{ferret is submerged} \right]\!\right] =$ (not final)

$\forall x[x \leq_{s\iota} y\ [\text{ferret}(y)] \rightarrow \text{submerged}(x)]$

This reflects the meaning paraphrased by 'every part of the ferret was submerged'. For all its simplicity and naturalness, this approach runs afoul of many of the observations in section 5.2. On its own it does not explain the determiner restrictions *whole* adjectives impose, and, as demonstrated below, it makes wrong predictions with respect to scope and discourse anaphora.

5.4.2 *Scope Problems*

Perhaps the chief prediction of an account of*whole* as a universal part quantifier is that it should give rise to scope ambiguities. As noted in section 5.2, this may be borne

out with respect to negation, but not with respect to existentials. So an account such as that reflected in (32) would account for the ambiguity of (11a), repeated here:

(32) The whole ferret wasn't submerged.

But such an account would also predict that the unambiguous (13a), repeated here, should be scopally ambiguous:

(33) The whole ferret was submerged by a child.

Similarly unwelcome predictions would be made with respect to licensing *different* as in (14–15).

5.4.3 *Discourse Anaphora Problems*

Another unwelcome consequence of a view of *whole* adjectives as universal part quantifiers is the prediction that they should render discourse referents introduced in their scope inaccessible to discourse anaphora. Again, as noted in section 5.2, *whole* adjectives do not have this effect.

In particular, they do not impede the DP in which they occur from introducing a discourse referent, as universally quantifying determiners do and as an account along the lines of (31) would predict. This is reflected in (17), repeated here:

(34) a. Every ferret was submerged. *It didn't seem pleased.

b. A entire ferret was submerged. It didn't seem pleased.

Nor do they impede discourse referents from being introduced inside their scope, as in (19), repeated here:

(35) a. Every building with a faulty roof$_i$ was demolished. *It$_i$ had been leaking.

b. An entire building with a faulty roof$_i$ was demolished. It$_i$ had been leaking.

5.4.4 *A Conjunctive Revision to the Universal Part Quantification Approach*

As expressed in (31), then, the universal part quantification analysis does not seem to be getting much traction on the problem. A slight revision may help. Instead of taking *whole* adjectives to be merely universal part quantifiers, one might take them instead to have a fundamentally conjunctive meaning, contributing universal part quantification as a kind of addition to the semantics of the sentence. Thus, (1) might be interpreted as in (36):

(36) $[\![\textit{The} \left\{ \begin{matrix} \textit{whole} \\ \textit{entire} \end{matrix} \right\} \textit{ferret is submerged}]\!] =$ (not final)

$\text{submerged}(\iota y[\text{ferret}(y)]) \wedge \forall x[x \leq_s \iota y[\text{ferret}(y)] \rightarrow \text{submerged}(x)]$

This reflects a paraphrase such as ‘the whole ferret was submerged, and every part of the ferret was submerged (as well)’.

This revision immediately avoids the discourse anaphora trouble encountered earlier, because one of the conjuncts is interpreted as though *whole* were absent and therefore unable to block discourse anaphora, and discourse referents introduced in either conjunct of a coordinate structure are accessible beyond it (*The ferret$_i$ and Greta were submerged. It$_i$ smelled worse wet than she did.*).

This may initially appear to avoid the scope difficulties too, since any existentials that might occur in the first conjunct would not have any universal introduced by *whole* to interact with. But scope problems with indefinites would apparently remain, as (37) reflects:

(37) A painter moistened the whole canvas.

In order to confront the problem here, it is necessary to consider more seriously how denotations such as the one in (36) are to be assembled compositionally. First, it will be necessary to commit to a denotation for *the whole canvas*:

(38) $[\![\textit{the whole canvas}]\!] = \lambda P \,.\, P(\iota y[\text{canvas}(y)]) \wedge \forall z[z \leq_s \iota y[\text{canvas}(y)] \rightarrow P(z)]$ (not final)

This encodes the essentially conjunctive meaning, by requiring that *the whole canvas* apply to a property that holds of the canvas and every part of it. Crucially, this is a generalized quantifier denotation. Thus to be interpreted, the whole canvas will have to QR from object position in (37), giving rise to LFs in which it takes scope somewhere below a painter, as in (39), or above it, as in (40):[5]

(39) *a painter$_1$ [the whole canvas$_2$ [t_1 moistened t_2]]*

a. $[\![{}_2\ t_1 \textit{ moistened } t_2]\!] = \lambda x_2 \,.\, \text{moisten}(x_2)(x_1)$

b. $[\![\textit{the whole canvas}_2\ t_1 \textit{ moistened } t_2]\!] = \text{moisten}(\iota y[\text{canvas}(y)])(x_1) \wedge \forall z[z \leq_s \iota y[\text{ferret}(y)] \rightarrow \text{moisten}(z)(x_1)]$

c. $[\![{}_1\ \textit{the whole canvas}_2\ t_1 \textit{ moistened } t_2]\!] = \lambda x_1 \,.\, \text{moisten}(\iota y[\text{canvas}(y)])(x_1) \wedge \forall z[z \leq_s \iota y[\text{ferret}(y)] \rightarrow \text{moisten}(z)(x_1)]$

d. $[\![\textit{a painter}_1\ \textit{the whole canvas}_2\ t_1 \textit{ moistened } t_2]\!] = \exists x[\text{painter}(x) \wedge \text{moisten}(\iota y[\text{canvas}(y)])(x) \wedge \forall z[z \leq_s \iota y[\text{ferret}(y)] \rightarrow \text{moisten}(z)(x)]]$

---

[5] For ease of exposition, I do not consider the possibility that the scope configuration in (39) may be attained by QRing *the whole canvas* to a position below the lowest subject position without movement of the subject (say, VoiceP or *v*P), since this would require making additional commitments about how subjects may be introduced into semantic composition (cf. Kratzer 1996). This will not affect the argument made here.

(40) *the whole canvas$_2$ [a painter moistened t_2]*

a. $[\![{}_2$ *a painter moistened* $t_2]\!] = \lambda x_2 . \exists x[\text{painter}(x) \wedge \text{moisten}(x_2)(x)]$

b. $[\![$ *the whole canvas$_2$ a painter moistened* $t_2]\!] = \exists x[\text{painter}(x) \wedge$
$\text{moisten}(\iota y[\text{canvas}(y)])(x)] \wedge \forall z[z \leq_s \iota y[\text{ferret}(y)] \rightarrow \exists x[\text{painter}(x) \wedge$
$\text{moisten}(z)(x)]]$

The denotation in (39), in which *the whole canvas* has narrow scope with respect to the painter, does not present any obvious problems—it requires that there be some painter who moistened the canvas and who moistened every part of it. But the denotation in (40) is too weak—it requires only that there be a painter who moistened the canvas, and that each part of the canvas had been painter-moistened. On the reading in (40), then, this sentence would be true if a team of painters collectively moistened every part of the canvas, but no single painter moistened every part of it. Certainly, this is not a possible reading of (37).

It is not clear how this undesirable result could be avoided. So long as a *whole* adjective is attributed any quantificational force of its own with scope over the matrix predicate, even in this conjunctive way, the DP in which it occurs will have to denote a generalized quantifier and consequently be susceptible to QR, and to the possibility of taking wide scope over an indefinite as in (40). The option of simply stipulating that DPs with *whole* adjectives do not QR is unavailable, since generalized quantifiers often must QR to be interpreted. One could, however, stipulate that a *whole* DP can QR only to positions in which its scope with respect to other quantifiers is unaltered. But this does not seem like an especially promising path to pursue, beginning as it does with the bizarre stipulation that QR, a grammatical device which exists chiefly to account for variation in scope possibilities, should for a particular class of expressions be both obligatory and obligatorily

restricted to movements that do not give rise to new scope possibilities.[6]

5.4.5 *Moltmann's Intensional Approach*

Moltmann (1997) proposes an analysis of whole along different lines, cast in her general theory of part structure. This general theory makes available a rich set of tools for manipulating what sort of part structure an individual is construed as having, subject to restrictions various linguistic expressions and the context can impose. She proposes that the grammar is intricately sensitive to reference situations, or situations that determine the part structure an individual is construed as having. These are pervasive in the grammar; every predicate and every argument position is relativized to a reference situation.

Whole, on this view, restricts an individual's perceived part structure—it triggers an interpretation in which an individual is not construed as what she calls an 'integrated whole'

---

[6] Relegated to this footnote is an apparent but not actual counterexample to taking *whole* adjectives to be universal quantifiers. In (i), one might claim, there is a *whole* adjective but the quantification is not quite universal:

(i) a. The whole town was pregnant.
b. I cleaned the whole house.

In (ia), for example, it need not be the case that the men be pregnant, much less that physical parts of the town such as trees and buildings be pregnant. In (ib), under normal circumstances I need not have cleaned the plumbing system of the house. These, however, appear to be ordinary domain of quantification effects, if in an unfamiliar place. Corresponding to (i) are the sentences in (ii), which involve uncontroversially universal quantifiers:

(ii) a. Everyone in the town was pregnant.
b. I cleaned everything in the house.

In both of these cases, quantification is understood to be restricted in roughly the way it is in (i). Both of these examples also manifest tell-tale sensitivity to pragmatic considerations:

(iii) a. The whole town was wiped out.
b. A hurricane destroyed the whole house.

In these cases, the universal quantification could be understood as extending over a far broader domain, one including men and physical structures in (iiia) and plumbing in (iiib). As will be seen in section 5.5, though, none of this requires assuming that the quantifier being restricted is introduced by the *whole* adjective.

(one in which 'the connections among parts do not hold that would define it as an integrated whole' (p. 81)). More precisely, she suggests whole applies to a pair of a reference situation and a situated generalized quantifier—a generalized quantifier interpreted with respect to a reference situation—and yields a situated generalized quantifier:

(41) Let Q be a situated generalized quantifier and s' a situation: (Moltmann 1997, p. 81)

$[\![\textit{whole}]\!](Q,s') =$

$\lambda s\ \lambda P\ [Q^{s''}(P^s)=1(\forall x)[P^s(x,s')=1 \rightarrow$

$\text{BASED-ON}(\langle P^s, \langle x,s'' \rangle \rangle, <\langle \neg \text{INT-WH})^{s''}, \langle x,s'' \rangle \rangle) \wedge \forall x'[x'<x \rightarrow x'<_s x]]]$,

where s'' is the situation that differs minimally from s' in that for all $x \in D(s'')$, $\neg \text{INT-WH}(x,s'')$

Superscripts indicate reference situations with respect to which predicates are evaluated; (INT-WH(x,s) = 1 iff x is an integrated whole in s; BASED-ON($<P^s,<x,s'>>,<Q^{s"},<y,s'''>>$) = 1 iff $Ps(<x,s'>)$ 'forms a basis of' $Q^{s"}(<y,s'''>)$.

Paraphrasing in English what (41) is intended to express is not entirely straightforward, partly because it fits into a larger theory with a lot of moving parts. But its effect is, roughly, that *whole* applied to the ferret with respect to a reference situation s' will yield a situated generalized quantifier true of is submerged with respect to a reference situation s if the non-integrated-whole version of the ferret in the reference situation s'—that is, the ferret in a situation that differs minimally from s' in that there are no integrated wholes in it—is submerged in s, and the non-integrated-whole version of the ferret in s' is submerged in s on the basis of this non-integrated-whole version of the ferret not being an

integrated whole, and the submersion-situation s contains all the 'actual parts' of the ferret. The motivation for all the components of this denotation will not be recapitulated here, though it is worth noting that the 'based-on' requirement is one that Moltmann suggests figures in the semantics not only of *whole* but also of other 'part-structure-sensitive perspective shifters' such as *together* and *as a whole*. The most important aspect of (41) for current purposes is that it gives rise to an interpretation in which an individual is construed as having a part structure that licenses distributivity—essentially, it is construed as an amalgamation of parts, so that anything predicated of the whole will distribute to the parts.

Although Moltmann does not make this explicit, this appears to make the right predictions with respect to the scope and discourse anaphora facts. The crucial feature of (41) that makes this possible is the absence of any universal part quantifier with free variables in its scope. Because of this, no discourse referents could be introduced inside the scope of a universal part quantifier contributed by *whole*, where they would wrongly be predicted to be inaccessible. Similarly, no scope ambiguities are expected with respect to other quantifiers, because they could not occur inside the scope of any universal part quantifier. No scope ambiguities are predicted with respect to negation, either, but though this appears problematic in light of the observations in section 5.2, another independent explanation of these effects is in principle available, as section 5.5.13 below will suggest.

One immediately striking feature of the denotation in (41) is that, without further assumptions, it is not compositional. *Whole* applies directly to a DP denotation, though the overt position of *whole* adjectives in English is lower, below the D level. To ensure a smooth mapping from syntax to semantics, it would be necessary to assume movement of the *whole* adjective to a position above DP (an adjoined position, say). This is not an unreasonable assumption to have to make, but it is an additional commitment. Moreover, without further elaboration, the ban on inherently quantificational determiners *whole* adjectives impose remains mysterious on this approach.

Neither of these problems seem intractable. More generally, this denotation for *whole* is a small component of an approach that provides an explicit and integrated theory of part structure, and thus is considerably more ambitious in its aims. However, the intensional machinery necessary to execute this program is powerful, and its consequences for the grammar pervasive. In light of that, it might be desirable to consider alternatives in which the lexical semantics of *whole* need not be crucially intensional, and the necessary theoretical commitments need not be so many and so profound.

5.4.6 *Summary*

Treating *whole* adjectives as universal part quantifiers makes undesirable predications regarding scope and discourse anaphora, and does not obviously lead to an account of the determiner restrictions *whole* adjectives impose. A highly intensional approach like Moltmann's fares considerably better, but requires making a number of profound additional commitments.

5.5 *Whole* Adjectives as Maximizing Modifiers

5.5.1 *Nonmaximal Interpretations: Brisson (1998)*

An alternative intuition about *whole* adjectives is available, which has already been expressed in another form with respect to plurals introduced by all. Rather than supposing *whole* adjectives universally quantify over parts or trigger reconceptualization of part structure, one can instead suppose they restrict exception tolerance. Brisson (1998) develops an analysis rooted in this intuition of plural all which may provide the basis for an understanding of *whole* adjectives.

On this view, all is a ‘maximizing modifier’ rather than a universal quantifier, in that it eliminates the ability of plurals to tolerate nonmaximal interpretations. Under normal circumstances, this tolerance is readily extended:

(42) The boys are building a raft.

One would naturally judge (42) true if only 98 of 100 boys are building a raft—this is a ‘nonmaximal’ interpretation. But in (43), the presence of *all* eliminates this possibility:

(43) All the boys are building a raft.

For (43) to be true, every boy must be building a raft. No exceptions are tolerated.

5.5.2 *Covers*

To capture this effect, Brisson develops a model of nonmaximality in terms of a notion of a ‘cover’ of the domain of discourse (Schwarzschild 1996). A cover is essentially a means of structuring the domain, so that various individuals are grouped together in a way consistent with the context. This is useful in accounting for interpretations of predicates in which they distribute only to certain pragmatically-determined non-atomic parts. On such a reading, (44)—due to Gillon (1987) and cited by both Schwarzschild and Brisson—is true of the Broadway musical writers Rodgers, Hammerstein, and Hart:

(44) The men wrote musicals.

Rodgers and Hammerstein wrote musicals together, as have Rodgers and Hart. None of them, however, seems to have written a musical alone, so the reading in question is not one

that distributes to singular individuals.[7] Nor is it the case that all three of the them wrote a musical together, so the reading in question is not a collective one, either, at least in the most prototypical sense of the term. Covers provide a means of representing such readings. On this reading, (44) is interpreted with respect to a cover which groups together Rodgers and Hammerstein, as well as Rodgers and Hart, but no other combinations of the three.

To put it more precisely, a cover groups the individuals in a domain into cells. Thus, in the cover required to render (44) true, there is a cell containing only Rodgers and Hammerstein, and one containing only Rodgers and Hart, but none containing, for example, only Rodgers or only Hammerstein and Hart. The true reading of (44) now can be understood to be a distributive one, but not in the conventional sense that involves distributivity to singular individuals; rather, distributivity here is to cells of the cover.

Schwarzschild expresses this in terms of a theory in which it is of central importance that plurals simply denote sets, rather than plural individuals in the sense of Link (1983). The notion of a cover, though, does not seem inextricably bound up in this decision, and it will be useful here to attempt separating them. Putting things in more Linkian terms, or at least in terms more compatible with a Link-style approach, one might suppose that each 'cell' of the cover is an individual, frequently a plural individual. A cover, then, is a set of individuals, singular or plural, which has the same supremum as the domain:

(45) Cov is a cover of D iff *sup*(Cov)=*sup*(D)

That is, that the smallest plural individual composed of all the cells of the Cov must also be the smallest plural individual composed of all the members of D. Consequently Cov

---

[7] I will avoid the term 'atomic individuals' here. Since *whole* adjectives will require reference to parts of singular individuals, they cannot be atomic at least with respect to that part relation. They are, of course, atomic with respect to the plural part relation.

merely imposes a structure on D, adding nothing and leaving nothing out. Putting this to use, a cover such as the one in (46) will be required to render (44) true given the domain D indicated:[8]

(46) {Rodgers, Hammerstein, Hart, Greta, Hilda}⊆D, and D contains nothing else apart from all parts and all sums of its members
Cov = {Rodgers+Hammerstein, Rodgers+Hart, Greta, Hilda}

Pursuant to the discussion in section 5.3, the nature of the part relation that undergirds (45–46) will be left unexplored, though for the moment it will do no harm to suppose that it is Link's individual part relation.

To derive the desired reading for (44), these cells in the cover will have to be distributed over. Thus Schwarzschild proposes a VP distributivity operator Part, which is restricted by a cover variable Cov. Reformulating slightly to accord with the view of plurals assumed here:[9]

(47) $[\![\mathit{Part}_{\mathit{Cov}}]\!]^g = \lambda P\ \lambda x\ \forall y[[y \in g(\mathit{Cov}) \wedge y \leq x] \rightarrow P(y)]$ (not final)

What this requires is that a predicate to which this distributivity operator applies be satisfied by every cell in the cover which is a part of the individual distributed over. The relevant reading of (44), then, could be derived as in (48):

---

[8] That is, D is closed under the part relation and sum formation ($\forall x \forall y[[x \in D \wedge y \leq x] \rightarrow y \in D]$ and $\forall x \forall y[[x \in D \wedge y \in D] \rightarrow x+y \in D]$).

[9] Another departure made here from the original conception of this operator is that it is being represented here in the object language and interpreted directly.

(48) {Rodgers, Hammerstein, Hart, Greta, Hilda}⊆D, and D contains nothing else apart from all parts and all sums of its members

Cov = {Rodgers+Hammerstein, Rodgers+Hart, Greta, Hilda}

The men wrote musicals.

$[\![sgPart_{Cov}]\!]([\![\textit{wrote musicals}]\!])([\![\textit{the men}]\!])$

$= [\![sgPart_{Cov}]\!]$(wrote-musicals)(Rodgers+Hammerstein+Hart)

$= \forall y[[y \in Cov \wedge y \leq$ Rodgers+Hammerstein+Hart$] \rightarrow$ wrote-musicals(y)]

Given this cover, it will in fact be the case that every cell that is a part of $[\![\textit{the men}]\!]$ satisfies $[\![\textit{wrote musicals}]\!]$.

5.5.3 *Nonmaximality and Ill-Fitting Covers*

In this framework, Brisson points out, nonmaximal readings are expected. In the cover in (48), there are no coed cells—the men and the women are separated. But this is not necessary. One might imagine another kind of cover:

(49) {Rodgers, Hammerstein, Hart, Greta, Hilda}⊆D, and D contains nothing else apart from all parts and all sums of its members

Cov = {Rodgers+Hammerstein, Rodgers, Hart+Greta, Hilda}

With respect to this cover, Hart need not have participated in any way in writing a musical for The men wrote musicals to be true. This is because Hart is not a member of the cover and so does not satisfy the restriction on the distributive operator. Of course, the plural individual Hart+Greta does, since this individual is a member of the cover—however,

Hart+Greta is not a part of ⟦ *the men* ⟧. Thus Hart may be an exception to the predication here, and a nonmaximal interpretation is permitted. Naturally, with only three men such an interpretation is pragmatically rather unlikely, for reasons Brisson discusses.[10] But it is possible, and even more clearly when more men are involved.

Nonmaximal interpretations, then, arise with covers that are, in Brisson's terminology, ill-fitting. A cover is ill-fitting with respect to an individual x if there are parts of x that are stranded in cells in the cover that are not themselves parts of x. More precisely, Brisson (1998, p. 97) proposes the definition like that in (50), again reformulating here to accord with current assumptions:

(50) For some cover of the universe of discourse Cov and some DP denotation x, Cov is *a good fit* with respect to x iff $\forall y[y \leq x \rightarrow \exists z[z \in Cov \wedge y \leq z \wedge z \leq x]]$

An ill-fitting cover is one that is not a good fit.

5.5.4 *Brisson's* All *and the Good-Fit Requirement*

What all does, Brisson suggests, is impose a good-fit requirement on the cover with respect to which it is interpreted. This, of course, has the effect of eliminating the possibility of nonmaximal interpretations. One might naturally consider extending this approach to *whole* adjectives.[11]

---

[10] In fact, the framework in principle allows indefinitely nonmaximal interpretations—with the right sort of cover, any part of an individual, however small proportionally, can be sufficient for that individual to satisfy a predicate on a nonmaximal interpretation. Brisson illustrates how pragmatic considerations (interacting with lexical meaning) impose practical limits on the use of such highly nonmaximal readings.

[11] Brisson herself regards apparent nonmaximality effects involving singular DPs as entirely different from nonmaximality effects involving plural DPs. She is skeptical about the parallel nonmaximality possibil-

For Brisson, the way *all* makes this contribution is to restrict the range of possible assignment functions to only those with good-fitting covers. This is, she suggests, not part of truth-conditional meaning, but rather a component of what she calls the 'domain-adjusting meaning' of a sentence, an independent kind of meaning that is 'derived alongside truth conditions' (Brisson 1998, p.201). An alternative approach to this, though, is to impose the good-fit requirement directly in the denotations of particular lexical items. So, in order to express the intuition that all maps a DP denotation onto itself and adds the requirement that it be interpreted with respect to a good-fitting cover, one might adopt a denotation such as (42b):

(51) $[\![all_{Cov}]\!]^g = \lambda x . \iota y[y=x \text{ and } g(Cov) \text{ is a good fit with respect to } x]$

Or, combining this with the definition in (50):

---

ities in (i), noted by Landman (1989):

(i) a. John touched the ceiling.
b. The boys touched the ceiling.

All that is required for (ia) to be true is that John touch the ceiling. All of his body need not do so. Likewise, if the boys are stacked in a pyramid so that only the topmost boy touches the ceiling, (i) is true. But Brisson constructs various examples in which only the plural counterpart apparently exhibits nonmaximality:

(ii) a. The soldiers of the fourth platoon were captured by the enemy.
b. Bill was captured by the enemy.

(iii) a. Polly graded the exams.
b. Polly graded the exam.

(iv) a. Jane knows the answers to these questions.
b. Jane knows the answer.

These do appear to demonstrate that is not the case that all singulars exhibit nonmaximality in all contexts. This does not mean, however, that singulars do not in principle exhibit nonmaximality effects. In particular, it is possible, and indeed expected, that a singular and a plural similar to it would manifest different nonmaximality effects in different contexts. Thus, just as pragmatics determines when and how a plural can exhibit nonmaximality, it can determine when and how a singular can. In (ii-iv), it seems quite clear why the singular would be different from the plural for pragmatic reasons.

(52) $[\![all_{Cov}]\!]^g = \lambda x \,.\, \iota y[y{=}x \wedge \forall v[v \leq x \rightarrow \exists z[z \in g(Cov) \wedge v \leq z \wedge z \leq x]]]$

This amounts to treating the good-fit requirement as a presupposition (a possibility Brisson considers). To give the denotation of all access to the Cov variable with respect to which the DP will be interpreted, it is subscripted with it here. This sort of access to the cover is apparently required in the domain-adjusting meaning approach too. It is not obvious how to ensure that the cover variable with which all is subscripted matches the cover variable with which the corresponding Part operator is subscripted. This will have to remain a stipulation, though one might certainly imagine that some binding-theoretic requirement is involved. Importantly, this denotation—as well as any built on Brisson's definition of a good fit—will require that the DP argument of all be of type *e*, since the part relation is defined for this type only.

5.5.5 *The Singular Good-Fit Requirement and* Whole *Adjectives*

To extend this approach to *whole* adjectives, a number of additional revisions will be necessary. The most basic of these will be to make reference to parts of singulars, per the discussion in section 5.3, in the definition a good fit. Two approaches seem sensible. One is a disjunctive approach in which a distinct notion of good fit is defined for singulars with reference to a singular part relation:

(53) For some cover of the universe of discourse Cov and some DP denotation x, Cov is *a singular good fit* with respect to x iff $\forall y[y \leq_s x \rightarrow \exists z[z \in Cov \wedge y \leq_s z \wedge z \leq_s x]]$

A singular counterpart would have to be defined for the Part operator as well, again by simply replacing the plural part relation with the singular one. A tentative definition of *whole*, then, could be as in (54):

(54) $[\![\textit{whole}_{Cov}]\!] = \lambda x \,.\, g(\textit{Cov})$ is a singular good fit with respect to x (not final)

Combining (53) and (54):

(55) $[\![\textit{whole}_{Cov}]\!] = \lambda x \,.\, \forall y[y \leq_s x \rightarrow \exists z[z \in g(\textit{Cov}) \wedge y \leq_s z \wedge z \leq_s x]]$ (not final)

Entire would have an identical denotation.

An alternative approach that initially may seem more appealing is not to distinguish between the singular and plural part relations in the definition of a good fit and of the Part operator. This is where a theory in which there is a single generalized part relation, such as Moltmann's (1997), would be useful. There are independent reasons for not taking this route, however, which will begin to emerge momentarily, where focus will turn to the Part operator in more detail.

5.5.6 *The Determiner Restriction*

Given (55) as it stands, there would be no reason to expect the incompatibility of *whole* adjectives with inherently quantificational determiners. This can be corrected. One might hope to achieve an account of the determiner restriction without recourse to the brute-force technique of simply assigning *whole* adjectives a high type. But it is not clear what such a solution might be, or that a high type in this instance might not actually be more explanatory.

A determiner restriction similar to the one noted here is imposed by epistemic adjectives such as *unspecified* or *unknown* (Abusch and Rooth 1997). Such adjectives have what they call propositional readings:

(56) a. The suspects were arrested at unspecified locations. (Abusch and Rooth 1997)
'The suspects were arrested at some locations and it was unspecified which locations they were arrested in.'

b. Fabienne put the money in an unexpected place.
'Fabienne put the money at some place, and it was not expected that she would put the money in that place.'

c. Fabienne put the money in the expected place.

These readings are impossible with inherently quantificational determiners:

(57) a. Solange has stayed in every unspecified hotel. (Abusch and Rooth 1997)

b. The campus police installed burglar alarms in most unknown buildings.

Their account of this distinction, however, does not appear to extend naturally to *whole* adjectives. What they propose (roughly characterized) is that on the propositional reading, an adjective such as unspecified takes an argument corresponding to the individual whose identity is (or is not) specified, and that this argument is picked up as a discourse referent, which inherently quantificational DPs do not introduce.[12] But for *whole* adjectives, there is no independent reason to adopt this view, and indeed it seems potentially problematic to assume that a good fit requirement on covers might be imposed through an anaphoric mechanism.

On the alternative, higher-type approach, rather than denoting properties as in (55), *whole* adjectives would take determiner denotations as arguments:

---

[12] This is simplified in that it does not reflect their elaborations of the account in light of the observation that inherently quantificational DPs may introduce group discourse referents.

(58) $[\![\textit{whole}_{Cov}]\!]^g = \lambda P\ \lambda D\ .\ \iota x[x=D(P) \wedge g(\textit{Cov})$ is a singular good fit with respect to x]

This yields the unique individual that the DP would have denoted had the *whole* adjective been absent, with the additional requirement that the DP be interpreted with respect to a good-fitting singular cover. Thus it must be the case that had the *whole* adjective been absent, the DP denotation would have been of type *e*. Inherently quantificational determiners, then, are ruled out.

This approach has at least three virtues, aside from empirical adequacy. First, it parallels very closely the denotation given for *all* in (52), repeated here as (59):

(59) $[\![\textit{all}_{Cov}]\!]^g = \lambda x\ .\ \iota y[y=x \wedge g(\textit{Cov})$ is a good fit with respect to x]

This is more than a superficial or accidental parallel—it amounts to the claim that *whole* adjectives are incompatible with inherently quantificational determiners for the same reason that all is incompatible with inherently quantificational DP complements. Yet the high type allows this to be captured without being forced into syntactic assumptions, such as movement of *whole* adjectives to a position above the DP. Second, because in (58) *whole* adjectives no longer denote properties, the correct prediction is made that they should not occur in predicative positions. Third, in supposing that *whole* adjectives take determiner denotations as arguments, something is reflected of the special relationship certain 'functional' kinds of adjectives may bear to determiners. Larson (1999) and Zimmerman (2000) have proposed that certain adjectives may incorporate with certain determiners. This is the case overtly in English *another*. In German, the *whole* adjective *ganz* 'whole' is actually restricted to certain determiners in an idiosyncratic way. It seems desirable, then, to be able to capture a close relationship between certain kinds of adjectives and determiners. Yet in English, there does not appear to be any evidence, aside from the

similarity to *all*, to suppose that any such incorporation takes place with *whole* adjectives, at least in a syntactic or morphological sense. Supposing that *whole* adjectives take determiner arguments provides a way of expressing this relationship, but it does not require or preclude making these other commitments.

In this case, then, the higher-type approach actually may be more elegant, and perhaps even in some sense more parsimonious, than a lower-type alternative might be. This may be among the rare occasions when brute force is exactly what's called for.

5.5.7 *Syntactic Position and Type Ugliness*

There is, however, an additional concern, which may actually hint at a means of alleviating—or at least localizing more narrowly—the the need for this brute force.

This high-type approach, whatever its merits, inherently attributes to *whole* adjectives a kind of exceptionality. Not just any word, after all, could have so high a type. So if *whole* adjectives have this type, they are unavoidably special. In this fact alone, there is some reason for discomfort in light of the need to refer to *whole* adjectives and not merely to *whole*, say—even if it's a class with only two members, it's a bit suspicious that both members should have this high type. Equally worrying, perhaps, is that in addition to the high type, both require a cover index—again, not a characteristic that just any word can plausibly have.

This difficulty may ultimately be aggravated by cross-linguistic facts in this domain. Only the vaguest, most inchoate remarks in this direction will be attempted here, since it is necessary to tread on this empirical turf exceedingly lightly, for several reasons. For one thing, the facts in this domain are far from clear even in English. For another, it is not only possible but actually very probable that there are subtle but crucial semantic

differences in these sorts of words from language to language that could not be apparent without fairly detailed examination. And even on a very crude level, the basic distinction between *whole* adjectives and *complete* adjectives is often not easy to discern, and one can be mistaken for the other. But, acknowledging this and leaping blindly into the abyss, it seems to be the case that the behavior observed here for English *whole* adjectives is not altogether unusual—not only in its basic semantics, which isn't terribly surprising, but also in apparently much more superficial properties like the connection between interpretation and position. In Polish, for example, *whole* adjectives seem to be impossible postnominally, though *complete* adjectives are fine there:[13]

(60) a. konferencji prasowych . . . transmitowanych po raz pierwszy na cały
conferences press.GEN transmitted for time first to whole

kraj przez telewizje
country by television

'press conferences . . . televised for the first time to the whole country'

b. konferencji prasowych . . . transmitowanych po raz pierwszy na kraj cały przez telewizje

'press conferences . . . televised for the first time to the structurally intact country'

(61) a. Cały rok był marny.
whole year was lousy

'The whole year was lousy.

---

[13] The example in (60a) is from: Anna Carewicz. 2003. 'Kto zabił prezydenta?' ('Who killed the president?'). *Nowy Dziennik*. http://www.dziennik.com/www/dziennik/inne/artykuly/kto-zabil-prezydenta.html.

b. #Rok cały był marny.

‘The complete year was lousy.’

Cały in (60a) appears to be a pretty straightforward *whole* adjective, and the sentence yields an interpretation in which the spatial comprehensiveness of the broadcast is asserted. In (60b), though, the far more natural reading and possibly the only one possible for *cały* is as a *complete* adjective, and the sentence yields an interpretation in which the country is claimed to be complete or intact.[14] Similarly, in (61b) postnominal *cały* is necessarily a *complete* adjective, rendering the sentence odd. If it turns out that behind facts like these lurks a deeper cross-linguistic connection between *whole* adjective-style semantics and a particular position near the left periphery of the DP, this cannot be captured by assigning these adjectives a high type and proceeding as suggested in section 5.5.6. Because it’s not impossible to reflect *whole* adjectives-style semantics without this high type, it would be simply a lexical accident that multiple languages should have *whole* adjectives with high-type denotations.

I have no firm solution to offer to this problem—indeed, offering one would be premature without a better sense of the cross-linguistic facts. But one analytical option that’s available and seems natural enough is to break apart the semantics proposed in section 5.5.6 into two distinct chunks—one a high-type functional element which might plausibly be cross-linguistic, the other a better-behaved lower-type lexical element. This amounts to a kind of lexical decomposition, as (62) reflects (where the feature [+WHOLE] serves as a label for this ‘functional element’):[15]

---

[14] Use of a *complete* adjective in this context is actually counterpragmatic (the country in question is not Poland, it bears pointing out) so a preference would actually be expected in the other direction.

[15] This is, of course, an instance of the mediated approach to introducing modifiers laid out in a more explicit and general fashion in chapter 1.

(62) a. $[\![\,[\text{+WHOLE}]_{Cov}\,]\!]^g = \lambda P\ \lambda R\ \lambda D\,.\,\iota x[x=D(P) \wedge R(x)(g(Cov))]$

b. $[\![\,\textit{whole}\,]\!]^g = \lambda x\ \lambda Cov\,.$ Cov is a singular good fit with respect to x

The idea here is that the [+WHOLE] feature takes as arguments an NP, a *whole* adjective, and a determiner, and does two things with them: it assembles them in the necessary way, and it introduces the contextually-supplied cover into the mix. Exporting the complicated work from the denotation of *whole* itself in this way, it is possible to treat *whole* more straightforwardly as denoting the good-fit relation itself. This is, of course, a more desirable denotation, and likely to be more learnable than the high-type alternative. This makes these words less uncomfortably exceptional. And since functional elements are less likely to vary cross-linguistically, this would help avoid treating the properties [+WHOLE] reflects as a lexical accident.

But in order for this to capture the facts, it's necessary to assume a fixed position for [+WHOLE]. The fact that it takes a determiner as an argument will ensure that it has to appear close enough to it. Assuming that this feature occurs—perhaps in multiple languages—on the modified noun itself or else in the functional structure above it in its extended projection, one arrives at an interpretation like that in (63):

(63) *the* $\left[\textit{whole}\ \underset{[\text{+WHOLE}]}{\textit{sandwich}}\right]$

$[\![\,[\text{+WHOLE}]_{Cov}\,]\!]^g([\![\,\textit{sandwich}\,]\!]^g)([\![\,\textit{whole}\,]\!]^g)([\![\,\textit{the}\,]\!]^g)$

$= \iota x[x=\iota y[\text{sandwich}(y)] \wedge g(Cov)$ is a singular good fit with respect to x]

This, then, alleviates some of the concerns encountered in this section, though at the price of a certain amount of complexity. That said, though this more sophisticated theory of how *whole* adjectives are introduced may be the sort of thing that is ultimately required, it can safely be set aside for expository reasons in the subsequent sections, which focus primarily

on further refining the core semantics of *whole* adjectives, in favor the earlier higher-type non-decomposed denotation in (58) and the simpler syntax behind it.

5.5.8 *Choice Functions*

There is another complication the denotation in (58) presents. It requires that DPs modified by *whole* adjectives denote individuals. This is straightforward for definites. But for indefinites, especially nonspecific ones, this requires making some further assumptions.

For specific and wide-scope indefinites, this can be achieved naturally through a choice function interpretation (Winter 1997, Reinhart 1997, Kratzer 1998, Matthewson 1999).[16] Thus, assuming the choice-function interpretation for *a sandwich* in (64a), *a whole sandwich* may have the denotation in (64b):

(64) a. *a sandwich*

$[\![a]\!] = \lambda P . f(P)$, where F is a partial function picking out an x such that P(x)

$[\![a]\!]([\![sandwich]\!]) = f(\text{sandwich})$

b. *a* $whole_{Cov}$ *sandwich*

$[\![whole_{Cov}]\!]^g = \lambda P\ \lambda D . \iota x[x=D(P) \wedge g(Cov)$ is a singular good fit with respect to x

$[\![whole_{Cov}]\!]^g([\![sandwich]\!])([\![a]\!]) = \iota x[x=D(P) \wedge g(Cov)$ is a singular good fit with respect to x

Choice functions can thus ensure that indefinites may denote individuals. Since the denota-

---

16 One might also attempt to draw the distinction between inherently quantificational DPs on the one hand and definites and indefinites on the other in Kamp-Heim terms (Heim 1982, Kamp 1981). I won't pursue this possibility here.

tion for *whole* in (58) requires that DPs modified by *whole* adjectives always denote individuals, this would require that any indefinite modified by *whole* be a choice-function indefinite. This requires particular assumptions about choice functions.

In particular, the choice function mechanism must be one that would permit indefinites modified by *whole* adjectives to take narrow scope, as in (65):

(65) Every ferret ate an entire sandwich.

This has a reading in which sandwiches eaten vary with ferrets. In order to accommodate this fact, it will be necessary to assume that choice functions may be existentially closed inside the scope of other quantifiers (Winter 1997, Reinhart 1997). Supposing choice functions receive only widest-scope existential closure (Matthewson 1999) would predict that, given this approach to *whole* adjectives, the indefinite in (65) should have only a wide-scope reading. Although it is not entirely straightforward to demonstrate, a parameterized choice function mechanism of the sort proposed by Kratzer (1998) may not suffice, either. Kratzer accounts for pseudoscope effects—cases where indefinites take unexpectedly wide scope (or, in some cases, unexpectedly narrow scope)—by proposing that indefinite determiners may denote choice functions with an implicit argument that can be bound by higher quantifiers, creating the illusion of unexpected scope. This could in principle account for (65). It might be the case that the indefinite determiner here denotes a choice function with an argument bound by every ferret, so that for each ferret there is a different way of picking out a particular sandwich. But Kratzer further proposes that choice functions have values provided by the context—essentially, that only specific indefinites may have choice-function interpretations. Yet one may probably utter (65) on the reading in which sandwiches may vary with ferrets in a context that does not make salient a particular function from ferrets to sandwiches. This judgment may be somewhat precarious, though, if only in that it is unclear how it could adequately be tested.

5.5.9 *A Reservation: Is There Really 'Singular Distributivity'?*

At this point, it is worth pausing to note one potential qualm with embarking on an analytical path that leads to such a close parallel between singular and plural semantics.

One may, perhaps, sensibly abstract away from some of the problems concerning the differences between the part relation required for singulars and for plural and mass nouns, even if only tentatively. But if singulars really are interpreted using a Part distributivity operator, even in a cover-theoretic framework in which the notion of distributivity does not entail universal quantification over atoms, some notion of distributivity, however weakened, is placed foursquare at the heart of the semantics of singulars. Whether this is justified is not clear.

Of course, a predicate sometimes holds of all the parts of a singular of which it is predicated. Inversely, a predicate sometimes holds of a singular of which it is predicated but not necessarily of its parts. This is all entirely consistent with the cover-theoretic notion of distributivity—the former case may be modeled with a cover in which the parts occupy different cells and hence are distributed over, and the latter with a cover in which the parts are grouped together in a single cell. But covers coupled with the Part distributivity operator can give rise to other interpretations as well, as the discussion in section 5.5.2 illustrated. One might reasonably wonder, then, if these sorts of interpretations are also attested for singulars. Since a more pressing concern with the same source is about to arise, it will suffice to simply leave this question unresolved here, as a kind of pebble in our analytical shoe.

5.5.10 *A More Acute Problem: Trouble with Scope*

A more clear-cut and dramatic sign of trouble is that the account as it stands makes the wrong predictions with respect to the scope of certain indefinites. Scope ambiguities with respect to negation could arise, on this account, by supposing that the Part operator can occupy a syntactic position above or below negation. The absence of scope ambiguities with respect to indefinites in monoclausal sentences might also be explained. But this account wrongly predicts certain narrow-scope interpretations of indefinites in embedded clauses that are not in fact possible.[17]

The crucial fact about the interaction of *whole* adjectives and indefinites is that there doesn't seem to be any—that is, no scope ambiguities arise. Yet the Part distributivity operator introduces a universal quantifier with a property variable free in its scope. This leaves a backdoor open through which existential quantifiers might sneak in and thereby perniciously achieve narrow scope. In monoclausal sentences, there might be independent ways to guard against this risk. On typical assumptions, QR can attach indefinites only to proposition or truth-value denoting expressions; a corresponding assumption might be maintained on the choice-function approach adopted above. The Part operator, on the other hand, applies only to property-denoting[18] expressions, so in monoclausal structures will necessarily occur below the lowest point at which an indefinite may on these assumptions

---

[17] Thanks especially to Roger Schwarzschild for making me think more clearly about scope with respect to indefinites.

[18] This oversimplifies slightly; in fact, a counterpart of the Part operator that applies to relations must also be defined to account for distributivity over object positions, and more generally it is probably necessary to assume that these are alternative members of a family of distributivity operators that are generalized, as suggested by Lasersohn (1998), to various types in the same general way that generalized conjunction (Partee and Rooth 1983) is; this is the approach Brisson (1998) adopts. What is crucial for the current point is that no distributivity operator applies to propositions.

be interpreted, thereby ensuring that no indefinite is interpreted inside the scope of the universal introduced by the Part operator.

But in complex sentences, this mode of explanation is unavailable, and the backdoor is wide open. Before facing the problem head-on, though, it may be helpful to demonstrate that indefinites introduced in embedded clauses indeed do not seem to be in the scope of a universal even when a *whole* DP occurs in the matrix clause. Contrasts in the available interpretations of *different* reflect this especially clearly:

(66) a. Every member of the committee suggested that Floyd submerge a different ferret.

b. The whole committee suggested that Floyd submerge a different ferret.

(67) a. Everyone in town suspects that a different lamp store is run by the mafia.

b. The whole town suspects that a different lamp store is run by the mafia.

Just as when *whole* and *different* occur in clausemate DPs, the particular reading of *different* that arises in the scope of universals is not present in the (b) sentences, as it is in the (a) sentences—(67a) for example, cannot be used to express what an out of the blue utterance of (67b) most naturally means. The embedded indefinites, then, do not appear to be in the scope of a universal.[19] Yet given the account of *whole* adjectives as it stands, they should. The Part distributivity operator introduces a universal quantifier with scope over the entire matrix VP. If the embedded indefinite scopes within its clause, it would then be inside the scope of this universal, and the relevant readings of *different* should be possible in (66b)

---

[19] Differences in available readings might be detectable here even without *different*, but *different* provides a convenient way of abstracting away from complications introduced by the independently present *de re-de dicto* ambiguity.

and (67b). Illustrating:[20]

(68) *the whole*$_{Cov}$ *committee [Part*$_{Cov}$ *[suggested that Floyd submerge a ferret]]*

a. $[\![$ *suggested that Floyd submerge a ferret* $]\!]^g$ =

$\lambda x\,.\,\text{suggest}(^\wedge\exists f[\text{submerge}(f(\text{ferret}))(\text{Floyd})])(x)$

b. $[\![$ *Part*$_{Cov}$ $]\!]^g = \lambda P\ \lambda x\,.\,\forall y\ [[y \in g(Cov) \wedge y \leq_s x] \rightarrow P(y)]$

c. $[\![$ *Part*$_{Cov}$ *suggested that Floyd submerge a ferret* $]\!]^g = \lambda x\ \forall y[[y \in g(Cov) \wedge y \leq_s x]$

$\rightarrow \text{suggest}(^\wedge\exists[\text{submerge}(f(\text{ferret}))(\text{Floyd})])(y)]$

d. $[\![$ *the whole*$_{Cov}$ *committee* $]\!]^g = \iota x[x = \iota z[\text{committee}(z)] \wedge g(Cov)$ is a singular good fit with respect to x]

e. $[\![$ *the whole*$_{Cov}$ *committee Part*$_{Cov}$ *suggested that Floyd submerge a ferret* $]\!]^g$ =

$\forall y[[y \in g(Cov) \wedge y \leq_s \iota x[x = \iota z[\text{committee}(z)] \wedge g(Cov)$ is a singular good fit with respect to x]] $\rightarrow \text{suggest}(^\wedge\exists f[\text{submerge}(f(\text{ferret}))(\text{Floyd})])(y)]$

The resulting reading is one true if the cover is a singular good fit, and every member of the cover consisting only of parts of the committee suggested that Floyd commit ferret-submersion. The indefinite is in the scope of the universal introduced by the Part operator, and it is a puzzle why the indefinites in (66–67) do not seem to behave as though they were in the scope of a universal.

---

[20] The embedded indefinite is interpreted here in the choice function style, with existential closure at the clause level. The collective *the whole committee* is treated as denoting a singular individual here, though the point being made in the text probably does not hinge on what the proper analysis of collectives is (see Schwarzschild 1996 for discussion of this question). For ease of readability, I depart in a purely notational way from the roughly Heim and Kratzer (1998) assumptions otherwise adopted here in using the cap operator.

5.5.11 *The Singular Part Operator*

What the scope problem and the uneasiness expressed above collectively suggest is that in fact there may be something wrong about thinking of singulars as involving distributivity, even in the cover-theoretic way. Perhaps, then, the locus of distributivity—the Part operator—is to blame.

We cannot simply throw out the rotten apple here, however. While distributivity may not be quite the notion that is required, the Part operator does more than that. It is also crucial to the mechanics of covers—it does the essential work of relating a predicate and an individual in a cover-sensitive way. What is necessary, then, is some means of doing this essential work without distributivity. Fortunately, this seems to be possible.

Rather than simply borrowing the singular Part operator as a kind of calque on the plural one, or possibly treating the singular and plural Part operators as somehow not fully distinct, one can define the singular Part operator on its own terms. To do this, one might begin by thinking afresh and perhaps a bit impressionistically about what exactly it means to predicate something of a singular individual. If I assert of a sandwich that it is eaten, I seem to be predicating consumption in a kind of collective way of a single individual, possibly an ‘integrated whole’ in Moltmann’s sense. Depending on the context, this single individual may not actually include every part of the sandwich, of course—it may simply be part of the sandwich, and in most contexts, it is a big part. But I am probably not predicating consumption in some independent, distributive way of the bread, the hummus, the cucumber slices, and the mysterious little seeds (cumin?) falling all over my plate. Indeed, if I had for some reason wanted to do this, I might have used a plural (the parts of the sandwich, possibly).

Perhaps, then, we can do justice to this vague intuition by supposing that the singular Part operator, unlike its plural counterpart, applies a predicate to only one cell

in the cover—namely, in this case, the one containing whatever I regard as the sandwich in the current context. Assuming this context is fairly normal, the bread, hummus, cucumber slices, and most of the seeds will probably count as part of the sandwich, but the seeds that have fallen on my plate probably do not. Thus the bread, hummus, cucumber slices, and non-fallen seeds will make up a single cell of the cover in this context; the fallen seeds will have to be included in some other cell (perhaps the one that also includes my plate) of which consumption will not be predicated. Given the domain D indicated, things might look like this:

(69) {the bread, the hummus, the cucumbers slices, the seeds that have fallen, the seeds that have not fallen, me, my plate, Clyde}⊆D, and D contains nothing else apart from all parts and all sums of its members

Cov = {the bread+the hummus+the cucumber slices+the seeds that have not fallen, the seeds that have fallen+my plate, me, Clyde}

This, then, models a nonmaximal reading—it separates the irrelevant parts of the sandwich—but does not involve distributivity. It is, in fact, precisely analogous to how collective predication works in plurals. In singulars, though, this seems to be the only interpretation possible. Use of a singular seems to presuppose that the cover contains a single cell of which something will be predicated (collectively, in this sense).

The singular Part operator can be redefined to reflect this requirement; for comparison, the plural Part operator is repeated here as well, and they are more clearly distinguished notationally:

(70) a. $[\![\mathit{sgPart}_{Cov}]\!]^g = \lambda P\ \lambda x\ .\ P(\iota y[y \in g(\mathit{Cov}) \wedge y \leq_s x])$

b. $[\![\mathit{plPart}_{Cov}]\!]^g = \lambda P\ \lambda x\ .\ \forall y[[y \in g(\mathit{Cov}) \wedge y \leq x] \rightarrow P(y)]$

The singular Part operator is no longer a distributivity operator. It now applies its first argument, a property, to the unique cell in the cover that is part of its second argument, an individual. The plural Part operator is just as it was before. It applies its first argument, a property, to every cell in the cover that is part of its second argument, an individual. So the singular Part operator is a more demanding version of the plural one, in that it presupposes a particular kind of cover.

The cover in (69) satisfies this presupposition, since there is a unique cell that contains only parts of the sandwich. A cover otherwise identical to it in which the seeds that have fallen are a cell of their own, however, would not, since there would be two sandwich-only cells. That this possibility is ruled out in singular predication—understood here as a failure of presupposition—is consistent with the intuition expressed above that singular predication is in a sense always collective.

Despite this quite fundamental revision, the role of *whole* adjectives remains unchanged. They can still contribute a good fit requirement, just as they did before.

Now that the definitions of the singular and plural Part operators are so clearly distinct, it becomes especially pressing that the right Part operator occur with the right sort of DP. There are a number of ways of approaching this. One might imagine that mismatches are ruled out by what varieties of individuals can stand in the singular or plural relation to each other. Alternatively, this could be a matter of syntactic agreement. If, as Schwarzschild suggests, Part operators are related to floated quantifiers, this would not be unexpected—floated quantifiers may in fact agree with their associated DPs, as they do in for example Hebrew (Shlonsky 1991) and German (Merchant 1996). A related understanding might view the relationship between the plural Part operator and floated quantifiers as fundamentally the same as the relationship between the singular Part operator and adverbs that seem to quantify over parts of a singular, such as mostly, wholly, or largely, or even the use of

all in expressions like all wet (Schwarzschild 1996, p. 162–170).[21]

This revision, then, has divorced the Brissonian cover-theoretic approach to nonmaximal interpretations from distributivity, allowing *whole* adjectives to render nonmaximal readings impossible without presupposing a theory of singular predication that requires universal part quantification.

5.5.12 *Scope of Indefinites Revisited*

In eliminating universal quantification from the semantics of the singular Part operator, the prediction of the unattested scope interactions with respect to indefinites has been eliminated as well.

The new interpretation of the previously problematic (68) illustrates this:

(71) *the whole$_{Cov}$ committee [sgPart$_{Cov}$ [suggested that Floyd submerge a ferret]]*

a. $[\![$ *suggested that Floyd submerge a ferret* $]\!]^g$ =

λx . suggest($^\exists$f[submerge(f(ferret))(Floyd)(x)

b. $[\![$ *sgPart$_{Cov}$* $]\!]^g = \lambda P \lambda x$. P(ιy[y$\in$g(*Cov*) $\wedge$ y$\leq_s$x]

c. $[\![$ *sgPart$_{Cov}$ suggested that Floyd submerge a ferret* $]\!]^g$ =

λx . suggest($^\exists$f[submerge(f(ferret)(Floyd)])(ιy[y$\in$g(*Cov*) $\wedge$ g(*Cov*) is a singular good fit with respect to x])

d. $[\![$ *the whole$_{Cov}$ committee* $]\!]^g$ = ιx[x=ιz[committee(z)] $\wedge$ g(*Cov*) is a singular good fit with respect to x]

---

[21] Thanks to an anonymous SALT reviewer for pointing out the relevance of examples like this.

e. ⟦ *the whole*$_{Cov}$ *committee sgPart*$_{Cov}$ *suggested that Floyd submerge a ferret* ⟧^{g} =

suggest($^{\wedge}\exists$f[submerge(f(ferret))(Floyd)])(ιy[y$\in$g(*Cov*) $\wedge$

y$\leq_{s}\iota$x[x=ιz[committee(z)] $\wedge$ g(*Cov*) is a singular good fit with respect to

x]])

The existential introduced by the indefinite was previously in the scope of the universal quantifier over cells in the cover. Now, there is no such universal for it to interact with, and no scope interactions with the Part operator are expected.

More generally, an explanation has now emerged for why *whole* adjectives do not give rise to scope ambiguities with respect to indefinites. *Whole* adjectives are not themselves universally part-quantifying, and the singular Part operator does not contribute any universal quantifiers that might give rise to such scope interaction, either.

5.5.13 *Scope of Negation:* Not*, Scope, and Not Scope*

At this point, one might be concerned about the apparent scope interactions between *whole* adjectives and negation observed in section 5.2 above. In redefining the singular Part operator, we have eliminated scope interactions with indefinites, but in doing so, have we also lost an account of the interaction of *whole* adjectives and negation?

In one sense, yes. On a theory in which the singular Part operator has universal force, one might suppose that multiple interpretations are possible depending on whether negation occurs immediately above or below the syntactic position of the Part operator, and therefore with scope over or under the universal it contributes. But with the revised singular Part operator, such an account is no longer available. The relative position of Part and negation could make no difference:

(72) The whole ferret isn't submerged.

(73) *the whole$_{Cov}$ ferret isn't [sgPart$_{Cov}$ submerged]*

a. $[\![\textit{submerged}]\!] = \lambda x . \text{submerged}(x)$

b. $[\![\textit{sgPart}_{Cov} \textit{ submerged}]\!]^g = \lambda x . \text{submerged}(\iota y[y \in g(Cov) \wedge x \leq_s x])$

c. $[\![\textit{isn't sgPart}_{Cov} \textit{ submerged}]\!]^g = \lambda x . \neg\text{submerged}(\iota y[y \in g(Cov) \wedge y \leq_s x])$

d. $[\![\textit{the whole}_{Cov} \textit{ ferret}]\!]^g = \iota x[x = \iota z[\text{ferret}(z)] \wedge g(Cov)$ is a singular good fit with respect to x]

e. $[\![\textit{the whole}_{Cov} \textit{ ferret isn't sgPart}_{Cov} \textit{ submerged}]\!]^g =$ $\neg\text{submerged}(\iota y[y \in g(Cov) \wedge y \leq_s \iota x[x = \iota z[\text{ferret}(z) \wedge g(Cov)$ is a singular good fit with respect to x]])

(74) *the whole$_{Cov}$ ferret sgPart$_{Cov}$ [isn't submerged]*

a. $[\![\textit{isn't submerged}]\!] = \lambda x . \neg\text{submerged}(x)$

b. $[\![\textit{sgPart}_{Cov} \textit{ isn't submerged}]\!]^g = \lambda x . \neg\text{submerged}(\iota y[y \in g(Cov) \wedge y \leq_s x])$

c. $[\![\textit{the whole}_{Cov} \textit{ ferret sgPart}_{Cov} \textit{ isn't submerged}]\!]^g = \neg\text{submerged}(\iota y[y \in g(Cov) \wedge$ $y \leq_s \iota x[x = \iota z[\text{ferret}(z)] \wedge g(Cov)$ is a singular good fit with respect to x]])

An identical denotation is arrived at in (73) and (74), so an explanation in terms of the relative scope of negation and Part is not an option.

There is, however, another way of looking at things. The interpretation a sentence such as (72) receives on this analysis is a weak, scopally 'underspecified' one. What is denied in (72) is that the ferret, construed maximally, is submerged; it is not asserted that the ferret, construed maximally, is unsubmerged. For current purposes, this could actually be paraphrased in quantificational terms: the interpretation of (72) corresponds to a reading

in which negation has scope over a universal part quantifier, not vice versa. Of course, such a weaker interpretation does not rule out the possibility that the stronger one could also be true. To deny that the ferret, construed maximally, is submerged is not to assert that the ferret, construed maximally, is unsubmerged. Perhaps, then, the weaker reading is all the semantics provides, and the stronger reading arises in some contexts pragmatically.

One reason this seems a sensible route to take is that some version of it would probably be necessary anyway in a scope-based account, as the interpretation of (75) illustrates:

(75) Floyd really didn't want to talk about that whole sensitive topic.

The most natural interpretation here is the strong one, in which this sentence is taken to claim that there was no part of the topic Floyd wanted to talk about. But even on a scope-based account, only the weaker interpretation (in which Floyd's desire was for discussion of not all of the topic) could be provided by the semantics. Both the *whole* adjective and the relevant singular Part operator are embedded under negation here, which would yield the weak reading. And because negation and the *whole* DP are not clausemates, even if some quantificational force could be assumed for the DP, it is not clear how it could achieve scope over the matrix-clause negation. It is independently necessary, then, to assume that stronger readings can sometimes arise for sentences in which the semantics provides only the weaker ones.

Another reason for attributing the stronger interpretations to pragmatic strengthening is their sensitivity to lexical properties of predicates. The stronger interpretations arise most naturally with what Yoon (1996) (see also Kennedy and McNally 1999 and Rotstein and Winter 2001) calls partial predicates—very roughly, predicates such as *dirty*,

wet, or *touch* that typically hold of an individual if they hold of any part of it. These contrast with total predicates such as *clean*, *dry*, or *eat*, which typically hold of an individual only if they hold of all of it.[22] With total predicates, the stronger reading does not as naturally arise:

(76) a. *Partial predicate*: Clyde didn't touch this whole sandwich.
b. *Total predicate*: Clyde didn't eat this whole sandwich.

(77) a. *Partial predicate*: Floyd wouldn't mention that entire problem.
b. *Total predicate*: Floyd wouldn't solve that entire problem.

Out of the blue, it is quite natural to take the partial predicate sentences as involving the complete absence of sandwich-touching and problem-mentioning. But the most likely interpretation of the total predicate sentences involves incomplete sandwich-consumption and

---

[22] The relation between partial and total predicates and the cover approach to nonmaximal interpretations may warrant more investigation than it can be given here. One way of understanding partial predicates in this light is as predicates that in most contexts tolerate extremely ill-fitting covers; total predicates can be understood to tolerate only covers that fit relatively well. Kennedy and McNally (1999) develop a fine-grained theory of essentially the total-partial distinction in adjectives, with special attention to degree modification, in terms of the scale structure. Some connection between a cover-theoretic and a scale-structure representation of this contrast might be expected. Both approaches have at their heart the interaction of context-sensitivity with aspects of lexical semantics that can circumscribe it. Empirically, there is a connection, too. Certain part-structure adverbs that may be thought of as overt relatives of the singular Part operator also have readings that involve not proportions of an individual, but, loosely put, proportions of a scale:

(i) a. The campus is completely nauseating.
b. Oklahoma is wholly Republican.
c. Gretchen is mostly Portuguese.
d. The table is largely dry.
e. Her committee is entirely opposed to actually trying to use that title.

These are ambiguous. The campus in (ia), for example, may be claimed to have no non-nauseating parts, or else to cause a maximal degree of nausea. Oklahoma in (ib) may have been claimed to contain only Republicans, or else to be Republican to a maximal degree. Gretchen in (ic) may be claimed to have been assembled in such a way that most of her parts came from Portugal (suppose she is a mannequin, say); or else to be Portuguese to a maximal degree. Further examination of this empirical connection, of the possible usefulness of covers in this domain, and of how this may relate to adjectival scale structure will have to be left to future research.

incomplete problem-solving. The outlines of an explanation of this contrast may lie in the observation that on the weak reading, the partial-predicate sentences deny that something independently fairly unlikely. In (76a), it would be denied that Clyde made full physical contact with all of the sandwich, perpetrating upon it either systematic poking or some sort of exhaustive rubbing. In (77b), it would be denied that Floyd's mere mention of the problem exhausted the subject, which would require either that the problem be so small that a discussion of all of it could be characterized as a mention, or else that Floyd kept mentioning the problem in separate acts that ultimately exhausted it. Given the relative unlikelihood of these readings, it seems reasonable that the stronger ones would be in general more salient. The weak readings of the total-predicate sentences, however, deny something relatively less pragmatically odd. Of course, none of this is evidence against a scope ambiguity account directly—on such a view, these pragmatic factors would still select among the possible readings in the same way. But again, this is a case in which independent pragmatic factors probably must be appealed to on either approach.

All this suggests, then, that an understanding of the interaction of *whole* adjectives and negation should be largely pragmatic—an unsurprising conclusion, perhaps, since covers model essentially pragmatic, or in any case contextual, phenomena. There is more to be said about these issues, though, and some speculation about possible research directions will be ventured in section 5.6 below.

5.5.14 *Discourse Anaphora*

The view that *whole* adjectives are maximizing modifiers rather than universal part quantifiers seems to provide a natural explanation of their discourse anaphoric properties. Since *whole* adjectives merely enforce a requirement on covers, with respect to discourse

anaphora DPs containing them should behave though they were absent. This correctly predicts that *whole* adjectives should have no effect on discourse anaphora possibilities.

As noted in section 5.2, DPs with *whole* adjectives may introduce discourse referents, while universally quantified DPs may not. Repeating (17–18):

(78) a. Every ferret was submerged. *It didn't seem pleased.
b. A entire ferret was submerged. It didn't seem pleased.

(79) a. Cockroaches ate each sandwich. *It had been sitting out too long.
b. Cockroaches ate a whole sandwich. It had been sitting out too long.

This is now expected, since there will be no universal quantifier in DPs modified by *whole* adjectives.

For exactly the same reason, DPs with *whole* adjectives do not impede DPs in their scope from introducing discourse referents accessible outside it. Repeating (19–20):

(80) a. Every building with a faulty roof_i was demolished. *It_i had been leaking.
b. An entire building with a faulty roof_i was demolished. It_i had been leaking.

(81) a. Each recent movie was marred by an irritating soundtrack_i. *It_i will probably be released on CD.
b. The whole movie was marred by an irritating soundtrack_i. It_i will probably be released on CD.

Again, since *whole* adjectives are not universal quantifiers, and indeed are nonquantificational, it is expected that they should have no effect on discourse anaphora possibilities.

This approach to *whole* adjectives does raise a potentially interesting possibility: might the good-fit requirement on covers *whole* adjectives impose persist in discourse anaphora? That is, might we expect pronouns with antecedents modified by *whole* adjectives to also impose the good-fit requirement? (This question could also be asked with respect to all, incidentally.) Empirically, this does not seem to be the case:

(82) Clyde ate the whole sandwich. It tasted funny.

The second sentence of (82) does not require that the whole sandwich have tasted funny. (It does not preclude this possibility, but nothing can be concluded from this—this possibility is normally not precluded.) The current analysis of *whole* adjectives predicts this result, for two independent reasons. First, it need not be the case that the cover used in the first sentence in the interpretation of the whole sandwich is the same as the cover used in the second sentence in the interpretation of it. Second, the denotation of a DP modified by a *whole* adjective will always contain a portion consisting of the denotation of the DP without the *whole* adjective, so the discourse referent introduced by the DP without the *whole* adjective will always be accessible.

5.6 Some Further Questions and a Final Word

5.6.1 *Further Questions*

A few puzzles associated with *whole* adjectives have not been addressed here, but bear pointing out briefly. One of them is a mysterious quirk of the syntax of *whole* in English with respect to numerals:

(83) a. A whole three thousand ferrets were submerged.

b. Three thousand whole ferrets were submerged.

Both of these orders are in one way or another potentially problematic. In (83a), the position of the *whole* adjective is just below the determiner, as the account developed here would predict. But it is puzzling what an overt determiner is doing in this position at all. In general, English plurals are incompatible with *a*:

(84) *A three thousand ferrets were submerged.

Yet the DP in (83a) is indisputably grammatically plural, as both the plural morphology on the noun and the plural agreement morphology on the verb reflect. It seems possible that this is related to a more general (but nonetheless mysterious) phenomenon in which certain adjectives can license a on plurals—e.g. *An incredible 2 billion people watched the World Cup final*. Semantically, it is not entirely clear what *whole* contributes in such sentences, either, or indeed whether it receives the *whole* adjective or the *complete* adjective reading. Whatever the syntax and semantics of this sort of structure might be, it does not seem to follow directly from anything said here. In (83b), another sort of problem seems to emerge. The position of *whole* is below a numeral, which would be puzzling under this analysis unless the numeral is analyzed as a determiner rather than an adjective. Perhaps more vexing, the DP is plural, and it is not obvious that *whole* here receives the irrelevant *complete* adjective reading.

A more general puzzle that may be connected to these observations concerns the relationship between *whole* adjectives and polarity. This is a topic explored in den Dikken (to appear), who examines Dutch *heel* ‘whole’, which shares many of the properties of English *whole*. He notes an intriguing peculiarity of this *whole*: it has a negative

polarity item incarnation with intricate and unusual licensing conditions. A taste of the effect:

(85) Ik ken die hele vent *(niet) (den Dikken to appear)
I know that whole.INF bloke not

'I don't know that bloke at all.'

Interestingly, this *heel* occurs most easily with demonstratives, which may correspond to an effect in English in which demonstratives give rise to strengthened readings with respect to negation more easily than do definites or indefinites:

(86) Floyd doesn't want to talk about $\left\{\begin{matrix}\text{this}\\\text{the}\\\text{an}\end{matrix}\right\}$ entire subject.

In German, polarity effects and the syntactic strangeness in (83) seem to converge (Angelika Kratzer, p.c.). As in English, both the pre-numeral and post-numeral position is possible for a *whole* adjective in German, but interestingly, the pre-numeral position requires positive polarity and receives a special interpretation:

(87) Hans/niemand hat ganze zwei Dissertationen betreut.
Hans/no one has whole two dissertations supervised

'Hans/no one has supervised two dissertations.'

(Gives rise to the feeling that two is not a lot.)

These issues will be left for future research. It may be worth noting, though, that the use of covers may help in understanding the connection between *whole* adjectives and polarity. Covers are essentially a kind of domain restriction, and the good-fit require-

ment can be understood as an intellectual cousin to the sort of domain-widening Kadmon and Landman (1993) suggest is responsible for the polarity sensitivity of English *any*.[23] Since an intimate connection between manipulation of contextually-provided domains and polarity sensitivity is independently expected, an approach to *whole* adjectives in terms of restrictions on covers may provide the foundation for an understanding of their polarity properties.

5.6.2 *A Final Word*

The central argument made here is that *whole* adjectives are not universally quantifying, but maximizing modifiers in the sense of Brisson (1998). This accounts for their essential semantic contribution, the restrictions they impose on determiners, their scope properties, and the anaphoric possibilities they license. In the course of pursuing this approach, a general model of nonmaximal interpretations of singulars was developed and some consequences for choice functions were encountered. To the extent that this approach managed to explain nonmaximal interpretations of singulars and plurals similarly using the same sorts of theoretical tools, it suggests that there is something right about the analytical intuition, expressed especially by Moltmann (1997), that part structure in singulars and in plurals should be understood in highly parallel terms.

---

[23] Thanks to Ana Arregui pointing me in this direction.

REFERENCES

Abney, Steven. 1987. *The English Noun Phrase in Its Sentential Aspect*. Doctoral dissertation, MIT.

Abusch, Dorit. 1997. 'Sequence of tense and temporal *De Re*'. *Linguistics and Philosophy* 20:1.

Abusch, Dorit and Mats Rooth. 1997. 'Epistemic np modifiers'. In Lawson, A., ed., *Proceedings of SALT VII*. CLC Publications, Ithaca, New York.

Alexiadou, Artemis. 1997. *Adverb Placement: A Case Study in in Antisymmetric Syntax*. John Benjamins, Amsterdam.

Atlas, Jay D. 1984. 'Comparative adjectives and adverbials of degree: an introduction to radically radical pragmatics'. *Linguistics and Philosophy* 7(4):347.

Baker, Mark C. 2003. *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge University Press, Cambridge.

Barker, Chris. 2003. 'A gentle introduction to Type Logical Grammar, the Curry-Howard correspondence, and cut elimination'. Ms., University of California, San Diego.

Bartsch, Renate. 1976. *The Grammar of Adverbials*. North-Holland, Amsterdam.

Beck, Sigrid. 2000. 'The semantics of 'different': Comparison operator and relational adjective'. *Linguistics and Philosophy* 23(2):101.

Bellert, Irena. 1977. 'On semantic and distributional properties of sentential adverbs'. *Linguistic Inquiry* 8(2):337.

Bernstein, J. 1993. *Topics in the Syntax of Nominal Structure Across Romance*. Doctoral dissertation, City University of New York.

Bierwisch, Manfred. 1989. ‘The semantics of gradation’. In Bierwisch, Manfred and Ewald Lang, eds., *Dimensional Adjectives*. Springer-Verlag, Berlin.

Bittner, Maria and Ken Hale. 1996. ‘The structural determination of case and agreement’. *Linguistic Inquiry* 27:1.

Blutner, Reinhard. 1998. ‘Lexical pragmatics’. *Journal of Semantics* 15:115.

Bolinger, Dwight. 1967. ‘Adjectives in English: Attribution and predication’. *Lingua* 18:1.

Bosque, Ignacio and Carme Picallo. 1996. ‘Postnominal adjectives in Spanish DPs’. *Journal of Linguistics* 32:57.

Bouma, Gosse. 2003. ‘Verb clusters and the scope of adjuncts in Dutch’. In Seuren, Pieter A. M. and Gerard Kempen, eds., *Verb Constructions in German and Dutch*. Benjamins, Amsterdam and Philadelphia.

Bresnan, Joan and Jane Grimshaw. 1978. ‘The syntax of free relatives in English’. *Linguistic Inquiry* 9:331.

Brisson, Christine. 1998. *Distributivity, Maximality, and Floating Quantifiers*. Doctoral dissertation, Rutgers University.

Bunt, H. 1979. *Mass Terms and Model-Theoretic Semantics*. Cambridge University Press, Cambridge.

Carlson, Greg. 1981. ‘Distribution of free-choice *Any*’. In *Proceedings of CLS 17*.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris Publications, Dordrecht.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. Praeger, New York.

Cinque, Guglielmo. 1994. ‘Evidence for partial N-movement in the Romance DP’. In Cinque et al., ed., *Paths Toward Universal Grammar: Essays in Honor of Richard S. Kayne*. Georgetown University Press, Washington D.C.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, New York.

Cinque, Guglielmo. 2003. ‘The dual source of adjectives and XP- vs. N-raising in the Romance DP’. Handout, North Eastern Linguistic Society (NELS) 35.

Cooper, Robin. 1983. *Quantification and Syntactic Theory*. D. Reidel, Dordrecht.

Corver, Norbert. 1990. *The Syntax of Left Branch Extractions*. Doctoral dissertation, Tilburg University.

Cresswell, Max J. 1976. ‘The semantics of degree’. In Partee, Barbara H., ed., *Montague Grammar*. Academic Press, New York.

Davidson, Donald. 1967. ‘The logical form of action sentences’. In Rescher, Nicholas, ed., *The Logic of Decision and Action*. University of Pittsburgh Press, Pittsburgh. Republished in Donald Davidson, *Essays on Actions and Events*, Oxford University Press, Oxford, 1980.

den Dikken, Marcel. to appear. ‘Direct and parasitic polarity item licensing’. *The Journal of Comparative Germanic Linguistics* .

Di Sciullo, Anna Maria and Edwin Williams. 1987. *On the Definition of Word*. MIT Press, Cambridge, Massachusetts.

Dixon, R.M.W. 1982. *Where Have All the Adjectives Gone?* Mouton, Berlin.

Dowty, David R. 1979. *Word Meaning in Montague Grammar*. D. Reidel Publishing Co., Dordrecht, Holland.

Elliott, Dale E. 1974. ‘Toward a grammar of exclamations’. *Foundations of Language* 11:231.

Emonds, Joseph. 1978a. *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations*. Academic Press, New York.

Emonds, Joseph. 1978b. ‘The verbal complex V′–V in French’. *Linguistic Inquiry* 9:151.

Emonds, Joseph. 1987. ‘The invisible category principle’. *Linguistic Inquiry* 18:613.

Ernst, Thomas. 1984. *Toward an Integrated Theory of Adverb Position in English*. Doctoral dissertation, Indiana University.

Ernst, Thomas. 2002. *The Syntax of Adjuncts*. Cambridge University Press, Cambridge.

Faller, Martina. 2000. 'Dimensional adjectives and measure phrases in vector space semantics'. In Faller, M., S. Kaufmann, and M. Pauly, eds., *Formalizing the Dynamics of Information*. CSLI Publications, Stanford.

Fillmore, Charles. 1994. 'Under the circumstances (place, time, manner, etc.)'. In *Proceedings of the 20th Annual Berkeley Linguistics Society*.

Geuder, Wilhelm. 2000. *Oriented Adverbs: Issues in the Lexical Semantics of Event Adverbs*. Doctoral dissertation, Universität Tübingen.

Gibson, Kean. 1992. 'Tense and aspect in Guyanese Creole with reference to Jamaican and Carriacouan'. *International Journal of american Linguistics* 58:49.

Gillon, Brendon S. 1987. 'The readings of plural noun phrases in English'. *Linguistics and Philosophy* 10(2):199.

Giorgi, Alessandra and Fabio Pianesi. 1997. *Tense and Aspect: From Semantics to Morphosyntax*. Oxford University Press, Oxford.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press, Chicago.

Grice, H. Paul. 1975. 'Logic and conversation'. In Cole, Peter and Jerry Morgan, eds., *Syntax and Semantics 3: Speech Acts*. Academic Press, New York.

Grimshaw, Jane. 1979. 'Complement selection and the lexicon'. *Linguistic Inquiry* 19(2):279.

Grimshaw, Jane. 1991. 'Extended projection'. In *Lexical Specification and Lexical Insertion*. Erlbaum, Hillsdale, NJ.

Groenendijk, Jeroen and Martin Stokhof. 1984. 'On the semantics of questions and the pragmatics of answers'. In Landman, Fred and Frank Veltman, eds., *Varieties of Formal Semantics*. Foris, Dordrecht.

Guttiérrez Rexach, Javier. 1997. 'Questions and generalized quantifiers'. In *Ways of Scope Taking*. Kluwer Academic Publishers, Dordrecht.

Hackl, Martin. 2000. *Comparative Quantifiers*. Doctoral dissertation, MIT.

Hamblin, Charles. 1973. 'Questions in Montague English'. *Foundations of Language* 10(1):41.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation, University of Massachusetts.

Heim, Irene. 1994. 'Comments on Abusch's theory of tense'. In Kamp, Hans, ed., *Ellipsis, Tense, and Questions*. University of Amsterdam.

Heim, Irene and Angelika Kratzer. 1997. *Semantics in Generative Grammar*. Blackwell Publishers, Oxford.

Hetzron, Robert. 1978. 'On the relative order of adjectives'. In Seiler, H., ed., *Language Universals*. Gunter Narr Verlag, Tübingen.

Hitzeman, Janet. 1992. 'The selectional properties and entailments of *Almost*'. In *Proceeding of the Chicago Linguistics Society 28*.

Hopf, Jens-Max, Josef Bayer, Markus Bader, and Michael Meng. 1998. 'Event-related brain potentials and case information in syntacticc ambiguities'. *Journal of Cognitive Neuroscience* 10:264.

Horn, Laurence R. 1972. *On the Semantic Properties of Logical Operators in English*. Doctoral dissertation, UCLA.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, Massachusetts.

Jackendoff, Ray. 1977. *X-Bar Syntax: A Study of Phrase Structure*. No. 2 in Linguistic Inquiry Monographs. MIT Press, Cambridge, Massachusetts.

Jacobson, Pauline. 1999. 'Towards a variable-free semantics'. *Linguistics and Philosophy* 22(2):117.

Jacobson, Polly. 2000. 'Pied piping semantics in relative clauses'. Handout for a talk at the Workshop on the Syntax and Semantics of Relative Clauses at Tel Aviv University.

Johnson, Kyle. 1991. 'Object positions'. *Natural Language and Linguistic Theory* 9:577.

Kadmon, Nirit and Fred Landman. 1993. ‘Any’. *Linguistics and Philosophy* 16(4):353.

Kamp, Hans. 1975. ‘Two theories about adjectives’. In Keenan, Edward L., ed., *Formal Semantics of Natural Language*. Cambridge University Press, Cambridge, England.

Kamp, Hans. 1981. ‘A theory of truth and semantic representation’. In Groenendijk, Jeroen A., Theo Janssen, and Martin Stokhof, eds., *Formal Methods in the Study of Language*. Foris, Dordrecht.

Kamp, Hans and Barbara Partee. 1995. ‘Prototype theory and compositionality’. *Cognition* 57(2):121.

Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Modeltheoretic Semantics in Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer Academic Publishers, Dordrecht.

Karttunen, Lauri. 1977. ‘Questions revisited’. Unpublished manuscript, The Rand Corporation.

Kay, Paul and Charles J. Fillmore. 1999. ‘Grammatical constructions and linguistic generalizations: the ‘what’s X doing Y?’ construction’. *Language* 75:1.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. The MIT Press, Cambridge, Massachusetts.

Kennedy, Christopher. 1997. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Doctoral dissertation, University of California, Santa Cruz. Published in 1999 by Garland, New York.

Kennedy, Christopher and Louise McNally. 1999. ‘From event structure to scale structure: Degree modification in deverbal adjectives’. In Mathews, T. and D. Strolovitch, eds., *Proceedings of SALT IX*. CLC Publications, Ithaca, New York.

Kennedy, Christopher and Louise McNally. 2004. ‘Scale structure and the semantic typology of gradable predicates’. Under review for *Language*.

Keyser, Samuel Jay. 1968. ‘Review of Sven Jacobson, Adverbial position in English’. *Language* 44:357.

Klein, Ewan. 1991. ‘Comparatives’. In von Stechow, Arnim and Dieter Wunderlich, eds., *Semantik: Ein internationales handbuch der zeitgenössischen forschung*. Walter de Gruyter, Berlin.

Klein, Wolfgang. 1994. *Time in Language*. Routledge, London.

Koopman, Hilda. 1996. ‘The spec head configuration’. In *UCLA Working Papers in Linguistics 1*. Los Angeles.

Kratzer, Angelika. 1996. ‘Severing the external argument from its verb’. In Rooryck, Johan and Laurie Zaring, eds., *Phrase Structure and the Lexicon*. Kluwer Academic Publishers, Dordrecht.

Kratzer, Angelika. 1998. ‘More structural analogies between pronouns and tenses’. In *Proceedings of SALT VIII*. CLC Publications.

Kratzer, Angelika. 1999a. Lecture notes, seminar on the semantics of clausal embedding. University of Massachusetts Amherst.

Kratzer, Angelika. 1999b. ‘Beyond *Ouch* and *Oops*: How descriptive and expressive meaning interact’. Ms., University of Massachusetts Amherst; presented at the Cornell Conference on Theories of Context Dependency.

Kratzer, Angelika. 2000. Lecture notes. University of Massachusetts Amherst.

Kratzer, Angelika. 2002a. *The Event Argument and the Semantics of Verbs*. Book manuscript.

Kratzer, Angelika. 2002b. ‘Telicity and the meaning of objective case’. In Guéron, Jacqueline and Jacqueline Lecarme, eds., *The Syntax of Time*. MIT Press, Cambridge, Massachusetts.

Krifka, Manfred. 1989. ‘Nominal reference, temporal consitution, and quantification in event semantics’. In Bartsch, Renate, Johan van Benthem, and Peter van Emde Boas, eds., *Semantics and Contextual Expression*. Foris Publications, Dordrecht.

Krifka, Manfred. 1998. ‘The origins of telicity’. In Rothstein, Susan, ed., *Events and Grammar*. Kluwer, London.

Kural, Murat. 1996. ‘Two types of bare measure phrases’. In *Proceedings of the Western Conference on Linguistics (WECOL) 1996.*

Laenzlinger, Christopher. 2000. ‘French adjective ordering: Perspectives on DP-internal movement types’. In *Generative Grammar in Geneva*, vol. 1.

Landman, Fred. 1989. ‘Groups I’. *Linguistics and Philosophy* 12:559.

Larson, Richard. 1985. ‘Bare-NP adverbs’. *Linguistic Inquiry* 16:595.

Larson, Richard. 1999. ‘Semantics of adjectival modification’. Lecture notes, LOT Winter School, Amsterdam.

Lasersohn, Peter. 1998. ‘Events in the semantics of collectivizing adverbials’. In Rothstein, Susan, ed., *Events and Grammar*. Kluwer Academic Publishers, Dordrecht.

Lasersohn, Peter. 1999. ‘Pragmatic halos’. *Language* 75.

Lasnik, Howard and Nicholas Sobin. 2000. ‘The who/whom puzzle: On the preservation of an archaic feature’. *Natural Language and Linguistic Theory* 18:343.

Lewis, David. 1981. ‘Ordering semantics and a precise semantics for counterfactuals’. *Journal of Philosophical Logic* 10:217.

Link, Godehard. 1983. ‘The logical analysis of plurals and mass terms: A lattice-theoretical approach’. In Bauerle, et al, ed., *Meaning, Use, and Interpretation of Language*. DeGruyter, Berlin.

MacDonald, Lorna. 1990. *A Grammar of Tauya*. Mouton de Gruyter, Berlin.

Matthewson, Lisa. 1996. *Determiner Systems and Quantificational Strategies: Evidence from Salish*. Doctoral dissertation, University of British Columbia. Published in 1998 by Holland Academic Graphics, The Hague.

Matthewson, Lisa. 1999. ‘On the interpretation of wide-scope indefinites’. *Natural Language Semantics* 7(1):79.

Matthewson, Lisa. 2001. ‘Quantification and the nature of cross-linguistic variation’. *Natural Language Semantics* 9:145.

McCawley, James D. 1972. ‘Kac and Shibtani on the grammar of killing’. In *Syntax and Semantics I*. Academic Press, New York.

McCawley, James D. 1988. ‘Adverbial NPs: Bare or clad in see-through garb?’ *Language* 64.

McCawley, Noriko. 1973. ‘Boy, is syntax easy!’ In Corum, C., T. Cedric Smith-Stark, and Ann Weiser, eds., *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*. CLS.

McConnell-Ginet, Sally. 1982. ‘Adverbs and logical form: A linguistically realistic theory’. *Language* 58:144.

McNally, Louise and Gemma Boleda Torrent. 2003. ‘Relational adjectives as properties of kinds’. In *Proceedings of the Colloque de Syntaxe et Sémantique à Paris (CSSP 2003)*.

Merchant, Jason. 1996. ‘Object scrambling and quantifier float in German’. In Kusumoto, Kiyomi, ed., *Proceedings of NELS 26*. GLSA, University of Massachusetts, Amherst.

Michealis, Laura and Knud Lambrecht. 1996. *Conceptual Structure, Discourse, and Language*. CSLI, Stanford.

Milsark, Gary L. 1976. *Existential Sentences in English*. Doctoral dissertation, MIT.

Mitchell, Erika. 1991. ‘Case and the Finnish object’. In *Cornell Working Papers in Linguistics*, vol. 9. Ithaca, New York.

Mithun, Marianne and Elisabeth Ali. 1996. ‘The elaboration of aspectual categories: Central Alaskan Yup’ik’. *Folia Linguistica* 30:111.

Moltmann, Friederike. 1991. ‘Measure adverbials’. *Linguistics and Philosophy* 14:6.

Moltmann, Friederike. 1997. *Parts and Wholes in Semantics*. Oxford University Press, New York.

Moortgat, Michael. 1996. ‘Categorial type logics’. In *Handbook of Logic and Language*. The MIT Press, Cambridge, Massachusetts.

Morgan, Jerry. 1969. ‘On arguing about semantics’. In *Papers in Linguistics*, vol. 1.

Morrill, G. 1994. *Type Logical Grammar: Categorial Logic of Signs*. Kluwer Academic Publishers, Dordrecht.

Morzycki, Marcin. 2001a. '*Almost* and its kin, across categories'. In Jackson, Brendan, Rachel Hastings, and Zsofia Zvolenszky, eds., *Proceedings from Semantics and Linguistic Theory XI*. CLC Publications, Ithaca, New York.

Morzycki, Marcin. 2001b. 'Interpreting measure DP adverbials'. In Megerdoomian, Karine and Leora A. Bar-el, eds., *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL) XX*. Cascadilla Publications, Somerville, Massachusetts.

Morzycki, Marcin. 2002. 'Wholes and their covers'. In Jackson, Brendan, ed., *Proceedings from Semantics and Linguistic Theory XII*. CLC Publications, Ithaca, New York.

Nilsen, Øystein. 2001. 'Adverb order in type logical grammar'. In van Rooy, Robert and Martin Stokhof, eds., *Proceedings of the Amsterdam Colloquium*.

Nilsen, Øystein. 2003. *Eliminating Positions: Syntax and Semantics of Sentence Modification*. Doctoral dissertation, Universiteit Utrecht.

Nunberg, Geoffrey. 1995. 'Transfers of meaning'. *Journal of Semantics* 12:109.

Parsons, Terence. 1990. *Events in the Semantics of English: A Study in Subatomic Semantics*. MIT Press.

Partee, Barbara. 1986. '*Any*, *Almost*, and superlatives (the airport squib)'. Ms, University of Massachusetts Amherst.

Partee, Barbara H. 1987. 'Noun phrase interpretation and type-shifting principles'. In Groenendijk, Jeroen, Dick de Jongh, and Martin Stokhof, eds., *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Foris Publications, Dordrecht.

Partee, Barbra H. 2001. 'Privative adjectives: Subsectives plus coercion'. Ms, University of Massachusetts Amherst. To appear in T. E. Zimmerman, ed. *Studies in Presupposition*.

Paxton, Norman. 1986. *German Grammar: A Modern Guide*. Hodder Arnold.

Pereltsvaig, Asya. 2000. 'On accusative adverbials in Russian and Finnish'. In *ZAS Papers in Linguistics 17*.

Pollock, Jean-Yves. 1989. ‘Verb movement, universal grammar, and the structure of IP’. *Linguistic Inquiry* 20(3).

Potts, Christopher. 2002. ‘No vacuous quantification constraints in syntax’. In *Proceedings of NELS 32*.

Potts, Christopher. 2003. *The Logic of Conventional Implicatures*. Doctoral dissertation, University of California, Santa Cruz.

Quine, W.V.O. 1960. *Word and Object*. MIT Press, Cambridge, Massachusetts.

Rapp, Irene and Arnim von Stechow. 1999. ‘*Fast* ‘almost’ and the visibility parameter for D-adverbs’. *Journal of Semantics* 16.

Rawlins, Kyle. 2003. ‘A study in some adverb denotations’. B.A. honors thesis, University of Massachusetts Amherst.

Reinhart, Tanya. 1997. ‘*Wh*-in-situ in the framework of the minimalist program’. *Natural Language Semantics* 6(1):29.

Rizzi, Luigi. 1997. ‘The fine structure of the left periphery’. In *Elements of Grammar*. Kluwer, Dordrecht.

Roeper, Thomas, Elaine Ramos, Harry Seymour, and Lamya Abdul-karim. 2000. ‘Language disorders as a window on universal grammar: An abstract theory of agreement for IP, DP, and V-PP’. Ms, University of Massachusetts Amherst.

Rotstein, Carmen and Yoad Winter. 2001. ‘Partial adjectives vs. total adjectives: Scale structure adn higher-order modification’. In *Proceedings of the Amsterdam Colloquium*.

Rullman, Hotze. 1995. *Maximality in the Semantics of Wh-Constructions*. Doctoral dissertation, University of Massachusetts.

Runner, Jeff. 1995. *Noun Phrase Licensing and Interpretation*. Doctoral dissertation, University of Massachusetts Amherst.

Russell, Bertrand. 1905. ‘On denoting’. *Mind* 14:479.

Sadock, Jerry. 1981. ‘*Almost*’. In Cole, Peter, ed., *Radical Pragmatics*. Academic Press, New York.

Sauerland, Uli. 2002. ‘The present tense is vacuous’. *Snippets* 6.

Schlenker, Philippe. 2002a. ‘Conditionals as definite descriptions: A referential analysis’. *Journal of Language & Computation* Special issue on ‘Choice Functions in Natural Language’, Ruth Kempson and Klaus von Heusinger, eds.

Schlenker, Phillipe. 2002b. ‘Sequence phenomena and double access readings generalized: Two remarks on tense, person, and mood’. Ms., UCLA and Institut Jean-Nicod.

Schwarzschild, Roger. 1996. *Pluralities*. Kluwer Academic Publishers, Dordrecht.

Schwarzschild, Roger and Karina Wilkinson. 2002. ‘Quantifiers in comparatives: A semantics of degree based on intervals’. *Natural Language Semantics* 10(1):1.

Seuren, Pieter A.M. 1973. ‘The comparative’. In *Generative Grammar in Europe*. Reidel, Dordrecht.

Sevi, Aldo. 1998. *A Semantics for* Almost *and* Barely. Master’s thesis, Tel-Aviv University.

Shlonsky, Ur. 1991. ‘Quantifiers as functional heads: A study of quantifier float in Hebrew’. *Lingua* 84:159.

Sobin, Nicholas. 1997. ‘Agreement, default rules, and grammatical viruses’. *Linguistic Inquiry* 28(2):318.

Sproat, Richard and Chilin Shih. 1988. ‘Prenominal adjectival ordering in English and Manderin’. In Blevins, J. and J. Carter, eds., *Proceedings of NELS 18*. University of Massachusetts GLSA, Amherst.

Stowell, Tim. 1981. *Origins of Phrase Structure*. Doctoral dissertation, MIT.

Thomason, Richard and Robert C. Stalnaker. 1973. ‘A semantic theory of adverbs’. *Linguistic Inquiry* 4:195.

Travis, Lisa. 1988. ‘The syntax of adverbs’. In *McGill Working Papers in Linguistics*.

Valois, D. 1991. *The Internal Syntax of DP*. Doctoral dissertation, UCLA.

von Fintel, Kai. 1992–1993. ‘Exceptive constructions’. *Natural Language Semantics* 1(2):123.

von Fintel, Kai. 1994. *Restrictions on Quantifier Domains*. Doctoral dissertation, University of Massachusetts Amherst.

von Fintel, Kai. 1995. 'The formal semantics of grammaticalization'. In Beckman, Jill, ed., *Proceedings of NELS 25*. GLSA, Amherst, Massachusetts.

von Fintel, Kai. 1997. 'The presupposition of subjunctive conditionals'. In Percus, Orin and Uli Sauerland, eds., *MIT Working Papers in Linguistics 25*.

von Stechow, Arnim. 1984. 'Comparing semantic theories of comparison'. *Journal of Semantics* 3:1.

Wechsler, Stephen and Yae-Sheik Lee. 1996. 'The domain of direct case assignment'. *Natural Language and Linguistic Theory* 14:629.

Westerståhl, Dag. 1985. 'Determiners and context sets'. In van Bentham, Johan and Alice ter Meulen, eds., *Generalized Quantifiers in Natural Language*. Foris, Dordrecht.

Winter, Yoad. 1997. 'Choice functions and the scopal semantics of indefinites'. *Linguistics and Philosophy* 20(399–467).

Winter, Yoad. 2001. 'Measure phrase modification in vector space semantics'. In Megerdoomian, Karine and Leora A. Bar-el, eds., *Proceedings of WCCFL XX*. Cascadilla Publications, Somerville, Massachusetts.

Wunderlich, Dieter. 1997. 'Argument extension by lexical adjunction'. *Journal of Semantics* 14:95.

Wyner, Adam. 1994. *Boolean Event Lattices and Thematic Roles in the Syntax and Semantics of Adverbial Modification*. Doctoral dissertation, Cornell University.

Wyner, Adam. 1998. 'Subject-oriented adverbs are thematically dependent'. In Rothstein, Susan, ed., *Events in Grammar*. Dordrecht: Kluwer.

Yoon, Yoongeun. 1996. 'Total and partial predicates and the weak and strong interpretations'. *Natural Language Semantics* 4(217-236).

Zanuttini, Raffaella and Paul Portner. 2003. 'Exclamative clauses: At the syntax-semantics interface'. *Language* 79(1):39.

Zimmerman, Malte. 2000. ‘Pluractional quantifiers: The *Occasional* construction in English and German’. In Jackson, B. and T. Matthews, eds., *Proceedings of SALT X*. CLC Publications, Ithaca, New York.

Zwarts, Joost. 1997. ‘Vectors as relative positions: A compositional semantics of modified PPs’. *Journal of Semantics* 14:57.

Zwarts, Joost and Yoad Winter. 2000. ‘Vector space semantics: A modeltheoretic analysis of locative prepositions’. *Journal of Logic, Language and Information* 9:171.